The commercial use of segmentation and predictive modeling techniques for database marketing in the Netherlands

Peter C. Verhoef a,*, Penny N. Spring b, Janny C. Hoekstra b, Peter S.H. Leeflang b

aDepartment of Marketing and Organization, School of Economics, Erasmus University Rotterdam, Office H15-17, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
bSubdepartment of Marketing and Marketing Research, Department of Economics, University of Groningen, Groningen, The Netherlands

Accepted 28 February 2002

Abstract

Although the application of segmentation and predictive modeling is an important topic in the database marketing (DBM) literature, no study has yet investigated the extent of adoption of these techniques. We present the results of a Dutch survey involving 228 database marketing companies. We find that managers tend to rely on intuition and on the long-standing methods RFM and cross-tabulation. Our results indicate that the application of segmentation and response modeling is positively related to company and database size, frequency of customer contact, and the use of a direct channel of distribution. The respondents indicate that future research should focus on models applicable for Internet marketing, long-term effects of direct marketing, irritation from direct marketing offers, and segmentation and predictive modeling techniques.

1. Introduction

Among business practitioners and marketing scientists today, there is an increasing interest in customer relationship management (CRM) [13,14]. Customer databases and the analysis of the data they contain are essential ingredients of CRM [27]. The use of statistical techniques for analyzing customer data, thereby providing information for marketing decisions, is an important element of database marketing (DBM) and data mining (e.g., Ref. [21]). Theoretical work in the field of DBM has focused on the development, improvement, and comparison of (new) statistical techniques (e.g., Ref. [6]). These techniques are mainly used in segmentation or response modeling. Segmentation in DBM (also called list segmentation) serves to group customers into clusters, which are internally homogenous and mutually heterogeneous, implying that the members of a segment react to (direct) marketing actions similarly, but differently than members of another segment [31,p,3] Predictive modeling in DBM refers to the prediction of a response to a mailing or contact, for example, order, sales volume, and so on [28,p,187]. In the DBM literature, a number of techniques, such as crosstabulations, CHAID, probit analysis, and neural networks, have been examined that can be used in
segmentation and/or predictive modeling (e.g., Refs. [5,6,9,22,23,28,34,35]). Despite the wide availability of this literature, no study has yet investigated the extent and nature of the commercial use of segmentation and predictive modeling (techniques). Hence, given the increasing importance of customer relationship management and DBM, we report on the commercial use of segmentation and predictive modeling techniques in DBM. Using a survey among companies performing DBM in the Netherlands, we specifically address the following seven outstanding issues:

1. Which characteristics of customers and prospects are practitioners saving in databases?
2. To what extent is segmentation employed and how do database marketers perform it?
3. How prevalent is the use of response modeling for selecting consumers in DBM?
4. Which techniques do database marketers use for segmentation and/or predictive modeling?
5. What is the relation between the use of analytical segmentation and predictive modeling techniques for database marketing and firm characteristics?
6. Does the stated performance of the companies database analysis activities depend on the employed methods?
7. Which research issues do practitioners deem most important for study in the future?

The structure of this paper is as follows. A description of our data collection is given in Section 2. Subsequently, we present the results of our analysis in Section 3. We conclude with a summary of our findings, research limitations and future research issues in Sections 4 and 5.

2. Data collection

We define the target population for our research as all companies in business-to-consumer markets that use DBM techniques with a customer database in the Netherlands. In the Netherlands, direct mail accounts for approximately 30% of the advertising expenditures, which indicates the importance of DBM [29]. A commercial list of a direct marketing (DM) services provider was used as the basis for our sampling frame. Self-administered questionnaires were mailed to the persons responsible for database marketing in 1678 companies in October 1999. A personalized cover letter explaining the purpose of the study and stressing the confidentiality of the response was included. The mailing results in responses from 290 companies, which is a 17.3% response rate. Sixty-two questionnaires are excluded from the analysis because the companies are not active in the business-to-consumer market, do not have a customer database, or did not fill in the questionnaire correctly. This results in a final usable sample of 228 companies (13.5%). We tested for non-response bias by comparing the last 50 respondents with the rest of the sample [1]. As these tests do not reveal any significant differences, we conclude that there is no non-response bias.

The survey instrument contains 43 structured, closed-end questions concerning the company’s database marketing activities, the content of their customer and/or prospect database, the use of database marketing segmentation and selection techniques, and the respondents’ opinions on important issues for future research. In order to assure good wording of the questions, drafts of the questionnaire were tested with a panel of four marketing academics and seven DBM practitioners.

The sample can be described as follows. The most-represented branch in the responder universe is financial services, which is well known for the application of DBM [8]. Charities comprise 18.4% of the responding institutions. A majority of the respondents holds a marketing function (76.2%). The median number of employees at the responding companies is between 50 and 100. The majority of the companies (81.1%) use direct marketing as a channel of distribution. Most responding companies (71.1%) have started their DM operations more than 5 years ago. Hence, we have a relatively experienced responder group. Direct mail is the most frequently employed direct marketing medium (95.2%). Telemarketing is utilized by 64% of our sample, while catalogues are sent by 27.6%. Relatively few responders (25.4%) use e-mail as a direct medium. These media are mainly used to sell products (78.6%), to maintain relationships (62.7%), to provide information (53.6%), and to attract new customers (40.6%). Sixty-two percent of the organizations in our sample have a customer database consisting of more than 50,000 customers.
Based on these figures, we conclude that the sample consists of companies with sizeable customer data-bases utilizing all types of DM media and striving for a number of DM objectives.

3. Results

In this section, we describe our empirical results. Thereby, we will subsequently deal with all stated research questions.

3.1. Available customer information

To a great extent, the precision and depth of a company’s database determines the potential of database analysis to increase profitability [7]. From a modeling point of view, a lack of customer data can result in the omission of relevant variables and thus might lead to incorrect interpretations and poor predictions [15, Chapter 15, and 21]. Companies can collect data themselves or buy or rent data at an individual or zip code level from external suppliers. Data at the individual level are naturally more specific. However, individual data are also more expensive and are not always directly available [23].

Organizations should make a trade off between the cost of renting additional data and performing more effective modeling [16]. The majority of the companies we studied do not rent or buy external data for their current customers (Table 1). The most popular externally supplied information on customers is socio-demographics. For prospects, 64% of the organizations rent lists for the purpose of obtaining names and addresses. As in the case of prospects, individual level data are more costly and less available, companies report a higher incidence of zip level than individual level purchase or rental.

Table 2 shows that name and address are the most widely stored characteristics. Furthermore, approximately 60% of the companies keep valuable purchase history data [24], such as the type of product purchased, the date of the first purchase, and the amount of following purchases. Data that must be purchased or gathered by questionnaires, such as socio-demographic data, lifestyle data, satisfaction data, and purchase data from other companies, have very low coverage, presumably due to their high cost. The lack of data on purchase behavior at competitors implies that companies are not able to calculate customer share, which is considered to be important in CRM [20]. The source of the customer and the channel of purchase, which can be used to evaluate the performance of the utilized channels and media, are stored by more than half of the respondents [11].

3.2. The use of segmentation

One hundred and sixty respondents (70.1%) employ segmentation in their database marketing
activities. Specific questions referring to segmentation strategy in the questionnaire regard three issues:

- The purpose of the segmentation: how are consumers in different segments treated differently?
- The criterion variable: how should consumers in different segments react differently to the offers they receive?
- The segmentation variables: which customers’ characteristics determine into which segment the customer will be allocated?

It is remarkable that 90.0% of the companies employ segmentation for the purpose of target selection (Table 3). Fewer respondents report its use for varying the treatment or timing of treatment of customers. Almost 27% of all responding companies use segmentation to build predictive models per segment. Not unexpectedly, most companies seek segmentation rules that discriminate groups with differing response rates. Many companies also classify customers such that purchase amount is homogenous within a segment. Discriminating profitability is the purpose of segmentation at only 42.8% of the companies, but profitability is highly correlated with both response rate and purchase amount. The discrimination of levels of creditworthiness interests only few companies as a purpose of segmentation. From a historical perspective, RFM variables are the most employed characteristics on which to base segmentation [2]. Respondents value recency as the most discriminating characteristic for segmentation. Only 32.1% report segmenting on purchase amount, but 61.6% report discrimination in purchase amount as a purpose of segmentation. In line with the availability of customer characteristics, lifestyle and socio-demographic variables are employed to a lesser extent.

With cross-tabulations and chi-square tests for association, we explored whether the use of segmentation differs between companies with different objectives for DM for selling products discriminate significantly more on purchase amount ($p=0.00$). For companies using DM to maintain relationships, this is also true ($p=0.00$), but they also discriminate on customer profitability more often ($p=0.00$). In line with this result, these companies use segmentation for varying treatment more frequently ($p=0.00$). We could not find any significant differences for companies disseminating information with DM. However, our results show that companies attracting customers with DM are more likely to use segmentation ($p=0.04$), and that their purpose of segmentation is more likely to be the selection of target customers ($p=0.01$).

### 3.3. Use of predictive modeling

Almost half of the respondents report that they predict the responses to their marketing activities. We asked two specific questions with regard to response modeling.

- Are tests performed on random samples of the customer base for the purpose of fitting response models?
- Which criterion variable is predicted?

Selecting random samples of the customer base for testing purposes can be costly. However, almost 71% of the respondents who fit predictive models report testing for the purpose of generating a sample for model building. Nearly all of the companies that model attempt to predict primary response. Purchase quantity or donations are predicted to a lesser extent.
(Table 4) [12]. Very few companies estimate secondary response such as payment or reject. Just as we examined differences in the use of segmentation between companies with different DM objectives, we explored whether the use of predictive modeling is different between companies with various DM objectives. However, the analysis reveals no significant differences in the latter regard.

Table 4
Use of predictive modeling

<table>
<thead>
<tr>
<th>Predict. (.n=110)</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary response</td>
<td>98.2</td>
</tr>
<tr>
<td>Purchase amount</td>
<td>43.6</td>
</tr>
<tr>
<td>Payments</td>
<td>2.7</td>
</tr>
<tr>
<td>Rejects</td>
<td>3.6</td>
</tr>
</tbody>
</table>

3.4. Employed statistical techniques

In our questions on the employed statistical techniques, we included those techniques that are cited in the literature as being effective. Despite the fact that literature advocates more sophisticated techniques [18], cross-tabulation is the most widely utilized method followed by RFM analysis (Fig. 1). Although linear regression performs worse than CHAID and logit or probit models [18], it is the third most popular technique. CHAID or CART is employed by approximately 16% of the respondents employing segmentation and/or modeling, while logit or probit models are used less. As Bult and Wansbeek [6] report that CHAID does not predict as well as probit regression, it is surprising to find that CHAID is also used for predictive modeling purposes. Results of discriminant analysis are not reliable in the case of low response

Fig. 1. Use of different statistical methods for segmentation and predictive modeling. The differences in sample sizes in Tables 4 and 5 and this figure are the result of firms leaving the entire section on modeling and segmentation techniques blank. The lack of response in these sections might be due to secrecy of the firms, or due to the fact that no techniques are used at all to develop segmentation rules. Since we cannot know the motive behind the skipping of this part of the questionnaire, it was decided to omit these respondents from this section of the analysis.
rates [28], common to DBM. Hence, the low adoption of this technique is not surprising. Zahavi and Levin [35] report that neural networks perform worse than logit or probit models. In practice, we only find a few companies using neural networks. Probably, this is due to the mentioned disappointing performance, but also the black-box character and the low explanatory power of neural networks might be an explanation [30]. A few companies also use genetic algorithms.

Many respondents reported being unfamiliar with logit/probit models, neural networks, and genetic algorithms. It is interesting to note that the ranking of techniques is the same for both segmentation and predictive modeling, with the exception of factor and cluster analysis, which only appear in the section on segmentation techniques. While factor analysis and cluster analysis can be employed as exploratory or preparatory analyses for predictive modeling, their outcomes do not directly yield response predictions. They are performed for segmentation by approximately one quarter of the companies.

Of companies reporting the utilization of segmentation and predictive modeling, 30.6% and 35.5%, respectively, indicate no method. A number of these respondents used the fill-in line for the segmentation and modeling technique questions to write in “gut feeling” or “experience.” Despite the increasing availability of data and computer power in the last two decades, segmentation and selection are still often performed on the basis of intuition by a large group of companies [5,p,17].

3.5. Company characteristics and modeling

We continue with a discussion and empirical examination of the relationship between company characteristics and the use of predictive modeling and segmentation techniques. The use of these techniques is justified if the expected benefits exceed the expected costs. We note that the expected benefits of the models depend on the decision type and context [10]. Increased expected benefits with regard to segmentation and modeling in DBM can be found in more efficient targeting. Expected costs of modeling and segmentation models are comprised of initial development costs, maintenance costs, costs inherent to model use, and costs of marketing data. The trade off between costs and benefits varies with the number of times a model is used [15,Chapter 20]. In a DM environment, the number of applications for which a model can be used is related to the frequency of direct contacts. If a firm can reuse a model for more applications, the total costs will rise less rapidly. Hence, we expect that larger companies using DM instruments more frequently will have lower incremental modeling costs. Moreover, these large companies benefit from economies of scale to invest in these models [3,32]. We also expect that the benefits of DBM models are related to the size of the customer database and the years of experience in DM. For companies with a larger customer database, more efficient modeling will result in absolute larger increases in profitably. Likewise, firms exclusively utilizing the direct channel can expect greater gains in profit from modeling than those utilizing both direct and indirect channels. Experience also influences the use of models, as companies new to DBM will not assign modeling and segmentation a first priority.

Based on the above expectations, we estimate three models to explore the relationship between modeling and firm characteristics. In these models, we relate (i) use of segmentation, (ii) use of predictive modeling, and (iii) level of sophistication of statistical models to the following firm characteristics: company size (number of employees), frequency of DM contacts, database size, and years of experience in DM. For channel type and frequency of DM, dummy variables are employed. The classes defined for company size, database size, and number of years of DM experience are replaced with midpoints of these classes for all analyses.

We employ logistic regression to model the probability of use of segmentation and predictive modeling. The second and third columns of Table 5 show the results of the two regression models. The chi-square statistic indicates that the model predicting incidence of segmentation is not significant ($p=0.39$). Thus, the company characteristics included in the questionnaire cannot explain the incidence of segmentation in DBM. The logistic model explaining predictive modeling is significant ($p<0.01$). Two significant predictors of predictive modeling are the number of customers in the database ($p<0.01$) and the exclusive use of a direct channel of distribution ($p<0.05$). Other significant ($p<0.10$) predictors are company size, years of experience in DBM, and the frequency of DM activities.
The objective of the third model is to explain the level of statistical sophistication among companies. The companies included in this model are restricted to only those that perform segmentation and/or predictive modeling (n = 159). Within these companies, we distinguish three groups of users: those who do not use models (29%), those employing only the simple models RFM or cross-tabulations (23%), and those using more complex models such as regression analysis, CHAID, etc. (48%). It should be noted that the first group is comprised mainly of companies indicating experience and gut feeling as techniques. Some companies report using both simple and complex models, presumably employing the simpler models for exploratory analysis, then the more complex for the final model. The firms are categorized by the highest level of sophistication of the models they use. This classification of techniques portrays the progression from intuition to sophisticated modeling. RFM and cross-tabulations can be performed without statistical know-how or a statistical computer package. Those techniques we call complex require both statistical software and a minimum level of statistical understanding. We consider the modeling sophistication classification to be ordinally scaled. Therefore, ordered regression analysis is an appropriate tool to investigate the impact of firm characteristics on the level of sophistication of DBM modeling and segmentation in companies. The fourth column of Table 5 shows that three firm characteristics significantly affect the level of sophistication of DBM modeling and segmentation in companies. Both company size and database size are positive predictors of sophistication (p<0.05). The frequency of DM activities is overall a positive predictor of sophistication.

Table 5
Model results for the effect of company characteristics on the use of segmentation, predictive modeling, and statistical techniques

<table>
<thead>
<tr>
<th>Variables</th>
<th>Segmentation (n = 201)</th>
<th>Modeling (n = 201)</th>
<th>Sophistication (n = 159)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company size</td>
<td>4.52x10^{-4}</td>
<td>7.42x10^{-4}</td>
<td>6.34x10^{-4}</td>
</tr>
<tr>
<td>Database size</td>
<td>3.04x10^{-1}</td>
<td>5.27x10^{-1}</td>
<td>2.02x10^{-1}</td>
</tr>
<tr>
<td>Years of experience in DBM</td>
<td>0.005</td>
<td>0.029^ab</td>
<td>-0.013</td>
</tr>
<tr>
<td>Direct channel only</td>
<td>0.425</td>
<td>1.318^b</td>
<td>0.281</td>
</tr>
<tr>
<td>Both channels</td>
<td>0.310</td>
<td>0.471</td>
<td>-0.152</td>
</tr>
<tr>
<td>Frequency DM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;1 per week</td>
<td>0.542</td>
<td>1.068^c</td>
<td>0.590^c</td>
</tr>
<tr>
<td>1 per week</td>
<td>0.426</td>
<td>0.114</td>
<td>1.229^b</td>
</tr>
<tr>
<td>1–3 per month</td>
<td>0.358</td>
<td>0.774</td>
<td>0.940^a</td>
</tr>
<tr>
<td>1–2 per quarter</td>
<td>0.329</td>
<td>-0.204</td>
<td>0.367</td>
</tr>
<tr>
<td>Intercept 1</td>
<td>-0.016</td>
<td>-1.925^b</td>
<td>0.143</td>
</tr>
<tr>
<td>Intercept 2</td>
<td>N.A.</td>
<td>N.A.</td>
<td>0.840^b</td>
</tr>
<tr>
<td>Chi-square/LR statistic</td>
<td>9.50</td>
<td>50.56^a</td>
<td>35.92^a</td>
</tr>
</tbody>
</table>

1 Twenty-seven cases were listwise deleted because of missing values in independent variables.
2 Estimated with logistic regression.
3 Parameter significant at 0.01 level; 4parameter significant at 0.05 level; 5parameter significant at 0.1 level.
4 Estimated with ordered regression in E-views.
5 Sample restricted to companies that segment and/or model.

The objective of the third model is to explain the level of statistical sophistication among companies. The companies included in this model are restricted to only those that perform segmentation and/or predictive modeling (n = 159). Within these companies, we distinguish three groups of users: those who do not use models (29%), those employing only the simple models RFM or cross-tabulations (23%), and those using more complex models such as regression analysis, CHAID, etc. (48%). It should be noted that the first group is comprised mainly of companies indicating experience and gut feeling as techniques. Some companies report using both simple and complex models, presumably employing the simpler models for exploratory analysis, then the more complex for the final model. The firms are categorized by the highest level of sophistication of the models they use. This classification of techniques portrays the progression from intuition to sophisticated modeling. RFM and cross-tabulations can be performed without statistical know-how or a statistical computer package. Those techniques we call complex require both statistical software and a minimum level of statistical understanding. We consider the modeling sophistication classification to be ordinally scaled. Therefore, ordered regression analysis is an appropriate tool to investigate the impact of firm characteristics on the level of sophistication of DBM modeling and segmentation in companies. Both company size and database size are positive predictors of sophistication (p<0.05). The frequency of DM activities is overall a positive predictor of sophistication.

Table 6
The subjective performance of database analysis methods

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean score of stated performance of database analysis methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>We are able to have better results with our DM campaigns by using database analysis tools (1= totally disagree, 5 = totally agree)</td>
</tr>
<tr>
<td>Total sample (n = 228)</td>
<td>3.90</td>
</tr>
<tr>
<td>Segmentation (N= 228)</td>
<td>4.01</td>
</tr>
<tr>
<td>Predictive modeling (N= 228)</td>
<td>4.23</td>
</tr>
<tr>
<td>Sophistication (n = 172)</td>
<td>3.80</td>
</tr>
<tr>
<td>No techniques</td>
<td>3.97</td>
</tr>
<tr>
<td>Sophisticated techniques</td>
<td>4.22</td>
</tr>
</tbody>
</table>

Numbers in bold denote significant differences in mean values according to t-test or F-test.
3.6. Performance of analysis

Because the literature suggests that more complex models outperform simpler ones (e.g., Ref. [5]), we also question whether companies using the more sophisticated models believe that they attain better results by using them. We gave the respondents two Likert-type statements on which they could indicate their opinion on a 5-point scale ranging from totally agree to totally disagree. The mean scores indicate that the companies attribute better performance to their analysis (3.9), but they still are not completely satisfied with the results (3.07) (see Table 6). With regard to segmentation, companies applying segmentation indicate a significantly better performance (4.01 vs. 3.61; \( p = 0.01 \)) and they are more satisfied with their results (3.17 vs. 2.82; \( p = 0.02 \)). Modelers have a significantly better performance than non-modelers (4.23 vs. 3.58; \( p = 0.00 \)). They also appear more satisfied (3.17 vs. 2.97), although the latter result is not significant \( (p = 0.14) \). Finally, in line with results in the literature, we find that companies using more sophisticated techniques indicate a significantly better performance \( (p = 0.04) \). Note that this performance increases per sophistication level. In spite of the perceived higher performance, the sophisticated modelers are not significantly more satisfied with the results of their campaigns \( (p = 0.92) \). Hence, it seems that they seek still better models.

3.7. Research issues

We ended the survey with a question to the respondents as to which issues they believe require further research, from their point of view as practitioners. Answers to this question can be used to guide further scientific research that bridges the gap between academic research and business practice. We have selected issues from a review of the literature in the questionnaire development phase. The most prevalent research issues are the role of DBM in an Internet environment and the long-term effect of DM campaigns (see Table 7). The importance of the latter is also reflected in the interest for topics such as influencing the lifetime value of customers and the calculation of lifetime value. It is remarkable that although 42.7% of the respondents report irritation to be an important issue, only 25.6% state that privacy in DBM is an issue that calls for more research attention.

4. Conclusion

In this paper, we presented a research project that studies how companies practice DBM. This is of particular interest now that the growing interest in CRM has resulted in an increasing importance of databases in the marketing strategy of firms. We investigated the contents of the customer database and the use of external data. The results show that although the large majority of companies store transaction type data, socio-demographic data and attitudinal data are stored limitedly. In line with the literature on conjoint analysis and UPC-scanner data (e.g., Refs. [4,33]), we have investigated the commercial use of segmentation and predictive modeling techniques in DBM. The most important conclusions are the following. First, although the majority of the companies perform segmentation, predictive modeling is used to a lesser extent. Second, few managers use the models and insights generated by scientific research. The statistical techniques employed by most are still relatively simple. Many companies report that they base their target selections on intuition and “gut feeling”, thereby using simple heuristics, such as “mail all customers, who recently purchased product X”, or “have an income above Y”. The fact that models and insights generated by research are still limitedly used by managers is an important finding of our research. It also has implications for the scientific community. In order to further disseminate the developed models, the scientific community may wish to...
take specific steps, such as presenting research at practitioner conferences. Moreover, as CRM and DBM are recent developments in marketing and have not widely gained attention in the curricula of marketing programs at business schools, there is a need to develop marketing courses on CRM and DBM focusing on the application of the considered techniques. Third, the sophistication of these models depends on company size, size of the database, and the contact frequency. This result emphasizes that economies of scale are important in applying DBM techniques. Fourth, companies applying more sophisticated techniques appear to have a better performance than companies not applying these techniques. Hence, the utilization sophisticated models indeed seems to pay-off from a practitioners’ perspective. However, despite the improved performance, these companies still seek techniques that will further improve results. Finally, from a practitioners’ perspective, important research issues concern the long-term impact of DM media, LTV and irritation effects of DM.

5. Research limitations and future research

This research has three limitations. First, we restrict our study to companies in the business-to-consumer market. Future research might also consider firms in business-to-business markets. Second, our study is limited to the Netherlands. However, due to the effect of sharing of methods by multinationals, we expect the differences to be smaller than mere population and size of database would dictate. Future research could replicate and extend this study in other European countries as well as in the US. These multi-country studies would enable one to make cross-country comparisons. These comparisons would probably offer interesting insights, such as which market characteristics (i.e., market size, use of DM) explain the divergence in the use DBM techniques between countries. Third, although we do not find any non-response bias, the use of a self-administered questionnaire might have led to self-selection effects. Besides, the noted issues future research could consider the research issues raised by practitioners. Three issues are especially relevant from a scientific perspective. First, the prediction of customer lifetime value needs further attention in the literature. Recently, Rust et al. [25] have stressed the importance of customer lifetime value as a new performance measure. Despite this importance, only a few studies have modeled customer lifetime value (e.g., Ref. [26]). Second, due to increasing importance of Internet, new types of data, such as click data, are collected. What we do not know is the value of these data? Furthermore, models and systems should be developed that fit into the interactive nature of the Internet. These models are perhaps not based on traditional DBM models. Third, while there is a lot of knowledge on the long-term impact of for example sales promotions (e.g., Ref. [19]), the long-term impact of CRM instruments has not been investigated. Future research should elaborate on this issue. Finally, there remains value in research that investigates the effective use of customer databases within customer relationship management.

References


**Peter C. Verhoef** is a post-doctoral researcher affiliated with the Department of Marketing and Organization of the School of Economics at the Erasmus University Rotterdam. He is also a member of Erasmus Research Institute in Management (ERIM). In 2001, he received his PhD in marketing from the Erasmus University. In his dissertation, he investigated how companies can affect the lifetime value of customers. His research interests include customer relationship management, waiting times, and private labels. His work has been or will be published in the Journal of Consumer Psychology, Journal of Retailing, Journal of the Academy of Marketing Science, Decision Support Systems, and various other journals.

**Penny, N. Spring** is a former Ph.D. student of University of Groningen (The Netherlands), Department of Economics, Sub-department Marketing and Marketing Research. She defended her dissertation entitled “Statistical Methods in Database Marketing” in October 2002. She is currently Data Science Director for Reader’s Digest Europe. Her research has been published in the Journal of Market Focused Management and the Journal of Targeting for Measurement and Analysis in Marketing.

**Janny C. Hoekstra** received her PhD in economics from the University of Groningen in 1987. From 1994, she held a part-time position as a Professor in Direct Marketing at the Erasmus University Rotterdam, The Netherlands. As from June 2000, she holds this position at the University of Groningen. Her research interests include the development of the marketing concept, market orientation, customer lifetime value, direct marketing, and e-commerce. She published papers in among others Industrial Marketing Management, the European Journal of Marketing, the Journal of Direct Marketing, the Journal of Market-Focused Management, and the Journal of Retailing.
**Peter S. H. Leeflang** studied Econometrics at the Erasmus University of Rotterdam. He obtained his PhD in 1974. At present, he is Professor of Marketing at the Department of Economics, University of Groningen, The Netherlands. He wrote several books and published in the Journal of Marketing, the Journal of Marketing Research, Management Science, Applied Economics and the International Journal of Research in Marketing.