university of groningen

University Medical Center Groningen

University of Groningen

Quantitative approaches for profit maximization in direct marketing

van der Scheer, H.R.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
1998

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
van der Scheer, H. R. (1998). *Quantitative approaches for profit maximization in direct marketing*. s.n.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 03-06-2022

Chapter 7

Frequency of direct mailings

7.1 Introduction

One of the most prominent questions in direct marketing is how to define a strategy to maximize the lifetime value (LTV) of an individual. Pearson (1994) defines LTV as: 'the net present value of the stream of contributions to profit resulting from the revenues from customer transactions and allowing for the costs of delivering products, services and promised rewards to the customer'. Similar definitions are given by Jackson (1994) and Kestnbaum (1992), among others. Focusing on the LTV means that the success of a strategy is not defined in a short-term criterion such as the response rate, but in a long-term criterion. Only if current purchases are the sole consideration, as with an encyclopedia, is a short-term criterion appropriate. However, often even these purchases can be followed by purchases of related products, so-called *cross-selling*, and hence also the long-term aspects become of interest. One element of a strategy to maximize the LTV, which we examine in this chapter, is the frequency with which an individual should receive a mailing. Before we turn to the mailing frequency, we discuss the LTV concept in somewhat more detail.

Three aspects characterize the traditional way of calculating and employing the LTV (e.g. Courtheoux in Roberts and Berger 1989, p. 411, Hughes and Wang 1995, and Kestnbaum 1992). First, the LTV is used for decision making in the areas of customer reactivation and customer acquisition. Secondly, the calculation is solely based on variables of the organization's database, e.g. on past purchase behavior. Thirdly, the LTV is determined for a group of individuals.

As Hoekstra en Huizingh (1996) argue, the traditional way of employing the LTV is too narrow within direct marketing. With regard to the use of the

LTV, it should be employed - apart from customer acquisition - for relationship building. That is, the LTV should be used to choose media for communication with customers, develop loyalty programs and assess the strength of the relationship. Fully exploiting the possibilities of the LTV for relationship building has several consequences for the other two aspects of the traditional way of employing the LTV.

First, when focusing on relationships rather than transactions, the calculation of the LTV should not only be based on past purchase behavior but also on data concerning future situations. That is, instead of only using the organization's database, which is more or less the by-product of the order process, additional data should be collected, such as customer satisfaction and (repeat) purchase intention, the so-called *forward-looking data*. These data are also useful to capture the dynamics in the relationship between the consumer and the organization. Secondly, in order to build one-to-one relationships, the LTV should be determined at the individual level rather than at the group level. Consequently, the data should be collected and analyzed at the individual level, and the strategy should be based on these individual results.

Thus, in order to employ the LTV to its full extent, it should be used for decisions regarding creating, developing and maintaining relationships. The ultimate goal should be an individual-based normative direct marketing strategy. The strategy should depend on past purchase behavior and on forward-looking data. The normative aspect implies that the strategy indicates when an individual should receive a mailing and what kind of mailing that should be.

The first aspect, 'when', deals with the timing and sequencing of the mailings. The timing relates to the choice of the day of the week, month or season. Common wisdom in direct marketing suggests that, for obvious reasons, the last days of the week are preferable (Fraser-Robinson 1989, p. 106, Vögele 1992, pp. 292-293). Sequencing relates to the period between two consecutive mailings. Clearly, the optimal period differs between individuals. Furthermore, the optimal period may differ between the various mailings for one particular individual. Bitran and Mondschein (1996) derive a heuristic for a catalog retailer for when and how often to mail. Gönül and Shi (1996) extend this approach by explicitly modeling the consumers' response behavior. They use a structural dynamic programming model to determine a strategy for a catalog retailer. Both papers conclude that it is optimal to mail individuals who have purchased a small number of times, and to mail those individuals who did not purchase anything for a long time. Note that these findings may be typical for a catalog retailer, since the response probability is not zero if an individual

does not receive a mailing. That is, the possibility remains that the individual will order from the last catalog received.

The second aspect of the strategy, ‘what mailing’, relates to the type of mailing an individual should receive. It includes the design of the mailing and the offer submitted. The importance of a carefully designed mailing is demonstrated in chapter 6. Furthermore, we showed in chapter 6 that it is useful to take the interaction between the target and mailing characteristics into account in order to choose a mailing design that maximizes the expected profits per individual. The offer submitted in the mailing includes the product itself, the price and the other elements for positioning the product. For example, the product could be submitted with an additional promotion. This promotion could be directed to reactivate an individual or to award individuals for their loyalty. In order to maximize the expected profits, the additional promotion should be determined per individual.

Note that target selection is part of such an individual-based strategy. That is, an individual is selected for a certain mailing at a particular time if that maximizes its LTV. An individual is excluded from the mailings list if his LTV is negative for all possible direct marketing activities. Furthermore, note that the decision of ‘when’ and ‘what’ are not independent. For example, the sequencing of the mailings of a catalog retailer differs between the choice of sending a reminder or a new catalog.

It goes without saying that a fine-tuned normative individual strategy is difficult to formulate and that it suffers from several practical drawbacks. The major difficulty is that, even in a rather simple setting, the structure of individual characteristics (including past purchase behavior and forward looking data), mailing and offer characteristics and situational factors, is very complex. The dynamical aspects in the analysis are enhanced by the selection effects resulting from the chosen strategy. That is, the chosen strategy is not randomly assigned to individuals but is based on the individuals’ behavior.

One of the main practical drawbacks is data availability. In order to determine individual values of the LTV on data concerning future situations, such as purchase intentions, it implies that organizations should collect additional data at the individual level. The collection of these data will be an expensive and difficult task. Not surprisingly, Hoekstra and Huizingh (1996) find that organizations with many customers have significantly less of this kind of information at the individual level than organizations with a small number of customers. Another practical problem is that a fine-tuned individual strategy may not be cost-efficient. That is, it may well be that the cost of implementing

the strategy (data storage, computing, handling of the mailings and the like) exceeds the potential benefits.

To sum up, although a fine-tuned normative individual strategy based on the LTV should be the ultimate objective of a direct marketing organization, it will be quite hard to realize. Two main reasons are that the normative strategy should result from a complex structure and that there is a lack of specific individual data. There is, however, considerable room for improvement of various aspects of such a strategy. One of these aspects, which we examine in this chapter, is to determine the optimal frequency with which an organization should send its customers a mailing for frequently purchased consumer goods.

Roberts and Berger (1989, pp. 241-242) argue that this question has to be answered by experimentation. An experiment can for instance be performed by considering several (random) groups of the mailing list that receive the direct mailings with different frequencies. The objective of this experiment is to determine the frequency that generates the highest average profit of the addressees. Such an experiment, however, has several drawbacks:

1. The group that receives mailings with a low frequency and the group that receives mailings with a high frequency consists of "good" and "moderate" targets. This causes inefficiencies since the organization has to send some of the best (moderate) targets mailings with a much lower (higher) frequency than is preferable. This is especially harmful since the groups must be large and the experiment should be continued for some time to obtain useful and reliable results.
2. Given the length of the experiment, it takes some time before the organization has reliable results enabling it to determine the mailing frequency.
3. The results may suffer from non-stationarity. This means that one year's optimal frequency differs from the optimal frequency a few years later.

This chapter, which is based on Van der Scheer and Leeflang (1997), proposes a method to determine the optimal frequency of direct marketing activities for frequently purchased consumer goods (e.g. books, compact discs), which does not suffer from these drawbacks. The underlying idea of this method is that *consumers' purchase behavior* should be the basis for the supplier's direct marketing behavior. The method is operationalized by specifying a conceptual framework for the purchase behavior of the individual, i.e. a model that describes the decisions an individual has to make before the purchase of the direct marketing (DM) product. The model is used to simulate the decisions of an individual for a range of frequencies of DM activities. As a result, the

optimal frequency can be determined. The model decomposes the purchase behavior in the timing of purchases, which involves interpurchase time and purchase acceleration, and the DM-product choice. Thus, the model takes *all* the purchases of the individuals in the product category into account, and specifies explicitly the individual's decision to buy the product through DM or in a regular store.

The order of discussion is as follows. In section 7.2 we present the maximization problem of the direct marketing organization and discuss the model describing the consumers' purchase behavior. The specification of the components of this model is discussed in section 7.3. In section 7.4 we discuss ways to obtain data for the input parameters for the simulation. The calibration of the proposed model requires that data are collected which satisfy specific criteria. We investigated whether DM organizations collect these data on a continuous basis. We were only able to find organizations that collect the specified data on an ad-hoc basis. To calibrate our model we collected data among undergraduate students. The student population is not representative of a whole population and the outcomes of this study should therefore be interpreted with care. Thus the empirical part of this chapter is no more than an empirical illustration and application of our model. A sketch of the simulation is given in section 7.5. We present the empirical results in section 7.6, and discuss additional research issues in section 7.7.

7.2 The model

Under consideration is a direct marketing organization whose goal it is to maximize expected profits over a given period by deciding the number of direct mailing campaigns for a frequently purchased consumer good. We will not address the problem of target selection and assume that the DM organization selects a group of individuals for these direct mailings. We determine the frequency with which these individuals should receive a mailing by simulating the *consumers' purchase behavior*. First, we formally define the maximization problem of the organization. Then, we specify the model describing the consumers' purchase behavior.

Let $i,\ i = 1, \ldots, I$, be the individuals and $m,\ m = 1, \ldots, M$, be the mailings. The organization's maximization problem is

$$\max_{M} \sum_{m=1}^{M} \sum_{i=1}^{I} \left(w R_{im} - c \right), \tag{7.1}$$

subject to the constraint that the M mailings are sent in the specified period. R_{im} is the random variable given by

$$R_{im} = \begin{cases} 1 & \text{if individual } i \text{ responds to mailing } m \text{ by} \\ & \text{purchasing the DM product} \\ 0 & \text{if individual } i \text{ does not respond to mailing } m; \end{cases} \tag{7.2}$$

w are the revenues to a positive reply, and c is the cost of a mailing. Since it suffices to consider only a short period, as will become clear later on, we do not include a discount factor for future revenues. The key element in expression (7.1) is R_{im}, in particular $\mathrm{P}(R_{im} = 1)$, i.e. the probability that i will respond to mailing m.

Consider an individual that purchases products of a certain category. The individual has to decide *when* to purchase the product, and, in case of a purchase, *where* to purchase (cf. Gupta 1988). The first decision depends, among other things, on the time elapsed since the last purchase and the distribution of the *interpurchase times*. An interpurchase time of individual i, T_i, is the period between two consecutive purchases. Interpurchase times are random variables which are inversely related to the frequency with which the individual purchases the product. The distribution of interpurchase times enables us to determine the probability that the time between two purchases is at least of length t: $\mathrm{P}(T_i > t)$. If the individual decides to purchase the product, he has to choose where to buy it. Here we assume that this can be either in a (regular) store or through a direct marketing organization.

Let t_m be the time between the last purchase and the next mailing, and let θ_i be the interval in which an individual i takes the DM product into consideration. We allow this interval to vary across individuals. Some individuals decide immediately after they receive the mailing whether or not they want to buy the product, whereas other individuals keep the mailing to decide later on. Consequently, we can define a period $(t_m, t_m + \theta_i)$ that the DM product is considered as an alternative. Thus,

$$\begin{aligned} &\mathrm{P}(R_{im} = 1) \\ &\quad = \mathrm{P}(R_{im}=1 \mid t_m \le T_i \le t_m + \theta_i)\mathrm{P}(t_m \le T_i \le t_m + \theta_i) \\ &\qquad + \mathrm{P}(R_{im}=1 \mid T_i < t_m \vee T_i > t_m + \theta_i)\mathrm{P}(T_i < t_m \vee T_i > t_m + \theta_i) \\ &\quad = \mathrm{P}(R_{im} = 1 \mid t_m \le T_i \le t_m + \theta_i)\mathrm{P}(t_m \le T_i \le t_m + \theta_i), \end{aligned} \tag{7.3}$$

where the last step holds, since a DM product can only be purchased in the period that a category purchase is made, i.e. $\mathrm{P}(R_{im} = 1 \mid T_i < t_m \vee T_i >$

$t_m + \theta_i) = 0$. In (7.3), $P(R_{im} = 1)$ is the product of the probability that i will purchase the DM product conditional on the purchase of a product from the category, which we call the *DM-product choice*, and the probability of a category purchase in $(t_m, t_m + \theta_i)$. We now elaborate on the latter probability. To this end, we introduce the concept of *planned* interpurchase time T^* (cf. Bucklin and Lattin 1991). It is an unobserved random variable that denotes the planned or intended period between two consecutive purchases. The individual purchases the product as planned, i.e. in accordance with the planned interpurchase time, except when the market behaves differently than expected. Unexpected behavior may be brought about by e.g. a price discount or a direct marketing activity. Thus, the observed interpurchase time is the planned interpurchase if the market behaves as expected; otherwise they differ.

Write $P(t_m \leq T_i \leq t_m + \theta_i)$ conditional on the various intervals of the planned interpurchase time, i.e.

$$\begin{aligned} &P(t_m \leq T_i \leq t_m + \theta_i) \\ &\quad = P(t_m \leq T_i \leq t_m + \theta_i \mid T_i^* < t_m) P(T_i^* < t_m) \\ &\qquad + P(t_m \leq T_i \leq t_m + \theta_i \mid t_m \leq T_i^* \leq t_m + \theta_i) P(t_m \leq T_i^* \leq t_m + \theta_i) \\ &\qquad + P(t_m \leq T_i \leq t_m + \theta_i \mid T_i^* > t_m) P(T_i^* > t_m) \\ &\quad = P(t_m \leq T_i^* \leq t_m + \theta_i) + P(t_m \leq T_i \leq t_m + \theta_i \mid T_i^* > t_m) P(T_i^* > t_m). \end{aligned} \tag{7.4}$$

Relation (7.4) holds because

$$P(t_m \leq T_i \leq t_m + \theta_i \mid T_i^* < t_m) = 0,$$

since individuals are not aware of the mailing at the time they plan to purchase the product, and

$$P(t_m \leq T_i \leq t_m + \theta_i \mid t_m \leq T_i^* \leq t_m + \theta_i) = 1$$

because the planned interpurchase time T_i^* and the interpurchase time T_i coincide.

If the individual receives the mailing before the planned purchase, this may influence his behavior and cause him to decide to purchase the product at an earlier time than planned. This is called *purchase acceleration* (or *forward buying*), of which the probability is given by $P(t_m \leq T_i \leq t_m + \theta_i \mid T_i^* > t_m)$.

To summarize, expressions (7.3) and (7.4) define the model describing the consumers' purchase behavior. The essential components of this model are:

1. the distribution of planned interpurchase times, which determines $P(t_m \leq T_i^* \leq t_m + \theta_i)$ and $P(T_i^* > t_m)$;
2. purchase acceleration, which determines $P(t_m \leq T_i \leq t_m + \theta_i \mid T_i^* > t_m)$;
3. DM-product choice, which determines $P(R_{im} = 1 \mid t_m \leq T_i \leq t_m + \theta_i)$.

We will specify these components in the next section.

7.3 Specification of the model's components

In order to operationalize the model, we have to define the specifications underlying the (planned) interpurchase time, purchase acceleration and DM-product choice. For the time being we assume that all the information required to estimate the unknown parameters of these specifications is available.

Interpurchase time

To obtain a density function of interpurchase times we can use a parametric or a non-parametric approach. In the parametric approach we postulate the underlying function of the interpurchase times. The parameters of this function are estimated from the observed data. A standard approach to do so is with a hazard model (e.g. Gupta 1991, Kiefer 1988, Vilcassim and Jain 1991). However, even in frequently researched product categories there is no single function that adequately characterizes individuals' interpurchase times (Jain and Vilcassim 1991). Moreover, there is no theory that specifies the probability function. For those reasons, Jain and Vilcassim recommend the use of a very general specification.

A non-parametric probability function is such a general specification. It has the advantage that we do not have to specify a function a priori. An obvious choice is the kernel density estimator defined by

$$f(T) = \frac{1}{hN} \sum_{n}^{N} \mathcal{K}\left(\frac{T - t_n}{h}\right), \tag{7.5}$$

where $\mathcal{K}(\cdot)$ is the kernel, and h is the *smoothing parameter*; t_n denotes the observed interpurchase times of all purchases of all the individuals, $n = 1, \ldots, N$. We use the Gaussian kernel with $h = 1.06\sigma N^{-1/5}$, where σ is the standard deviation of the interpurchase times (Silverman 1986, p. 45).

Purchase acceleration

Let t_i^*, $t_i^* > t_m + \theta_i$ be a realization of T_i^*. We focus on purchase acceleration due to a direct marketing activity. We assume that the actual interpurchase time, t_i, will be $t_m + \theta_i$ if the individual decides to accelerate a purchase. Hence, the difference in time between the planned purchase and the accelerated purchase is $\delta_i = t_i^* - (t_m + \theta_i)$.

For a number of purchase situations, $j = 1, \ldots, J$, with $t_i^* > t_m + \theta_i$, we wish to determine whether or not i accelerates the purchase. Let $y_{ij} = 1$ if i accelerates the purchase in situation j, and $y_{ij} = 0$ otherwise. We denote *willingness* of i to accelerate the purchase in situation j by the latent variable y_{ij}^* that satisfies a linear model,

$$y_{ij}^* = \alpha_i + \beta' x_j + \rho \delta_{ij} + e_{ij}, \tag{7.6}$$

where x_j is a vector of covariates of the offer in situation j (e.g. a price discount), δ_{ij} is the time between t^* and $t_m + \theta_i$ for the jth purchase situation; β and ρ are unknown parameters, and $e_{ij} \sim N(0, 1)$, independently of x_j and δ_{ij}. The heterogeneity across individuals is represented by the α_i, assumed to satisfy $\alpha_i \sim N(\alpha, \sigma_\alpha^2)$, with α and σ_α^2 unknown parameters. We do not observe the willingness but the actual decision

$$y_{ij} = \begin{cases} 1 & \text{if } y_{ij}^* > 0 \\ 0 & \text{otherwise,} \end{cases}$$

which is the so-called random effects probit model (e.g. Hsiao 1986). The parameter ρ gives the effect of the length of the period between the planned purchase and the accelerated purchase. It is expected that the willingness decreases in δ_{ij}, i.e. ρ has a negative sign.

The probability of purchase acceleration for given x, t_i^*, and $t_m + \theta_i$, so $\delta_i = t^* - t_m + \theta_i$, is given by

$$\begin{aligned} \mathrm{P}(t_m \leq T_i \leq t_m + \theta_i \mid t_i^* > t_m + \theta_i) &= \mathrm{P}(y_i^* > 0 \mid x, \delta_i) \\ &= \Phi(\alpha_i + \beta' x + \rho \delta_i), \end{aligned}$$

where $\Phi(\cdot)$ is the standard normal distribution function.

Accelerated purchases can be either *incremental* or *borrowed*. It is said to be (completely) incremental if the successive planned interpurchase time does not change. It is (completely) borrowed if successive planned interpurchase time increases by δ_i. Most situations will probably fall in between these two extremes. This means that the subsequent interpurchase time increases by $\kappa \delta_i$ $(0 \leq \kappa \leq 1)$, where κ is defined as the *borrowing rate*.

DM-product choice

Given that the individual purchases the product in $(t_m, t_m + \theta)$, he has to choose whether to purchase it through DM or in the store. For this decision we also use a random effects probit model, viz.

$$R^*_{ij} = \mu_i + \gamma' x_j + u_{ij}, \tag{7.7}$$

where γ is a vector of unknown parameters, and $u_{ij} \sim N(0, 1)$ is the disturbance term. Heterogeneity across individuals is represented by μ_i, assumed to satisfy $\mu_i \sim N(\mu, \sigma^2_\mu)$, with μ and σ^2_μ unknown parameters. For convenience of notation we assume that the same x_j occur in (7.6) and (7.7), but this is innocuous since elements of β and γ can a priori be set at zero. We do not observe R^*_{ij}; we only observe whether the individual decides to purchase the DM product ($R_{ij} = 1$) or not ($R_{ij} = 0$), i.e.

$$R^*_{ij} = \begin{cases} 1 & \text{if } R^*_{ij} > 0 \\ 0 & \text{otherwise,} \end{cases}$$

given that $t_m \leq T_i \leq t_m + \theta_i$. The latent variable, R^*_{ij}, can be interpreted as the difference in utility between buying the DM product and buying the product in a store. If the utility of the former is larger, i will buy the DM product (Domencich and McFadden 1975). The probability of choosing the DM product is given by

$$\mathrm{P}(R_{im} = 1 \mid t_m \leq T_i \leq t_m + \theta_i) = \Phi(\mu_i + \gamma' x),$$

where $\Phi(\cdot)$ is the standard normal distribution function.

7.4 Data

In order to obtain estimates for the unknown parameters in the specified models and hence to employ the simulations, we need several kinds of information. The information should relate to: (1) interpurchase times, (2) purchase acceleration, and (3) DM-product choice. We consider two ways to obtain this information. First, it can be obtained from a household panel. Secondly, information can be collected by questionnaires or experiments.

Apart from demographic and geographic information, a household panel typically provides information on all the households' purchases. Hence, it

provides sufficient information, in principle, to derive a distribution of interpurchase times. In order to examine the purchase acceleration, we also need for each purchase information on the confrontations of the individuals with promotions that, at least in principle, may influence their purchase decision. Using a hazard model with covariates, the effect of promotions on the interpurchase time can be determined (e.g. Jain and Vilcassim 1991). To examine the effect of a promotion in a store versus direct marketing, we also need to know where the product is purchased, i.e. through direct marketing or in a regular store. The DM-product choice can be analyzed when information is available on all the direct mailings that the household received in the product category. Information on promotions and direct mailings is, however, not usually available through household panel data.

The other way to obtain information on the relevant variables is by a questionnaire and/or experiments. Ideally, this should be collected in addition to a household panel. A questionnaire could be used to obtain information on, for example, the period that a household takes a DM product into account. Experiments, like conjoint analysis, can be used to derive information on purchase acceleration and DM-product choice. The advantage of conjoint analysis is that the individuals are confronted with more or less real-life situations. That is, all the aspects that are expected to play a role in the individual's decision can be incorporated in the choice sets. In particular the method of pairwise comparison is useful since it elicits choices from the individuals, in contrast with other conjoint methods. Obviously, an experiment will not give an exact picture of the individual's behavior but it will yield helpful information to obtain the input parameters.

It is important to realize that we do not need information for all the households on the mailing list. In principle information of a small sample is sufficient. If this sample is not representative for the individuals on the mailing list, the findings could without difficulties be adjusted with the use of the geographic and demographic variables.

7.5 Simulation of individuals' decisions

In this section we specify the set-up for the simulation of individuals' decisions. We choose the number of mailings (M) in the given period. In our simulation we take these equally spread out over time, with a random start. Hence, given the time of the first mailing we know the (calendar) times of all the other mailings.

Situation 1: The planned purchase is before the direct mailing

Situation 2: Planned purchase is in the interval $(t_m, t_m + \theta_i)$

Situation 3: Direct mailing is before the planned purchase

Figure 7.1: Purchase situations in simulation

The ms denote the time mailings are sent, and p denotes the last purchase (indicated by the dot). The time between the last purchase and the mailing is defined by t_m, and in the interval $(t_m, t_m + \theta_i)$ the DM-product is taken into consideration. Let t^* be the realization of the planned interpurchase time. In situations 1 and 2 the product will be bought in accordance with the planned interpurchase time. In situation 2 this could be the DM-product. In situation 3 the actual interpurchase time is $t_m + \theta_i$ if the consumer accelerates the purchase by δ_i.

We describe the simulation of the purchases for one particular individual. Hence, it should be employed I times.

a Draw $\theta \sim \hat{g}(\theta)$, $\alpha \sim N(\hat{\alpha}, \hat{\sigma}_{\alpha}^2)$, and $\mu \sim N(\hat{\mu}, \hat{\sigma}_{\mu}^2)$, where $\hat{g}(\theta)$ is the estimated (kernel) density of θ, and $(\hat{\alpha}, \hat{\sigma}_{\alpha}^2)$ and $(\hat{\mu}, \hat{\sigma}_{\mu}^2)$ are the estimated parameters of the normal distribution of the α and μ, respectively.

b Draw $t^* \sim \hat{f}(T)$, a realization of T^*; $\hat{f}(T)$ is the estimated kernel density of T. Determine t_m: time period between the last purchase and the next mailing. Three situations can occur:

1 $t^* < t_m$, then $R_{im} = 0$.

2 $t_m \leq t^* \leq t_m + \theta_i$, then $t_m \leq t \leq t_m + \theta_i$. Draw u from the uniform density on (0, 1). $R_{im} = 1$ if $u < \Phi(\mu_i + \gamma' x)$; else $R_{im} = 0$.

3 $t^* > t_m + \theta_i$, then $\delta_i = t^* - (t_m + \theta_i)$. Draw u from the uniform density on (0, 1). $R_{im} = 1$ if $u < \Phi(\alpha_i + \beta' x + \rho\delta_i)\Phi(\mu_i + \gamma' x)$; else $R_{im} = 0$. If $R_{im} = 1$, then the next interpurchase time will be increased by $\kappa\delta_i$ (where κ is the borrowing rate).

c Repeat until R_{iM} is simulated.

d Determine profits for i: $\sum_{m=1}^{M}(wR_{im} - c)$.

In step **a** the individual effects and the interval in which i takes the DM product into consideration are determined. These parameters do not change during the simulation. The planned interpurchase times are drawn in step **b**. The three situations which are possible are depicted in figure 7.1. The ms indicate the (calendar) times of the DM activities; p, indicated by the dot, is the time of the last purchase. Hence t_m is the period between p and the next m. The interval $(t_m, t_m + \theta_i)$ indicates the period in which the DM product is considered as an alternative. In situation 1 the planned purchase is before the direct mailing, so the product will be bought as planned and it is not the DM product. In situation 2 the product will also be bought as planned and this could probably be the DM product. The probability that it is the DM product is $\Phi(\mu_i + \gamma' x)$. In situation 3 the planned interpurchase time is longer than $t_m + \theta_i$. The DM product is bought if the individual accelerates its purchase and chooses the DM product (conditional on purchase acceleration); this probability is $\Phi(\alpha_i + \beta' x + \rho\delta_i)\Phi(\mu_i + \gamma' x)$. If the DM product is purchased at $t_m + \theta_i$, the successive interpurchase time will increase by $\kappa\delta_i$ (κ is the borrowing rate). This is repeated (step **c**) until R_{iM} is simulated. Employing this for I individuals we obtain the overall profits. We examine this for different values of M, which gives the optimal frequency.

Figure 7.2: Probability density function of interpurchase times (in weeks)

7.6 Empirical illustration

We applied the proposed model and simulation to compact disc purchases among students. The students filled out a questionnaire with questions on DM-product choice, purchase acceleration and compact disc purchases (in that order). Of the 146 students who received the questionnaire, 141 filled it out completely. First we briefly discuss the components of the individual response model, then we discuss the simulation results.

Data collection

Using a time line the respondents indicated the periods and prices of their CD-purchases of last year. Generally this would be difficult, but given the age and the students' background this is not impossible. The data obtained in this way give us indications about the interpurchase times. To determine the kernel density of the interpurchase times, which is depicted in figure 7.2, we considered those respondents who bought at least four CDs at a price higher than NLG 25. We considered these respondents because they would be an interesting group for a DM organization. We chose the critical price level of NLG 25 because we want to consider CDs which have been bought at a more or less standard price, which is about NLG 40. Consequently, we used 69

Figure 7.3: Probability density function for θ (in weeks)

respondents with, in total, 321 interpurchase times to determine the kernel density. The average interpurchase time of these respondents is 8.45 weeks. The smoothing parameter, based on the formula given in section 7.3, equals 2.52.

The distribution of the θ_is is obtained as follows. In the questionnaire we asked the respondents how long they kept a brochure with compact discs promotions, given that they were interested in the price discount but had not planned to purchase a compact disc. They had to choose from several digit preferences (e.g. a day, two or three days, . . ., two weeks). The average θ is 0.75 week. Figure 7.3 shows the estimated probability density function based on the Gaussian kernel. Using the smoothing parameter defined in section 7.3, which equals 0.275, we obtain a kernel that shows too much of the fixed answers in the questionnaire. Therefore we use a larger smoothing parameter (0.4), which resulted in the depicted curve.

We used a conjoint experiment, with the method of paired comparison, to obtain purchase acceleration data. The attributes we included are price (NLG 40, 35, and 30), when (direct, in 1, 2 or 3 weeks), and whether or not there was a *savings plan* (also called *patronage awards*). The latter refers to a type of promotion in which an individual saves up for a free compact disc. We kept several attribute levels fixed between the different comparisons, to keep

the task of the respondents tractable. A fractional factorial design resulted in eight comparisons. We gave the respondents one additional comparison, to be used for validation. All respondents provided information about the same nine pairs. We operationalized planned purchase in the following way: "Assume that you plan to buy a compact disc **in two weeks** because you receive your monthly grant (or because it is your friend's birthday). However, you can buy the CD with a price discount **now**." Then the respondents were asked to choose between two options.

Of the 141 respondents, 63 indicated that they 'always' preferred forward buying. This information is useful but it complicates our simulation. Moreover, we cannot use it for estimation of the random effects probit model, because these respondents would have an infinitely large positive intercept. Hence, we estimated the model on the basis of the 78 respondents with variation in their choices. This group is indicated by G_1. The group of respondents that do not have any variation in their answers is indicated by G_2. Each respondent made eight choices. Hence, $J = 8$, $I = 78$, and the covariates in x_j are 'price discount' and 'savings plan'. Price discount is the difference between the regular price (NLG 40) and the price offered. The time to a planned purchase, δ_i, takes only three values: one, two and three weeks.

Table 7.1 gives the estimated coefficients. Since σ_α^2 was not significantly different from zero, we used the standard probit model. The signs of the coefficients are as expected. The price discount and the savings plan have a positive effect. Hence, an individual is more inclined to accelerate his purchases if there is a promotion. The time until the planned purchase has a negative effect on purchase acceleration. This shows that individuals are less likely to accelerate their purchase by, say, two weeks than by one week. To check the internal validity we used the additional paired comparison. For this paired comparison the predicted probability of purchase acceleration is 0.35. The actual percentage that chose purchase acceleration was 0.37.

In our simulation we need a probability of purchase acceleration for each value of δ. Since a probit model defines the shape of the distribution function, we are able to do so by interpolation and extrapolation for all values of δ for respondents in G_1. However, we do not have a model for G_2. We approach this problem by making the bold assumption that the parameter values, except for the intercept, are equal for both groups. For G_2 we choose the intercept in such a way that the probability of purchase acceleration up to three weeks equals one. For the simulation this means that we have a mixture model with

Table 7.1: Probit model estimates, with standard errors between parentheses, for purchase acceleration (8 observations for each of the 78 respondents)

Constant	-0.278	(0.230)
Price discount	0.370	(0.026)
Time to planned purchase (δ)	-0.978	(0.087)
Savings plan DM	0.545	(0.127)
log-likelihood	-277.29	

two segments (G_1 and G_2). This is incorporated in our simulation by assigning each individual to G_1 or G_2 with probability $\frac{78}{141}$ and $\frac{63}{141}$, respectively.

We also used a conjoint experiment with paired comparisons for the DM-product choice. In each comparison, the choice is between making a purchase in a store and through direct marketing. The price in the store equals NLG 40. The attributes in the experiment are price of DM product, delivery time, savings plan of the store, and savings plan of DM organization. Table 7.2 presents the estimates of the random effects probit model. The signs of the coefficients are as expected. The choice of a DM product is positively affected by a price discount and a savings plan of the DM organization. Delivery time and a savings plan of the store have a negative effect on DM-product choice. The DM-organization's savings plan has a stronger effect than the store's savings plan. This may indicate that the switching behavior between DM organizations is smaller than that between stores. Thus, a savings plan of a DM organization is more valuable to the individual. The large value of σ_μ^2 indicates that there is much heterogeneity among individuals. That is, some individuals are much more inclined to buy the DM product than others. The respondents had to evaluate a particular paired comparison twice (albeit presented in different ways). Of the 141 respondents 86.5% chose the same option both times. This means that the internal validity of the approach is satisfactory. We distinguished two groups in the purchase acceleration experiment, one with variation in their choices (G_1) and one without (G_2). To see whether these groups differ with respect to DM-product choice, we estimated the random effects probit model for the two groups separately. On the basis of a likelihood ratio test we concluded that there was no significant difference.

Table 7.2: Random effect probit model estimates for DM-product choice (9 observations for each of the 141 respondents), with standard errors between parentheses

Constant	-0.908	(0.032)
Price discount	0.221	(0.030)
Delivery time	-0.442	(0.014)
Savings plan DM	0.320	(0.019)
Savings plan store	-0.117	(0.024)
σ_μ^2	2.463	(0.024)
log-likelihood	-436.54	

Simulation results

Figure 7.4 shows a typical simulation result of the integrated model ($w =$ NLG 20, $c =$ NLG 2.5, $I = 1000$). It depicts the average expected profit per individual for different frequencies and price discounts. The frequency varies from one to ten mailings per year. As expected, profits first increase and later decrease with frequency. When the price discount increases from five to ten guilders, the optimal frequency increases from four to five mailings. If the price discount equals zero, the optimal frequency is equal to zero. However, we may not conclude from figure 7.4 that a rise in the price discount always causes an increase in the optimal frequency. The fact is that there are two effects. First, the probability of buying the DM product increases. Higher probabilities imply higher frequencies. Secondly, the probability of purchase acceleration increases, which implies a decrease in the optimal frequency. The net result of these two effects is not unambiguous.

Another promotion tool for the DM organization is the savings plan. We also performed a simulation with this variable. The results are similar to the outcomes of a simulation with a price discount. The only difference is the absolute value of expected profits. The similarity is not an unexpected result, since the coefficients of a savings plan, in the probability model for purchase acceleration and for DM-product choice, have the same sign as those of a price discount. To examine the effect of the borrowing rate we ran simulations for the two extremes, viz. borrowing rate equal to zero and one, respectively. Simulation results indicate that there is hardly any difference between these

Figure 7.4: Effect of discount on expected profits for different frequencies (per year)

two extremes. In other words, the lack of information is not harmful in this case.

We assumed that there is an interval in which respondent i takes the DM product into consideration (θ_i). There are two ways to implement this interval in the simulation: 1) the same value is taken for each individual, i.e. the average value; 2) for each individual we draw a θ_i from the estimated probability density function. The optimal frequency and the expected profit differ depending on the approach taken. Expected profits are slightly larger for a fixed θ. An additional simulation showed that a rise in the average θ implies that the optimal frequency increases. Even though the optimal frequency changes with another parameter value or assumption of θ, the simulations indicate that the effect on profits is small. In other words, the results are not very sensitive to this parameter value or assumption.

7.7 Discussion and conclusion

Traditional direct mail research focuses on optimization of aspects associated with one particular direct mail campaign. A major drawback of this research is that it focuses on a short-run criterion rather than on a long-run criterion

such as the lifetime value. In order to employ the LTV to its full extent, it should be used for decisions regarding creating, developing and maintaining relationships. The ultimate goal should be an individual-based normative direct marketing strategy. The strategy should depend on past purchase behavior and on forward-looking data. The normative aspect implies that the resulting strategy indicates when an individual should receive a mailing and what kind of mailing that should be. Unfortunately, such a fine-tuned strategy based on the LTV will be quite hard to realize, since (1) this strategy should result from a complex structure, and (2) there is a lack of specific individual data. There is, however, considerable room for improvement of various aspects of such a strategy. One of these aspects, which we examine in this chapter, is the question with what frequency the organization should employ its direct marketing activities.

We proposed a method that seeks the optimal frequency of direct mailings for frequently purchased consumer goods. The method is based on the idea that the consumers' purchase behavior is the basis of the supplier's direct marketing behavior. The method is operationalized by specifying a model that describes the decisions an individual has to make before the purchase of the DM product. The proposed model decomposes the purchase behavior in the timing of purchases, which involves interpurchase time and purchase acceleration, and the DM product choice. The model takes all the purchases of the individuals in the product category into account, and explicitly specifies the individual's decision to buy the product through DM or in a regular store. The model is used to simulate the decisions of an individual for various frequencies of DM activities. As a result, the optimal frequency can be determined. We illustrated the method with an application for which the input parameters were obtained by a questionnaire and conjoint analyses. It demonstrates that the proposed method is a relatively easy way to determine the optimal frequency of direct mailings.

There are various limitations that should be recognized. These limitations refer to: (a) specification of the maximization problem; (b) specification of the consumers' purchase behavior, and (c) data collection.

There are several aspects related to the specification of the maximization problem. We consider the simple situation in which the DM organization solely focuses on the optimal frequency for a fixed group of individuals. Thus, target selection does not play a role, and each individual will receive the mailings with the same frequency. An obvious extension would be to identify segments and to determine the optimal frequency for each segment. Another stringent

assumption is that the individual chooses between purchasing the product either in the store or through DM. Hence, DM activities are not used to boost the incentive to purchase the product in a store (e.g. with coupons). Finally, we considered only the purchase of one product at a time. Thus, we did not pay attention to the quantity decision or to mailings that offer various products (e.g. catalogues).

A drawback of our model specification is that it does not take into account the possible long-term impact of DM activities. In other words, DM-product choice is independent of the number of mailings the individual has received. Furthermore, impulse purchases are only accounted for to a limited extent.

The structure of the proposed model can be modified in various ways. For example, it is possible to specify other distribution functions and to incorporate heterogeneity into the model in an alternative manner. In the present model, heterogeneity is included by an individual specific intercept in the probit model and in the interval in which the DM product is taken into account. However, we assumed that these forms of heterogeneity are independently distributed. Moreover, we did not assume heterogeneity in the slope parameters.

Some of the limitations of the data collection of our illustration have already been discussed. In particular the way in which we obtained the data of the past purchases is open for discussion. However, it should be realized that our objective was solely to collect data on interpurchase times.

The choices of the various aspects of these three components obviously play a crucial role in the simulation. Consequently, the reliability of the results is enhanced by the validity of the assumptions. We wish to emphasize, however, that several aspects of the maximization problem and the specification and implementation of the model are easy to adapt. This implies that it is straightforward, at least in principle, to apply the proposed method to more complex situations. Moreover, it implies that the sensitivity of the assumptions can be easily explored. We have demonstrated this by examining the effect of (1) the borrowing rate and (2) the distribution of the interval in which the individual takes the DM product into consideration.