A DYNAMIC SEGMENTATION APPROACH FOR TARGETING AND CUSTOMIZING DIRECT MARKETING CAMPAIGNS

THOMAS REUTTERER, ANDREAS MILD, MARTIN NATTER, AND ALFRED TAUDES

An important aspect of customer relationship management is the targeting of customer segments with tailored promotional activities. While most contributions focus on the selection of promising customers for targeting, only few authors address the question of which specific differential offers to direct to the selected target groups. We focus on both issues and propose a flexible, two-stage approach for dynamically deriving behaviorally persistent segments and subsequent target marketing selection using retail-purchase histories from loyalty-program members. The underlying concept of behavioral persistence entails an in-depth analysis of complementary cross-category purchase interdependencies at a segment level. The effectiveness and efficiency of the proposed procedure are demonstrated in a controlled field experiment involving the targeting of several thousands of customers enrolled in the loyalty program of a "do-it-yourself" retailer. Our empirical findings provide evidence of significant positive impacts on both profitability and sales for segment-specific tailored direct marketing campaigns.

JOURNAL OF INTERACTIVE MARKETING VOLUME 20 / NUMBER 3–4 / SUMMER / AUTUMN 2006

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/dir.20066

THOMAS REUTTERER
is Associate Professor at the Institute for Retailing and Marketing, Vienna University of Economics and Business Administration (VUEBA), Augasse 2–6, A-1090 Vienna, Austria; e-mail: thomas.reutterer@wu-wien.ac.at

ANDREAS MILD
is Associate Professor at the Institute for Production Management, VUEBA, Austria; e-mail: andreas.mild@wu-wien.ac.at

MARTIN NATTER
is the Hans Strothoff Professor of Retail Marketing at the Johann Wolfgang Goethe-University Frankfurt, Germany; e-mail: natter@wiwi.uni-frankfurt.de

ALFRED TAUDES
is Professor at the Institute for Production Management, VUEBA, Austria; e-mail: alfred.taudes@wu-wien.ac.at

The authors thank the editors and the reviewers for their helpful comments on an earlier version of the article. We also thank the participants of the NYU conference for valuable insights.

INTRODUCTION

The retail industry is characterized by fierce price competition among companies that offer rather similar product assortments and pursue aggressive promotional policies within given retail formats (Corstens & Corstens, 1995; Kahn & McAlister, 1997). Retailers are collecting huge amounts of personally identifiable point-of-sale (POS) transaction data dissembling rich information about customers' purchasing habits (e.g., sizes, spending values, or compositions of shopping baskets). The individual purchase histories collected from the customers enrolled in the program can be linked back to store and marketing data, sociodemographic background characteristics, and additional survey or feedback information (if available). Within advanced concepts of customer relationship management (CRM), this database of consolidated data sources plays a central role in analyzing and planning targeted direct marketing actions (Winer, 2001).

As our brief review in the next section will show, considerable advances have been achieved in the field of target segment selection for direct marketing purposes; however, the authors could not detect any contributions to the academic marketing literature that convincingly address the question of which specific differential offers (in terms of merchandise types or product categories to be featured or subjected to rewards) to direct to the customer segments that turn out to be worth targeting.

This article attempts to utilize the multicategory nature of choice decisions made by individual shoppers throughout their shopping trip histories to assist direct marketers in selecting *who* to target with *what specific offer(s)*. In doing so, both the process of segment formation and the customization of targeted cross- and upselling campaigns are based on a measure that quantifies a customer's "interest" in particular (combinations of) product categories. As in the case of the Tesco *Clubcard* program (Humby & Hunt, 2003) or the indications provided by Pearson and Gessner (1999), resolution of this issue is sometimes claimed by practitioners, but there is a lack of a more thorough treatment in the academic literature. Furthermore, specification of such "interest measures" and especially consideration of cross-category effects are mostly accomplished on a rather ad hoc basis or guided by pure managerial intuition.

In contrast, we advocate a more general and data-driven approach for quantifying a specific customer's tendency to symptomatically (re)purchase distinctive combinations of product categories included in a retail assortment. The latter is denoted as "behavioral persistence," which will be evaluated based on an in-depth exploratory analysis of shopping basket histories. The remainder of this article proceeds as follows: Following a brief discussion of current practices in target market selection within loyalty programs and their specific deficiencies, an overview of previous research on analyzing cross-category purchase interdependencies based on shopping basket data is provided. Next, we propose a flexible and dynamic approach to derive segments of customers who are behaviorally persistent in the aforementioned sense. The proposed two-stage modeling framework includes a data-compression step of the observed shopping basket data and resolves the subsequent target group selection. In an empirical application study, targeting effectiveness and efficiency of the proposed procedure are evaluated in a controlled field experiment involving segment-specific adapted direct mailings to a large customer sample of a "do-it-yourself" (DIY) retailer. Finally, we discuss conclusions and outline some suggestions for future research.

CONCEPTUAL BACKGROUND AND RELATED LITERATURE

Current Issues and Practices in Target Marketing Selection

Due to the typically noncontractual loyalty-program settings in retailing, shoppers are free to enroll in any competing programs within the same retail industry and experience much lower (if any) switching costs as compared to financial institutions or telecommunication providers (Reinartz & Kumar, 2000; Rust, Zeithaml, & Lemon, 2000). This has led to an increasingly competitive environment with different companies vying for retention of the same customer base. Furthermore, many programs appear to be launched as a defensive marketing strategy rather than as a sophisticated CRM initiative (Kumar & Shah, 2004; Reinartz & Kumar, 2002).

Traditionally, loyalty programs have spent too much focus on purchase frequency or spending without considering profitability (Reinartz & Kumar, 2000). Especially in retailing, many companies still appear to try to maximize loyalty-program members, forcing them to engage in loyalty wars with their competitors. Since customers might exhibit "cherry-picking" behavior among the most attractive rewards they are being offered without contributing enough to enhanced overall profitability, such programs are running the risk to erode already slim profit margins (Rust et al., 2000, p. 102). Indeed, the supermarket chain Safeway based in the United Kingdom had to abandon its *ABC Card* loyalty program due to overwhelming communication and operation costs as more members were enrolling. From past experience, it thus becomes obvious that without any clear differentiation or value proposition, loyalty programs run the risk of imminent failure (Reichheld, 1996).

Although the value of purchase-history data for targeting purposes is well recognized in the literature (Rossi, McCulloch, & Allenby, 1996), behavioral patterns are still frequently analyzed at the aggregate level only and converted into direct mailings or other CRM actions with little or even no differentiation across the customer portfolio. To recognize differences in purchase behavior and, as a consequence, in profitability contributions to the company, the imperative questions to be efficiently answered by targeted direct marketing efforts are *who* to target (i.e., which individual customers or segments) with *what offer* (i.e., which items or product categories) using *what action* (i.e., which media and rewards). According to an extensive literature review by Prinzie and Van den Poel (2005), however, the vast majority of contributions to the relevant marketing literature are limited to the issue of target selection based on behavioral loyalty patterns.

In particular, the Recency, Frequency, and Monetary value (RFM) framework has seen widespread industry applications (Berry & Linoff, 2004; Hughes, 1996). In the classic RFM approach, allocation of direct marketing resources to customers is determined as a function of how recently, how frequently, and what current and past amounts of spending are exhibited by customers (Colombo & Jiang, 1999). Various extensions to the conventional set of predictors are additional behavioral measures, mailing features, or household demographics (Haughton & Oulabi, 1993; Morwitz & Schmittlein, 1998; Zahvai & Levin, 1997). Furthermore, timing and frequency issues of direct marketing policies have been modeled simultaneously with the target selection problem in the catalog retail industry (Britan & Mondschein, 1996; Elsner, Krafft, & Huchzermeier, 2004). As a more sophisticated concept to assist target selection for direct marketing investments, in recent years, the concept of Customer Lifetime Value (CLV) has received increasing attention (for an overview, see Berger & Nasr, 1998; Jain & Singh, 2002; Mulhern, 1999; Rust, Lemon, & Zeithaml, 2004). Whereas traditional RFM-based approaches primarily focus on past behavioral loyalty, CLV is designed as a forward-looking metric that explicitly takes the various drivers of customer profitability into account. Furthermore, new modeling approaches that link RFM with the CLV paradigm are becoming available to the marketing community (Fader, Hardie, & Lee, 2005).

Exploratory Analysis of Shopping Basket Data

The problem with the previously mentioned concepts is that they do not provide any clear indications on which (type of) products to feature in targeted cross-/upselling campaigns. A more detailed analysis of shopping basket data might help in this respect. A market or shopping basket registered at a retailer's electronic POS check-out systems is representing the outcome of a customer's multicategory-choice decisions among items or categories[1] included in the assortment offered by the retailer of her or his choice. Once personalized transaction data are available, the composition of sequences of such shopping baskets realized by loyalty-program members reflects valuable behavioral information about the dynamic aspects of customer needs. As claimed in the introduction of this article, insights into the interdependency structure of cross-category correlations underlying the collected patterns of joint-category purchases could potentially leverage both the issue of target segment selection and assist in designing tailored offers in direct marketing campaigns (also see Elrod et al., 2002).

[1] For simplicity reasons, we restrict our exposition to the product category level; however, the analogue applies to any subcategory level of retail assortments under study. Furthermore, we leave the modeling of purchase quantities associated to the category-choice decisions as a topic for further research.

Manchanda, Ansari, and Gupta (1999) identified three types of influencing factors that might lead to the combination of two or more categories in one shopping basket: (a) the complementary nature of consumption or usage (e.g., bolts, nails, and tool kit) of the categories involved, (b) they may be subject to similar (i.e., coinciding) repurchase cycles or seasonal effects (e.g., plants and paints in spring), and (c) observed or unobserved consumer heterogeneity (e.g., different stages in the "life cycle" of a do-it-yourselfer). Depending on their methodological provenience, the various approaches for studying cross-category effects more or less explicitly distinguish between these sources for category co-occurrences. In contrast to predictive models of the type proposed by Manchanda et al. (1999), the task of exploratory approaches to market basket analysis is to discover pronounced cross-category interrelationships based on observed frequency patterns of jointly purchased product categories. In the marketing literature, this also is referred to as "affinity analysis" (Russell et al., 1999). Multidimensional scaling or various hierarchical clustering techniques are typically employed to assist this task. Inference of such interdependency structures also is one of the primary objectives of various data mining techniques such as association rule mining, as applied by database marketers for quite a while (Berry & Linoff, 2004). The common aim of the various techniques for exploratory shopping basket analysis is to parsimoniously represent the observed cross-category correlations in a meaningful fashion. Thus, they also can serve as a data-compression step prior to modeling cross-category effects in response to marketing actions (Boztug & Reutterer, 2006).

With a few exceptions, however, the majority of (both predictive and exploratory) approaches proposed so far examine cross-category purchase effects on the aggregate level of household demand (State-of-the-art reviews are provided by Russell et al., 1999; Seetharaman et al., 2005; or in the context of market structure analysis by Elrod et al., 2002). Using a finite mixture modeling framework, Russell and Kamakura (1997) and in a similar approach Andrews and Currim (2002) identified household segments with homogeneous purchase behavior across product categories. Individual-level approaches based on collaborative filtering (CF) techniques are available as well (Mild & Natter, 2003; Mild & Reutterer, 2003). While CF methods are designed to derive personalized item recommendations and are mostly used in an interactive online media context, they seem to be less suited for the specific managerial requirements of more traditional direct marketing campaigns. Additionally, efficiency considerations are still delimiting perfect mass customization of conventional "offline" direct mailings in retail marketing practice. In principle, recent improvements of powerful Markov chain Monte Carlo simulation methodologies can help to successfully alleviate the estimation problems, multivariate logit (Hruschka, Lukanowicz, & Buchta, 1999; Russell & Petersen, 2000) or probit models are obviously confronted with when the number of product categories to be analyzed is increasing. In the case of predictive modeling, the contributions of Ainslie and Rossi (1998), Seetharaman, Ainslie, and Chintagunta (1999), and Chib, Seetharaman, and Strijnev (2002) demonstrated significant progress. Nevertheless, since real-world retail assortments are typically consisting of dozens or hundreds of product categories, further progress in this direction remains to be seen.

In an exploratory context, however, the idea of representing cross-category purchase effects at a more disaggregate level was introduced to the marketing community only recently (Decker, 2005; Decker & Monien, 2003; Schnedlitz, Reutterer, & Joos, 2001). The authors employed various neural network architectures with unsupervised learning rules as a data compression device that results in a mapping of binary-valued category incidence vectors (the latter representing retail transactions) onto a set of so-called prototypes. In their empirical applications, they illustrated that each of these prototypes is post hoc responsible for a specific class of market baskets with internally more pronounced (complementary) cross-category purchase interdependencies as compared to the aggregate case. We will elaborate on these approaches in the next section, where we develop the proposed segmentation procedure.

METHODOLOGY

The conceptual framework of the proposed analytical model for dynamic customer segmentation is outlined in Figure 1. The process of segment construction is required to warrant that past purchase behavior of customers assigned to a distinctive segment is characterized by distinguished patterns of cross-category

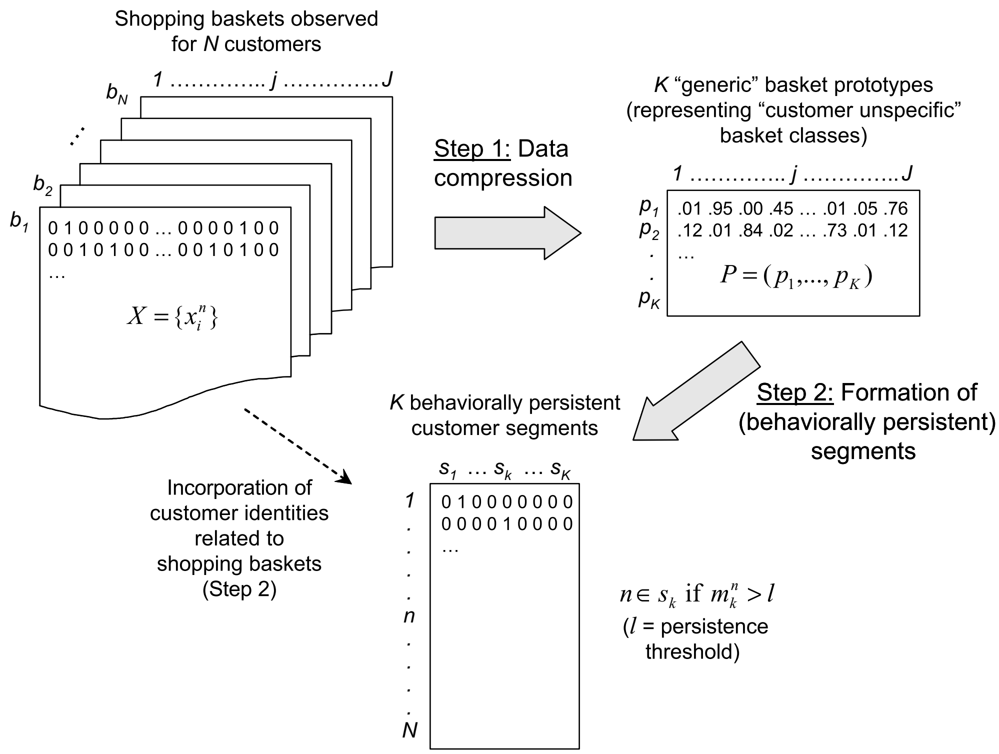

FIGURE 1

Conceptual Framework of the Proposed Two-Stage Procedure

interdependencies with a certain minimum degree of persistence. More specifically, the method proceeds in the following stepwise manner: Using a methodology similar to the attempts provided by Schnedlitz et al. (2001) or Decker and Monien (2003), shopping baskets from a customer transaction database are first compressed onto a set of so-called "generic" prototypes (Step 1). As a result of this data compression step, the system of prototypes constitutes a classification of the observed shopping baskets, which is characterized by more distinguished complementary cross-category coincidences within each of the derived shopping basket classes; however, this classification is "generic" in the sense that with proceeding time periods, customers are consciously allowed to freely "fluctuate" across the partition of basket classes depending on their respective "best-fitting" patterns of multicategory choices.

Only in the second step of the procedure, the customer identities associated with the observed shopping baskets are introduced for determination of corresponding behaviorally persistent customer segments. Using the distribution of basket class memberships of each single customer's purchase history, segment memberships are derived for arbitrary levels of behavioral persistence. In the following subsections, we comment on the analytical steps in more detail.

Step 1: Construction of Market Basket Prototypes

Consistent with prior work, personalized market baskets are considered as "pick-any/*J*" data (Manchanda et al., 1999; Russell & Petersen, 2000). Hence, each shopping basket is represented as a *J*-dimensional binary data vector $x_i \in \{0,1\}^J$, with i being a pointer to the elongated arrangement $\{b_1, b_2, \ldots, b_N\}$ of the "stacked" sequences of purchase transactions and J representing the retailer's assortment of categories. This data format implies that utilization of the customer-specific provenance of shopping baskets (indicated by x_i^n for the b_n transactions realized by Customer n) is postponed to a later stage of the analysis.

The task of finding a partition of the data into a fixed number of K "generic" basket classes $C = \{c_1, c_2, \ldots, c_K\}$ with more distinguished complementary joint-purchase incidences within classes requires resolution of the following objective function ("minimum dispersion criterion" of the principal point problem; cf. Bock, 1999; Jain & Dubes, 1988):

$$\sum_k \sum_{i \in c_k} d(x_i, p(x_i)) \rightarrow \min_{C,P} \tag{1}$$

where $P = (p_1, p_2, \ldots, p_K)$ denotes a set of prototypes or "centroids" with $p_k \in \mathbb{R}^J \, \forall k$ and $d(\cdot)$ being a distance measure. For any optimum configuration (C^*, P^*), the condition $p^*(x_i) = \arg\min\{d(x_i, p_k^*), k = 1, \ldots, K\}$ holds and warrants that each Basket x_i is mapped onto its minimum distant prototype. Furthermore, when using the Euclidean distance metric, it can be shown that the prototypes $p_k^* = \bar{x}_{c_k^*}$ are equal to class-specific means for the corresponding partition as generated by the optimal prototypes under stationarity conditions (Bock, 1999).

Since purchase incidences are encoded as (typically extremely sparse) binary vectors and we aim at detecting complementary cross-effects, the well-known asymmetric Jaccard coefficient giving more weight to joint purchases than to common zeros (i.e., nonpurchases) is preferred here as a distance measure. A simple extension of the Jaccard coefficient to measure the distance between a binary vector x_i and a real-valued prototype p_k is as follows:

$$d(x_i, p_k) = 1 - \frac{(x_i, p_k)}{\|x_i\|^2 + \|p_k\|^2 - (x_i, p_k)} \quad \forall k = 1, \ldots, K, \tag{2}$$

where (x_i, p_k) denotes the scalar product of vectors x_i and p_k. Notice that the subtrahend in Equation (2) is the Tanimoto similarity coefficient (Anderberg, 1973).

Probably the most famous approach for solving the principal point problem is the iterative K-means clustering algorithm. Based on a given initial partition, the method minimizes Criterion (1) recursively with respect to $C(t) \rightarrow P(t) \rightarrow C(t + 1) \rightarrow P(t + 1) \ldots$ and converges after a finite number of t iterations. Albeit any arbitrary distance measure can be embedded in the algorithm (Anderberg, 1973; MacQueen, 1967), it is predominantly implemented using Euclidean distances (hence, the term K-means). Though convergence to the next local minimum is generally guaranteed, the quality of the final cluster solution is known to heavily depend on the starting partition. To cope with this "algorithmic variability," generation of several solutions for different random initializations and subsequent search for the "best fitting" partition or heuristics for obtaining "proper starting values" is recommended (Gordon & Vichi, 1998; Hornik, 2005). Such selection strategies, however, make K-means methods computationally expensive and impractical when the number of data points becomes very large and high-dimensional, which is the case for shopping basket data derived from several hundred thousands of retail transactions and large assortment sizes.

Fortunately, there are other methods available to solve the principal point problem. Descending from the field of machine learning, numerous online versions of K-means type clustering are available, known as competitive learning or vector quantization (VQ) algorithms (Hastie, Tibshirani, & Friedman, 2001; Ripley, 1996; in a marketing context, also see Elrod et al., 2002). In contrast to offline K-means methods, the VQ approach minimizes (1) via stochastic approximation by directly manipulating the prototype system in an iterative updating scheme. In the context of market basket analysis, probably the most appealing property of VQ-type partitioning techniques is that only one single data point (e.g., a shopping basket accruing at electronic retail POS check-out systems) is required per iteration. Hence, they are suitable to process datasets of practically unlimited size. The generalized VQ algorithm adopted here for quantization of shopping baskets included in a given retail transaction database proceeds as follows:

1. Start with a random initialization of the set of Prototypes P by drawing K "seed points" from the input dataset of market baskets.
2. Compute the distances between a randomly chosen market basket Vector x_i and each Prototype p_k according to Equation (2).
3. Determine the minimum distant ("winning" or "best fitting") prototype $d(x_i, p_{k^*}) = \min\{d(x_i, p_k), k = 1, \ldots, K\}$ to x_i.
4. Update the "winning" prototype according to the following "learning rule:"

$$p_{k^*} := p_{k^*} + \alpha(t) \cdot (x_i - p_{k^*}) \tag{3}$$

where $\alpha(t)$ represents a "learning factor," which to fulfill the conditions for stochastic approximation is conceived as a monotonically shrinking function of time [i.e., $\lim_{t\to\infty}\alpha(t) = 0$].

5. Repeat Steps 2 to 4 until convergence (i.e., if prototype improvements are becoming very small) or the prespecified maximum number of iterations is reached.

The VQ procedure proposed here differs from more conventional implementations in the following respect: Both due to data sparsity and for conceptual reasons (i.e., focus on complementary cross-effects), we are using Jaccard distances for determination of the "winner" p_{k^*} but perform a Euclidean-like updating following the learning rule (3); hence, the notion "generalized" VQ. The important practical reason for doing so is that after convergence, we obtain prototypes that coincide with respective class means $\bar{x}_{c_k^*}$ and therefore can be easily interpreted as empirical expectations of observing a value of unity (Leisch, 2006).[2] Consequently, in the present context, each j-element of an optimal prototype vector p_{k^*} indicates the purchase incidence probability of the corresponding product category within the "generic" shopping basket class c_{k^*}. Exceptionally (un)marked combinations of these class-conditional probabilities are indicative for (weaker) stronger cross-category purchase complementarities at the basket class level and will serve as a basis for further investigation.

Step 2: Determination of Behaviorally Persistent Customer Segments

The second step of the proposed procedure is responsible for assigning customers dynamically to segments that exhibit a certain level of behavioral persistence as predefined by the analyst. Let us first clarify the notion of behavioral persistence that is argued here: The outlined data compression step produces a K partition C^* of market baskets. Due to the absence of any constraints, the single transactions of a particular customer's purchase history are not necessarily (or even unlikely to be) mapped onto one and the same prototype. Consequently, the shopping basket classes c_{k^*} generated by the set of "generic" prototypes represent baskets with distinctive patterns of category co-occurrences observed in the pooled dataset. The term "generic" suggests that they do not yet recognize the customer identities behind the individual shopping baskets; however, disaggregate information on the customer- (or time period, store-, etc.) specific identification of transactions assigned to the various "generic" basket classes is still available after the data reduction step. Consequently, the relative frequency distribution of basket class assignments over a customer's purchase history can be exploited to construct behaviorally persistent segments.

Consider, for example, a transaction database that includes $b_n = 10$ purchase transactions from Customer n throughout a (fixed) time interval of interest. Assume further that the current partition consists of $K = 3$ classes, and seven of Customer's n shopping baskets x_i^n $(i \in b_n)$ are assigned to class 1, three transactions to class 2, and none to class 3. Then there seems to be a strong indication that the prototype representing class 1 is best summarizing the past purchase behavior of Customer n. Using the terminology introduced earlier, we would qualify this consumer to expose a considerable degree of "behavioral persistence" with respect to the purchase pattern represented by shopping basket class 1. More formally, construction of behaviorally persistent customer segments is based on a simple majority "voting" for best-fitting class assignments. For each Customer n, we therefore calculate the number of basket class k assignments:

$$m_k^n = \sum_{i=1}^{b_n} 1_{\{x_i^n \in c_k\}} \tag{4}$$

The logical expression $1_{\{x_i^n \in c_k\}}$ equals 1 if Transaction i from Customer n is a member of basket class k, or zero otherwise. Notice that the set of voting measures m_k^n/b_n result in relative indices that can be interpreted as fuzzy memberships. Taking the maximum value for each customer would provide a segmentation of the customer base with respect to the relative basket class assignment frequencies; however, since the number of

[2] Notice that once the issues of dataset size and local optima are resolved, standard offline expectation-based K-means clustering using Jaccard distances will result in virtually the same partition of the dataset as the proposed VQ technique. This is warranted by the identical objective function. Of course, different partitions are to be expected when Euclidean distance metrics are used. Due to the "excessive zeros problem" inherent to market basket data, however, the Jaccard coefficient proved to be a robust distance metric in a number of experiments, as reported by Mild and Reutterer (2001).

transactions b_n is varying across customers, this would not warrant that customers with diverging purchase frequencies were targeted differently. The latter could be considered only in an additional step by sorting out more frequent from infrequent customers or by using any other measures of interest (e.g., RFM or CLV metrics). Our approach for assisting marketers in customizing the content of direct marketing campaigns can be combined easily with more traditional methods for target segment selection.

To convert our understanding of behavioral persistence into a segment-specific concept in the absence of any such further information for segment selection, Customer n is assigned to a behaviorally persistent Segment s_k only if her or his absolute values m_k^n are exceeding a user-defined threshold value l:

$$s_k = \{n \in N | m_k^n > l\} \tag{5}$$

This assignment rule has the natural consequence that not all customers will be treated as behaviorally persistent and thus worth targeting with segment-specific customized marketing actions. Conditional that threshold level l will be held constant, with evolving time periods more and more customers will be included in persistence segments. Hence, the segmentation is dynamically updated. As it is illustrated in the following empirical application, determination of meaningful threshold levels can be controlled by management.

EMPIRICAL APPLICATION

To illustrate the procedure described earlier, we use transaction data of about 470,000 active loyalty program members of a DIY retailer. We used $J = 153$ product categories (PCs) which have been purchased by at least 2% of the active customers during a 1 1/2-year observation period. For convenience reasons, in the present application, individual shopping baskets were aggregated across quarters. Hence, each customer is characterized by a set of a maximum of six such pooled “shopping baskets.” Purchase rates are rather heterogeneous across customers; on average, 3.1 purchase incidences were observed per quarter. In the present application, quarterly aggregation of customers’ transactions is in accordance with the planning horizons of marketing management and does not affect the conceptual framework underlying our approach. The only notable difference to the previous exposition is that the concept of behavioral persistence now applies to customers’ shopping habits at the DIY retailer under study within subsequent quarters instead of single purchase occasions.

Derived Basket Prototypes and Customer Segments

As any other clustering task in marketing, the appropriate choice of the number of shopping basket classes K also is an issue with the employed VQ procedure. Numerous heuristics were proposed to assist the analyst in this respect (Milligan & Cooper, 1985). On the other hand, many authors also express doubts about the existence of “quasinatural” groupings in empirical datasets (Aldenderfer & Blashfield, 1984; Dubes & Jain, 1979). Even though one is willing to accept this assumption, it is very unlikely that such a presumably “natural” grouping is detectable with an efficiently manageable and managerially acceptable number of clusters in light of the excessively large and high-dimensional dataset of joint category purchase incidences at hand.

Evaluating the combined information available on statistical measures for internal cluster validity and managerial considerations, a solution with $K = 20$ basket classes was considered to be a decent and adequate representation of the data. As a result of the data compression step of our procedure, each basket class now can be best characterized by its corresponding profile of prototypical category purchase probabilities, with combinations of particularly outstanding values signaling stronger degrees of cross-category purchase complementarities. Due to space limitations, we restricted our attention on two rather distinctive basket classes that will be of further interest in the subsequent report on results from a direct marketing experiment. Details on the composition of other basket classes are available from the authors upon request.

Consider the pictorial representation of the class-conditional prototypical profile of category choice probabilities representing basket class 2 according to the solid line in the upper graph of Figure 2. Instead, the gray bars are representing the unconditional probabilities of category purchase incidences. Quite obviously, this prototype vector can be characterized by very high purchase frequencies in only a few categories whereas other categories are indicated to be

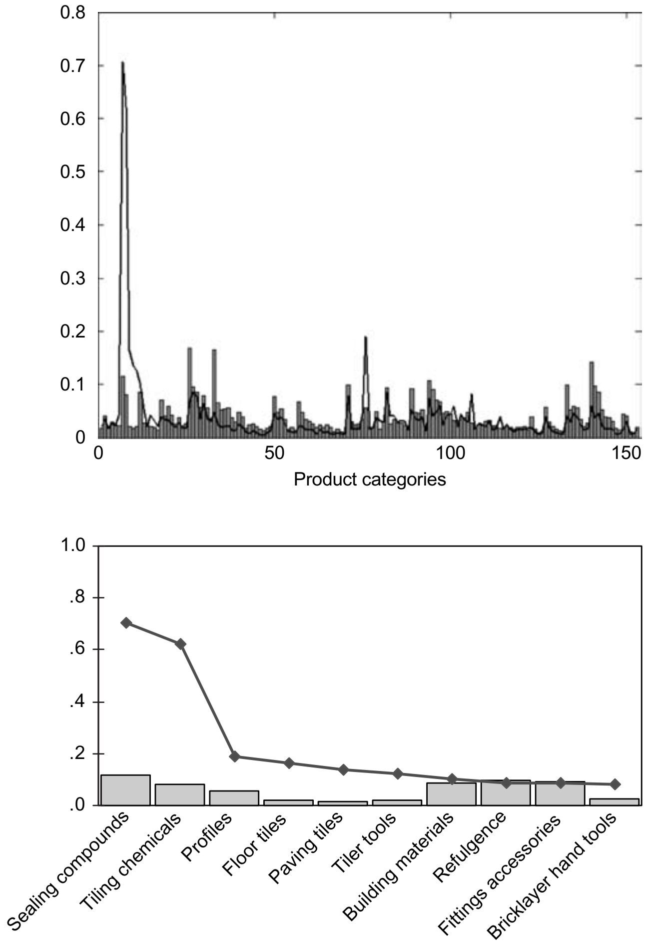

FIGURE 2

Prototype 2 Profile of Category Choice Probabilities

purchased significantly below average. From the lower graph (emphasizing the top-10 categories in terms of class-conditional probabilities), it becomes obvious that the purchase behavior of this "generic" class of shopping baskets is clearly dominated by remarkably high purchase incidences of categories required for conducting tiling projects including various accessories [viz., "sealing compounds" (.706), "tiling chemicals" (.622), "profiles" (.190), "floor tiles" (.164), etc.] and only moderate class-conditional choice probabilities in the remaining categories. Thus, management labeled this class as the "tiling shopping basket class."

A completely different picture yields inspection of the prototype profile corresponding to basket class 19 depicted in Figure 3. Again, remarkably high conditional purchase incidences can be observed for a selected number of categories such as "substrates" (.567), "bed and group plants" (.383), "seeds" (.276), "hardy plants" (.157), and so on, which altogether are related to gardening activities. Hence, this prototype can be considered to represent a "gardening shopping basket class." Other basket classes are characterized by their own prototypical basket compositions (i.e., cross-category purchase interdependencies) that are clearly distinctive from those further investigated here.

Based on the "generic" shopping basket classification, the second stage of our procedure for target-segment selection was conducted. Table 1 provides an overview of the resulting segmentation of loyalty program members (in terms of absolute frequencies) when different levels of behavioral persistence are imposed. Management considered a threshold value of $l = 3$ to be appropriate for determination of behaviorally persistent segments and further targeting. Hence, customers assigned to the highlighted shopping basket class 2 or 19, respectively, in at least 3 of the 6 quarters under observation are considered as behaviorally persistent in the aforementioned sense.

TABLE 1 Basket Class Frequencies for Different Levels of Behavioral Persistence

		NO. OF CUSTOMERS WHO WERE MEMBERS IN BASKET CLASS FOR AT LEAST . . .					
BASKET CLASS	**NO.**	**ALL QUARTERS**	**LAST 4 QUARTERS**	**LAST 3 QUARTERS**	**3 OF 4 LAST QUARTERS**	**LAST 2 QUARTERS**	**3 OF 6 QUARTERS**
Tiling	2	11	31	93	431	693	*2,999*
Gardening	19	22	45	247	1,116	4,033	*8,012*

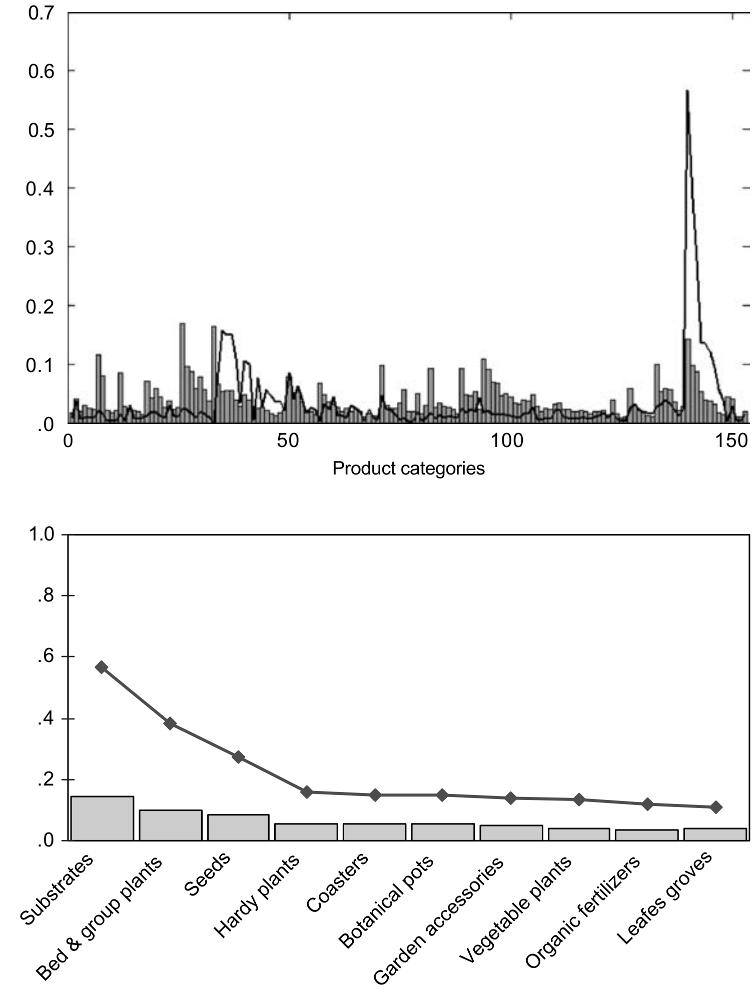

FIGURE 3

Prototype 19 Profile of Category Choice Probabilities

A Direct Marketing Experiment

To test the empirical performance of the proposed data-driven approach for segment construction and to evaluate both the effectiveness and efficiency of target marketing actions implied by the outcome of the procedure, a controlled field experiment was conducted. Discussions with managers resulted in the selection of the two different target segments as already illustrated (see Table 1, last column). One segment comprising a total number of 2,999 customers was considered to exhibit a substantial degree of behavioral persistence with respect to categories including various tiling products. The second segment consisted of 8,012 loyalty club members with considerable interest in gardening-related product categories. To customize the direct marketing campaigns, two suitably adapted cover letters and segment-specific flyers were designed that featured specific items of interest recruited from the most distinguished tiling- and gardening-related product categories as depicted in the previous section, respectively.

The experiment was conducted separately for both target segments. Furthermore, it was designed to allow for an evaluation of differential effects due to targeting media (e-mail vs. conventional mailing) and the potential impact of an additional coupon accompanied to the direct mailing. For identification of the effects of the various treatment factors, each segment was accordingly split into subsamples and treated differently. One of these subsamples received the segment-specific flyer via either e-mail or standard mail. This group was split again, with one split group receiving an additional

TABLE 2 Experimental Groups and Sample Sizes

TILING	GARDENING	TREATMENT CHARACTERISTICS OF SUBSAMPLE
902	2,444	Untreated segment members (control group)
92	190	E-mail & coupon sent to segment members
92	190	E-mail sent to segment members
903	2,444	Mail & coupon sent to segment members
903	2,444	Mail sent to segment members
964	2,596	Mail & coupon sent to random sample
964	2,596	Mail sent to random sample

coupon offering a 10% discount on one item of the customer's choice (including items not featured in the flyer). The other group of the selected segment received no mailing and served as a control group. In addition, to separate the effect of the direct mailing from the effect of choosing the target segment, a group of randomly chosen nonsegment members received the same offers. Table 2 summarizes the characteristics of the various groups and quotes the sizes of the resulting subsamples. Note that the unequally smaller sizes of the respondent groups receiving e-mails were due to limited availability of addresses. Furthermore, we were not able to send e-mail messages to randomly selected customers.

For evaluation of the experiment, we used a complete record of 1-year purchase transactions realized by the customers under treatment on a weekly basis. The DIY retailer usually maintains a biweekly schedule for mass mailings, which also was the case for the coupon duration of the direct mailing campaigns. For the purpose of evaluation, we therefore attribute 2 weeks (Weeks 33 & 34) of respondents' sales and profit contributions as being influenced by the previously described actions. To measure the impacts on sales and profits, we formulated a generalized linear model. An RFM criteria based ABC-classification of customers, the sample membership (random or target segment), and a set of dummy-coded variables that control for monthly seasonality in the data serve as independent variables. To measure the campaign effectiveness for the differently treated subsamples, we included first order interaction effects between sample membership and the treatment factors "Mailing," "E-mail," and "Coupon." Sales and profits pooled across each of the independent factors act as response variables. With exemption of the seasonal parameters (not reported here), Table 3 shows the parameters and *t*-test statistics for the resulting four submodels.

TABLE 3 Experimental Results

	TILING				GARDENING			
	SALES		PROFIT		SALES		PROFIT	
PARAMETER	COEFF.	*t*	COEFF.	*t*	COEFF.	*t*	COEFF.	*t*
Constant	0.07	*3.2*	0.02	*2.2*	0.46	*38.4*	0.25	*60.3*
ABC = A	17.24	*1102.5*	6.56	*1145.0*	15.61	*1402.5*	6.09	*1553.7*
ABC = B	5.69	*404.6*	2.21	*428.9*	5.39	*726.5*	2.15	*823.4*
Random sample	−0.02*	*−1.2*	0.12	*25.1*	−0.03	*−4.2*	−0.13	*−56.2*
Random sample*mailing	0.70	*9.9*	0.17	*6.8*	−0.10	*−2.7*	−0.07	*−4.8*
Target segment*mailing	2.70	*37.5*	1.07	*40.7*	0.95	*23.8*	0.24	*17.0*
Random sample*coupon	0.29	*4.2*	−0.09	*−3.4*	1.85	*47.6*	0.39	*28.2*
Target segment*coupon	−0.18	*−2.6*	0.07	*2.7*	1.58	*41.1*	0.57	*42.2*
Target segment*e-mail	1.29	*6.1*	0.22	*2.9*	1.91	*14.4*	1.21	*25.8*
R-squared	0.850		0.858		0.778		0.818	

All parameters except one marked with * are significant at the 95% confidence level.

TABLE 4 Sales and Profits Effects for Combinations of Direct Marketing Actions

		TILING EXPERIMENT		GARDENING EXPERIMENT	
		SALES (%)	PROFITS (%)	SALES (%)	PROFITS (%)
Random Sample	Mailing only	18.4	12.1	−2.3	−3.6
	Mailing & coupon	26.1	6.1	38.9	17.4
Target segment	Mailing only	71.1	74.7	21.1	13.0
	Mailing & coupon	66.4	79.3	56.4	44.1
	E-mail only	33.9	15.4	42.6	65.8
	E-mail & coupon	29.2	20.1	77.9	97.0

As expected, both sales and profits are considerably higher for the premium segment of customers (ABC = A). The random sample-specific parameters for the mailing show a positive impact on both sales and profits in the tiling and a small negative effect in the gardening experiment as compared to the baseline of the segment control group (which was not targeted). If a coupon was added to the mailings, sales were enhanced in both experiments whereas profits were slightly decreased in the tiling experiment. The opposite is true for the increase in profits in the gardening experiment. The target-segment-specific interaction parameters, however, clearly reveal the importance of a proper target-group selection. Obviously, the mailing exerted a much higher impact when sent to the specified target group as compared to a randomly selected group of customers. Regarding the additional effect of a coupon, negative relative sales effects with respect to the random sample can be reported. Albeit the e-mail campaign shows relatively strong positive impacts for both experimental groups, these comparative effects could not be thoroughly evaluated for the e-mail media since e-mail messages where exclusively sent to (a small number of) target-segment members, but not to the random groups. Hence, the e-mail variable covers both effects.

To evaluate the effectiveness of the various combinations of direct marketing actions under study, the respective overall effects on profits and sales are compared to the base case of no campaign during the corresponding month including Weeks 33 and 34. The resulting figures are given in Table 4 as percentage deviations from a zero-effects model for the untargeted segment control groups. It can be seen that the proposed target-segment selection clearly outperformed the random group in all instances of both experiments. Again, due to the restricted availability of valid e-mail addresses, our design does not permit separation of the effect of target-group selection for the e-mail experiments. Nevertheless, a clear positive effect on sales and profits can be attributed to the e-mail campaign. Comparing the effectiveness of conventional mailings and e-mail messages, no method outperforms the other in both experiments. Given the low cost of targeted e-mail messages, this instrument does turn out to be a valuable amendment to more conventional direct marketing media.

Table 5 provides the results from calculating two measures for an additional comparison; namely, Return on Advertising Spend (ROAS) as the additional revenue generated by the direct marketing media divided by its cost and Return on Investment (ROI) as the additional profits generated by the media divided by its cost. These two measures can serve as an evaluation of the efficiency of the conducted variations of direct marketing campaigns. The full cost for each segment-specific mailing was calculated as 0.65 EUR whereas the cost of a customized e-mail was 0.07 EUR.

As can be seen in Table 5, the customized mailings to random samples never exhibit an ROI above 1 for both experimental groups and therefore turn out to be

TABLE 5 ROAS and ROI for Combinations of Direct Marketing Actions

		TILING EXPERIMENT		GARDENING EXPERIMENT	
		ROAS	ROI	ROAS	ROI
Random Sample	Mailing only	1.1	0.3	−0.2	−0.1
	Mailing & coupon	1.5	0.1	2.7	0.5
Target segment	Mailing only	4.2	1.7	1.5	0.4
	Mailing & coupon	3.9	1.8	3.9	1.2
	E-mail only	18.4	3.2	27.3	17.3
	E-mail & coupon	15.8	4.1	49.9	25.5

not profitable. Generally, an ROI above 1 can be reported for mailings directed to the target segments; however, a notable exemption is the case of the subsample from the gardening segment, which did not receive an additional coupon. Nevertheless, compared to the random sample counterpart, at least an improvement can be registered. Due to the still relatively high cost of conventional direct mailings, the effects of such campaigns should be measured and controlled diligently, and the targeted groups need to be chosen carefully. But as a result of the combination of high effectiveness and low cost, e-mail communication can extremely leverage profitability. Our results show strong positive results for the e-mail subsamples in terms of both the ROAS and the ROI measures; however, these results should be taken with caution because of the small sample sizes.

CONCLUSION

We introduced a novel approach for assisting retail marketing managers in planning segment-specific, customized direct marketing campaigns. In contrast to most existing approaches, both target-segment selection and customizing the content of a direct mailing are addressed. The proposed two-stage procedure consists of a data compression step that serves for deriving prototypes of distinguished cross-category interdependencies among the categories included in a retail assortment. Due to their adaptive nature, these prototypes can be updated continuously for evolving time periods. Equally, the second stage entailing the construction of behaviorally persistent segments is flexible enough to be dynamically adjusted. Furthermore, the degree of behavioral persistence responsible for segment formation can be controlled easily by management via suitable choice of a threshold parameter. The empirical performance of our approach is demonstrated for two different target segments selected from a DIY retailer's customer database. Both segments exhibiting a certain minimum degree of behavioral persistence were selected for targeting in a controlled direct marketing experiment. The empirical findings support the usefulness of the proposed procedure in terms of impacts on both sales and profits. Compared to a randomized customer sample, our recommended segments were targeted more effectively and efficiently.

A research agenda for further validation of the empirical performance of the presented methodology should include the following tasks: First, an extension of segment-specific targeting campaigns to segments other than the two illustrated in the present application study is advisable. Second, the approach should be evaluated for different target segments resulting from varying degrees of the behavioral persistence measure proposed in this article. This also could include a combination with more traditional methods for target-segment selection such as RFM or CLV metrics. Third, the linkage with predictive approaches to market basket analysis of the type proposed by Manchanda et al. (1999) or Russell and Petersen (2000) possibly could further enhance the managerial implications of our approach to segment-specific category selection for target marketing. Some nonexperimental findings in this direction are available in a recent work by Botzug and Reutterer (2006). Finally, a comparison to the impacts of one-to-one targeting strategies and applications to other retail or nonretail industries would be helpful.

From a more conceptual perspective, our proposed measure of behavioral persistence is restricted to the diagonal elements of a shopping basket classes' switching matrix for the purchase histories of the available customer database. Hence, further research endeavors also could be devoted to studying customers' latent basket class switching behavior. In addition to behaviorally persistent segments, identification of significant switching paths across the derived partition of prototypical shopping basket classes could serve as a valuable basis for targeting switching segments in accordance with their basket class transition probabilities.

REFERENCES

Ainslie, A., & Rossi, P.E. (1998). Similarities in Choice Behavior Across Product Categories. Marketing Science, 17, 91–106.

Aldenderfer, M.S., & Blashfield, R.K. (1984). Cluster Analysis. Beverly Hills, CA: Sage.

Anderberg, M.R. (1973). Cluster Analysis for Applications. New York: Academic Press.

Andrews, R.L., & Currim, I.S. (2002). Identifying Segments With Identical Choice Behaviors Across Product Categories: An Intercategory Logit Mixture Model. International Journal of Research in Marketing, 19, 65–79.

Berger, P.D., & Nasr, N.I. (1998). Customer Lifetime Value: Marketing Models and Applications. Journal of Interactive Marketing, 12, 17–30.

Berry, M.J.A., & Linoff, G. (2004). Data Mining Techniques for Marketing, Sales, and Customer Relationship Management. New York: Wiley.

Bock, H.-H. (1999). Clustering and Neural Network Approaches. In W. Gaul & H. Locarek-Junge (Eds.), Classification in the Information Age. Proceedings of the 22nd annual conference of the German Classification Society (pp. 42–57). Heidelberg, Germany: Springer.

Boztug, Y., & Reutterer, T. (2006). A Combined Approach for Segment-Specific Analysis of Market Basket Data. SFB 649 Discussion Paper 2006–06, Humboldt-Universität, Berlin.

Britan, G., & Mondschein, S.V. (1996). Mailing Decisions in the Catalog Sales Industry. Management Science, 42, 1364–1381.

Chib, S., Seetharaman, P.B., & Strijnev, A. (2002). Analysis of Multi-Category Purchase Incidence Decisions Using IRI Market Basket Data. In P.H. Franses & A.L. Montgomery (Eds.), Econometric Models in Marketing (Vol. 16, pp. 57–92). Amsterdam: Elsevier Science.

Colombo, R., & Jiang, W. (1999). A Stochastic RFM Model. Journal of Interactive Marketing, 13, 2–12.

Corstens, J., & Corstens, M. (1995). Store Wars. The Battle for Mindspace and Shelfspace. Chichester, England: Wiley.

Decker, R. (2005). Market Basket Analysis by Means of a Growing Neural Network. International Review of Retail, Distribution and Consumer Research, 15, 151–169.

Decker, R., & Monien, K. (2003). Market Basket Analysis With Neural Gas Networks and Self-Organizing Maps. Journal of Targeting, Measurement and Analysis for Marketing, 11, 373–386.

Dubes, R., & Jain, A.K. (1979). Validity Studies in Clustering Methodologies. Pattern Recognition, 11, 235–254.

Elrod, T., Russell, G.J., Shocker, A.D., Andrews, R.L., Bacon, L., Bayus, B.L., Carroll, D.J., Johnson, R.M., Kamakura, W.A., Lenk, P., Mazanec, J.A., Rao, V.R., & Shankar, V. (2002). Inferring Market Structure From Consumer Response to Competing and Complementary Products. Marketing Letters, 13, 219–230.

Elsner, R., Krafft, M., & Huchzermeier, A. (2004). Optimizing Rhenania's Direct Marketing Business Through Dynamic Multilevel Modeling (DMLM) in a Multicatalog-Brand Environment. Marketing Science, 23, 192–206.

Fader, P.S., Hardie, B.G.S., & Lee, K.L. (2005). RFM and CLV: Using Iso-Value Curves for Customer Base Analysis. Journal of Marketing Research, 42, 415–430.

Gordon, A.D., & Vichi, M. (1998). Partitions of partitions. Journal of Classification, 15, 265–285.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). The Elements of Statistical Learning. New York: Springer.

Haughton, D., & Oulabi, S. (1993). Direct Marketing Modeling With CART and CHAID. Journal of Direct Marketing, 11, 7–19.

Hornik, K. (2005). Cluster Ensembles. In C. Weihs & W. Gaul (Eds.), Classification—The Ubiquitous Challenge. Proceedings of the 28th annual conference of the German Classification Society (pp. 65–72). Heidelberg, Germany: Springer.

Hruschka, H., Lukanowicz, M., & Buchta, C. (1999). Cross-Category Sales Promotion Effects. Journal of Retailing and Consumer Services, 6, 99–105.

Hughes, A.M. (1996). The Complete Database Marketer. New York: McGraw-Hill.

Humby, C., & Hunt, T. (2003). Scoring Points. How Tesco Is Winning Customer Loyalty. London, Sterling, VA: Kogan Page.

Jain, A.K., & Dubes, R.C. (1988). Algorithms for Clustering Data. Englewood Cliffs, NJ: Prentice Hall.

Jain, D., & Singh, S.S. (2002). Customer Lifetime Value Research in Marketing. A Review and Future Directions. Journal of Interactive Marketing, 16, 34–46.

Kahn, B., & McAlister, L. (1997). Grocery Revolution. The New Focus on the Consumer. New York: Addison-Wesley.

Kumar, V., & Shah, D. (2004). Building and Sustaining Profitable Customer Loyalty for the 21st Century. Journal of Retailing, 80, 317–330.

Leisch, F. (2006). A Toolbox for k-Centroids Cluster Analysis. Computational Statistics and Data Analysis (forthcoming).

MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. In L.M. Le Cam & J. Neyman (Eds.), Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (Vol. 1, pp. 281–297). Berkeley: University of California Press.

Manchanda, P., Ansari, A., & Gupta, S. (1999). The "Shopping Basket." A Model for Multi-Category Purchase Incidence Decisions. Marketing Science, 18, 95–114.

Mild, A., & Natter, M. (2003). Collaborative Filtering or Regression Models for Internet Recommendation Systems? Journal of Targeting, Measurement and Analysis for Marketing, 10, 303–313.

Mild, A., & Reutterer, T. (2001). Collaborative Filtering Methods for Binary Market Basket Data Analysis. Lecture Notes in Computer Science [Special Issue]. Active Media Technology, 2252, 302–313.

Mild, A., & Reutterer, T. (2003). An Improved Collaborative Filtering Approach for Predicting Cross-Category Purchases Based on Binary Market Basket Data. Journal of Retailing and Consumer Services, 10, 123–133.

Milligan, G.W., & Cooper, M.C. (1985). An Examination of Procedures for Determining the Number of Clusters in a Data Set. Psychometrika, 50, 159–179.

Morwitz, V.G., & Schmittlein, D.C. (1998). Testing New Direct Marketing Offerings. The Interplay of Management Judgment and Statistical Models. Management Science, 44, 610–628.

Mulhern, F.J. (1999). Customer Profitability Analysis. Measurement, Concentration, and Future Directions. Journal of Interactive Marketing, 13, 25–40.

Pearson, M.M., & Gessner, G.H. (1999). Transactional Segmentation of Slow Customer Defections. Marketing Management, 8, 16–23.

Prinzie, A., & Van den Poel, D. (2005). Constrained Optimization of Data-Mining Problems to Improve Model Performance. A Direct-Marketing Application. Expert Systems With Applications, 29, 630–640.

Reichheld, F.F. (1996). The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value. Boston: Harvard Business School Press.

Reinartz, W.J., & Kumar, V. (2000). On the Profitability of Long-Life Customers in a Noncontractual Setting. An Empirical Investigation and Implications for Marketing. Journal of Marketing, 64, 17–35.

Reinartz, W.J., & Kumar, V. (2002). The Mismanagement of Customer Loyalty. Harvard Business Review, 80, 86–94.

Ripley, B.D. (1996). Pattern Recognition and Neural Networks. Cambridge, England: Cambridge University Press.

Rossi, P.E., McCulloch, R.E., & Allenby, G.M. (1996). The Value of Purchase History Data in Target Marketing. Marketing Science, 15, 321–340.

Russell, G.J., & Kamakura, W.A. (1997). Modeling Multiple Category Brand Preference With Household Basket Data. Journal of Retailing, 73, 439–461.

Russell, G.J., & Petersen, A. (2000). Analysis of Cross Category Dependence in Market Basket Selection. Journal of Retailing, 76, 367–392.

Russell, G.J., Ratneshwar, S., Shocker, A.D., Bell, D., Bodapati, A., Degeratu, A., Hildebrandt, L., Kim, N., Ramaswami, S., & Shankar, V.H. (1999). Multiple-Category Decision-Making: Review and Synthesis. Marketing Letters, 10, 319–332.

Rust, R.T., Lemon, K., & Zeithaml, V.A. (2004). Return on Marketing: Using Customer Equity to Focus Marketing Strategy. Journal of Marketing, 68, 109–127.

Rust, R.T., Zeithaml, V., & Lemon, K.N. (2000). Driving Customer Equity. How Customer Lifetime Value Is Reshaping Corporate Strategy. New York: Free Press.

Schnedlitz, P., Reutterer, T., & Joos, W. (2001). Data-Mining und Sortimentsverbundanalyse im Einzelhandel. In H. Hippner, U. Küsters, M. Meyer, & K.D. Wilde (Eds.), Handbuch Data Mining in Marketing. Knowledge Discovery in Marketing Databases (pp. 951–970). Wiesbaden, Germany: Vieweg.

Seetharaman, P.B., Ainslie, A., & Chintagunta, P.K. (1999). Investigating Household State Dependence Effects Across Categories. Journal of Marketing Research, 36, 488–500.

Seetharaman, P.B., Chib, S., Ainslie, A., Boatwright, P., Chan, T., Gupta, S., Mehta, N., Rao, V., & Strijnev, A. (2005). Models of Multi-Category Choice Behavior. Marketing Letters, 16, 239–254.

Winer, R. (2001). A Framework for Customer Relationship Management. California Management Review, 43, 89–105.

Zahvai, J., & Levin, N. (1997). Applying Neural Computing to Target Marketing. Journal of Direct Marketing, 11, 5–22.