Tijdschrift voor Economie en Management
Vol. XLVIII, 3, 2003

Predicting Mail-Order Repeat Buying: Which Variables Matter?

by D. VAN DEN POEL*

Dirk Van den Poel
Department of Marketing, Ghent University, Gent

ABSTRACT

In this study, we propose a customer-oriented conceptual model of segmentation variables for mail-order *repeat* buying behavior. We investigate (1) from a theoretical perspective what customer-related variables should be included in response models for modeling repeat purchasing, and (2) empirically validate how these variables perform for predictive purposes. We use binary logit modeling. Our results confirm that all three traditionally-used R(ecency), F(requency) and M(onetary value) variables are very important in predicting who is going to purchase during the next mailing period, with frequency being the most important one. In total, they account for 50 % of the 'room for improvement' in terms of AUC performance. However, next to the RFM variables, our findings suggest that at least three other variables significantly increase the predictive performance of the models: (1) credit usage, (2) length of relationship, and (3) general mail-order buying behavior. Depending on the context of the specific company use of these additional variables may translate into millions Euro of additional profit. Furthermore, we conclude that buying additional data from external sources is not economically justified when predicting *repeat* purchasing behavior.

* The author would like to thank Joseph Leunis and Marnik Dekimpe for their insightful comments on a previous version of this text.

I. INTRODUCTION

The success of a database-driven (mail-order) marketing campaign mainly depends on the customer list to which it is targeted (Bult and Wansbeek (1995); Bult et al. (1997)). Response modeling for database marketing is concerned with the task of modeling the customers' purchasing behavior. The information at the level of the individual consumer is typically used to construct a response score (based on the estimated model), which is subsequently used to rank the customers. Customers are then mailed up to a certain mailing depth (i.e., percentage of the total mailing list of the customers ranking highest).

The substantive relevance of response modeling comes from the fact that an increase in response of only one percentage point can result in substantial profit increases (Baesens et al. (2002)). Given a tendency of rising mailing costs (Hauser (1992)) and increasing competition[1], we see an increasing importance for response modeling. Improving the targeting of the offers may indeed counter these challenges by lowering non response. Moreover, from the perspective of the recipient of the (direct-mail) messages, mail-order companies do not want to overload consumers with catalogs. The importance of response modeling to the mail-order industry is further illustrated by the fact that the issue of improving targeting was among the top three concerns with 73.5 % of the catalogers in the sample mentioning this issue (DMA (1998), p.99).

In this article, we present an empirical study including many segmentation variables for mail-order repeat purchasing behavior. First, we focus on the traditionally-used R(ecency), F(requency), and M(onetary value) variables (cf. infra). We investigate the predictive performance of the combined use of the three RFM variables in predicting repurchase behavior (RQ1a[2]), followed by a study of the relative importance of recency, frequency, and monetary value (RQ1b). Second, we go beyond the traditional RFM variables to include other segmentation variables that could be included in models predicting mail-order repeat purchase behavior (RQ2). This research question is justified by the absence of a conceptual framework of segmentation variables for mail-order buying behavior in the existing literature. It structures the variables which have been 'loosely suggested' by several authors, but which have never been thoroughly investigated in terms of their *additional* predictive power over the traditional (RFM) variables.

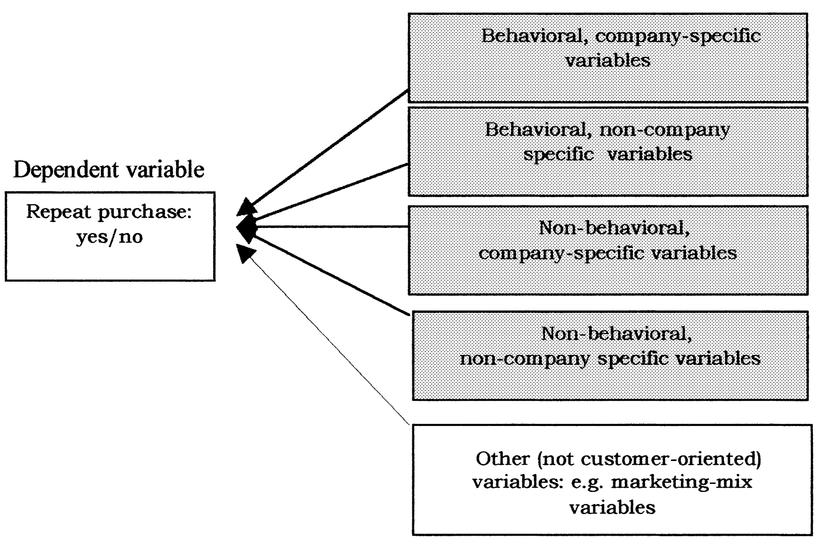
This study is further justified by the fact that many of the additional variables have to be purchased or collected at a substantial cost.

II. DIRECT-MAIL PATRONAGE BEHAVIOR: A CONCEPTUAL MODEL OF SEGMENTATION VARIABLES

The objective of this study is to develop a general model of *repeat* purchasing behavior in a direct-marketing setting. In Figure 1, we present a general overview of the model we propose for predicting mail-order repeat buying. We try to predict, using four types of independent variables, whether a customer will purchase during the next mailing period, i.e., we try to predict purchase incidence within a fixed time interval. In doing so, we address the fundamental mailing problem discussed in the introduction.

FIGURE 1

Predictive modeling of purchase incidence



An overview of the possible predictor variables is provided in Table 1. We only include customer-oriented variables in our conceptual model of segmentation variables (greyed boxes in Figure 1), which implies that offer-specific features and marketing-mix variables such as promotions or general advertising are not incorporated. Some customer-oriented variables related to marketing-mix variables (e.g. price sensitivity) are not explicitly taken into account.

We distinguish between two dimensions in the categorization of predictors (cf. Figure 1): (1) company versus non-company specific variables; (2) behavioral versus non-behavioral predictors of direct-mail consumer behavior. We use these two dimensions for the following reasons: (1) non-company specific variables generally have to be purchased from external vendors. Therefore, it is likely that the cost will be higher than company-internal information. Since the benefits have to be traded off with their costs, this represents a relevant dimension when classifying possible predictors; (2) behavioral variables usually correlate more strongly with future purchase behavior than non-behavioral predictors.

TABLE 1

Overview of customer-oriented segmentation variables for modeling mail-order repeat buying

	Behavioral	Non-behavioral
Company Specific	*Recency* *Frequency* *Monetary value* Length of relationship Type/category of product Source of customer Customer/company interaction	Customer satisfaction
Non-company Specific	General mail-order buying behavior	Benefit segmentation Socio-demographics

Cullinan (1977) is generally credited for identifying the three (RFM) variables most often used in database marketing modeling: recency, frequency and monetary value (Bauer (1988)). Since then, the literature has accumulated so many uses of these three variables, that there is overwhelming evidence from existing studies (cf. Table 2) that the RFM variables are the most important set of predictors for modeling mail-order purchasing. These traditional variables fall into the category of behavioral and company-specific variables, shown in the upper left-hand box of Table 1.

Apart from the traditional RFM variables, other variables fall into the category of behavioral and company-specific variables, which can also be derived from company-internal records. These will be discussed in Section VI.A. Several studies have shown (cf. Table 2) that company-specific behavioral characteristics provide very good predictors for modeling mail-order repeat purchasing. This fact combined with the reality that most behavioral company-specific variables can be gathered at very low cost from company-internal records, makes them the most interesting set of predictors. In contrast, non-behavioral variables and behavioral but not company-related information have to be purchased externally or collected by sending questionnaires to the customers. Given this substantial cost, the important question arises whether these variables add any value in predicting repurchase behavior. These variables will be discussed in Sections VI.B – VI.D.

III. RESEARCH QUESTIONS

When reviewing the vast amount of relevant literature (cf. Table 2), it becomes evident that only very limited attention has been devoted to selecting the right set of variables to include into the model of mail-order repeat buying. In fact, the focus of these articles lies entirely on selecting the appropriate *modeling technique*. In contrast, this study focuses on the issue of what variables to include in predicting repeat purchase behavior by mail-order. Research questions RQ1a and RQ1b focus on the traditionally-used RFM variables, while RQ2 addresses the issue of including other predictors into the response model.

The first research question addresses the issue of how good a model performance can be achieved by the RFM variables. This is motivated by the fact that before adding new variables, we have to determine what performance can be obtained by the RFM variables that are traditionally used in order to benchmark the additional performance. Even though most studies (cf. Table 2) include some combination of the recency, frequency and monetary value (RFM) variables, no study to date investigates the importance of the combined use of RFM variables. Formally:

RQ1a: What is the total performance of the combined use of the three RFM variables in predicting repurchase behavior?

Second, the *relative* importance of the three R(ecency), F(requency), and M(onetary value) components has never been

TABLE 2

Academically reviewed studies about mail-order repeat-purchase modeling

Reference	Independent variable						Dependent variable[3]			Context of application		
	R	*F*	*M*	*Length of relationship*	*Other behavioral var.*	*Socio-demographic*	*Binary*	*Binary & Amount*	*Binary & Timing*	*Fund-raising*	*Catalog*[4] *(general)*	*Catalog or specialty mailing*
Baesens et al. (2002)	X	X	X	X	X	X	X				X	
Berger and Magliozzi (1992)	X	X	X			X	X				X	
Bitran and Mondschein (1996)	X	X	X					X			X	
Bult and Wittink (1996)	X		X			X		X		X		
Bult (1993a)	X	X					X					X
Bult (1993b)			X		X	X	X					X
Bult et al. (1997)	X	X	X	X	X	X	X			X		
Gönül and Shi (1998)	X	X					X				X	
Kaslow (1997)	X	X	X	X	X	X			X			
Levin and Zahavi (1998)	X	X	X		X	X	X	X				X
Magliozzi and Berger (1993)	X	X	X				X				X	
Magliozzi (1989)	X	X	X				X					
Thrasher (1991)	X	X			X	X	X					X
Van den Poel and Leunis (1998)	X	X		X	X	X			X			X
Van den Poel (1998)	X	X	X	X		X	X				X	
Van der Scheer (1998)			X			X		X		X		
Zahavi and Levin (1997)	X		X		X	X	X					X

thoroughly investigated. Since these variables measure different aspects of the same underlying behavioral customer transactions, we may question whether they should (all three) be included in a model. As such, we may wonder whether one (or even two) of them is redundant. In the literature, we find some general statements such as "in general the frequency variable is the most important of the set of three RFM predictors" (Nash (1994)), but due to the proprietary nature of most studies, no detailed results are reported (Bass and Wind (1995)). This hiatus leads to the following research question:

RQ1b: What is the relative importance of recency, frequency and monetary value in predicting repurchase behavior?

Finally, several variables have been added to the RFM variables in specific implementations, but they have never been systematically investigated in one setting. Given that some of these variables have to be purchased from external sources justifies the question whether these variables add any predictive power to the response model.

RQ2: How much predictive power do additional, i.e. non-RFM, variables offer in modeling mail-order repeat purchasing?

IV. CHOICE OF DEPENDENT VARIABLE

In this problem formulation, the focus is predicting, based on (a subset[5] of) the variables suggested in Table 1, whether a customer will purchase during the next mailing period from *any* product category, i.e. one tries to predict the purchase incidence within a fixed time interval (Bult (1993a)). This choice is motivated by the fact that the majority of previous research in the direct marketing literature focuses on the purchase incidence problem (cf column about dependent variable choice in Table 2), because this is exactly the setting that mail-order companies are typically confronted with: they have to decide whether or not a specific offering will be sent to a (potential) customer during a certain mailing period.

V. RFM VARIABLES FOR PREDICTING MAIL-ORDER PATRONAGE

When considering the overview in Table 2, it is clear that there is overwhelming evidence for the use of all three RFM variables in purchase-incidence models. Due to space constraints we refer to Van den Poel (1999) for details regarding the specific operationalizations of the variables.

Recency has been found to be inversely related to the probability of the next purchase (Cullinan (1977); Shepard (1995)), i.e., the longer the time delay since the last purchase, the lower the probability of a next purchase. It is clear that this relationship can only hold for broad assortment catalogs including sufficient product categories with varying interpurchase times (Ganzach (1993)).

There is overwhelming evidence that heavier buyers show greater loyalty as measured by their repurchase probabilities (Morrison (1966); Lawrence (1980)). Also in the specific context of direct mail, it has generally been observed that multi-buyers (buyers who already purchased several times, i.e., *frequency*) are more likely to repurchase than buyers who only purchased once (Bauer (1988)). Although no detailed results are reported because of the proprietary nature of most studies, the frequency variable is generally considered to be the most important of the RFM variables (Nash (1994)).

The volume of purchases a consumer makes with a particular mail-order company (measured in monetary terms, which enables us to obtain a global measure across product categories, i.e. *monetary value*) is a measure of usage which has been an important behavioral segmentation variable in several studies (Kotler (1994))[6]. In the direct-marketing literature, the general convention is that the more money a person has spent with a company, the higher his likelihood of purchasing the next offering (Levin and Zahavi (1996)).

VI. NON-RFM VARIABLES FOR PREDICTING MAIL-ORDER PATRONAGE

In this section, we investigate what other variables could represent important information in the prediction of future repeat-purchase behavior. In the following Sections VI.A – VI.D we give an overview of the variables which, on theoretical grounds, could be considered as predictors for modeling mail-order repeat buying. The variables are grouped according to the two dimensions introduced earlier in Table 1: (1) company specific or not; and (2) behavioral or non-behavioral.

A. *Company Specific, Behavioral Variables*

The RFM variables themselves are company-specific, behavioral predictors that summarize a customer's purchase history in three key statistics. In this section, we investigate the theoretical grounds for other variables derived from this same purchase history, which could potentially add to the forward-looking, predictive power of the model.

1. Length of the Relationship

Relationships have been throughly investigated in diverse scientific disciplines: (1) Social psychology focuses on (intimate) interpersonal relationships (Rusbult et al. (1982)); (2) Economics investigate, among others, interorganizational relationships (Williamson (1975)); and (3) the organizational behavior literature studies, e.g., employment relationships (Doeringer and Piore (1971); Schmittlein and Morrison (1981) and (1983)). All three areas use interdependence theory (Kelley and Thibaut (1978)) or the derived 'investment model' (with the notion of transaction-specific investments) to explain the continuation/dissolution of relationships. From these diverse areas, we obtain evidence that the duration of a relationship may have predictive power with regard to the continuation of the relationship. In the social-psychology literature, Simpson ((1987), p.684) states that "Relationship duration also ought to prognosticate relationship stability". We argue that the 'investment' argument also holds for the relationship between a customer and the general mail-order company because a customer has to learn how to interact (to his benefit) with the company.

2. Type/Category of Product

Kestnbaum (1992) suggests to replace RFM by the new acronym FRAC (frequency, recency, amount, and category of product), in order to explicitly include the type of product bought before as an additional predictor of mail-order patronage behavior. However, very little empirical evidence is publicly available to corroborate this extended focus.

3. Source of the Customer

How a prospect became a customer of the mail-order company often continues to have an influence in the later stages of the buying process (Nash (1995)). The source effect is also supported by Bhattacharya's (1998) findings in the context of membership programs.

Customers can be acquired in a number of different ways by a mail-order company. Based on company records, we were able to distinguish between the following types: 1. member introduces member; 2. child from a member parent; 3. spontaneous requests; 4. internal mailing lists (other catalogs); 5. rented mailing lists. We hypothesize that referrals by word-of-mouth (categories 1 and 2) and customers who already have a relationship with the company (category 4) have a higher probability of repurchase than customers lacking this information (categories 3 and 5).

4. Customer/Company Interaction

Contact-information consists of several different types: (1) information inquiries (pre-purchase information), (2) orders (purchasing), and (3) complaints (post-purchase). It is on the latter kind of information we will focus in this section. Complaint handling takes on a central role in maintaining customer loyalty, and is a key element of many quality-improvement programs (Blodgett et al. (1997)). In particular, Schibrowsky and Lapidus (1994) argue that complaint management is a key element in building and maintaining customer relationships for direct-marketing companies, because of the lack of a personal relationship through face-to-face contact with salespeople.

The consumer-complaint behavior (CCB) literature has focused mainly on the determinants of complaint intentions and complaining behavior (Jacoby and Jaccard (1981); Morel et al. (1997)). Nevertheless, some studies investigate the relationship between complaining behavior and repeat-purchase behavior (Richins (1983)). Based on these studies, we hypothesize that both information requests and complaints are signs of more intense relationships, which result in a higher probability of repurchase.

Customer dissatisfaction may not only be revealed by formal complaints, but also by returning items. Consumers returning goods can be considered a form of actual complaint behavior. In analogy with the findings for complaint behavior, we hypothesize a positive impact of returns on repurchase probability, because we consider them to be signs of a more intense relationship[5].

B. *Company Specific, Non-Behavioral Variable: Customer Satisfaction*

After discussing company-specific *behavioral* variables in the previous section, we now focus on customer satisfaction as a company-specific, but *non-behavioral* variable. Behavioral learning theory (Rothschild and Gaidis (1981)) states that behavior that is positively reinforced is more likely to recur than non-reinforced behavior. When applied to direct marketing, we can state that the probability of repeat behavior will increase if the total buying experience meets or exceeds the expectations of the consumer with respect to the performance (Parasuraman et al. (1985)). One way of measuring the extent to which the purchasing behavior was positively reinforced is tracking customer satisfaction. In the domain of relationship marketing (and the related concept of customer retention), customer satisfaction data has become an essential part of the marketing information system (Grönroos (1990) and (1995)).

All theoretical and empirical evidence suggests that a positive relationship exists between a satisfying experience and the probability of repurchase (e.g. LaBarbera and Mazursky (1983); Bolton (1998)). Following all abovementioned findings, we hypothesize a positive relationship between customer satisfaction and likelihood of repeat purchase.

C. *Non-company Specific & Behavioral Variable: General Mail-Order Buying Behavior*

When company-specific data are sparse, e.g. when the person only recently became a customer at a particular mail-order company, knowledge about the customer's general mail-order buying behavior (at competing companies) may be valuable in predicting future purchasing behavior, because it may provide an indication to the mail-order company of the sales potential. We hypothesize that the higher the general level of mail-order buying from any catalog company, the higher the probability of repeat purchase at the specific mail-order company.

D. *Non-company Specific & Non-Behavioral Variables*

1. Benefit Segmentation

The general concept of benefit segmentation was introduced in the 1960s (e.g., Yankelovich (1964)). Advocates of this concept argue that the benefits people seek in products are the basic reasons for the heterogeneity in their choice behavior. Therefore, benefits are relevant bases for segmentation. Other studies have shown that benefit segments are identifiable and substantial (Myers (1976)), and differ in brand purchase behavior (Wilkie (1970)).

Customers differ in terms of the benefits sought by purchasing through a direct-marketing channel. Convenience is often cited as one of the major driving forces for direct-marketing patronage behavior (Gehrt et al. (1996)). The fact that credit can be obtained in an easy, non-bureaucratic way at most mail-order companies is an important example of financial convenience. Once a customer is approved, a credit line is automatically opened. The customer still has the possibility to pay the whole amount, but no questions are asked if he/she keeps paying the monthly amount which is automatically proposed by the mail-order company. Several studies have shown that the existence of an easy credit line (provided by the company or by credit cards) does facilitate spending and also increases the amounts being spent (Feinberg (1986)). In the domain of direct marketing, Cunningham and Cunningham (1973) revealed that active in-home shoppers have a more positive attitude toward the use of credit. Nevertheless, most attention in this literature has been devoted to identifying variables to predict default, as in Peltier and Schibrowsky (1997). No research to date has investigated the influence of the use of a credit line for a customer on his future purchasing behavior with the direct-mail company. Managers of the mail-order company emphasized the importance of this predictor. In the empirical part of this study, we only had information on this construct at our disposal.

2. Socio-Demographic Variables

Several studies have been conducted to characterize the demographic profile of the non-store shopper. Burnett and McCollough (1994) report in their overview article that the following variables were confirmed to have a significant impact on non-store shopping behavior: wife employed, number of children at home, household income, occupation, gender, city location, marital status, home ownership. The age and education variables did not result in unambiguous conclusions, as both significantly positive and negative effects have been reported in the literature. Akaah et al. (1995) find no significant relationship between age and intention to purchase by mail-order, nor between income and purchase intention by mail. Previous studies did not investigate whether these demographic variables provide any additional predictive value over transactional data.

Indeed, given the availability of very detailed transaction data, we can wonder how important demographic data still can be. Previous research does not offer conclusive evidence. For brand choice, for example, several publications including Kalyanam and Putler (1997) have shown some added value of including demographic characteristics in the analysis, while Rossi et al. (1996) have shown that demographic variables contain very little additional information in a target couponing setting. The inconclusive previous findings may be due to the fact that socio-demographic variables are only indirect explanatory variables, which may result in little net effect on the predictive performance of repeat-purchasing behavior.

VII. METHODOLOGY

In order to address the research questions of Section III, we specify and justify in this section: (1) the specific modeling technique for purchase-incidence modeling, (2) the model structure and the level of parameterization, (3) the evaluation criteria used to assess the improvement in predictive accuracy, and (4) the procedure for variable introduction.

A. *Modeling Techniques for Purchase Incidence*

We address the research questions in the context of a purchase-incidence model. In essence, purchasing or not is a binary decision problem (two-class classification), which we model by means of binary logit model. This choice is justified by the following reasons: (1) logit modeling is well-known, conceptually simple and frequently used in marketing (Bucklin and Gupta (1992)) both at the aggregate market level (Bultez and Naert (1975)), at the segment level (Mela et al. (1997)), and at the level of the individual consumer (Jones and

Zufryden (1980)); (2) the ease of interpretation of logit is an important advantage over other methods (e.g. neural networks); (3) logit modeling has been shown to provide good and robust results in general comparison studies; and (4) more specifically in database marketing, it has been shown by several authors (e.g., Van den Poel (1998)) that logit modeling may even outperform more sophisticated methods.

The binary logit model is used to approximate a probability P_i by the following expression (Aldrich and Nelson (1984)):

$$P_i = \frac{1}{1 + e^{-\sum_{j=1}^{n} b_j X_{ij}}} \qquad (1)$$

Whereby:

- P_i represents the *a posteriori* probability[8] of a repeat purchase for customer i;
- X_{ij} represents independent variable j for customer i;
- b_j represent the parameters (to be estimated);
- n represents the number of independent variables.

Once the parameters b_j are estimated using maximum likelihood, this expression allows us to obtain a conditional probability estimate of purchase.

B. *Model Structure and Level of Parameterization*

We combine predictors in an additive way in our model of repurchase incidence. We introduce the variables into the model as they are presented in Sections V–VI.D, i.e. no additional transformations are considered. Moreover, we pool data across individual customers. This may introduce some 'hidden' aggregation bias as discussed in Bass and Wittink (1978). Jones and Landwehr ((1988), p.55) present an approach to account for unobserved heterogeneity in logit models, but they argue that their procedure is less appropriate for predictive purposes. Despite its disadvantages, we pool across individual customers because (1) it is the most commonly used procedure, (2) insufficient data is available to estimate a separate model for each customer, and (3) Brodie and de Kluyver (1987) argue that intermediate levels of parameterization (e.g. where models are estimated by segment) only seldomly improve predictive performance[9].

C. *Evaluation Criteria*

This section introduces and justifies the choice of two performance criteria:

- Percentage correctly classified (accuracy) at the 'economically optimal' cutoff purchase probability (PCC);
- Area under the receiver operating characteristic curve (AUC).

Both criteria are *predictive* in nature, as opposed to more traditional evaluation criteria such as (1) the hypothesis-testing approach (investigating the statistical significance of parameter estimates); or (2) the resubstitution rate (PCC based on the estimation sample). The choice of two 'predictive' evaluation criteria (PCC & AUC) is justified by the following arguments: (1) we are mainly interested in knowing what the additional predictive power is of an additional variable over a model which does not contain this variable, because of research questions RQ1b, RQ2 (cf. Section III); (2) the primary goal of response modeling is to predict purchase behavior (Berger and Magliozzi (1992)); (3) when using correlated predictors, the hypothesis testing approach may suffer as multicollinearity may cause inflated variance of the estimates (Mason and Perreault (1991)), which may in turn lead to insignificant parameters; the latter authors emphasize that the predictive approach does not suffer from this deficiency; (4) the predictive approach to model selection has gained substantial importance in econometrics at the expense of the hypothesis-testing approach (Geisser and Eddy (1979)); the increased availability of larger datasets has played a catalyzing effect in this respect, because in large samples almost all coefficients become statistically significant (Granger (1998)); and (5) using the predictive evaluation criterion entails the use of a separate test (holdout) sample, which follows Morrison's (1969) warnings against the upward bias that results from classifying the same individuals used to calculate the classification model (i.e. in the estimates of the resubstitution rate).

1. Percentage Correctly Classified

The classification method is used to rank the potential buyers according to their predicted likelihood of repurchase (from most likely to least likely buyer). This measure is provided by the (logit) modeling technique as the 'a posteriori' probability. When an absolute cutoff value is chosen, all customers with an 'a posteriori' probability of repurchase higher than this absolute cutoff are classified as buyers, and all customers with a lower likelihood of repurchase are labeled as non-buyers. The result of a classification can be summarized in the following classification table or confusion matrix (Morrison (1969)):

TABLE 3

Confusion matrix

		Predicted status	
		Buyer	Non-buyer
True Status	Buyer	True Positive (TP)	False Negative (FN)
	Non-buyer	False Postitive (FP)	True Negative (TN)

The following meaningful measures can be extracted from this table (Bradley (1997)):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{3}$$

$$Specificit\,y = \frac{TN}{TN + FP} \tag{4}$$

Following these definitions, we observe that s*ensitivity* represents the proportion of event observations that the model predicts to be events (number of true positives divided by the number of events, i.e., a measure of accuracy for predicting events), and *specificity* is defined as the proportion of non-event observations that the model predicts to be non-events (number of true negatives divided by the total number of non-events, i.e., a measure of accuracy for predicting non-events).

It is important to notice that the abovementioned performance measures (PCC, accuracy, sensitivity and specificity) are dependent on the chosen cutoff value, and do not provide an indication as to how these measures vary as the cutoff value is varied. Since all these measures require the choice of a cutoff value, we need a criterion to determine an absolute or relative cutoff level[10]. We use an absolute cutoff value because sound economic reasons can be given for its use.

For a specific company/catalog setting, we can have an *a priori* judgement about the optimal absolute cutoff point by considering the past economics of the mailing. When the objective is to maximize total profits, we know from microeconomics that the optimal decision rule is to mail up until the point where the incremental revenue derived from the mailing equals the incremental cost incurred by sending this additional mailing. The latter can be assumed to be constant within certain ranges, and can be easily estimated, because these costs are clearly identifiable, and mainly consist of the variable mailing production cost and postage cost. The former can be decomposed into the probability of a purchase multiplied by the average contribution given that a purchase takes place. If we assume the latter to be constant (Bult and Wansbeek (1995)), we can derive the economically optimal estimate for the cutoff value of the probability of purchase[11]:

$$\begin{gathered} \textit{Cutoff value} = \textit{Minimal}\ \text{probability of purchase} = \\ \frac{\textit{Incremental}\ \text{cost of one additional mailing}}{\textit{Average}\ \text{contribution of a buyer}} \end{gathered} \tag{5}$$

The percentage correctly classified at this 'economically motivated' specific cutoff point will represent our first measure of goodness of fit when comparing the performance of models containing alternative predictors. However, this measure does have the following disadvantages:

The specific estimates of incremental cost and incremental revenue (given that a purchase is made) are estimates (usually based on previous year quantities), and may not exactly correspond to the economics during the actual campaign. Moreover, this approach does not take into account the heterogeneity in average contribution of customers. This, therefore, results in only an estimate or approximation of the truly optimal cutoff value. However, a sensitivity analysis shows that PCC is rather robust with regard to changes in the cutoff value.

There is a large heterogeneity in cutoff values used both across mail-order companies and within the same company for different catalogs. Given the proprietary nature of this information, no such data are publicly available.

2. Receiver Operating Characteristic Curve and AUC

Invariance of the performance criterion with respect to the selected cutoff value can be achieved by considering the curve which plots sensitivity (vertical axis) versus one minus specificity (horizontal axis) for all possible cutoff values. The former is also called hit percentage, the latter is also called the 'false-alarm probability' (Green and Swets (1966)). This curve is named a receiver operating characteristics (ROC) curve. We refer to Green and Swets (1966) and Swets (1979) for more details. These authors have shown that the predictive accuracy of a classification procedure (such as logit modeling) can be measured by the area under the ROC curve (AUC)[12]. Hanley and McNeil (1982) provide an intuitive interpretation of the AUC, which is based on the equivalence of the AUC to the Mann-Whitney or Wilcoxon statistic. They show that the AUC represents the probability that a randomly chosen positive example (rated as a buyer) is correctly rated (or ranked) higher than a randomly selected negative example (rated as a non-buyer). This again illustrates that this performance measure is not dependent on the choice of a cutoff value. The AUC statistic ranges from a lower limit of 0.5 for chance (null model) performance to an upper limit of 1.0 for perfect performance (Green and Swets (1966)).

Several authors (including Hanley and McNeil (1982); DeLong et al. (1988)) have investigated the statistical characteristics of ROC curves in great detail, for example to derive tests that can assess the statistical difference between two or more AUCs. When two or more empirical ROC curves are constructed based on different sets of variables reflecting the same individuals, statistical tests for the difference between these curves should take into account the correlated nature of the data (DeLong et al. (1988)). Hanley and McNeil (1982) derive a test statistic (critical ratio z) to assess whether the areas are significantly different. In this study, we use the DeLong et al. (1988) non-parametric test to determine whether the areas under the ROC curves are significantly different, because this test does not necessitate to make distributional assumptions.

Since neither of the two criteria (AUC and PCC) is clearly superior, we will use both performance measures, and assess their convergent validity. One disadvantage is shared by both performance criteria, namely, the assumption of equal opportunity costs of misclassification. Both PCC and AUC weigh the opportunity cost of misclassifying a buyer as a non-buyer and the cost of misclassifying a non-buyer as a buyer equally. It is easier to incorporate the issue of unequal misclassification costs into the PCC criterion than into AUC. The former is adapted by changing percentages into expected values, i.e., the probability of a misclassification is multiplied by the cost of misclassification (and this for both buyers and non-buyers). Both performance criteria (AUC and PCC) will be calculated on a test or holdout sample, which only consists of observations not used during model estimation, and which is half the size of the total sample.

D. *Procedure for variable introduction*

The importance of a variable has to be evaluated given the presence of other predictors in the model. Given the large number of potentially important variables (see Table 1), we follow a sequential procedure which is conceptually similar to the one in Robertson and Wind (1980). More specifically, we will first investigate models containing the following predictors: R(ecency), F(requency), M(onetary value), RF, RM, FM, RFM. This allows us to investigate the first research question (RQ1a – RQ1b). We start by introducing the RFM variables before any other variable (1) because most previous research (cf. Table 2) cites them as being most predictive, and (2) because they are internally available at very low cost. Once the 'best' RFM model has been identified, its composition is not varied any more. Then, we introduce the non-RFM variables one by one, and investigate whether they provide additional predictive value (cf. RQ2).

We do realize that the one-by-one introduction of variables may lead us to miss some effects, e.g., some variables may only have an impact if others are also present[13]. However, this would lead to a lengthy procedure introducing additional variables two-by-two, three-by-three, etc..

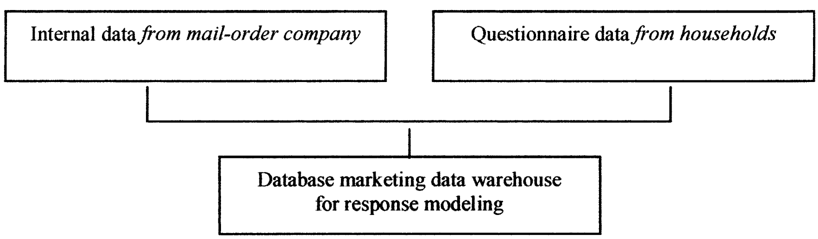
Finally, we investigate the additional effect of introducing multiple predictors simultaneously, but only for those variables, which have shown to contribute to predictive performance in addition to the RFM variables. To avoid testing too many combinations, we introduce them in decreasing order of their additional effect over the RFM variables.

VIII. DATA

We have data on the same customers from two different data sources: (1) From a major European mail-order company, we obtained Belgian data on past purchase behavior at the order-line level, i.e., we know when a customer-purchased what quantity of a particular product at what price as part of what order. This data source enables us to derive detailed company-specific behavioral variables. Moreover, the mail-order company provided credit usage data. (2) We have questionnaire information on each customer, which includes benefit segmentation variables, customer satisfaction, and general mail-order purchasing.

FIGURE 2

Summary of data sources



The transaction information spans several years of past purchase data: July 1993 – December 1997. The model uses the purchase information on the four-year period from July 1993 to June 1997 to calculate the independent variables derived from the purchase history data. The questionnaire was mailed towards the end of June 1997 (together with the new catalog) to all previous customers (about 360,000) measuring variables of Sections VI.B, VI.C, and VI.D, of which 2982 returned a useful questionnaire (response rate of less than 1 %)[14]. Of this group, 168 responding households had never purchased before at the mail-order company. Because of the focus of this study on customer retention, these households were deleted from the sample. This resulted in a total sample size of 2814 customers. Purchase information from the period July 1997 - December 1997 is used as the dependent variable. During this period, 2066 of the 2814 customers (73.4 %) had purchased at least one product. The corresponding percentage is only 37.9 % for the total customer file

(207,059 of the 546,118 customers). This means that our sample contains almost twice as many buyers relative to the population[15]. We therefore adopted a weighted, rather than unweighted logit estimation. In this specific case, the weight of a non-buyer is fixed to 1 (which are under-represented in our sample) and the weight of a buyer is taken to be 0.221 (= (((748 / 0.621) - 748) / 2066), the 2066 buyers of the sample have to be scaled back to represent 37.9 % of the 1205 sample size[16] because the buyers are given a weight of 1).

An analysis of the mailing results of the previous (6 month) mailing period allows us to determine the quantities necessary to estimate the 'economically optimal' cutoff value introduced in Section VII.C.1. Based on data from the previous mailing period, the numerator (in Expression 5) is estimated to be 59.90 EURO, the denominator amounts to 16.42 EURO. This results in an economically optimal cutoff value of 0.274 (= 16.42 / 59.90), which means that only those customers should receive a mailing during the next mailing period which have a purchase probability of 0.274 or higher. In this context, we assume that both the average revenue of a buyer and the marginal cost of sending an extra mailing are constant.

IX. EMPIRICAL FINDINGS

Section IX.A addresses the research questions RQ1a-RQ1b with regard to the predictive performance of the RFM variables. The additional predictive power of *other* variables (RQ2) is investigated in Section IX.B.

A. *Research Question 1: RFM Variables*

RQ1a: What is the total performance of the RFM variables in predicting repurchase behavior?

To evaluate the performance of a given model, we can compare its predictive power with two benchmarks:

- The performance of a *null model*. In terms of PCC, the performance of the null model is determined by the 'proportional chance criterion' (Morrison (1969))[17]. The AUC performance of a null model is 0.5.
- The performance of an ideal or *perfect model*. In terms of PCC and AUC, such model would result in a 100 % correct classification, and an AUC of 1.

The difference between these two benchmark models gives the "room for improvement", and we will express the predictive power in terms of the fraction of the gap between the null model and the perfect model already covered by the model under investigation.

In terms of the AUC criterion, the best RFM operationalization obtains a value of 0.754 on the test sample. This compares to the 'null model' AUC-performance of 0.500 and the 'ideal model' performance of 1.000. This implies that the best RFM operationalization already covers 50.8 % of the "room for improvement" ((0.754 – 0.500) / (1.000 – 0.500)). The PCC performance of the best RFM operationalization is 64.9 %. This compares to the 'null model' proportional chance criterion of 52.9 % (Morrison (1969)), and an ideal-model performance of 100.0 %. This implies that the best RFM operationalization already achieves 25.5 % of the difference between the null-model and the ideal-model performance ((0.649 – 0.529) / (1.000 – 0.529))[18]. These favorable results for the RFM variables corroborate the many academic references to these variables (Table 2). To the best of our knowledge, this specific finding is the first attempt to quantify the importance of the RFM variables.

We can therefore conclude with regard to RQ1a that in terms of AUC about half of the maximum possible improvement is already achieved by the 'best' RFM operationalization[19]. The PCC performance, in contrast, only achieves about one quarter of the maximum possible improvement.

RQ1b: What is the relative importance of recency, frequency and monetary value in predicting repurchase behavior?

First, we report findings with regard to this research question for models using only a single predictor. Second, we discuss models containing multiple predictor variables, and investigate the cumulative effect of different RFM variables. Table 4 reports the findings for both the AUC and PCC criteria. Frequency is the most important of the three RFM variables (an AUC-increase of 0.243 over the null-model performance of 0.500, and a PCC-increase of 0.149 over the null model of 0.529), closely followed by monetary value (an AUC-increase of 0.208, and a PCC-increase of 0.063), and finally, by the recency variable (an AUC-increase of 0.125, again vis-à-vis the 'null model', and a PCC-*decrease* of 0.112). The latter result in terms of PCC highlights that the best recency operationalization is not even able to match the proportional chance criterion.

TABLE 4

Predictive performance of single best predictor

Type	*AUC*		*PCC*	
	Best single predictor	*Value*	*Best single predictor*	*Value*
***R**ecency*	Recency	0.625	Log(recency)	0.417
***F**requency*	No. of different purchases over the last year excluding returned items	0.743	No. of orderlines over the whole history including returned items	0.678
***M**onetary*	Monetary value over the whole history excluding returned items	0.708	Log(Monetary value over the whole history excluding returned items)	0.592

Multiple predictors

The relative importance of each of the RFM variables is also revealed by investigating multiple-predictor models. Table 5 summarizes the results of a search for the best model containing 1, 2, and 3 variables.

TABLE 5

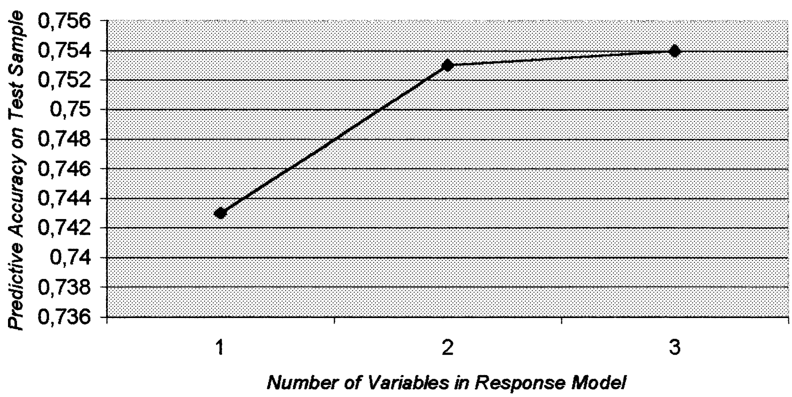
Overview of incremental best-performing models in terms of AUC

Best model containing .. variables	*R, F or M*	*List of variables*	*AUC*
1	F	No. of different orders during the last year excluding returned items	0.743
2	F & M	No. of different orders over the whole history excluding returned items & Log(Monetary value over the last year excluding returned items)	0.753
3	R, F & M	Recency, No. of different order over the whole history excluding returned items & Log(Monetary value over the last year excluding returned items)	0.754

These results again emphasize the importance of the frequency variable (because it is present in all three 'best' models), followed by the monetary value variable, and as such confirm both Nash's (1994) and Shepard's (1997) findings. The AUC performance is also shown in Table 5. and reveals that the recency variable is clearly of lower importance, because the increase in AUC performance is very limited when we move from the two-variable model to the three-variable model.

FIGURE 1

AUC performance of incremental best-performing models



In terms of PCC, we observe in Table 6 that these same models do not show any noticeable increase in performance. This leads to the conclusion that at the optimal mailing depth, the PCC performance of these models is not very sensitive to the variable selection, as long as the frequency variable is included in the predictive model.

TABLE 6

Overview of incremental best-performing models in terms of PCC

Best model containing .. variables	*R, F or M*	*List of variables*	*PCC*
1	F	No. of orderlines over the whole history including returned items	0.678
2	F & M	No. of different purchases over the whole history excluding returned items & Log(Average monetary value excluding returned items)	0.675
3	R, F & M	Recency, No. of different purchases over the whole history excluding returned items, Log(Monetary value over the last year excluding returned items)	0.650

Based on our single predictor and multiple predictor analysis, we conclude with respect to RQ1b that frequency is the most important predictor of repeat mail-order buying behavior, followed by monetary value. Finally, the recency variable is the least important of the RFM variables.

B. *Research Question 2: Non-RFM Variables*

RQ2: How much predictive power do additional, i.e. non-RFM, variables offer in modeling mail-order repeat purchasing?

We summarize the results in Table 7[20]. We observe that three variables result in significant improvements when added to the RFM variables:

1. Financial convenience: credit usage
2. Length of relationship: log(number of days)
3. General mail-order buying behavior: frequency

All predictors are of the behavioral type[21] and all but one (No.3) are company specific. The sequence corresponds to the increase in AUC and PCC, which is achieved when adding each of these variables.

TABLE 7

Table summarizing results of additional non-RFM variables

Variable	*Expected sign of the relationship with the probability of purchase*	*Coefficient estimate*	*Best AUC (PCC) performance (variable + RFM) on test sample*	*Difference in AUC (variable + RFM)*
Length/age of relationship	*Positive*	*Positive Significant*	*0.758 (0.652)*	*Significantly higher*
Type/category of product	Positive	Positive Significant	0.756 (0.625)	Not significant
Source of customer	complex	Not Significant	0.752 (0.640)	Not significant
Customer/Company interaction	Positive	Information negative not sign. Complaint positive not sign.	0.754 (0.636) 0.754 (0.628)	Not significant
Returns	Positive	Positive significant	0.754 (0.626)	Not significant
Customer satisfaction	Positive	Positive significant	0.755 (0.649)	Not significant
General mail-	*Positive order buying*	*Positive Not*	*0.755 (0.626) signicant*	*Significantly higher*
Benefit: financial convenience	*Positive*	*Positive significant*	*0.764 (0.687)*	*Significantly higher*
Socio-demographic variables	complex	No significant var.	0.754 (0.651)	No significant var.

When these three variables are combined in a model together with the RFM variables, we obtain the results shown in Table 8. We add the variables according to the sequence which corresponds to decreasing incremental performance in predictive accuracy (over the RFM model). Given that the financial convenience variable adds most to the predictive accuracy, we add this variable first, followed by the length of the relationship, and finally followed by general mail-order buying behavior. This is graphically shown in Figure 4.

TABLE 8

Predictive performance of 'best' multiple predictor models

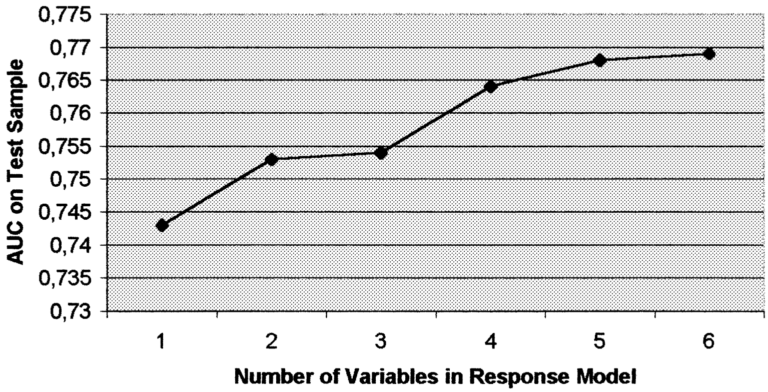
Best model containing .. variables	*List of variables*	*AUC*	*PCC*
4	Best RFM & Credit	0.764	0.687
5	Best RFM & Credit & Log(Age of Relationship)	0.768	0.690
6	Best RFM & Credit & Log(Age of Relationship) & Gen. MO buying	0.769	0.688

The results in Table 8 allow us to conclude that in terms of AUC performance, the best model including all 6 predictors achieves 53.8 % ((0.769-0.500)/(1.000-0.500)) of the possible improvement up from 50.8 % for the best RFM only model, and in terms of PCC the best model including 5 predictors achieves 34.6 % of the possible improvement up from 25.5 % for the best RFM only model[22].

Figure 4 shows the incremental gains in predictive accuracy on the test sample when additional variables are included in the model. The first three variables represent the RFM-variables (cf. Figure 3) and detailed results can be found in Table 6. By including additional variables (over the RFM variables) both the AUC and PCC performance increase substantially, but the incremental effect of additional variables diminishes. When performing the DeLong et al. (1988) statistical test comparing AUCs we find that all models differ significantly at the 0.10 significance level.

FIGURE 3

Internationale diffusie van mobiele telefonie



X. CONCLUSIONS AND ISSUES FOR FURTHER RESEARCH

In this research, we have investigated what variables should be added to a model of mail-order repeat buying. This is the first extensive study both in terms of the types of variables included in the empirical study as well as in terms of the number of alternative operationalizations of these variables. We find that the variables 'recency, frequency & monetary value' (RFM variables) are the most important predictors for modeling repeat buying, because they account for more than 50 % of the maximum possible "room for improvement". The collection of other, additional, data items only resulted in an improvement to 53.8 %. We found that only three variables are capable of significantly improving the predictive accuracy of the RFM model:

- credit usage as a proxy for financial convenience;
- length of the relationship; and
- general mail-order buying behavior.

All these predictors are of the behavioral type, and all but one[23] are company specific. More specifically, the increase in predictive performance due to the "financial convenience: credit usage" variable illustrates the importance of making it easy to obtain credit and/or of establishing not only a transactional but also a financial relationship. The second variable, "length of relationship" shows that relationships which have stood the test of time have a built-in tendency to continue. Finally, the larger the tendency of customers to engage (frequently) in mail-order transactions in general seems to positively affect the probability of repeat purchase. Even though the increase in ‘room for improvement’ does not seem large (from 50 to 53.8 %), other studies (e.g., Baesens (2002)) have shown that even modest increases in predictive performance may result in million euros of additional profit. The exact increase is, of course, dependent on the specific context.

In summary, we can conclude that the added value of using company-external data for the prediction of repurchase probabilities is very limited[24]. Since gathering these external data usually involves substantial costs, their acquisition is highly questionable *if they are only used for predictive purposes in repeat-buying applications*. However, it may be that these variables are of use for other purposes. A first alternative use could be to profile the existing customer base for customer-acquisition purposes (Mitchell and McGoldrick (1994)). An alternative use would be to create benefit segments based on motivational aspects to tailor communication messages to specific groups (Harvey (1990)).

We have to realize that our conclusions are dependent on: (1) the choice of the dependent variable, (2) the logit modeling technique, (3) the evaluation criteria, (4) the choice of the independent variables and their operationalizations, (5) the sequence of variable introduction, (6) the context of a general-merchandise mail-order company. It is clear that each of these issues may limit the generalizability of our results. Moreover, our empirical validation was performed on the data from one specific mail-order company. Therefore, we have to be careful in generalizing the conclusions. Further research could address each of these issues.

NOTES

1. The increase in competition is revealed by a decreasing market share of the top 30 catalog companies even though total catalog sales are growing (DMA (1998) p. 81)
2. RQ1a stands for research question 1a.
3. We discuss the choice of dependent variable in more detail in Section IV.
4. This column refers to applications for general merchandise catalogers offering a wide range of product lines. This is in contrast to the last column, which refers to rather narrow product lines.
5. Most studies in Table 2 only included a (small) subset of the variables proposed in our conceptual model.
6. In the short run, we may observe inventory effects, which lead to lower repurchase probabilities after large purchases. However, long mailing periods and sufficiently diverse product categories may attenuate these inventory effects.

7. We do realize that extreme levels of return behavior may be signs of the discontinuation of a relationship, but its frequency of occurrence in our dataset is too low to justify a separate treatment.
8. The *a posteriori* probability represents the probability of an event (repeat purchase) given all the values for the explanatory variables.
9. Ghosh et al. (1984) have shown that an intermediate level of parameterization, where response parameters are grouped into categories, does not improve the predictive performance in the context of market-share models.
10. A cutoff value is absolute when the specific value is used, i.e. all customers with values lower than the absolute score will not receive the mailing. An absolute cutoff value should only be used for modeling techniques (such as logit modeling) that provide scores that can be interpreted as probabilities, i.e. for which a value has a specific meaning. A relative cutoff value refers to a specific mailing depth (fraction of people selected), e.g. a mail-order company may decide (e.g. for budget reasons) only to mail to the best 60 % of its customers.
11. Expression 5 does not specify whether one should calculate the average contribution per mailing, per mailing period, or even the life-time value of a customer. Which of the three time periods should be chosen depends on the objectives of the company, whether the company maximizes respectively mailing profitability, short-term (in our case: six month period) profitability or long-term profitability.
12. This performance measure is very similar to the way relative concentration is measured by the gini coefficient used by Lorenz (1905). The gini coefficient is defined as a ratio of two areas which are determined by curves which are similar in shape to the ROC curve and which is closely related to the Pareto coefficient (Douglas (1975)). The Lorenz curve graphically shows the concentration of a distribution (Morgan (1962)).
13. We refer to Wittink ((1988) p. 64-68) for an illustrative example of this effect.
14. This low percentage can be explained by the fact that a lot of leaflets are included in the package bundled with the catalog, and no special attention was drawn to the questionnaire.
15. This confirms our expectations, because customers who answer questionnaires are usually more involved.
16. 748 non-buyers represent one minus 0.379 % of the total population, i.e. the total (rescaled) population is of size 1205 (= 748 / 0.621).
17. The proportional chance criterion is defined as: $C_{pro} = \alpha^2 + (1 -)^2$, whereby represents the actual proportion of repeat buyers. In this empirical application: $\alpha = 0.379$, because we have 37.9 % buyers in the population.
18. If we estimate the AUC (PCC) performance for all 242 alternative RFM operationalizations (instead of just for the ‘best’ operationalization), we obtain 45.8 % (9.6 %) of the available ‘room for improvement’.
19. We refer to Van den Poel (1999) for a complete list of operationalizations.
20. Variables that reveal a significant improvement over the best RFM model are shown in bold.
21. We do realize that we introduced financial convenience as part of a benefit segmentation variable, which is a non-behavioral variable in TABLE (p. 8). However, no data for a ‘non-behavioral’ operationalization was available.
22. If we estimate the AUC (PCC) performance for all 242 alternative RFM operationalizations (instead of just for the ‘best’ operationalization) added to the three non-RFM predictors, we obtain (an average of) 49.2 % (28.5 %) of the available ‘room for improvement’. This compares to 45.8 % (9.6 %) for the RFM-only models (cf. Section IX.A.5).
23. General mail-order buying behavior is a non-company specific variable.
24. Only the “General mail-order buying behavior” variable has to be purchased from external sources or collected by questionnaires.

Akaah, I.P., Korgaonkar, P.K. and Lund, D., 1995, Direct Marketing Attitudes, *Journal of Business Research*, 34, 211-219.

Baesens, B., Viaene, S., Van den Poel, D., Vanthienen, J., Dedene, G., 2002, Bayesian Neural Network Learning for Repeat Purchase Modelling in Direct Marketing, *European Journal of Operational Research* 138, 1, 191-211.

Bass, F.M. and Wind, J., 1995, Introduction to the Special Issue: Empirical Generalizations in Marketing, *Marketing Science*, 14, G1-G6.

Bass, F.M. and Wittink, D.R., 1978, Pooling Issues and Methods in Regression Analysis: Some Further Reflections, *Journal of Marketing Research*, 15, 277-279.

Bauer, 1988, A Direct Mail Customer Purchase Model, *Journal of Direct Marketing*, 2, 3, 16-24.

Berger, P. and Magliozzi, T., 1992, The Effect of Sample Size and Proportion of Buyers in the Sample on the Performance of List Segmentation Equations Generated by Regression Analysis, *Journal of Direct Marketing* 6, 1, 13-22.

Bhattacharya, C.B., 1998, When Customers are Members: Customer Retention in Paid Membership Contexts, *Journal of the Academy of Marketing Science*, 26, 1, 31-44.

Bitran, G.R. and Mondschein, S.V., 1996, Mailing Decisions in the Catalog Sales Industry, *Management Science* 42, 9, 1364-1381.

Blodgett, J.G., Hill, D.J. and Tax, S.S., 1997, The Effects of Distributive, Procedural, and Interactional Justice on Postcomplaint Behavior, *Journal of Retailing* 73, 2, 185-210.

Bolton, R.N., 1998, A Dynamic Model of the Duraction of the Customer's Relationship with a Continuous Service Provider: the Role of Satisfaction, *Marketing Science* 17, 1, 45-65.

Bradley, A.P., 1997, The Use of the Area under the ROC Curve in the Evaluation of Machine Learning Algorithms, *Pattern Recognition*, 7, 1145-1159.

Brodie, R.J. and De Kluyer, C.A., 1987, A Comparison of the Short Term Forecasting Accuracy of Econometric and Naive Extrapolation Models of Market Share, *International Journal of Forecasting*, 3, 423-437.

Bucklin, R.E. and Gupta, S., 1992, Brand Choice, Purchase Incidence and Segmentation: an Integrated Modeling Approach, *Journal of Marketing Research*, 29, 201-215.

Bult, J.R., 1993a, Target Selection for Direct Marketing, PhD. Dissertation, (Groningen University).

Bult, J.R., 1993b, Semiparametric Versus Parametric Classification Models: an Application to Direct Marketing, *Journal of Marketing Research*, 30, 380-390.

Bult, J.R., Van der Scheer, H. and Wansbeek, T., 1997, Interaction Between Target and Mailing Characteristics in Direct Marketing, with an Application to Health Care Fund Raising, *International Journal of Research in Marketing*, 14, 301-308.

Bult, J.R. and Wittink, D.R., 1996, Estimating and Validating Asymmetric Heterogeneous Loss Functions Applied to Health Care Fund Raising, *The International Journal of Research in Marketing*, 13, 215-226.

Bultez, A.V. and Naert, Ph.A., 1975, Consistent Sum-Constrained Models, *Journal of the American Statistical Association*, 70, 529-535.

Burnett, J.J. and McCollough, 1994, Assessing the Characteristics of the Non-Store Shopper, *The International Review of Retail, Distribution and Consumer Research*, 443-463.

Cohen, W.A., 1984, Direct Response Marketing: an Entrepreneurial Approach, (John Wiley & Sons, New York).

Cox, D.F. and Rich, S.U., 1964, Perceived Risk and Consumer Decision Making – the Case of Telephone Shopping, *Journal of Marketing Research*, 1, 32-39.

Courtheoux, R.J., 1995, Database marketing: miles to go, editorial, *Journal of Direct Marketing* 9, 4, 2-4.

Cullinan, G.J., 1977, Picking them by their Batting Averages' Recency – Frequency – Monetary Method of Controlling Circulation, Manual Release 2103, (Direct Mail/Marketing Association, N.Y.).
Cunningham, I. and Cunningham, W.H., 1973, The Urban In-Home Shopper – Socioeconomic and Attitudinal Characteristics, *Journal of Retailing*, 49, 3, 42-50.
DeLong, E.R., DeLong, D.M. and Clarke-Pearson, D.L., 1988, Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: a Nonparametric Approach, *Biometrics*, 44, 837-845.
Doeringer, P.B. and Piore, M.J., 1971, Internal Labor Markets and Manpower Analysis, (Lexington, Heath, M.A.).
Douglas, E., 1975, Economics of Marketing, (Harper & Row publishers, New York, NY).
DMA, 1998, Statistical Fact Book 1998, (Direct Marketing Association, New York, NY), 20^{th} edition.
Feinberg, R.A., 1986, Credit Cards as Spending Facilitating Stimuli: a Conditioning Interpretation, *Journal of Consumer Research*, 13, 3, 348-356.
Ganzach, Y., 1993, Frequency of Purchase and the Prediction of Buying Behavior in Direct Mail, *Journal of Direct Marketing*, 7, 3, 7-26.
Gehrt, K.C., Yale, L.J. and Lawson, D.A., 1996, The Convenience of Catalog Shopping: is There More to it Than Time?, *Journal of Direct Marketing*, 10, 4, 19-28.
Geisser, S. and Eddy, W.F., 1979, A Predictive Approach to Model Selection, *Journal of the American Statistical Association*, 74, 365, 153-160.
Ghosh, A., Neslin, S. and Shoemaker, R., 1984, A Comparison of Market Share Models and Estimation Procedures, *Journal of Marketing Research*, 21, 202-210.
Gönül, F. and Shi, M.Z., 1998, Optimal Mailing of Catalogs: a New Methodology Using Estimable Structural Dynamic Programming Models, *Management Science*, 44, 9, 1249-1262.
Granger, C.W.J., 1998, Extracting Information from Mega-Panels and High-Frequency Data, *Statistica Neerlandica*, 52, 3, 258-272.
Green, D. and Swets, J.A., 1966, Signal Detection Theory and Psychophysics, (John Wiley & Sons, NY).
Grönroos, C., 1990, Relationship Approach to Marketing in Service Context: the Marketing and Organizational Behavior Interface, *Journal of Business Research*, 20, 3-11.
Grönroos, C., 1995, Relationship Marketing: the Strategy Continuum, *Journal of the Academy of Marketing Science*, 23, 252-254.
Hanley, J.A. and McNeil, B.J., 1982, The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve, *Radiology*, 143, 29-36.
Harvey, J.W., 1990, Benefit Segmentation for Fund Raisers, *Journal of the Academy of Marketing Science*, 18, 1, 77-86.
Hauser, B., 1992, List Segmentation, in Nash E.L.,The Direct Marketing Handbook, 233-247.
Jacoby, J. and Jaccard, J.J., 1981, The Sources, Meaning, and Validity of Consumer Complaint Behavior: a Psychological Analysis, *Journal of Retailing*, 57, 3, 4-24.
Jones, J.M. and Landwehr, J.T., 1988, Removing Heterogeneity Bias from Logit Model Estimation, *Marketing Science*, 7, 1, 41-59.
Jones, J.M. and Zufryden, F.S., 1980, Adding Explanatory Variables to a Consumer Purchase Behavior Model: an Exploratory Study, *Journal of Marketing Research*, 42, 323-334.
Kalyanam, K. and Putler, D.S., 1997, Incorporating Demographic Variables in Brand Choice Models: an Indivisible Alternatives Framework, *Marketing Science*, 16, 2, 166-181.
Kaslow, G.A., 1997, A Microeconomic Analysis of Consumer Response to Direct Marketing and Mail Order, PhD. Dissertation, (California Institute of Technology).
Kelley, H.H. and Thibaut, J.W., 1978, Interpersonal Relations: a Theory of Interdependence, (Wiley, N.Y.).

Kestnbaum, R.D., 1979, Customer Record Predictive Model for List Segmentation, Manual Release 620.2, (Direct Marketing Association, New York).
Kestnbaum, R.D., 1992, Quantitative Database Methods, in Nash E.L., The Direct Marketing Handbook, 588-597.
Kotler, Ph., 1994, Marketing Management: Analysis, Planning, Implementation, and Control, 8th edition, (Prentice-Hall), 801 p.
LaBarbera, P.A. and Mazursky, D., 1983, A Longitudinal Assessment of Consumer Satisfaction/Dissatisfaction: the Dynamic Aspect of the Cognitive Process, *Journal of Marketing Research*, 20, 393-404.
Lawrence, R.J., 1980, The Lognormal Distribution of Buying Frequency Rates, *Journal of Marketing Research*, 42, 212-220.
Levin, N. and Zahavi, J., 1996, Segmentation Analysis with Managerial Judgment, *Journal of Direct Marketing*, 10, 3, 28-47.
Levin, N. and Zahavi, J., 1998, Continuous Predictive Modeling: a Comparative Analysis, *Journal of Interactive Marketing*, 12, 2, 5-22.
Lorenz, M.O., 1905, Methods of Measuring the Concentration of Wealth, *Quarterly publications of the American Statistical Association*, 9, 70.
Magliozzi, Th. L., 1989, An Empirical Investigation of Regression Analysis Meta-Strategies for Direct Marketing List Segmentation Models, PhD. Dissertation, (Boston University).
Magliozzi, Th. L. and Berger, P.D., 1993, List Segmentation Strategies in Direct Marketing, *Omega International Journal of Management Science* 21, 1, 61-72.
Mason, Ch.H. and Perreault, W.D., 1991, Collinearity, Power and Interpretation of Multiple Regression Analysis, *Journal of Marketing Research*, 28, 268-280.
Mela, C.F., Gupta, S. and Lehmann, D.R., 1997, The Long-Term Impact of Promotion and Advertising on Consumer Brand Choice, *Journal of Marketing Research* 34, 2, 248-261.
Mitchell, V.W. and McGoldrick, P.J., 1994, The Role of Geodemographics in Segmenting and Targeting Consumer Markets: a Delphi Study, *European Journal of Marketing* 28, 5, 54-72.
Morel, K.P.N., Poiesz, T.B.C. and Wilke, H.A.M., 1997, Motivation, Capacity and Opportunity to Complain: Towards a Comprehensive Model of Consumer Complaint Behavior, *Advances in Consumer Research*, 24, 464-469.
Morrison, D.G., 1966, Interpurchase Time and Brand Loyalty, *Journal of Marketing Research*, 3, 281-291.
Morrison, D.G., 1969, On the Interpretation of Discriminant Analysis, *Journal of Marketing Research*, 6, 156-163.
Myers, J.H., 1976, Benefit Structure Analysis – a New Tool for Product-Planning, *Journal of Marketing*, 40, 4, 23-32.
Nash, E.L., 1994, Direct Marketing: Strategy, Planning, Execution, 3rd edition, (McGraw-Hill, New York).
Parasuraman, A., Zeithalm, V.A. and Berry, L.L., 1985, A Conceptual Model of Service Quality and its Implications for Future Research, *Journal of Marketing*, 49, 41-50.
Peltier, J.W. and Schibrowsky, J.A., 1997, The Use of Need-Based Segmentation for Developing Segment-Specific Direct Marketing Strategies, *Journal of Direct Marketing* 11, 4, 53-62.
Richins, M.L., 1983, Negative Word-of-Mouth by Dissatisfied Consumers: a Pilot Study, *Journal of Marketing*, 47, 68-78.
Robertson, T.S. and Wind, Y., 1980, Organization Psychographics and Innovativeness, *Journal of Consumer Research*, 7, 24-31.
Rossi, P.E., McCulloch, R.E. and Allenby, G.M., 1996, The Value of Purchase History Data in Target Marketing, *Marketing Science* 15, 4, p. 321-340.

Rothschild, M.L. and Gaidis, W.C., 1981, Behavioral Learning Theory: Its Relevance to Marketing and Promotions, *Journal of Marketing*, 45, 70-78.
Rusbult, C.E., Zembrodt, I.M. and Gunn, L.K., 1982, Exit, Voice, Loyalty, and Neglect: Responses to Dissatisfaction in Romantic Involvements, *Journal of Personality and Social Psychology*, 43, 6, 1230-1242.
Schibrowsky, J.A. and Lapidus, R.S., 1994, Guidelines for Direct Marketers to Aggregate and Analyze Third-Party Complaints, *Journal of Direct Marketing* 8, 4, 40-50.
Schmittlein, D.C. and Morrison, D.G., 1981, On Individual-Level Inference in Job Duration Research: a Reexamination of the Wisconsin School Superintendents Study, *Administrative Science Quarterly*, 26, 84-89.
Schmittlein, D.C. and Morrison, D.G., 1983, Modeling and Estimation Using Job Duration Data, *Organizational Behavior and Human Performance*, 32, 1-22.
Shepard, D., Batra, R., Deutsch, A., Orme, G., Ratner, B. and Sharma, D., 1995, The New Direct Marketing: How to Implement a Profit-Driven Database Marketing Strategy, (Irwin, Homewood, IL).
Simpson, J.A., 1987, The Dissolution of Romantic Relationships: Factors Involved in Relationship Stability and Emotional Distress, *Journal of Personality and Social Psychology* 53, 4, 683-692.
Swets, J.A., 1979, ROC Analysis Applied to the Evaluation of Medical Imaging Techniques, *Investigative Radiology*, 14, 109-121.
Thrasher, R.P., 1991, CART: a Recent Advance in Tree-Structured List Segmentation Methodology, *Journal of Direct Marketing* 5, 1, 35-47.
Van den Poel, D., 1998, Rough Sets for Database Marketing, in L. Polkowski and A. Skowron, eds, Rough Sets in Data Mining and Knowledge Discovery II, series Soft Computing, (Physica Verlag, Springer), 324-335.
Van den Poel, D., 1999, Response Modeling for Database Marketing using Binary Classification, Dissertation, (Katholieke Universiteit Leuven), No. 129.
Van den Poel, D. and Leunis, J., 1998, Database Marketing Modelling for Financial Services Using Hazard Rate Models, *The International Review of Retail, Distribution and Consumer Research* 8, 2, 243-257.
Van der Scheer, H.R., 1998, Quantitative Approaches for Profit Maximization in Direct Marketing, PhD. Dissertation, (Rijksuniversiteit Groningen).
Wilkie, W.L., 1970; An Empirical Analysis of Alternative Bases for Market Segmentation, doctoral dissertation, (Stanford University).
Williamson, O.E., 1975, Markets and Hierarchies: Analysis and Antitrust Implications, (Free Press, N.Y.).
Wittink, D.R., 1988, The Application of Regression Analysis, (Prentice-Hall, MA).
Yankelovich, D., 1964, New Criteria for Market Segmentation, *Harvard Business Review*, 42, 83-90.
Zahavi, J. and Levin, N., 1997, Applying Neural Computing to Target Marketing, *Journal of Direct Marketing* 11, 1, 5-22.