Singapore Management University
Institutional Knowledge at Singapore Management University

Research Collection Lee Kong Chian School Of Business

Lee Kong Chian School of Business

12-2003

To Pester or Leave Alone: Lifetime Value Maximization through Optimal Communication Timing

Xavier Dreze
Wharton School, University of Pennsylvania

Andre Bonfrer
Singapore Management University, andrebonfrer@smu.edu.sg

Follow this and additional works at: https://ink.library.smu.edu.sg/lkcsb_research

Part of the Marketing Commons

To Pester or Leave Alone:

Lifetime Value Maximization through Optimal Communication Timing

January 3, 2002

Xavier Drèze
University of California, Los Angeles
110 Westwood Plaza, Box 951481
Los Angeles, CA 90095-1481
(310) 206-0416
(310) 206-7422 Fax
xavier.dreze@anderson.ucla.edu

André Bonfrer
Singapore Management University
Andrebonfrer@smu.edu.sg

To Pester or Leave Alone:

Lifetime Value Maximization through Optimal Communication Timing

Abstract

The marketing literature has long acknowledged the importance of a customer's *lifetime value* in customer relationship management. More recently, researchers have turned their attention to the links between satisfaction and both customer *acquisition* and *retention* strategies. In this paper, we are interested in understanding the impact of communication *frequency* on customer retention and ultimately on lifetime value. We develop a theoretical framework for managing a customer database and addressing the tradeoffs between *value extraction* and *customer retention*. An empirical application of this framework is conducted for permission-based email marketing in the entertainment industry. This application recognizes the customization ability underlying one-to-one marketing, and yields decision rules for how a firm should interact with individual customers.

We find that *inter-communication* time has a dramatic impact on customer behavior. It affects both attrition and consumer surplus and thus has a critical impact on the value of a customer database. This impact is *asymmetric*, thus managers are advised to err on the side of longer rather than shorter inter-communication times. We further find that *retention* raises the value of a customer database in two ways. First, one can only derive revenue from customers who are active. Second, we demonstrate that the larger a firm's customer base, the lower its per-customer contact costs. These findings are supported by our empirical analysis.

Keywords: Customer relationship management, database marketing, customer lifetime value, customer retention.

1 Introduction

An underlying theme of most direct marketing research is that firms are trying to maximize the *Customer Lifetime Value* (Gupta, Lehmann, and Stuart 2001; Blattberg and Deighton 1991; Berger and Nasr 1998) of the names contained in their databases. While much of the recent research provides marketers with powerful analytical tools for selecting customers to receive a (direct) marketing communication (e.g. Bult and Wansbeek 1995, Gönül and Shi 1998), there has been little attempt made to treat such marketing activities endogenously when calculating (maximizing) the value of customer names. Consequently, the primary research objective of this paper is to examine the importance of the frequency of contact between a marketer and a customer, and show how a sub-optimal (too much or too little contact) communication frequency impacts the value of a customer database.

When attempting to maximize the customer lifetime value the extant research has taken the frequency of contact as a given. This is a natural thing to do when one is concerned with catalogs or newsletters as these vehicles have a natural periodicity that marketers adhere to. For instance, fashion catalogs are sent based on seasons (Winter, Spring, Summer, and Fall); newsletters are sent on a monthly or weekly basis.

However, given the recent move towards personalization and customization, we believe that contact frequency must be made endogenous when trying to maximize the value of a customer's name. Hence, the model we propose extends the current body of research on name value maximization (in the spirit of Bult and Wansbeek 1995, or Berger and Nasr 1998) by incorporating the decision of how often a firm should contact the members of its database. When doing so, it is crucial to take customer attrition into

account. Indeed, trying to maximize revenues might lead a firm to contact customers with such a high frequency that they decide to sever their relationship with the company. This would in effect drive the lifetime value to zero; a counter-productive effort! Similarly, as one starts manipulating the frequency of contacts, one will affect the customer response function. Indeed, consider the *Book of the Month Club*. One would not expect the response rate (i.e., proportion of members who purchase any given book offered) to be the same if the club became the *Book of the Year*, or the *Book of the Day*?

When making the contact frequency endogenous, we find that it has a critical impact on the value of the names held in a company's database. We find that *excessive* marketing communications lower the value of customer names through the loss of future revenue due to customer defection. In contrast, if a firm does not communicate with customers frequently enough, the firm loses out on opportunities to make sufficient money on these customer names. This latter effect is amplified by the fact that future earnings are less valuable then current ones.

2 Background and Motivation

There are two principal components to the lifetime valuation of a customer: the duration of the relationship and the value of each customer-firm interaction. In terms of the first component, duration, it has long been recognized that firms benefit more from maintaining long-term than from short-term customer relationships (e.g., see Bendapudi and Berry, 1997). There appears to be considerable anecdotal evidence on the value of these long-term customer relationships (it is often said that it is cheaper to keep a customer than to get a new one), but academic research has generated very few generalizable empirical results that can substantiate this hypothesized rising profitability of long term customers (except

2

for Reichfeld and Teal, 1996). In a recent and well-cited *Sloan Management Review* article, Dowling and Uncles (1997) challenge customer loyalty programs, and caution that long life (or loyal) customers are not necessarily more profitable. Reinartz and Kumar (2000) provide empirical evidence to support Dowling and Uncles contention.

In response to this emerging understanding of the role of customer relationship management, researchers have started studying customer retention (e.g., Thomas, 2001; Mittal and Kamakura, 2001; Bolton, 1998; Smith, Bolton, and Wagner, 1999). To date, this research has focused almost exclusively on collecting customer satisfaction data (mainly in service industries) to understand customer defection and the revenue streams arising from active customers. This emerging literature is beginning to recognize that there are key levers that may be used to influence the lifetime value of customers, for example, by raising customer satisfaction and thereby lowering defection rates, or by raising individual-level purchases.

The second component of customer lifetime value, the value of each customer-firm interaction, has also been the subject of academic inquiry. Bult and Wansbeek (1995) propose to optimally select the target customer list for a mailing to maximize the profitability of each send. This is done by equating the marginal cost of sending the catalogs to the marginal expected revenue from the list. Bitran and Mondschein (1996) study list selection decisions in an environment where budget constraints force the company to divide resources between sending costs and inventory costs. Further, Gönül and Shi (1998) make consumer response endogenous by linking the purchase decision to past actions. When maximizing expected revenues, these papers do not explicitly study the

duration of the relationship. They assume away retention; they also take the frequency of mailing as exogenous to the problem.

In this paper, we bring together these two research steams (duration and value maximization) by making the timing decision (i.e., how often should a firm contact its customers) and its impact on both retention and consumer response endogenous. To see why this is important, one can conduct the following thought experiment: look at the two extremes in customer relationship management. At one extreme, a firm contacts its clients so often that the clients sever their links to the company, and thus their names are of no value. At the other extreme, the firm never contacts its clients, and although the names have a *potential* value, this value is never *realized*. Clearly, there must be an intermediary situation where one maximizes the realized value from a name by optimally trading off the extraction of value in the present with the loss of future value due to customer defection. The framework presented in this study can be used to examine this tradeoff between short-term profitability and customer retention. In so doing, this study extends what we know about customer lifetime value (e.g., Bult and Wansbeek, 1995; Berger and Nasr, 1998), and customer relationship management (e.g. Sheth and Parvatiyar, 1995; Morgan and Hunt, 1994). This framework also yields practically applicable decision rules that will help marketers manage their customer databases at a micro level (micromarketing). In particular, we show how consumers respond to varying frequencies of communication by marketers, and how this then impacts the lifetime profitability of the focal consumer.

The next section develops the conceptual model used to study the timing decision and shows how this decision affects the overall value of a customer's name through customer retention and customer response. In sections four and five, we present an

4

empirical application of this framework in the context of sending email communication to a database of customer names. The final sections discuss our findings and managerial implications.

3 Mathematical Development

3.1 General Name Value Formulation

The starting point of our analysis involves defining an expression for the expected returns, generated by marketing activities, of a customer name held in the firm's database. We then calculate the net present value of this customer name, by summing the discounted returns over all marketing activities. Such marketing activities could conceivably encompass any of the marketing mix instruments (price, distribution, communication, product offers). However, for the purpose of this study, we focus only on the marketing communications[1] efforts of the firm.

To be more specific, we are interested in the timing of a sequence of marketing communications and customers' reactions to these communications. These customer reactions can take on multiple forms: the consumer may ignore the message, interact with the firm (to gain more information and/or to purchase), or even ask the firm not to contact him/her anymore. Given the set of potential reactions, the firm needs to determine the optimal timing of a sequence of communications so as to maximize the net present value of

---

[1] By marketing communication, we refer to any direct contact by the firm that might lead the consumer to purchase the product or service sold by the firm (e.g., telephone, email, catalog, direct mail, sales calls).

its income stream. In mathematical terms the firm is trying to maximize the name value of a representative customer[2]:

$$V = \sum_{i=0}^{\infty} \frac{1}{(1+r)^{t_i}} A_i P_i .$$ (1)

where V, the value of a name, is defined (in the spirit of Berger and Nasr, 1998) as the discounted sum of net surpluses over a sequence of marketing activities and (see Table 1 for a list of the variables used in this paper):

t_i is the time at which communication i is made (expressed in fractional years for accounting purposes),

r is the cost of money,

$\dfrac{1}{(1+r)^{t_i}}$ is the discount rate of money,

A_i is the net surplus expected from the ith communication (revenues minus the cost of goods sold, and promotional expenditures),

P_i is the probability that we retain a customer for the ith communications.

If the firm is sending its communications at a fixed time interval (τ) we can set $t_i = i \cdot \tau$. Further, we assume that all communications are similar in nature (as in Gönül and Shi, 1998), such that $A_i = A(\tau)$ and $P_i = p(\tau)^i$. Thus, we can rewrite (1) as:

$$V(\tau) = \sum_{i=0}^{\infty} \frac{p(\tau)^i}{(1+r)^{i\tau}} A(\tau).$$ (2)

---

[2] In the sense that all consumers are, in this abstract context, identical. To allow for heterogeneous consumers requires an integration of (1) over consumer types.

We can further rewrite this infinite series as (see Appendix 1 for details):

$$V(\tau) = \frac{(1+r)^\tau}{(1+r)^\tau - p(\tau)} A(\tau).$$

(3)

To say something about $V(\tau)$, it is necessary to study how τ affects $A(\tau)$ and

$p(\tau)$. For instance, in the case where shorter inter-contact times increases the likelihood

of defection (i.e. $\partial p / \partial \tau > 0$), then more frequent communications (smaller τ) reduces the

database faster, and we expect the optimal τ to increase. Similarly, if decision makers are

faced with a higher discount rate, the value of revenues earned from future

communications is lower, and therefore we expect the optimal τ to be lower. Hence we

need to look at the impact of τ (through r, $p(\tau)$, and $A(\tau)$) on $V(\tau)$.

The problem of the firm is to maximize the lifetime value of its database with

respect to τ, or:

$$\max_\tau V(\tau) = \max_\tau \frac{(1+r)^\tau}{(1+r)^\tau - p(\tau)} A(\tau).$$

(4)

Looking at the first order condition, the maximum value of V occurs where:

$$\frac{\partial V(\tau)}{\partial \tau} = \frac{\partial A(\tau)}{\partial \tau}\left((1+r)^\tau - p(\tau)\right) - A(\tau)\left(p(\tau)\ln(1+r) - \frac{\partial p(\tau)}{\partial \tau}\right) = 0.$$

(5)

We can re-express this first order condition as:

$$\frac{\dfrac{\partial A(\tau)}{\partial \tau}}{A}\left(\frac{(1+r)^\tau}{p(\tau)} - 1\right) = \ln(1+r) - \frac{\dfrac{\partial p(\tau)}{\partial \tau}}{p(\tau)}$$

$$\Rightarrow \frac{\dfrac{\partial A(\tau)}{\partial \tau}}{A}\left(\frac{(1+r)^\tau}{p(\tau)} - 1\right) = \frac{\dfrac{\partial (1+r)^\tau}{\partial \tau}}{(1+r)^\tau} - \frac{\dfrac{\partial p(\tau)}{\partial \tau}}{p(\tau)}.$$

(6)

Multiplying each side of (6) by τ and taking advantage of the fact that the time elasticity (ε_X) of X with respect to τ is $\dfrac{\partial X}{\partial \tau}\dfrac{\tau}{X}$ we can further re-express the first order condition as:

$$\varepsilon_A \left(\frac{(1+r)^\tau}{p(\tau)} - 1 \right) = \varepsilon_{(1+r)^\tau} - \varepsilon_p.$$ (7)

To solve (7), it is necessary to understand the behavior of ε_p, ε_A, and $\varepsilon_{(1+r)^\tau}$. These are discussed in detail in the following subsections.

3.2   Retention Elasticity (ε_p)

What causes someone to leave a database? A strong argument can be made that attrition is linked to two factors: message content and message frequency. In terms of message content, we expect that the better the content, the higher the retention. However, as content is assumed to be constant across campaigns (in quality and form if not in actual message), the effect of content will be held constant through our analysis and thus can be ignored.

In terms of inter-communication time, it is reasonable to expect that as τ nears zero, the probability of the customer leaving the firm approaches unity. Indeed, if one were to receive incessant phone calls, it is very likely that one would make efforts to cease receiving such contact, and request one's removal from the database. Hence, the retention probability is bound by zero (as $\tau \to 0$) and the functional form for $p(\tau)$ has to be initially increasing in τ (i.e., $\varepsilon_p > 0$). Further, one can easily imagine that as τ becomes very

8

large, one might see a reversal where $\varepsilon_p < 0$ (the *Book of the Year Club* may be of dubious interest).

In which situation would one see an optimal τ such that $\varepsilon_p < 0$? This is an a priori strange situation as a shorter τ then leads to a higher retention rate. Equation (7) implies that the only situation in which this can arise is if $\varepsilon_A > 0$. Indeed, since $\ln(1+r) > 0$, we have $\varepsilon_{(1+r)^\tau} > 0$. In other words, the firm is willing to endure a higher attrition rate if spacing its communications longer increases the size of the surplus that can be extracted from each communication.

When ε_p is positive, spacing out the communications (i.e., increasing τ) increases the firm's retention rate. Waiting longer is better from a database standpoint, but how long should one wait? This depends on the sign of ε_A. We examine three cases:

(i) $\varepsilon_A = 0$

If $\varepsilon_A = 0$, then an increase or decrease in the inter-contact time does not affect the size of the surplus extracted from each communication. In this case, τ^* will be at the point where the benefit of waiting in terms of retention rate is equal to the cost of waiting in terms of the discount rate of money ($\varepsilon_p = \varepsilon_{(1+r)^\tau}$).

(ii) $\varepsilon_A > 0$

If $\varepsilon_A > 0$, then $\varepsilon_p = \varepsilon_{(1+r)^\tau} - \varepsilon_A \left[\dfrac{(1+r)^\tau}{p(\tau)} - 1 \right] = \varepsilon_{(1+r)^\tau} - \phi, \phi > 0$. This means that, compared with (i), the firm is less willing to lose people, because these people become more valuable.

9

(iii) $\varepsilon_A < 0$

If $\varepsilon_A < 0$, then $\varepsilon_p = \varepsilon_{(1+r)^\tau} + \phi, \ \phi > 0$. This means that, compared with (i), the firm is willing to lose more people, because if it tried to keep more people, these people would become less valuable, and thus less worth waiting for.

The cases above suggest that it is important for a manager to know the sign of ε_A. For (i) it is optimal to choose the value of τ that equates the percentage change in the discount rate to the percentage change in the retention rate. In this case, since A does not vary with τ, there is no need to worry about how much the time taken to send the next communication affects the expected value of that customer. Everything that depends on τ, now depends only on the trade-off between the declining value of money the more the next communication is delayed (by τ), and the rising probability that we retain people in the database. The optimal point for τ occurs where the percentage change in discount rate just equals the percentage change in the retention. Since these two percentages affect $V(\tau)$ on the same basis, but in opposite directions, this result is very intuitive.

The second and third cases (ii and iii) deal with the situation where A is not independent of τ, and highlight the importance of understanding the sign of $\dfrac{\partial A}{\partial \tau}$. The tradeoff now depends on how A varies with τ. If waiting longer between communications increases the return from a given communication ($\varepsilon_A > 0$), then spacing out communications generates a gain on two fronts (higher retention and higher transaction value) and there will be a tendency to increase τ. By contrast, if spacing out the communications decreases transaction size ($\varepsilon_A < 0$) then although longer inter-contact

time increases the retention rate, we lose on two fronts (lower transaction value, higher discount rate), and thus there will be a tendency to decrease τ .

The last possible case regarding ε_p is the situation in which $\varepsilon_p = 0$ at the optimum. This may arise if the retention probability is (locally) independent from τ . For this to be the case, we need not only for ε_p to be null, but we also need $\varepsilon_A > 0$, and in particular,

$$\varepsilon_A = \frac{\varepsilon_{(1+r)^\tau}}{\left(\dfrac{(1+r)^\tau}{p(\tau)} - 1 \right)} \quad \text{per equation (7).}$$

3.3 Surplus Elasticity (ε_A)

To understand the behavior of ε_A it is necessary to consider the components of $A(\tau)$. So far, the simplifying assumption that $A_i = A(\tau)$ has been made. This assumption is now relaxed and we describe A_i in more detail. The value for A_i is the net surplus extracted during each contact (expected revenues - cost). It has two primary components. The first component is the expected revenue (B_i) to the firm from the reaction of the consumer with the marketing effort of the firm. The second component is the cost (C_i) incurred by the firm when sending the ith marketing communication. The net surplus is the difference between the two. For communication i, we rewrite A_i in (1) as:

$$A_i = B_i - C_i \tag{8}$$

and

$$\frac{\partial A_i}{\partial \tau} = \frac{\partial B_i}{\partial \tau} - \frac{\partial C_i}{\partial \tau} .$$

11

What is the behavior of $\dfrac{\partial B_i}{\partial \tau}$? We previously assumed that each message was identical in nature, but not necessarily in actual content. We will further assume that each message is self-contained (i.e., one can evaluate each message independently from the others) and that the process is memory-less in terms of which messages have been received in the past. The only factor that makes a customer more or less likely to respond to a message is the frequency of communication. Hence, we write $B_i = f(Z, \tau)$, where Z is a vector of consumer level characteristics that moderate consumer response to the type of messages received. Since customer heterogeneity is not taken into account in this theoretical development, we restrict ourselves to $B_i = B(\tau)$.

Conventional wisdom, and the basis for RFM models, is that $\dfrac{\partial B}{\partial \tau} < 0$. This means that there will be a natural tendency to decrease τ in order to maximize $B(\tau)$.

What about $\dfrac{\partial C_i}{\partial \tau}$? There are three types of costs associated with extracting value from a database of names. There are fixed operating costs linked to the infrastructure of the business (e.g., SG&A, database administration and maintenance). In addition, each marketing campaign has a fixed cost component (F) as well as a variable cost component. The fixed cost of the campaign includes any setup cost associated with the campaign, such as the cost of hiring a designer to generate artwork, pre-launch copy testing, and so on. These costs are fixed for any given campaign in that they are independent from the number of recipients of the campaigns, but they are variable to the company in that a lower τ implies that the firm deploys more campaigns in any given period of time and thus incurs more cost. Henceforth, we refer to this class of fixed costs as semi-variable costs. The

12

truly variable costs (s) are the costs associated with reaching each individual for each campaign, such as catalog printing costs, mailing costs, or telephone charges. In this study, we make both the semi-variable cost and the variable cost of sending endogenous; the true fixed costs are considered to be sunk and do not affect the name value and its maximization. Thus, $C_i = \dfrac{F}{N_i} + s$, where N_i is the size of the database when communication i is sent.

To further refine this expression of C_i, let us say that there are N people in the database at $i=0$ and that the firm can attract v% new registrants every year. Taking both customer attrition and acquisition into account, the per campaign average sending cost becomes:

$$C_i = \frac{F}{Np(\tau)^i (1+v)^{i\tau}} + s \tag{9}$$

Let us differentiate (9) with respect to i to see how the contact costs evolve over time:

$$\Delta C_i = C_i - C_{i-1} = \frac{F}{Np(\tau)^i (1+v)^{i\tau}} + s - \left(\frac{F}{Np(\tau)^{i-1}(1+v)^{(i-1)\tau}} + s \right)$$

$$\Rightarrow \Delta C_i = \frac{F}{Np(\tau)^{i-1}(1+v)^{(i-1)\tau}} \left(\frac{1}{p(\tau)(1+v)^\tau} - 1 \right) \tag{10}$$

Since the first term of (10) is always positive, it follows that the cost of sending communications will decrease from one communication to another if:

$$p(\tau)(1+v)^\tau > 1. \tag{11}$$

In other words, if the inter-communication growth rate is large enough to compensate for the attrition rate, the database grows over time, and thus the cost of

13

contacts decrease concurrently. A corollary of this is that as τ decreases, $(1+v)^\tau$ also decreases and it becomes more difficult for (11) to be satisfied. This result gives us the basis for the intuition that it is not profitable in the long run to spam (send a very high frequency).

To further develop the relationship between C_i and τ, let us look at $\dfrac{\partial C_i}{\partial \tau}$:

$$\frac{\partial C_i}{\partial \tau} = \frac{-Fi}{N(1+v)^{i\tau}}\left(\frac{\ln(1+v)}{p(\tau)^i} + \frac{\frac{\partial p(\tau)}{\partial \tau}}{p(\tau)^{i+1}} \right) \tag{12}$$

$$\Rightarrow \frac{\partial C_i}{\partial \tau} = \frac{-Fi}{Np(\tau)^i(1+v)^{i\tau}}\left(\ln(1+v) + \frac{\frac{\partial p(\tau)}{\partial \tau}}{p} \right)$$

$$\Rightarrow \frac{\partial C_i}{\partial \tau} = \frac{-Fi}{\tau Np(\tau)^i(1+v)^{i\tau}}\left(\varepsilon_{(1+v)^\tau} + \varepsilon_p \right). \tag{13}$$

All elements of (13) are positive except possibly for ε_p. Thus the sign of $\dfrac{\partial C_i}{\partial \tau}$ depends on the relationship between ε_p and $\varepsilon_{(1+v)^\tau}$. For all $\varepsilon_p > 0$, $\dfrac{\partial C_i}{\partial \tau}$ is negative (i.e., spacing the communication lowers the average per name cost of communication). It is only in the extreme case where $\varepsilon_p < -\varepsilon_{(1+v)^\tau}$ (i.e., by spacing communications further we lose more people through attrition than can be replaced through acquisition) such that $\dfrac{\partial C_i}{\partial \tau}$ becomes positive. Thus we expect to have two counter-balancing forces working within $\dfrac{\partial A_i}{\partial \tau}$ (recall

14

$\frac{\partial A_i}{\partial \tau} = \frac{\partial B_i}{\partial \tau} - \frac{\partial C_i}{\partial \tau}$). With $\frac{\partial B_i}{\partial \tau} < 0$ and $\frac{\partial C_i}{\partial \tau} < 0$, the net effect on $\frac{\partial A_i}{\partial \tau}$ and thus ε_A will

depend on whether the net growth rate can compensate for any loss in revenue due to lower

communication frequency.

The above analysis reveals that a crucial component in optimizing the value of

customer's name is the *growth rate* of a firm's database. This may seem counter-intuitive

at first. Why should a firm's current valuation of a customer's name depend on how many

new customers it will acquire in the future? As we have just shown, the size of the

database, and its growth, come into play through the semi-variable costs. It is then worth

spending a moment looking at the impact of the growth rate on the value of names in a bit

more detail. Merging (2), (8), and (9), the value of a name is now:

$$V = \sum_{i=0}^{\infty} \left(\frac{p(\tau)}{(1+r)^\tau} \right)^i (B(\tau) - C_i) = \sum_{i=0}^{\infty} \left(\frac{p(\tau)}{(1+r)^\tau} \right)^i B(\tau) - \sum_{i=0}^{\infty} \left(\frac{p(\tau)}{(1+r)^\tau} \right)^i \left(\frac{1}{p(\tau)^i (1+v)^{i\tau}} \frac{F}{N} + s \right)$$

$$\Rightarrow V = \sum_{i=0}^{\infty} \left(\frac{p(\tau)}{(1+r)^\tau} \right)^i (B(\tau) - s) - \sum_{i=0}^{\infty} \frac{1}{(1+r)^{i\tau} (1+v)^{i\tau}} \frac{F}{N}$$

$$\Rightarrow V = \frac{(1+r)^\tau}{(1+r)^\tau - p(\tau)} (B(\tau) - s) - \frac{(1+r)^\tau (1+v)^\tau}{(1+r)^\tau (1+v)^\tau - 1} \frac{F}{N}. \tag{14}$$

We see in (14) that variable and semi-variable costs are treated very differently.

The variable costs (s) are a straight offset to revenue and depreciate at the same rate, while

the semi-variable costs depreciate more or less depending on the growth rate (v) of the

database. Further, we have:

$$\frac{\partial V}{\partial v} = \frac{F}{N} \frac{(1+v)^{\tau-1} (1+r)^\tau \tau}{\left((1+v)^\tau (1+r)^\tau - 1 \right)^2} > 0. \tag{15}$$

Hence, the higher the acquisition rate the higher the value of the names in the database.

There are two special cases for v: $v = 0$ and $v = \dfrac{1}{p(\tau)}$. When $v = 0$, no new names are being acquired and thus the database shrinks over time at a rate determined by $p(\tau)$. The semi-variable cost allocation for each name now becomes:

$$\frac{(1+r)^\tau}{(1+r)^\tau - 1} \frac{F}{N}. \tag{16}$$

This is a subtle departure from our traditional multiplier of $\dfrac{(1+r)^\tau}{(1+r)^\tau - p(\tau)}$. Indeed, it means that, although attrition is taken into account when discounting future revenue, it is not taken into account when discounting semi-variable costs. At time 0, an allocation is made to each person for the cost of sending all future emails. Each message recipient is allocated an equal share of the costs since everyone is equally likely to defect. Further, as the database shrinks, so does the value of people in it as the semi-variable costs are reallocated to the survivors.

When $v = \dfrac{1}{p(\tau)}$, the defection rate is equal to the subscription rate. Hence, the database does not change in size over time. The semi-variable cost allocation becomes:

$$\frac{(1+r)^\tau \dfrac{1}{p(\tau)}}{(1+r)^\tau \dfrac{1}{p(\tau)} - 1} \frac{F}{N} = \frac{(1+r)^\tau}{(1+r)^\tau - p(\tau)} \frac{F}{N}. \tag{17}$$

Now the semi-variable cost allocation is treated on par with the other costs and revenues. Message recipients are allocated a cost that is reflective of how many communications they are expected to receive. Thus, the cost allocation depends on the retention rate. The

higher the retention rate, the higher the allocation and vice versa. This allocation does not change over time.

3.4   The Discount Elasticity ($\varepsilon_{(1+r)^\tau}$)

As r does not depend on τ, $\dfrac{\partial (1+r)^\tau}{\partial \tau}$ is always positive. Thus, the effect of $\varepsilon_{(1+r)^\tau}$ in (7) is a constant pressure to decrease τ (money earned now is worth more than money earned in the future). It can be seen as an intercept shift in an attempt to find an equilibrium between ε_p and ε_A. Nevertheless, different firms might face different discount rates. For instance, a pre-IPO firm may want to show a higher return, therefore increasing its valuation of r. This higher discount rate lowers the value of names since $\dfrac{\partial V}{\partial r} < 0$! Consequently, an increase in the discount rate makes the firm emphasize present rather than future earnings. And thus the firm might damage its future position by trying to satisfy unrealistic short-term goals!

3.5   Summary of Theoretical Findings

Before moving into the empirical illustration of the mathematical developments from the previous section, it is probably to summarize the findings up to this point. Starting with a general formulation of the lifetime value (3) we derived the first order condition that must be met to reach a maximum (7). The first order condition highlights, through ε_p, ε_A, and $\varepsilon_{(1+r)^\tau}$, the tension underlying the maximization of the value of names in a database. It shows that one *trades off* surplus (*A*) for retention (*p*). In other words, one trades off large

17

surpluses generated by each communication with surpluses generated from a large number of people.

This tradeoff is not as straightforward as one could imagine in that it is necessary to consider the *net growth* of the firm's database in order to optimize the tradeoff. Indeed, the net growth of the database (i.e., acquisition - attrition) dictates how the semi-variable costs (*F*) are allocated to each name. The higher the growth rate, the lower the allocation and thus the higher the surplus generated by each communication.

Combining conventional wisdom and our theory, we expect our empirical analysis to reveal that, at V^*, τ^* is such that $\frac{\partial B}{\partial \tau} < 0$, $\frac{\partial p}{\partial \tau} > 0$, and thus $\frac{\partial C_i}{\partial \tau} < 0$. More simply put: the database is growing in size; more surplus can be extracted from each communication via more frequent communications; but this would negatively affect the retention rate.

4   Empirical Analysis

In the previous section, we developed a general expression for the lifetime value of a database name. We started with a basic expression describing the income stream one might generate from such a name (equation (1)) and developed a framework that incorporates the benefits to the firm of customer interaction, as well as the likelihood of retaining that customer for future communications. We now describe an empirical analysis conducted to validate and illustrate the theories developed in this paper. In this section we discuss methods used to measure the components of the name value equation in (14). We do so in the context of marketing communications via permission-based email newsletters. Equation (14) gives us the final expression for the lifetime value of a name as a function of the inter-email time. It states that (refer back to table 1 for a list of the variables):

18

$$V(\tau) = \frac{(1+r)^\tau}{(1+r)^\tau - p(\tau)}(B(\tau) - s) - \frac{(1+r)^\tau(1+v)^\tau}{(1+r)^\tau(1+v)^\tau - 1}\frac{F}{N}. \tag{18}$$

Although $V(\tau)$ is not directly observed or estimated in our data, it is possible to estimate the components on the RHS of (18), then solve to obtain optimal τ. To estimate the parameters of (18) we use a database of emails sent by a large entertainment company to promote its products.

4.1 Description of the Data

The data used in the empirical application were provided by a large entertainment products company. The company is building an on-line distribution presence and expanding its communications efforts via electronic channels. A significant proportion of its on-line communication activities takes the form of email updates on newly released video titles, as well as promotions for existing titles. Data was collected for 62 separate email campaigns for a variety of different promotions. The overall open and click rates as well as other statistics about the dataset are reported in Table 2; a histogram of the inter-email time (τ, in weeks) observed in the data is shown in Figure 1.

Our observation series covers about five million emails that were sent over a period of thirteen months (from 08/08/2000 until 09/01/2001). Of the emails sent, 51.3% were "trackable." Each of the trackable emails included a code to identify which recipient opened the email and, if the email was opened, we were also able to observe if the user clicked on any link contained in the email.

Every email, whether it was trackable or not, included a link to a member center web page where users could remove themselves from the database to stop receiving email.

19

4.2 Base parameters r, F/N, v

Of all the parameters in (18) the discount rate of money (r), the variable cost of sending one email (s), the semi-variable cost of developing and sending an email campaign (F), the growth rate of the database (v), and the current size of the database (N) are defined exogenously and can be obtained from the company's accounting (or project management) records. The company that provided us with the data routinely uses a discount rate of 10% for its financial projections. We will thus use the same value for r.

There were no variable costs (s) incurred, as the company owned all the equipment used to send the emails. Semi-variable costs were limited to the costs of designing content, which the company outsourced to a series of design companies. To estimate F and N, we examine the cost of sending the campaigns included in our database and the number of emails sent in each one. By averaging costs over all campaigns, we reach an estimate of $5,000 for F per email campaign.

Finally, the growth rate of the database is estimated using the geometric mean of the monthly growth rate (i.e., number of new users in month i / number of active users at the beginning of month i). The annual growth rate is computed as $(1 + Average\ Rate)^{12} - 1$. This yields an annual growth rate, v, of 43.7%.

4.3 Email Message Value: B_τ

Of the 2,548,362 trackable emails that were sent, 354,449 were opened (i.e., Open Rate = 13.9%). Further, 51,262 of the emails generated a click-through (i.e., Click-Through Rate = 14.5%). This click-through rate based on the total number of emails sent is 2.01%, or almost 10 times larger than current click-through rates for banner advertising (0.28% for

September 2001 according to Iconocast, 2001) and reflects the relative effectiveness that

helps explain the current growth in the email industry. This high click-through rate can be

explained by the voluntary nature with which members belong to the newsletter. They

have asked to receive the email on a periodic basis. In contrast, advertising click-through

rates tend to be quite low, since in most cases, the recipient has not specifically requested

to receive such advertising messages, and more often than not will actively disregard

banner ads (Drèze and Hussherr, 1999).

To model the reaction by consumers to the email promotions, we assume that a

two-stage process takes place. In the first stage, the email is delivered to the mailbox, and

the recipient can choose to open the message (with probability p_o) or not (with probability

1-p_o). In the second stage, and conditional upon opening, the recipient can click on the

email (with probability p_c) or not (1- p_c). We associate a benefit to the firm of σ if the

user opens the email. Further, if the recipient of the message clicks on a link in the email,

we associate an additional return of x. Thus we write $B(\tau)$ as:

$$B(\tau) = p_o(\tau) * \left(\sigma + p_c(\tau) * x\right).^3 \tag{19}$$

Note that p_o and p_c are both dependent on τ. The value σ represents the expected

value of a customer who opens the email but does not respond to it. The value x represents

the additional value of a customer who clicks on the email. It is important to note that for

different firms, the relative magnitude of these two customer value metrics is likely to

vary. This is due to the likely varying content of the promotional messages. It also

---

[3] An alternative formulation is $B_\tau = p_o\left(\sigma(1 - p_c) + x_1 p_c\right)$, the two formulations are equivalent if we

have $x_1 = x + \sigma$.

depends on the objective of the firm. For a sales promotion, the objective of the firm is some direct call to action, in which case it is expected that $x \gg \sigma$. On the other hand, if the objective of the email communication is to raise awareness then this is likely to result in $x < \sigma$. For example, a campaign designed to generate awareness of a launch of a product but otherwise requires no direct response, is likely to have $x = 0$. An example of a campaign where $x > \sigma$ is one where a direct link to an on-line retailing site is included in the email content.

In our case, the company executives attach a relatively small value to an open without a click. They recognize that there will be value through heightened awareness, and that users might decide to buy the product at another online retailer (e.g., Amazon) or at an offline store (e.g., Blockbuster). Thus, they set σ at \$0.01[4].

Estimating x was made possible by looking at the profits generated by the various email campaigns. One should note that we are looking here at the purchases that are directly attributable to specific campaigns. That is, we only take into account purchases made at the company's online store when the user clicks, not at any purchases made on subsequent visits or at different retailers. From historical data, x is estimated at \$2.50.

The last pieces of the puzzle needed to complete our estimate for $B(\tau)$ are the open and click probabilities. As discussed earlier, these probabilities are likely to be dependent on τ. To model the relationship between inter-email time and both open and click probabilities, two binary logit regressions are fitted to the data. As we assume that the

---

[4] We have discussed the possibility to perform a post-email survey using a control group to measure the increase in purchase probability due to receive an email. Unfortunately, the company does not see this as a priority.

users do not know the content of the message before they open it, we model the click

probability as conditional on open and thus estimate two separate logit models rather than

one nested logit (statistical tests failed to support the need for a nested logit).

To account for the fact that consumer response is probably not linear in τ, we use

both τ and τ^2 as independent variables. In addition we use two sets of indicator variables

to help control for other observed factors that may account for changes in the response

probabilities. First, we use six indicator variables (W_1-W_6) to account for the day on

which the email was *sent*. Second, some of the database members (22%) have indicated

their gender when registering for the newsletter. As gender information exist for only a

portion of the database, we use two indicator variables (G_1-G_2) to account for the three

possible gender states (Male, Female, Unknown). To a limited extent, these control

variables allow us to account for heterogeneity in consumer responses.

The independent variables of our logit regressions are specified as follows:

$$X\beta_{p_O} = \gamma_{0p_O} + \alpha_{1p_O}\tau + \alpha_{2p_O}\tau^2 + \sum_{l=1}^{L_d}\gamma_{W_lp_O}W_l + \sum_{l=1}^{L_g}\gamma_{G_lp_O}G_l \tag{20}$$

$$X\beta_{p_C} = \gamma_{0p_C} + \alpha_{1p_C}\tau + \alpha_{2p_C}\tau^2 + \sum_{l=1}^{L_d}\gamma_{W_lp_C}W_l + \sum_{l=1}^{L_g}\gamma_{G_lp_C}G_l \;. \tag{21}$$

Where the α's and γ's are the parameters to be estimated. The regression results are

shown in Table 3 and will be discussed after the description of the retention probabilities.

4.4 *Retention Probability:* $p(\tau)$

172,498 of the 4,968,520 emails generated *unsubscribe* requests (i.e., retention rate =

96.5%). Just as we use a logit specification to model open and click rate, and consistent

with the premise that users remain subscribed to the newsletter as long as the expected utility is larger than some threshold, we use a logit specification to model retention rate. We use the same specification as (20) and (21) except that we want our model to account for the fact that if τ shrank to 0, the retention rate should also shrink to 0 (i.e., a user receiving an infinite amount of email will unsubscribe immediately). To that end, we include $1/\tau$ in the list of independent variables, and expect its coefficient to be negative (i.e., when $\tau \to 0, \frac{1}{\tau} \to \infty$ and $\frac{e^{-\infty}}{1+e^{-\infty}} \to 0$). Thus, the estimation of the retention probabilities is done using a logit regression with:

$$X\beta_{p_\tau} = \gamma_{0p_\tau} + \alpha_{0p_\tau}\frac{1}{\tau} + \alpha_{1p_\tau}\tau + \alpha_{2p_\tau}\tau^2 + \sum_{l=1}^{L_d}\gamma_{W_lp_\tau}W_l + \sum_{l=1}^{L_g}\gamma_{G_lp_\tau}G_l .$$ (22)

5 Empirical Results

5.1 Estimation of Retention, Open and Click Probabilities

Table 3 presents the estimation results for all three binary logit models (20), (21), and (22). The results generated by the estimation of these models are reported in the pairs of columns entitled "Retention", "Open" and "Click|Open". The rows report each of the individual covariates' parameter estimates and their standard errors. Note that the indicator variables were coded using a [1,-1] scheme. For instance, the two variables for Gender had value of (1,0) for Females, (0,1) for Males, and (-1,-1) if the gender of the user was not known.

      All variables are significant in all three models except W_2 (Monday) which is not significant in the probability of Open regression, and W_3 (Tuesday) which is not

significant in the probability of click regression. When looking at the effect of the day of the week indicators, it is important to not confuse this with the day of the week on which the reception of the message was recorded.   That is, the day of week variable effect is the effect on the probability of the action given that the email was sent on the focal day.

Caution must also be taken when interpreting the estimated coefficients for the Gender variables. One might expect that there are both Males and Females among the users for whom we do not know the gender. Thus, we should see Male and Female coefficient estimates of opposite sign so that users of unknown gender behave on average in between Male and Female. Our results, however, suggest that this is not the case for the probability of Click and Retention. This result may appear counter-intuitive, but it can be explained as follows: the users who provide gender information take an extra step that other users do not take. This is an extra 'cost' to the user and will only be incurred if the users believe that the newsletter is worth making an extra (albeit small) effort. In other words, users are self-selecting themselves not only in terms of providing gender information, but also in terms of how much effort they are willing to expend to be part of this newsletter. It is thus not surprising that these users are also less likely to defect, and more likely to click once an email is opened.

Examining the estimates for the *Open* and *Click|Open* probabilities we find that men are more likely to open an email communication, and women are less likely. We also find that men are much more likely to click on at least one of the email links, given they opened the email. The combination of the higher open rate and the higher click-through rate works in tandem, making men significantly more likely than women to respond to the email given they were sent one. Men are also much more likely than women to stay in the

database given that they were sent an email. All these differences are statistically significant (p<0.001). We will come back to this result when we look at the overall value of names.

Rows 2 to 4 of Table 3 report the estimates for the various τ covariates in each of the models. All estimates for each of the binary logit models are significant (p <0.0001). For the retention probability, the value for τ is positive, with $\tau^2 < 0$, resulting in an inverted-U shaped response function of retention probability to email frequency. While we were expecting the coefficient for $1/\tau$ to be negative, consistent with the hypothesis that as τ nears 0 the retention probability should near 0 also, we find a positive, but extremely small, coefficient. Does this mean that the retention rate shoots up to 1 as the inter-email times nears 0? No! The smallest τ we have in our dataset is one day (τ =1/365). Given the size of the coefficient for $1/\tau$ (3.97E-6) there are no substantive impact of this parameter on the retention probability. We would need to collect data for much smaller τ in order to be able to say anything meaningful about this end of the spectrum. However, the company that provided us the data is reluctant to send out emails at such high frequency.

5.2   Impact of τ on Retention, Open, and Clicks

To help interpret the coefficients of the logit regression, we plotted the fitted *Retention*, *Open*, and *Click* probabilities as a function of τ in figure 2[5]. It can be seen that the *Retention* rate is inverted-U shaped, rising (as the inter-email time increase) to a maximum

---

[5] Note that the fitted values presented in figure 2 are in-sample projection. We restricted τ in the graph to the range observed in our dataset.

at 64 days (i.e., $\dfrac{\partial p}{\partial \tau} > 0$ at $\tau = 0$). So, from a retention rate standpoint the optimal inter-email time is about 9 weeks. Unfortunately, this optimum almost exactly coincides with the minimum open rate (72 days). Indeed, both *Open* and *Click* probabilities initially decrease as τ increases (i.e., $\dfrac{\partial B}{\partial \tau} < 0$ at $\tau = 0$). They later increase, but by then the retention rate falls so much that the increased open and click rate hardly compensates for the loss in users.

Instead of trying to maximize retention rate on a campaign per campaign basis, one could also try to maximize the size of the database. Figure 3 shows what the relative size of the database would be after one year as a function of τ. We see here that the optimal inter-email time is 110 days. Therein lies the crux of the name value optimization problem. How does a company trade off database size for revenue? What decrease (increase) in *Retention* rate should a company be willing to accept (require) in exchange for increase (decrease) in *Open* or *Click* rate?

To answer these question in a satisfying manner, we need to go back to equation (18). Indeed, we have now estimated all the parameters necessary to compute the value of a name as a function of τ. Thus, we can look at the changes in lifetime value of a name as τ increases or decreases (see figure 4). Note that when making the projections, we used a weighted average effect for both *Gender* and *Day of Week* (See table 4 for the incidence of each gender and day of week categorical variables).

Our computations show that the optimal contact time is every *11 days*. If our company were to respect this schedule, it would generate a stream of profits from its database that can be valued in net present value terms at $3.37 per name. It is interesting

to note that V is asymmetrical in τ and that it is much less costly to err on the side of sending emails less frequently than more frequently. Notice also that with a τ of 11 days, the database is expected to shrink by 10% every year. In other words, the current acquisition effort expanded by the company is not large enough to sustain an email newsletter in the long run. If the company wants to have a stable database, it will need to increase τ to 16 days. This will reduce the number of emails sent per year from 33 to 23, and reduce the name value from $3.37 to $3.20.

Instead of decreasing the email frequency, the company can increase its name acquisition efforts. Currently, it uses sweepstakes, often with expensive prizes, to entice customers to enter their names. It could run more frequent sweepstakes, offer bigger prizes, or find other sources of names. When doing so, it must never forget that it would be foolish to spend more than $3.37 per name.

5.3 Benefits of Targeting

In the specification of the retention, open, and click probabilities, we included indicator variables that account for some of the heterogeneity across genders. We can now compute the optimal inter-email time (τ^*) and the corresponding name value (V^*) for each of our three genders. As can be seen in Figure 5, users who chose not reveal their gender are worth 9¢ less than females, and 19¢ less than males. Further, the optimal inter-email frequency is 12 days for males versus 11 days for females and unknown genders.

5.4 Simulation With Results

Having calibrated the lifetime value model, (14), we can do some static comparisons to illustrate some of the findings that were made during the theoretical part of this paper.

When we dissected the model we tried to give the reader a sense for the impact of each its components ($B(\tau)$, r, etc.) on the optimal τ and the resulting name value. We can now illustrate this with some concrete numbers.

Let us start with the discount rate of money. We showed earlier that $\frac{\partial V}{\partial r} < 0$, or the higher the discount rate the lower the value of names. This is reflected in figure 6. It shows how dramatic the impact of r is on τ^*. The higher the discount rate, the less the future is valued and the less a firm will wait until it sends the next email. This will increase defection rate, and thus reduce the value of each name (V^*). At the limit, for $r \to \infty$, we have $V(\tau) = B(\tau) - s - \frac{F}{N}$. This means that when the discount rate of money is infinite, one does not worry about retention rate; one only concentrates on maximizing the return of a single email (i.e., maximizing $B(\tau)$).

Changes in database growth rate have a more subtle impact. An increase in ν means that the firm has an easier time replacing the names it loses. Consequently, it will be willing to trade off retention for click and open rate. Hence, as the growth rate increases, the firm will decrease its inter-email time (figure 7). In this case, however, the decrease in τ^* is accompanied by an increase in V^*. For growth rate greater than 8%, an increase in growth rate is accompanied by an increase in overall database size at the end of the year (see figure 8). Nevertheless, it would require our company to grow its database at a rate of 85% per year in order to keep a constant number of users.

Similar to the above discussion of the discount rate, we can look at the boundary condition on v. When $v \to \infty$, we have $V(\tau) = \dfrac{(1+r)^\tau}{(1+r)^\tau - p(\tau)}\big(B(\tau) - s\big) - \dfrac{F}{N}$. That is, the optimization can be made ignoring the semi-variable costs.

6   Managerial discussion

Our empirical analysis clearly illustrates some of the findings of our theoretical framework. For instance at V^* we find ourselves in the declining portion of both p_O and p_C (i.e., $\dfrac{\partial B}{\partial \tau} < 0$). This corroborates the conventional wisdom that recency is important and must be taken advantage of. We also find that retention opposes the force that urges managers to reduce the inter-communication time (τ). Indeed, the optimal τ is in the range where $\dfrac{\partial p}{\partial \tau} > 0$. This supports our premise that too frequent communications have a detrimental impact of retention, and thus ultimately on name value. The negative impact is compounded by the observation that low retention not only reduces future earnings, but it also increases the allocation of semi-variable costs assigned to each name. Thus spacing out the communications both increases retention ($\dfrac{\partial p}{\partial \tau} > 0$) and decreases the cost allocation ($\dfrac{\partial C}{\partial \tau} < 0$).

We find that the V is highly sensitive to changes in τ. At V^*, *decreasing τ by 5 days* (i.e., going from 11 days to 6 days) decreases the value of each name by about *50%*, while *increasing τ* by the same amount decreases the value of each name by only *10%*. Thus managers would be well advised to err on the side of too much spacing out rather

than not enough. We also find that acquisition plays an important role in the value of names. New names not only represent a new source of revenue, they also raise the value of the firm's current name. Thus a firm always benefits by growing its customer base. We discuss this in further detail in the next section.

6.1 Acquisition Costs

How much effort should be spent on acquisition? As we have shown, higher growth rates lead to higher name value. And thus it is worth spending money on the acquisition of new customers. But, how aggressive should one be? Clearly, one should not spend more on acquiring customers than they are worth. To figure out how much to spend on acquisition, one must look at the cost of growth. How much would it cost to increase the current growth rate by 1%? The optimal growth rate is one that equates marginal cost to marginal revenue. In this case, the marginal revenue is the revenue generated by the higher name value due to the higher growth rate (see figure 7). The marginal cost is the increase in cost due to a higher growth rate sought.

6.2 Short Term Constraints

We have taken a long-term approach in our model development. The firm is assumed to maximize the net present value of *all* its future actions. This is probably a reasonable assumption for large established firms or for firms that have a diversified portfolio of products. For startup firms, this might be a luxury they cannot afford. Indeed, in order to enjoy the benefits of future actions, one must still be in business in the future. Thus, firms may face a constrained optimization, and maximize *V* subject to generating sufficient short-term sales to recover operating expenses.

To illustrate possible revenue constraints, we plotted Year 1 and Year 2 revenues (total non-discounted revenues) in figure 9[6]. Year 1 revenues are the highest if the emails are sent every 2 days. Unfortunately, this leaves no revenues for Year 2. Year 2 revenues are the highest at 10 days.

6.3 Attrition

What if one does not observe attrition? Catalogs, for instance, often only see partial attrition. That is, they may be notified that a person has moved. But they rarely observe the attrition from people who are not in the market anymore. In the absence of a process where customer clearly state that they do not wish to receive further contacts from the firm, one can use the framework developed by Schmittlein, Morrison and Colombo (1987) and validated by Schmittlein and Peterson (1994). Indeed, Schmittlein *et al* propose a methodology to estimate the probability that a customer is still alive, at a given point in time, given their past (in)activity.

6.4 Limitations

As with most research, several limitations are present in our study. First, a limitation of the empirical work is that we cannot distinguish between an email left unopened and an email that was opened by a non-HTML email-reader. This censorship will bias downward our Open and Click probability estimates thereby reducing our estimates of *B* and ultimately of *V*.

Second, we are only modeling whether or not an email was opened and/or clicked on. We do not attempt to model how many times an individual email was opened, or how

---

[6] The discontinuities in the revenue curves are due to the discrete nature of τ.

many times a user clicked on an email link. We assume that the magnitudes of σ and x are the same if a user opened the same email 10 times or only once. That is, we are interested in the number of emails that were opened at least once, rather than the number of emails required to generate the effect σ, since we do not have any further information about the individual effectiveness of multiple opens. A challenge for subsequent research is to study the impact of multiple opens of a single email (e.g. via the mere exposure effect) and how these multiple opens may impact the value of an open (σ) differently.

Third, subscribers to the newsletter were not told about the inter-communication time (τ) prior to subscribing. On the one hand, this makes it difficult to support the assumption that recipients know this τ and can react accordingly. On the other hand, this gives us the opportunity to generate empirical estimates on the effects of τ on the likelihood of open, unsubscribe and click-through (a natural field experiment) since we are able to observe (in a panel like environment) the effects of varying the frequency.

Further limitations of this study are the lack of carryover effects of multiple communications over time as well as the limited modeling of customer heterogeneity. These are two important problems to study. However, we did not want to confuse the theoretical issues with practical implementation problems. We do intend to further investigate these problems, giving them the full attention they deserve.

7 Conclusion

The framework developed in this paper is a natural extension of the current literature on lifetime customer value and direct marketing optimization. Our mathematical development incorporates concepts that have proven empirically to be important such as

33

acquisition and attrition (Bolton, 1998; Smith, Bolton, and Wagner, 1999) and weaves them in a rigorous theoretical analysis that takes profit maximization as its goal (Bult and Wansbeek 1995; Gönül and Shi 1998).   In doing so we are among the first to make the mailing decision endogenous.

Our analysis shows that it is important to take such a holistic approach. Indeed, we find that:

- *Inter-communication* time (τ) has a dramatic impact on customer behavior. It affects both attrition and customer surplus and thus has a critical impact on lifetime value (*V*).

- The impact of inter-communication time on lifetime value is highly *asymmetrical*. Managers are advised to err on the side of longer rather than shorter inter-communication times.

- The desire for high *retention* rates (*p*) and high per-communication *revenues* (*B*) work as opposing forces when computing the optimal inter-communication time. Retention rate begs for long inter-communication times, revenues beg for short inter-communication times.

- *Retention* benefits lifetime value in two aspects. First, one can only derive revenue from customers whom we can contact. Second, the larger a firm's customer base, the lower its per-customer contact costs (*C*).

- *Acquisition* (v) efforts must be taken into account when valuating names. Acquisition is not only beneficial in that it increases revenues; it also reduces average contact costs.

These findings are supported by our empirical analysis.

References

Bendapudi, Neeli and Leonard L Berry (1997), "Customers' Motivations for Maintaining Relationships with Service Providers," *Journal of Retailing*, 73 (1), 15-37.

Berger, Paul D. and Nada I. Nasr (1998), "Customer Life Time Value: Marketing Models and Applications," *Journal of Interactive Marketing*, 12 (1), 17-30.

Bitran, Gabriel R. and Susana V. Mondschein (1996), "Mailing Decisions in the Catalog Sales Industry," *Management Science*, 42 (9), 1364-1381.

Blattberg, Robert C. and John Deighton (1991), "Interactive Marketing: Exploiting the Age of Addressability," *Sloan Management Review*, Fall, 5-14.

Bolton, Ruth (1998) "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction," *Marketing Science*, 17,1 45-65.

Bult, Jan Roelf and Tom Wansbeek (1995), "Optimal Selection for Direct Mail," *Marketing Science*, 14 (4), 378-394.

Direct Marketing Association (2001), "Annual Report 2000."

Dowling, Grahame R. and Mark Uncles (1997), "Do Customer Loyalty Programs Really Work?," *Sloan Management Review*, Summer, 71-82.

Drèze, Xavier and François-Xavier Hussherr (1999), "Internet Advertising: Is Anybody Watching?," *ACR talk*.

Gönül, Füsun and Meng Ze Shi (1998), "Optimal Mailing of Catalogs: A New Methodology Using Estimable Structural Dynamic Programming Models," *Management Science*, 44 (9), 1249-1262.

Gupta, Sunil, Donald R. Lehmann, and Jennifer Ames Stuart (2001), "Valuing

    Customers," *working paper*, Columbia University.

Iconocast (2001), "Response Rates Are Down All Over!," November 13.

Morgan, Robert M and Shelby D. Hunt (1994), "The Commitment-Trust Theory of

    Relationship Marketing," *Journal of Marketing*, 58 (July), 20-38.

Reichfeld, Frederick F. and Thomas Teal (1996), The Loyalty Effect. Boston: Harvard

    Business School Press.

Reinartz, Werner J. and V. Kumar (2000), "On the Profitability of Long-Life Customers in

    a Noncontractual Setting: An Empricial Investigation and Implications for

    Marketing," *Journal of Marketing*, 64 (October), 17-35.

Schmittlein, David C., Donald G. Morrison and Richard Colombo (1987), "Counting Your

    Customers: Who Are They and What Will They Do Next?," *Management Science*,

    33 (1), 1-24.

Schmittlein, David C. and Robert A. Peterson (1994), "Customer Be Analysis: An

    Industrial Purchase Process Application," *Marketing Science*, 13 (1), 41-67.

Sheth, Jadish N. and Atul Parvatiyar (1995), "Relationship in Consumer Markets:

    Antecedents and Consequences," *Journal of the Academy of Marketing Science*, 23

    (4), 255-71.

Smith, Amy K., Ruth N. Bolton, and Janet Wagner (1999), "A Model of Customer

    Satisfaction with Service Encounters Involving Failure and Recovery," *Journal of*

    Marketing Research, XXXVI (August), 356-372.

Thomas, Jacqueline S. (2001), "A Methodology for Linking Customer Acquisition to

    Customer Retention," *Journal of Marketing Research*, 38 (May), 262-268.

Appendix 1:

Equation (2) states that:

$$V(\tau) = \sum_{i=0}^{\infty} \frac{p(\tau)^i}{(1+r)^{i\tau}} A(\tau).$$ (23)

If we assume that τ is constant through the life of the consumer, and thus independent from i, we can rewrite (23) as:

$$V(\tau) = \sum_{i=0}^{\infty} A(\tau) \frac{p(\tau)^i}{(1+r)^{i\tau}} = A(\tau) + \frac{p(\tau)}{(1+r)^{\tau}} \sum_{i=0}^{\infty} A(\tau) \frac{p(\tau)^i}{(1+r)^{i\tau}}$$

$$\Rightarrow V(\tau) = A(\tau) + \frac{p(\tau)}{(1+r)^{\tau}} V(\tau)$$

$$\Rightarrow V(\tau) \left(1 - \frac{p(\tau)}{(1+r)^{\tau}}\right) = A(\tau)$$

$$\Rightarrow V(\tau) = \frac{(1+r)^{\tau}}{(1+r)^{\tau} - p(\tau)} A(\tau).$$

Table 1: Notation

V	Name Value
p	Retention Rate
A	Profit generated by one contact
B	Revenue Generated by one contact
C	Cost of One contact
F	Fixed cost of a campaign
N	Size of the database
r	Discount rate of money
ν	Growth rate of the database
p_O	Probability of Open
p_C	Probability of Click
σ	Benefits of an open
x	Benefits of a click

Table 2: Data Description

|	
Number of names at the beginning of the test	627,713
Number of new names	485,961
Number of unsubscribe requests	172,498
Number of names at the end of the test	941,176
Number of email campaigns	62
Number of email sent	4,968,520
Number of trackable emails	2,548,362
Average number of emails per user	3.62
Number of opened emails	354,449
Number of clicks	51,262

Table 3: Logit estimation results for probability of Open, Click|Open and Retention.

Variable	Retention		Open		Click|Open	
	Parameter Estimate	Std. Err.	Parameter Estimate	Std. Err.	Parameter Estimate	Std. Err.
Intercept	4.04	0.007	-1.81	0.004	-1.30	0.010
$1/\tau$	3.97E-6	1.75E-6	-	-	-	-
τ	5.04	0.053	-1.21	0.036	-3.63	0.099
τ^2	-13.05	0.086	2.79	0.068	4.87	0.193
W1 Sun	-0.17	0.007	0.08	0.004	0.26	0.011
W2 Mon	-0.24	0.008	*-0.002*	*0.005*	-0.55	0.015
W3 Tues	0.76	0.016	0.17	0.010	*-0.01*	*0.028*
W4 Wed	-0.19	0.008	0.02	0.006	0.24	0.015
W5 Thurs	0.26	0.008	-0.05	0.005	-0.04	0.014
W6 Fri	0.13	0.007	-0.25	0.004	0.33	0.010
G1 Female	0.20	0.008	-0.32	0.004	0.12	0.009
G2 Male	0.41	0.008	0.13	0.003	0.40	0.007
N	4,968,520		2,548,362		354,449	
(ones)	(4,796,022)		(354,449)		(51,262)	
LR	194,327		18,928		15,037	
LR χ^2 Prob	<.0001		<0.0001		<0.0001	

Table 4: Incidence of Categorical Variables

Gender		Day of Week	
Female	9.4%	Sunday	14.1%
Male	12.6%	Monday	10.4%
Unknown	78.0%	Tuesday	5.6%
		Wednesday	10.9%
		Thursday	19.8%
		Friday	20.7%
		Saturday	18.5%

Figure 1: Empirical Inter-Email Times

Frequency

Number of Weeks since last email sent
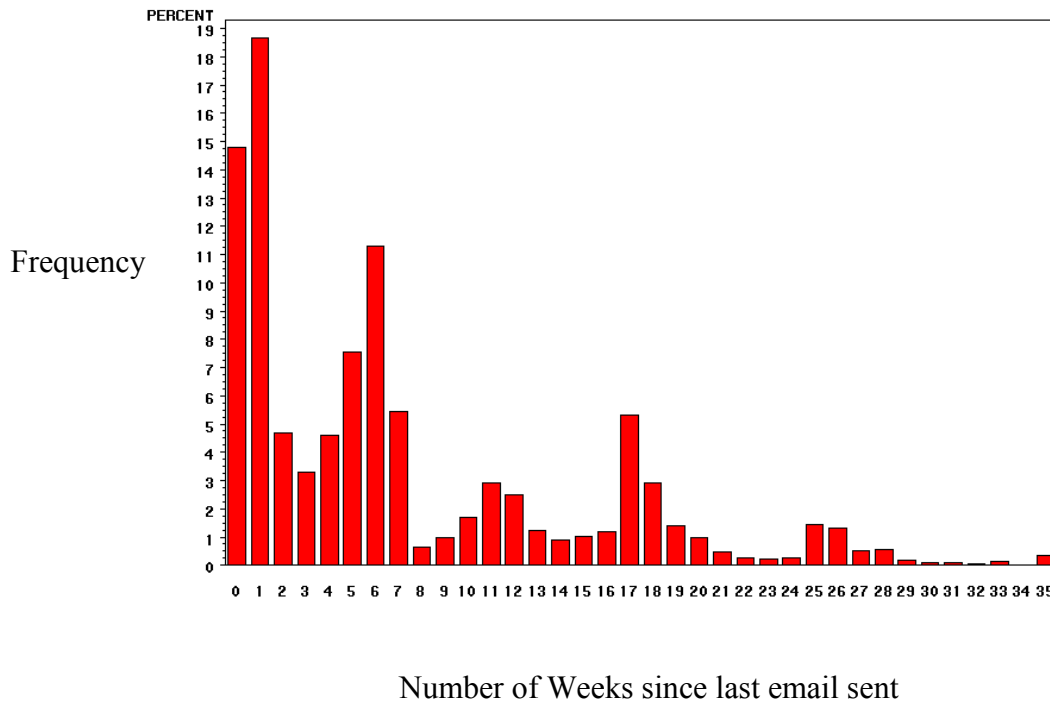
Figure 2: Retention, Open, and Click Probabilities

42
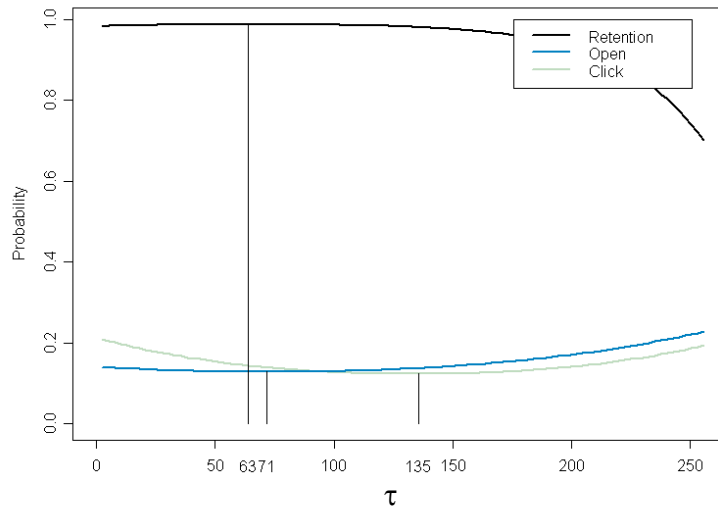
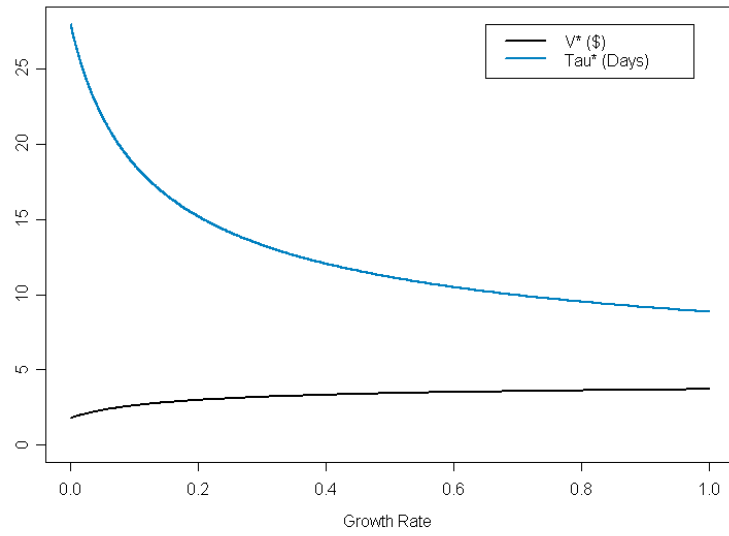**Figure 3: Database Size**
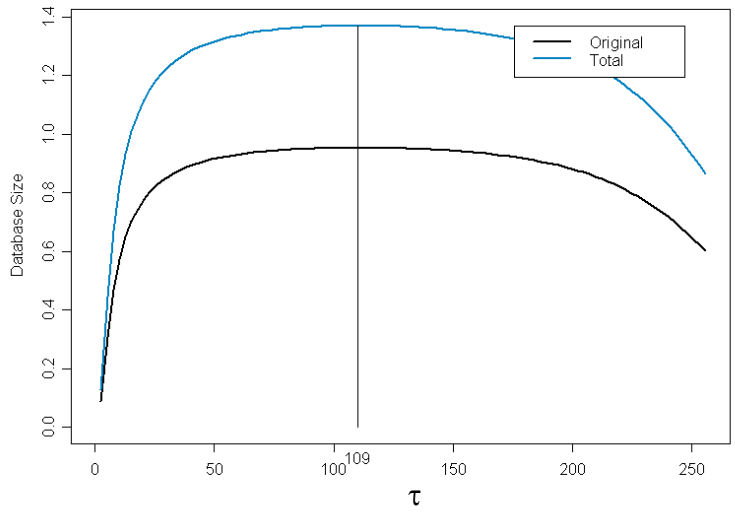
Figure 4: Name Value
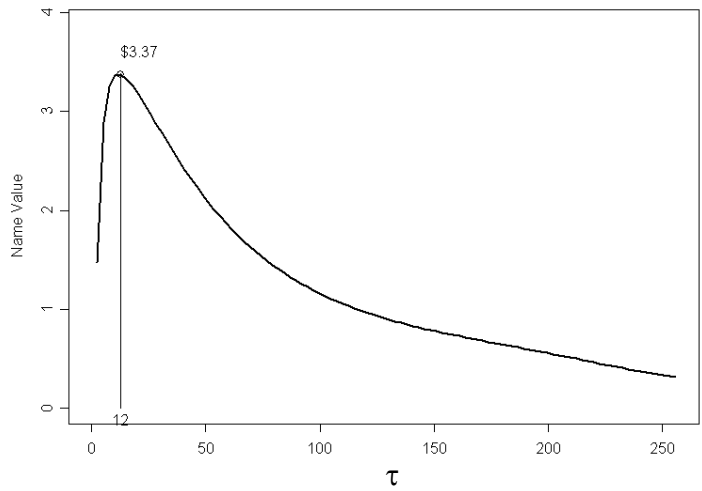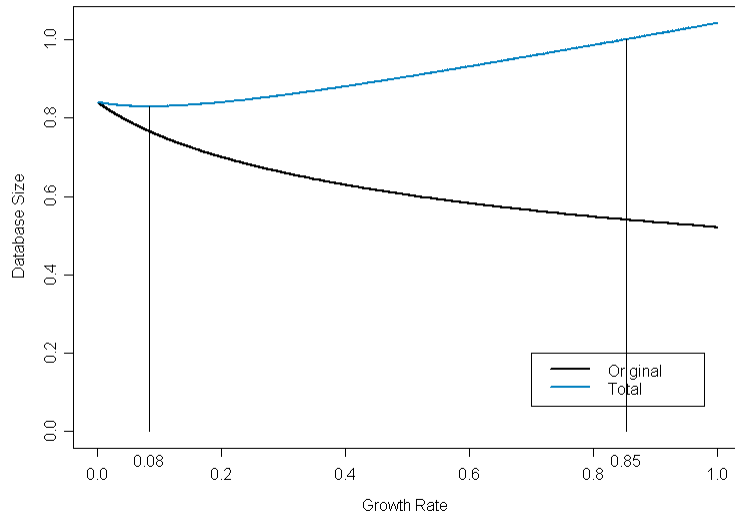
Figure 5: Gender Targeting
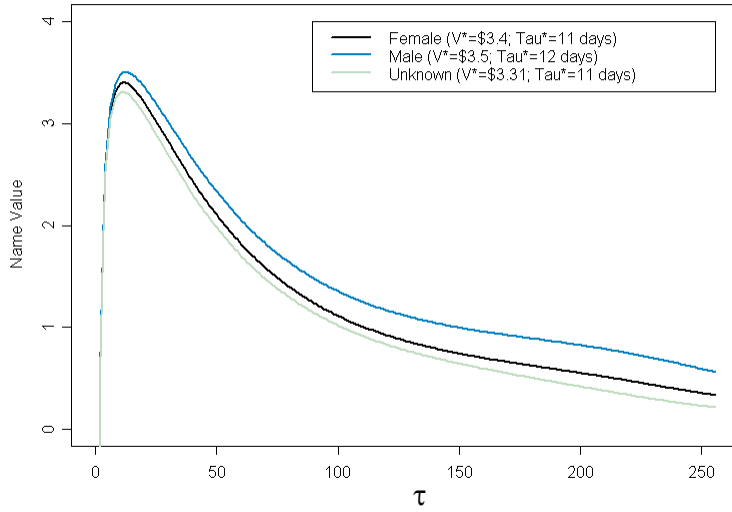
Figure 6: Impact of Discount Rate on Name Value

44
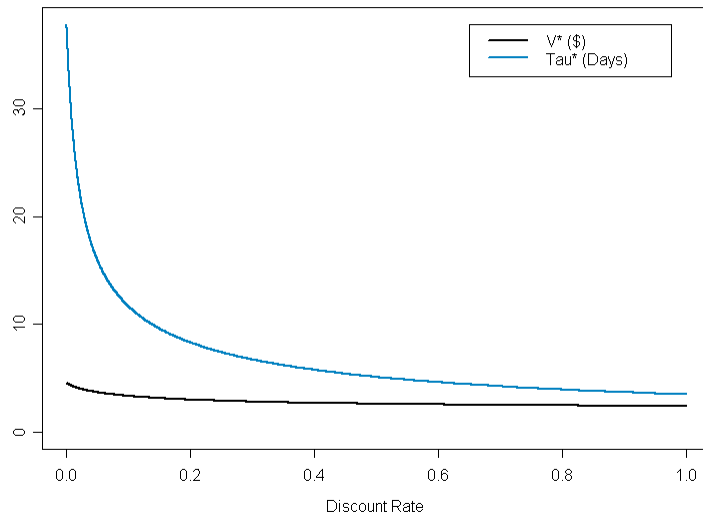
Figure 7: Impact of Growth Rate on Name value

Figure 8: Impact of Growth Rate on Database Size

Figure 9: Yearly Revenues