

A Data Mining Framework for Targeted Category Promotions

Work in progress - please do not quote or cite.

Reutterer Thomas^{*1}, Kurt Hornik^{†2}, Nicolas March¹ and Gruber Kathrin¹

¹Institute for Service Marketing and Tourism, WU Vienna University of Economics and Business

²Institute for Statistics and Mathematics, WU Vienna University of Economics and Business

Abstract

This research presents a new approach to derive recommendations for segment-specific, targeted marketing campaigns on the product category level. The proposed methodological framework serves as a decision support tool for customer relationship managers or direct marketers to select attractive product categories for their target marketing efforts, such as segment-specific rewards in loyalty programs, cross-merchandising activities, targeted direct mailings, customized supplements in catalogues, or customized promotions. The proposed methodology requires customers' multi-category purchase histories as input data and proceeds in a stepwise manner. It combines various data compression techniques and integrates an optimization approach which suggests candidate product categories for segment-specific targeted marketing such that cross-category spillover effects for non-promoted categories are maximized. To demonstrate the empirical performance of our proposed procedure, we examine the transactions from a real-world loyalty program of a major grocery retailer. A simple scenario-based analysis using promotion responsiveness reported in previous empirical studies and prior experience by domain experts suggests that targeted promotions might boost profitability between 15% to 128% relative to an undifferentiated standard campaign.

Keywords: cross-category purchases target marketing customized coupons clustering association rule mining

*thomas.reutterer@wu.ac.at

†kurt.hornik@wu.ac.at

1. Introduction

Even though their popularity might already have exceeded peak levels, with an average of more than twelve memberships per U.S. household and a reported half of U.S. adults being enrolled in at least one, loyalty programs continue to be a mainstay in customer relationship management (Kivetz and Simonson, 2003; Ferguson and Hlavinka, 2007; Berry, 2013). Many companies invest tremendous amounts of money both in their online and offline loyalty program environments to strive for building and preserving loyalty of their primary clientele. In the marketing literature, however, there is mixed evidence on the effectiveness of loyalty programs and many recent efforts to improve these concentrate on various program design components, tier or reward structures; excellent reviews are provided by Liu (2007); Liu and Yang (2009); Zhang and Breugelmans (2012).

In a nutshell, loyalty programs have been mainly criticized for their vanishing ability to gain a competitive advantage in an environment where almost all major competitors in a particular industry rival in “loyalty wars” for the most profitable clients (Kumar and Shah, 2004; Shugan, 2005; Singh et al, 2008). However, some smart companies have learned how to squeeze out valuable customer insights from the vast amount of data permanently accruing from monitoring customer interactions with their touch points and to benefit from personally identifiable purchase history data by customizing targeted marketing activities (Ailawadi et al, 2010; Liu, 2007; Bodapati, 2008). For example, the U.K.’s biggest and most profitable grocery chain Tesco pioneered data-driven loyalty programs by deriving a lifestyle segmentation of their customer base using behavioral data. At Tesco, these “lifestyle” segments are constructed by looking into the composition of their customers’ shopping baskets and used for deriving segment-specific targeted mailings, coupons and promotions (Humby et al, 2004). A similar approach is adopted by the French grocery retailer Carrefour.

In this paper, we take the perspective of multi-category retailers like Tesco or Carrefour, who need to manage their category level merchandising and target marketing decisions to increase sales generated by their existing customer base (Chen et al, 1999; Rowley, 2005). Such retailers frequently make use of targeted promotions to draw consumers’ attention to specific categories. One of the key challenges in such a setting is to decide which categories

to promote from among the hundreds or even thousands they offer and to whom, i.e., which customer segment(s) to target. We propose a procedure which addresses both the customer segmentation issue and the task to support managers with selecting attractive categories for deriving segment-specific, targeted category level promotion campaigns¹. In addition, we will explore the potential benefits of adapting the presented promotional decision support system and compare its empirical performance relative to that of a simple standardized promotion heuristic.

In the next section, we position the proposed framework against prior related research. Section 3 then presents the building blocks of the methodology to determine product categories to be featured in segment-specific promotional campaigns. Section 4 illustrates the empirical application of the framework by analyzing a real-world transaction dataset collected from a retailer's loyalty program. We also provide rough estimates of the expected outcome of our procedure to support target marketing campaigns using a simple scenario-based setting and compare the profitability implications with those anticipated from a standardized promotion heuristic. Finally, Section 5 discusses the results and provides an outlook on future enhancements of the proposed approach.

2. Literature Review and Research Contribution

In our research framework we consider targeted category level promotions in the same manner as Venkatesan and Farris (2012). These authors focus on retailer-initiated coupon campaigns, which are customized to customers' specific preferences (as reflected in their purchase histories) and are targeted to only a subset of the retailer's clientele. Such targeted promotions are typically offered by major retailers as part of their loyalty programs or are distributed by specialized target marketing services like Catalina Marketing for cooperating retailers and manufacturers (Zhang and Wedel, 2009; Pancras and Sudhir, 2007).

In recent years, the effectiveness of targeted promotions has increasingly been studied by marketing researchers (e.g., Rossi et al, 1996; Shaffer and Zhang, 1995; Zhang and Krishnamurthi,

¹In this paper we focus on category level customized promotions, a marketing instrument frequently used in loyalty programs offered by multi-category retailers (Drèze and Hoch, 1998; Osuna et al, 2016; Venkatesan and Farris, 2012). However, the research framework presented here could be easily adopted and/or extended to specific brands of products or even to an item-based level.

2004). There is empirical evidence that compared to conventional (i.e., undifferentiated) ones targeted promotions are capable to increase profits (Khan et al, 2009; Musalem et al, 2008). In an early contribution, Bawa and Shoemaker (1989) show for direct mailing coupons that consumers are more responsive to (coupon) promotions if their prior preference for the promoted brand is higher.

Using survey data, Shoemaker and Tibrewala (1985) also report that consumers indicate higher redemption intentions for brands they are loyal to. Zhang and Wedel (2009) show that profit differences are mainly the result of variations in redemption rates and in offline stores the incremental benefit of individual level targeting is relatively small compared to segment-level customization. In a retailer-customized setting, Venkatesan and Farris (2012) also provide support for higher redemption rates of targeted promotions. Furthermore, beyond a lift in coupon redemption the authors also find a mere exposure effect of customized coupon campaigns. This is consistent with a recent study by Sahni et al (2014); using data from a set of randomized field experiments the authors find significant carryover effects after the promotions expired and evidence for cross-category spillover to non-promoted items. Summing up, these findings suggest that promotions customized to the prior preferences of customer segments can boost company's profits.

Most prior research on designing targeted promotions has focused on how to detect interesting customer segments to target. For example, the direct approach by Bodapati and Gupta (2004) predicts whether a prospective customer exceeds a predetermined threshold on a defined outcome (e.g., grocery expenditures). Rossi et al (1996) assess the information content of various information using a target couponing problem that customizes coupons to specific households. Shaffer and Zhang (1995) analytical framework notes the effect of targeting coupons to selected households on firm profits, prices, and coupon face values. Zhang and Krishnamurthi (2004) provide recommendations about when to promote how much to whom, according to the time-varying pattern of purchase behavior and impact of current promotions on future purchases.

Notwithstanding its importance, most of this prior research neglects the selection of which category to promote for the derived customer segment. The approach presented in the next section aims to support decision makers in this respect. Our analytical framework shares some com-

mon notions with the approaches introduced by Reutterer et al (2006) and Boztug and Reutterer (2008). We also segment customers based on the multi-category choices observed in their past purchase history data with the focal company and we also derive the targeted categories based on their aptitude to stimulate cross-category spillover. However, beyond technical aspects, the major differences of the present contribution against these previous studies are as follows: For each customer segment our approach provides the decision maker with a list of candidate categories for segment level targeted promotion campaigns; the list is derived such that the included categories maximize the cross-category spillover effects for non-promoted categories. This task is accomplished by combining various data mining tools with optimization techniques in one integrated analytical framework.

Furthermore, in developing our approach we explicitly distinguish between two types of product categories: The first type contains categories purchased by a significant fraction of a specific company's customer base. Such "bestsellers" or top-selling categories show high purchase incidence rates and are very frequently bought compared to the rest of the assortment; we therefore denote these as high-frequency categories (HFCs). In a grocery retailing context, such HFCs typically include every day food categories like fresh milk, vegetables, bread or other fast moving consumer good categories. Using Drèze and Hoch (1998) terminology such "type 1 categories" are purchased by customers with the focal company on a regular basis whenever they visit the store.

Thus, such HFCs are perfect for traffic building and useful candidates for undifferentiated (or non-targeted) promotions to draw customers into the store. However, they are less useful for the targeted promotions we aim to derive, because their category expansion effects tend to be modest (Bell et al, 1999). For the purpose of deriving customer segments and selecting categories for segment-specific targeted promotions, we instead focus on a second type of categories we denote as low-frequency categories (LFCs), i.e., categories which show relatively low purchase incidences on the aggregate level but might be characterized by substantial variation across customers. The underlying rationale of considering such LFCs for target marketing purposes is related to the so-called "long tail effect" (Anderson, 2006; Elberse, 2008), which suggests that multi-category retailers can stimulate previously untapped demand by detecting and promoting specific category combinations that reflect distinctive tastes and preference

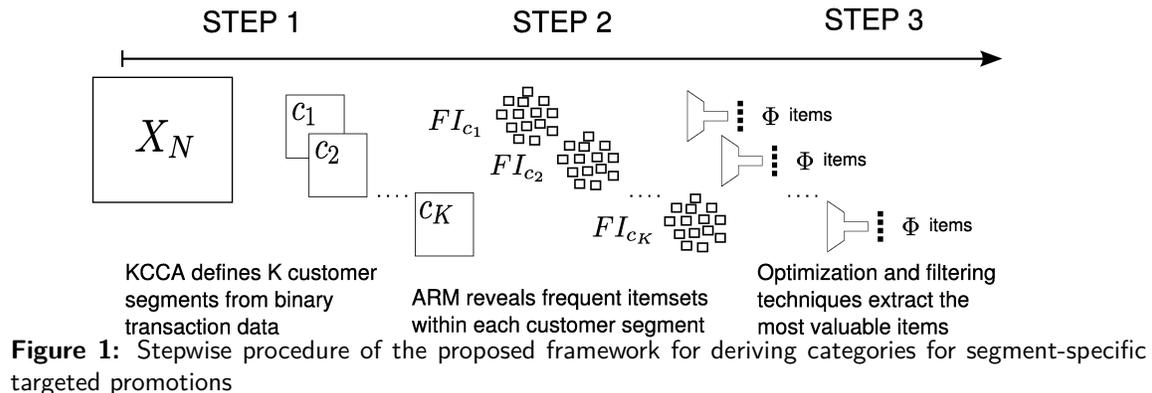
structures at the individual customer or segment level but are “averaged out” (i.e., vanish in relative small purchase incidences) on the aggregate level. More precisely, differentiating customer preferences and buying habits are more likely to be reflected in their specific multi-category choices in the “long tail” (i.e., LFCs) than in categories purchased by the vast majority of a company’s clientele. For example, a baby household and a young single household will probably both buy milk, bread and vegetables in combination. However, the latter household is not very likely to purchase any baby hygiene products. Instead, the shopping baskets of young singles might be significantly characterized by convenience food categories, frozen food, etc. Thus, we posit that using characteristic LFC combinations found in customers’ purchase histories might enhance the effectiveness of targeted promotions. The next section presents the technical details of the proposed procedure to derive targeted segment level promotions.

3. Methodological Framework

Figure 1 illustrates the stepwise procedure of our proposed framework for deriving recommendations of cross-category purchase sensitive items for targeted segment level marketing campaigns. To find a suitable customer segmentation which takes customers’ past purchase habits into account, step 1 employs a constrained K -centroids cluster algorithm (KCCA) as introduced by Leisch and Grün (2006). In step 2, an association rule mining (ARM) analysis identifies the segment-specific frequent itemsets in the pooled transactions for each segment detected in the previous step. In accordance with other association rule mining approaches, an additional filter measure separates the statistically interesting frequent itemsets from the less important ones and helps to reduce the number of considered cross-category associations. Finally, in step 3 the itemsets are used as input for an optimization procedure which recommends a list of categories maximizing profits with respect to their own profitability and a profit lift due to expected cross-category purchase associations.

3.1 STEP 1: Identifying Household Segments

It is common practice among marketing analysts conducting exploratory market basket analysis to assume that each customer transaction (shopping basket, market basket) reflects the output of a combined multi-category decision process made during a shopping trip (Manchanda et al,



1999; Russell and Petersen, 2000; Hruschka, 1991; Kwak et al, 2015). Following previous research this is considered as a “pick-any/ J ” decision task where each transaction can be represented as a J -dimensional binary vector $x_n \in \{1, 0\}^J$ of category purchase incidences, with J denoting the number of categories considered². A database of N transactions then gives the data set $X_N = \{x_n, 1 \leq n \leq N\}$.

Our approach employs a constrained K -centroids cluster algorithm introduced by Leisch and Grün (2006). In general, K -centroids methods (such as K -means, McQueen (1967)) partition data sets by finding a set of K centroids $P_K = \{p_k, 1 \leq k \leq K\}$ which optimally represent the data set, in the sense that the total distance between the data points x_n and their centroids $p(x_n) \in P_K$ become minimal. Formally, with $d(x, p)$ the distance between x and p , one aims at solving

$$\sum_{n=1}^N d(x_n, p(x_n)) \rightarrow \min_{P_K}$$

which implies that $p(x_n)$ should be taken as the p_k closest to x_n , and hence naturally provides a partition of the data points according to the closest centroids. I.e., the K segments $C_K = \{c_1, \dots, c_K\}$ obtained are such that c_k contains all x_n for which p_k is the closest centroid from P_K . Finding such optimal representations is typically based on heuristics which iterate between computing optimal centroids for the current partition and optimal partitions for the current centroids (see Bock (1999); Leisch (2006) for more information). We follow Leisch

²Following the discussion in the previous section, in our empirical illustration, we will only employ the LFCs for identifying household segments, in which case J is the number of LFCs as defined a priori by the analyst. However, note that from a purely technical perspective the proposed procedure is agnostic to any preselection and could also be applied for the complete set of categories.

(2006, Section 3.2) in taking d as the extended Jaccard distance, such that $d(x_n, x_m)$ is the relative frequency of categories purchased in only one of the transactions n and m (but not in both).

For personalized basket data, each transaction can be linked to a household it originates from. To identify household segments, one could follow Boztug and Reutterer (2008) to employ a two-step voting procedure, which first segments the transactions without taking the household information into account, and then assigns households to segments according to the majority of their transactions. Here, we follow a constrained clustering approach which already employs the household information when clustering the transactions via a so-called “must-link” constraint (Wagstaff et al, 2001; Basu et al, 2008), enforcing all transactions corresponding to one household to the same segment. This immediately yields segments of households with similar basket compositions.

The K -centroids approach very conveniently allows to impose such must-link constraints (Leisch and Grün, 2006). Write $X_{N,h}$ for the transactions in X_N corresponding to household h , and H for the number of households. Then for all h , all transactions x_n in $X_{n,h}$ should have the same centroid $p(X_{n,h}) \in P_K$, and constrained K -centroids clustering is performed via solving

$$\sum_{h=1}^H \sum_{n: x_n \in X_{N,h}} d(x_n, p(X_{N,h})) \rightarrow \min_{P_K}.$$

The thus obtained segmentation yields centroid vectors p_k which correspond to the prototypical “average” market basket for their corresponding segment c_k (and are typically similar to the vectors of category purchase frequencies of the respective segments (Leisch and Grün, 2006)).

A common challenge in the application of K -centroids based cluster methods is the determination of an appropriate value of K (Aldenderfer and Blashfield, 1984; Milligan and Cooper, 1985; Kaufman and Rousseeuw, 2005). Although this information is not available before the analysis, in most real-world situations K -centroids partitioning requires the analyst to predefine a priori the number of expected groups in the dataset or to use heuristics like the ‘elbow’ criterion (Thorndike, 1953; Gordon and Vichi, 1998), cluster validation indices (Dimitriadou et al, 2002) or index voting. Our approach to choosing K is based on the idea of increasing K until the corresponding partitions no longer markedly change. More precisely, we employ the

“corrected” Rand (1971) for measuring the agreement of two different partitions of the same data set. We compute the constrained K -centroids partitions for a suitable range of K values, and then inspect the Rand indices of the partitions using K and $K + 1$ centroids. We then choose K large enough to account for all large changes in the sequence of indices (ensuring that the obtained partition is rather stable with respect to increasing K).

3.2 STEP 2: Mining Segment-Specific Frequent Itemsets

Whereas the segment centroids p_k reflect the market basket structure of an average transaction of the segment, they do not provide any information on which categories are exactly bought in combination within the segments’ transactions. As such the set of centroid vectors merely informs the analyst about the specific “interests” of the various household segments in certain (combinations of) categories. For example, observing that a segment features rather frequent purchases of both white wines and red wines does not allow to conclude that these purchases occur together (i.e., in the same baskets). However, identifying interesting category associations clearly is a key ingredient to successful personalized target marketing of the type discussed in the previous section. This can be accomplished by employing transaction data mining techniques for finding frequent so-called itemsets and association rules (e.g., Brijs et al, 2004; Reutterer et al, 2007; Kamakura, 2012).

In our application context, *itemsets* correspond to sets of categories. We say that an itemset A is contained in a transaction x , symbolically $A \subseteq x$, if x features purchases of all categories in A . The basic measure of interestingness of an itemset is its *support*, which is the frequency of transactions containing the itemset. Formally, for transactions from segment c_k ,

$$\text{supp}(A) = \frac{|\{x_n \in c_k : A \subseteq x_n\}|}{|c_k|} \sim P(A),$$

where $|S|$ denotes the cardinality of a set S (i.e., the number of its elements). If the support value of an itemset is above a user-defined threshold (so-called *minimum support*), the itemset is referred to as “frequent” (Mannila, 1997). Even for very large transaction databases, frequent itemsets can efficiently be mined, for example, by using the APRIORI algorithm (Agrawal et al, 1993; Agrawal and Srikant, 1994; Bayardo and Agrawal, 1999; Zaki et al, 1997).

An *association rule* $A \rightarrow B$ splits an itemset $C = A \cup B$ into two non-empty disjoint itemsets A and B , the antecedent and the consequent of the rule. The strength of the association is typically measured by the *confidence* of the rule, which is the conditional frequency of transactions containing B within the transactions containing A . Formally, for transactions from segment c_k ,

$$\text{supp}(A \rightarrow B) = \frac{|\{x_n \in c_k : A \cup B \subseteq x_n\}|}{|\{x_n \in c_k : A \subseteq x_n\}|} = \frac{\text{supp}(A \cup B)}{\text{supp}(A)} \sim P(B|A).$$

As in general $\text{conf}(A \rightarrow B) \neq \text{conf}(B \rightarrow A)$, confidence provides an asymmetric measure of the statistical strength of the association between two itemsets A and B .

To separate the statistically attractive frequent itemsets from the ones less so, several measures of interest have been developed (Hettich and Hippner, 2001; Hahsler et al, 2006). A commonly employed symmetric measure of the overall strength of association within an itemset is the so-called *all-confidence* (Omiecinski, 2003), which computes the minimal confidence of all association rules that can be generated from the itemset. Formally,

$$\text{allconf}(C) = \min\{\text{conf}(A \rightarrow B) : \emptyset \subset A, B \subset C, A \cup B = C, A \cap B = \emptyset\}$$

(itemsets with at most one element have zero all-confidence). Employing all-confidence to measure statistical association is attractive within our application context because it promises particularly good results in transaction datasets which exhibit itemsets with markedly varying support values (Agrawal et al, 1993; Hui et al, 2006).

3.3 STEP 3: Optimization and Filtering Segment-Specific Itemsets

If associations are mined with a low minimum support in a dataset showing skewed purchase frequencies, the analyst has to be aware of finding many weakly related cross-support itemsets (Hui et al, 2006). This results from grouping customers with similar interest in purchasing certain categories of the assortment (e.g., customers of a “baby” cluster disproportionately often buy baby related products). As outlined in Section 3.2, these problems can be addressed using the all-confidence value suggested by Omiecinski (2003) to reduce the output of the APRIORI algorithm. We thus filter the frequent itemsets obtained from step 2 accordingly,

retaining those F frequent itemsets with the highest all-confidence values (and hence length at least two), for a suitable value of F . The remaining frequent itemsets are then transferred to the proposed optimization model, generating a list of single categories which should be promoted within the corresponding customer segment.

To select only the most valuable items for customized marketing, we use the generalized PROFSET model introduced by Brijs et al (2004), which determines the most profitable categories based on their profit lift into frequent itemsets of interest, by solving an all-binary optimization problem. The resulting categories imply a high monetary value *and* the ability to initiate cross-selling in the respective customer segment.

Let us write \mathcal{J} for the J categories to select from, and \mathcal{F} for the frequent itemsets of interest based on these categories (in our approach, \mathcal{F} consists of the F frequent itemsets with the highest all-confidence values). Write Q_j for the binary variable indicating the selection of category j (i.e., $Q_j = 1$ if j is selected, and zero otherwise). The PROFSET optimization selects the Φ best categories by solving

$$\sum_{A \in \mathcal{F}} V(A) * P_A - \sum_{j \in \mathcal{J}} COST_j * Q_j \rightarrow \max_{\{P_A\}, \{Q_j\}}$$

for the binary decision variables Q_j and P_A , subject to

$$Q_j \geq P_A \quad \forall A \in \mathcal{F}, \forall j \in A, \quad \sum_{j \in \mathcal{J}} Q_j = \Phi$$

where the first constraint ensures that an itemset can only be selected if all categories it contains are selected, and the second constraints ensures that exactly a prescribed number Φ of categories are chosen, $COST_j$ gives category-specific handling and inventory costs, and $V(A)$ is the “value” (profit margin) of the itemset A which is obtained by suitably aggregating the values of the transactions containing it.

The value of a single transaction x_n is given by

$$v(x_n) = \sum_{j \in x_n} (SP(j) - PP(j)) * f(j, x_n) \tag{1}$$

with $SP(j)$ and $PP(j)$ the sales and purchase prices, respectively, and $f(j, x_n)$ the number

of times j was purchased in transaction x_n . When aggregating the transaction values into itemset values, care must be taken to avoid that transactions contribute to several itemsets (e.g., all itemsets they contain). The original PROFSET approach thus takes $V(A)$ as the sum of the $v(x_n)$ over all transactions exactly matching A (i.e., featuring purchases of *exactly* the categories in A). When employing only *frequent* itemsets, this may result in excluding many transactions in the value aggregation, and hence under-estimating the actual values (see Section 3.3 in Brijs et al (2004)). This effect is particularly relevant in our approach which is based on using relatively small numbers of frequent itemsets with interesting cross-category associations as measured by all-confidence.

We thus generalize the PROFSET model as follows: For each transaction under consideration, we determine all frequent itemsets in \mathcal{F} it contains, and distribute the value of the transaction to the values of these itemsets weighted according to the support values of the itemsets, either by direct distribution, or alternatively by randomly selecting the itemsets according to these weights. For example, if a transaction contains exactly the two frequent itemsets $\{vegetables, water\}$ with support 0.02 and $\{bottled\ beer, water\}$ with support 0.07, then the weight of the first itemset is $0.02/(0.02 + 0.07) = 2/9$, and the weight of the second is $7/9$.

Employing the PROFSET approach requires the specification of a pre-defined number Φ of categories to be selected. Because both marketing budgets and advertising spaces are scarce resources in both on- and offline environments, loyalty program managers tend to be rather interested in focusing their target marketing efforts on a selected few product categories than dealing with a multitude of interrelated itemsets. After using e.g. a branch-and-bound algorithm to solve the all-binary PROFSET optimization problem, the solution determines Φ variables, which point to the categories to be selected for maximizing the objective function and subsequently used in target marketing actions.

4. Empirical Application

The application of the proposed framework to derive recommendations for segment-specific, category level targeted promotions is demonstrated below using a real-world data set. We obtained the data from a major grocery retailer who prefers to stay anonymous. The data set contains transactions realized by members of the loyalty program offered by the focal retailer.

Using a simple scenario-based analysis, we illustrate and discuss the benefits and comparative effectiveness of our proposed data-driven target marketing approach vis-a-vis undifferentiated standard promotions.

4.1 Data description

The data set at hand contains more than 1.4 million transactions made within one year by 56,000 households which are enrolled in the retailer’s loyalty program. The records available for the shopping baskets include prices, quantities and average gross profit margins for 268 categories. As illustrated by Figure 2, the supermarket’s assortment is dominated by a small range of categories that are bought very frequently. Therefore, and in accordance with our conceptual arguments in section 2, the assortment is separated into two distinct types of categories: the “bestsellers”, which are the 52 HFC and occur at least in 10% of all shopping baskets (left-hand side of the vertical dotted line in Figure 2), and the “long tail” range of the assortment which are the remaining 216 LFC (right-hand side of the vertical dotted line in Figure 2).

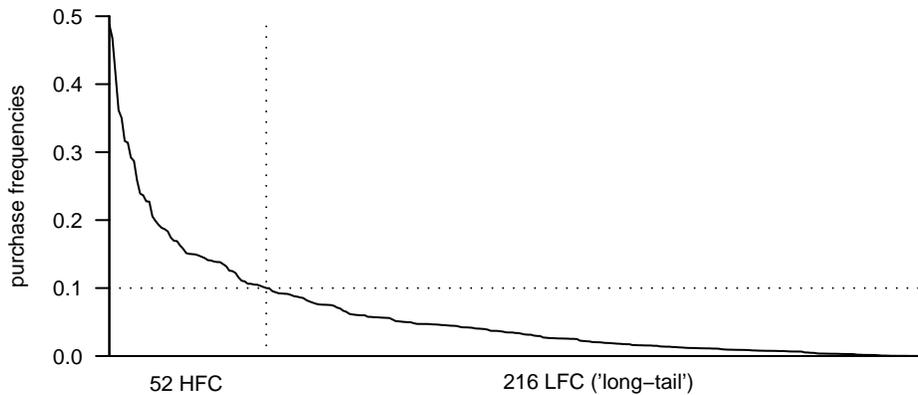


Figure 2: Relative purchase frequencies of all categories in the first sample in descending order

To illustrate the practical application of our proposed procedure, we drew two disjoint samples from the transaction data base, each containing the transactions of 3,000 randomly selected households. The first sample is used for selecting an appropriate KCCA segmentation model, i.e., to determine an appropriate value for K and the centroids P_K of the corresponding constrained cluster solution. The second sample is used for performance evaluation, using months 1–10 to update the KCCA segmentation obtained from sample 1 by performing KCCA

with the chosen K and the centroids initialized with P_K , and using months 11 and 12 as the hold-out sample for the profitability scenario-based analysis.

In grocery retailing it is typical that households have varying lengths of buying histories. On average, households in sample 1 made around 26 transactions with a basket size of six categories in the observation period from the LFC. To robustify the selection of an appropriate value for K , we only consider households with buying sequences that are sufficiently long but not extremely long. Specifically, we exclude those households with the smallest 20% and largest 5% numbers of transactions, leaving 2,250 households in sample 1 to use for selecting an appropriate KCCA model.

4.2 Identifying Household Segments and Extracting Itemsets for Targeting

According to step 1 of our proposed procedure (see Figure 1), the first goal is to partition the households of sample 1 into segments with the constrained cluster algorithm. To reduce the risk of getting stuck at a weak local optimum every partitioning task is repeated up to fifteen times (Gordon and Vichi, 1998; Aldenderfer and Blashfield, 1984; Hornik, 2005), retaining the partition minimizing the target function. The value of the Rand index apparently levels off after partitions with $K = 11$ and $K = 12$ (cf. Figure 3). It can also be seen that the arrangement of the data points would not change radically if a further cluster $K = 12$ (or $K > 15$) is added. Hence, we decide on $K = 11$, which also allows for sufficiently convenient interpretation by retail managers.

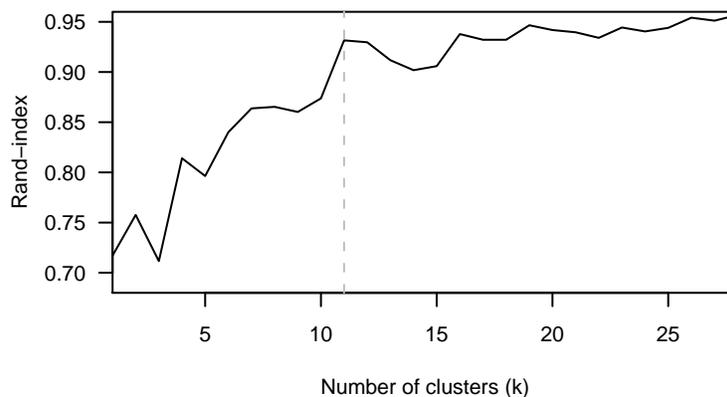


Figure 3: Cluster agreement by comparing K to $K + 1$ cluster solutions using the Rand index. Note that the higher the value the higher the pairwise similarity between two partitions.

We next derive the household segments of the second sample based on the transactions for months 1–10 by running the constrained cluster algorithm with the chosen $K = 11$ and the centroids initialized with those of the corresponding P_K (equivalently, the transactions of each household are initially simultaneously assigned to the best matching centroid from P_K). Figure 4 (dark-grey bars) shows the amounts of baskets from months 1–10 in the segments thus obtained. We can see that there exists one large segment ($k = 6$) containing about 20% of all baskets and a smaller segment ($k = 10$) containing less than 5%. The remaining baskets are assigned to the other nine segments quite equally. After clustering, the generated household segments are labeled according to the most frequent itemsets within each cluster (cf. Figure 4).

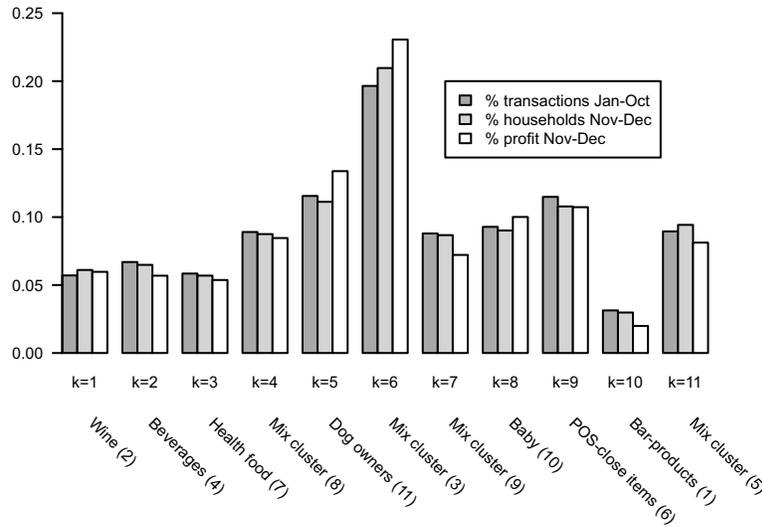


Figure 4: Number of household segments for the second sample of months 1–10 as well as the rank level of each cluster density in brackets (1: highest density, 11: lowest density) compared to the number of assigned households and the profit generated for months 11–12 in percentages.

To give an illustrative example of the cross-category purchase interrelationships resulting from the proposed algorithm, Figure 5 depicts the results for two of the clusters generated from the second sample obtained for the segments $k = 8$ and $k = 1$. The different peaks of light-gray bars on the left-hand side in Figure 5 indicate that the households in both groups show interests in quite different itemsets. To match these peaks to the corresponding categories the ten most frequently purchased itemsets in the corresponding segments are investigated (cf. Figure 5, right-hand side). The households in segment $k = 8$ (the “baby” cluster) seem to focus on baby food and baby care categories since these products are purchased at a substantially higher than

average rate. A typical household being represented by the baby cluster buys baby hygiene products with a probability of 35.24% and adds baby food in glass with a probability of 22.82% and baby food mush/powder with a probability of 17.62%. The households in segment $k = 1$ (the “wine” cluster) combine different kinds of — in particular — wine or other alcoholic beverages. In contrast to its overall purchase probability of 3.86%, red/rosé wines occur at a rate about ten times higher than average in a basket purchased by a wine cluster household (red/rosé wine’s group-specific purchase frequency is 32.21%).

Other clusters contain itemsets related to health food (such as cereals, organic products, whole meal products, frozen ice cream, etc.), meat (beef, chicken, other kinds of meat, etc.) or beverages (soda, lemonade, water, etc.) and are therefore equally easy to label with a generic term. However, some segments obtained do not contain itemsets with such an obvious interpretation. Therefore, to keep the application in this paper simple, these clusters are labeled as “mixed” clusters (cf. Figure 4), referring to household segments with cross-category associations which are not as straightforwardly interpretable as the generically labeled ones.

The number of mined associations depends on the pre-determined minimum support. Usually, analysts prefer low support thresholds to detect less obvious associations within the transaction datasets (Hui et al, 2006). Based on the APRIORI algorithm all frequent itemsets with a minimum support of 2% are revealed (cf. Section 3.2, step 2 in Figure 1). The resulting 70 frequent itemsets with the highest all-confidence value and a minimum length of two (ignoring circular associations) are passed on the PROFSET optimization model for further examination (cf. Section 3.3, step 3 in Figure 1). Table 1 includes the itemsets which are obtained for the wine, the health food and the baby segments in Figure 4 (in the present PROFSET application a value of $\Phi = 4$ was chosen). The categories included in these itemsets exactly correspond to the categories recommended by our proposed procedure to be featured in marketing actions targeted to the respective three segments. Analogous recommendations (which are not displayed here for space reasons) can be obtained for the remaining segments.

In comparison, the last column in Table 1 lists the top-four “bestselling” (in terms of generated sales values for the same households and observation period under study) categories from the set of HFCs. While these categories were excluded from our segmentation and subsequent frequent itemset mining procedure, they would represent promising candidates for an

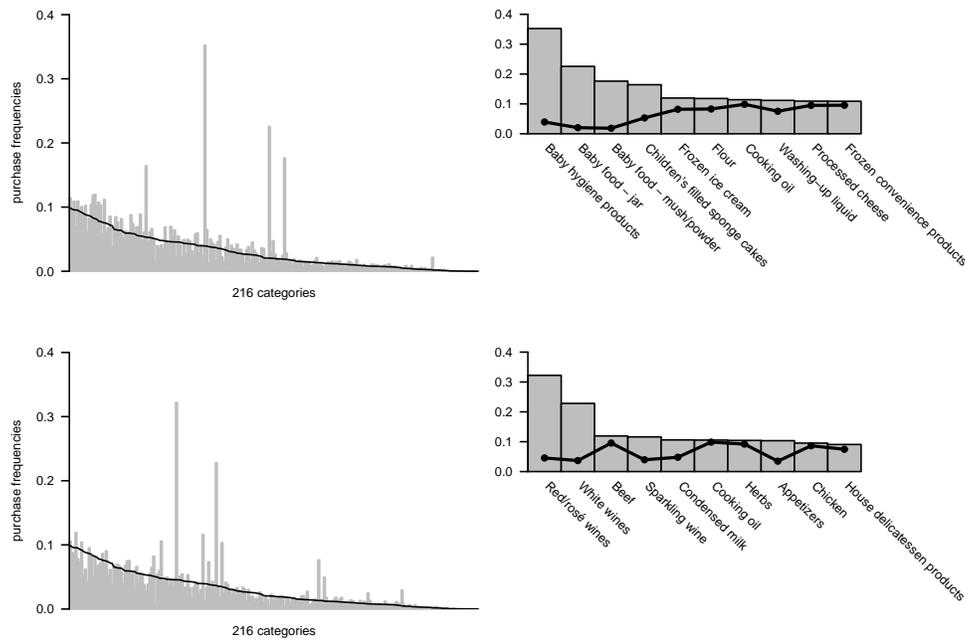


Figure 5: Graphical illustration of the market baskets of segments $k = 8$ (baby cluster, above) and $k = 1$ (wine cluster, below). Note: the black solid line (left-hand side) represents the sample's overall relative purchase frequencies, the light-gray bars correspond to the 216 categories purchase frequencies within the clusters; compared to the ten most frequently purchased itemsets of each segment (right-hand side).

undifferentiated, traffic-building promotional campaign. Next we further explore the potential performance of a segment-specific targeting approach against a standardized campaign.

4.3 Scenario Analysis to Evaluate Profitability Implications

After performing the stepwise procedure illustrated above, managers are provided with (i) a set of household segments, corresponding centroids and household assignments to segments as well as (ii) for each segment an itemset including the categories recommended for target

Segment 1 Wine	Segment 3 Health food	Segment 8 Baby	Standardized promotion
Red/rosé wines (32%)	Organic prod. (42%)	Baby hygiene prod. (35%)	Bottled beer (18%)
White wines (23%)	Wholemeal prod. (14%)	Baby food jars (26%)	Delicatessen (29%)
Sparkling wine (12%)	Organic beef (5%)	Baby food powder (18%)	Soft drinks (31%)
Beef (12%)	Frozen ice cream (10%)	Frozen ice cream (12%)	Vegetables (48%)

Table 1: Recommended categories for segment-specific targeting and categories selected for standardized promotion (with corresponding segment-specific or global relative purchase frequencies in brackets).

marketing actions. Now, suppose that a loyalty program manager considers to launch a targeted promotions campaign for the previously identified household segments. Of course, an evaluation procedure of first choice would be to run a series of (randomized) field experiments and to compare the effectiveness of targeted promotions relative to an undifferentiated approach (or doing nothing). Because we do not have access to such experimental data, we discuss some basic and preliminary considerations from a managerial perspective. In fact, prior to costly experimentation both analysts and – more importantly – managers typically wish to gain some initial notions on the prospective chances of success of such an approach and for which segments targeting is most likely to pay off. In doing a preliminary feasibility study, we next adopt a simple scenario-based evaluation of such a strategy by making some assumptions based on prior empirical findings.

4.3.1 Scenario Settings

The scenario analysis is pursued by estimating the profit margin generated with the itemsets recommended for the different campaigns (targeted vs. standard) using the empirical basket data set at hand. For this purpose we reutilize the transaction data included in the second sample used to illustrate the empirical application of our approach. Note that the households' transactions for months 1–10 were used to determine segment memberships and for deriving itemset recommendations. This data is now also used to calculate the expected profit margin for the following two months, which serve as a hold-out period. As profit margins are available on all categories, it is possible to determine the profits realized by the focal retailer with a certain itemset. This profit simulation is done using all 268 categories and all 3,000 households in the second sample.

Since we extend the cluster membership of a household determined for the calibration sample to the hold-out period, we also check whether the cluster size coincides relative to the number of included households (cf. Figure 4, light gray bars) and the generated percentage profit gains (cf. Figure 4, white bars) for months 11–12. Despite some smaller deviations, the three values obtained mostly correspond to each other. Thus we conclude that the size of the cluster approximately determines the profits generated by the households of the corresponding segments.

We assume that the retailer at hand considers to conduct a segment-specific promotional campaign for a predefined set of four categories per segment (see the examples in Table 1) in the hold-out period. For example, such a campaign could be effected by distributing targeted coupons among the household members for each segment. With targeted coupons, customers typically can earn a discount of a certain monetary amount (or a percentage value equivalent) if they bought at least one product in the mentioned category within a predefined period of time (Kalwani and Yim, 1992; DeVecchio et al, 2007). For evaluating the expected effectiveness of such a targeted marketing campaign we use a standardized promotion campaign as a benchmark (i.e., promoting categories with the highest revenue in months 1–10; see the HFC categories listed in the last column of Table 1). To compare the expected profits resulting from both campaign-types we use the expected gains in profit margins for the baskets along the complete set of categories (i.e., the 52 HFC with the previously identified 216 LFC). Following Brijs et al (2004) we only consider direct product costs and ignore category handling and inventory costs for sake of simplicity. Furthermore, we also do not consider any potential costs to implement a targeted marketing strategy.

To evaluate the relative performance of the two campaign-types in terms of profitability, their respective expected profit increases are calculated in months 11–12 conditional on the selected promotional campaign. Next we add up all the profit margins potentially accruing from two scenarios which we define according to Table 2 and discuss further below. Due to cross-category associations we also have to consider the purchase interrelationships between the promoted itemsets and the rest of the assortment in step 3 of the proposed framework. Therefore, after mining all association rules with a minimum support of 2% and a length of two (i.e., the major associations) the profits are multiplied with the confidence value of the corresponding rule, revealing the expected indirectly affected accumulated profits generated by selling the promoted categories. Finally, we estimate the percentage profit gain which is expected to be achieved by the corresponding promotion campaign compared to the real profit achieved in month 11–12. Thus, we implicitly assume stationary marketing activities in both the calibration and hold-out periods and do not account for any seasonal or stock-buying effects.

For evaluation purposes, we define the following two scenario settings: The first setting assumes a lower responsiveness for the segment-specific target marketing campaign and higher

values for the standardized promotion campaign, while the second setting does the opposite.

Campaign	Setting no. 1	Setting no. 2	Profit added to
	Gross-profit growth	Gross-profit growth	
Segment-specific promotion	10.00%	15.00%	promoted itemsets
Cross-merchandising	6.00%	10.00%	associated itemsets
Standardized promotion	5.00%	3.00%	promoted itemsets
Standardized promotion	5.00%	1.00%	associated itemsets

Table 2: Two scenarios of expected percentage profit growth in response to the corresponding campaigns featuring specific recommended itemsets.

The profit lift values expected for specific combinations of itemsets and campaign types under the two scenarios are included in Table. These values are based on prior empirical findings reported in the relevant marketing literature (e.g., Drèze and Hoch, 1998; Zhang and Wedel, 2009; Venkatesan and Farris, 2012; Sahni et al, 2014) and discussions with domain experts working with the focal retailer. For segment-specific targeted coupons, Drèze and Hoch (1998) report a 25% increase in the promoted categories after the program has been running for six months and taking costs into account. The overall profitability of the campaign depends on the length of the coupon's validity period; this applies even more, if the promoted categories are from the LFCs and average per-basket sales in these categories usually tend to increase over time (Drèze and Hoch, 1998). Thus, we assume a pessimistic value of 10% for setting no. 1. In contrast to Drèze and Hoch (1998) the promoted categories of our approach are matched to the purchasing behavior of the targeted households taking purchase interrelationships into account. Therefore, a more optimistic profit lift of as much as 15% in the promoted itemsets is assumed for setting no. 2. In addition, in a second study by Drèze and Hoch (1998) the authors applied cross-merchandising techniques (of the type "save a certain amount on category B products if you purchase category A products", Drèze and Hoch, 1998, see Section 3) and found sales increases in the targeted category ranging from 6% to 10%. Since the targeted items in cross-merchandising campaigns correspond to the right-hand side categories of rules derived from frequent itemsets, we adopt these values as proxies in our scenario-based analysis (cf. Table 2).

For the standardized promotion campaign, we refer to the meta-analyses by Tellis (1988) and Bijmolt et al (2005) which summarizes empirical research related to price elasticities on which our scenario assumptions are based on. Since the categories determined by the standardized promotion all come from the grocery domain the projected profit increase of 5% employed in

setting no. 1 is very optimistic. Nevertheless, to avoid overestimating the results from the segment-specific framework, a growth of 3% is still estimated for the accumulated profits of the promoted categories when the standardized promotion method is applied for the more pessimistic scenario setting no. 2³. It is also possible that the sales of associated itemsets could rise as much as the sales of the promoted categories (5% in setting no. 1), but in fact, the gain will likely be much smaller. Therefore, we assume as a pessimistic outcome that the profit in the associated itemsets will increase by only 1%.

4.3.2 Results

The bars in Figure 6 represent the expected profit margin gains in months eleven and twelve within each household segment derived for the two scenarios in Table 2. For the first scenario, the retailer would expect an overall profit margin gain of 15%. Figure 6(a) shows that the segment-specific targeted category level program outperforms the standardized promotion in only four out of the eleven household segments under investigation (in particular the wine and the baby clusters). In other words, only for these segment-specific targeted promotions as derived by our proposed framework using LFCs are the recommended option because of a higher expected gain in profitability compared to the standardized promotion campaign. For segments $k = \{2, 4, 6, 7, 9, 10, 11\}$ Figure 6(a) shows that the profit increase achieved by the standardized promotion will be twice as high as the projected profit lift for the segment-specific case. For these groups, the expected gain in profit by adopting a targeted marketing program are unlikely to compensate for the profit potential not realized by conventional standard promotion techniques. However, segment-specific targeting would still be profitable for 32.2% of the targeted households, while about two thirds of all households would still need to be addressed with the standardized promotion.

The situation changes for setting no. 2, which recommends the segment-specific target marketing approach as the preferred one (cf. Figure 6(b)). Under the conditions described for this scenario setting, our proposed segment-specific campaign clearly outperforms the standardized promotion and can expected to be much more profitable in every household segment. On aggregate, the additional profit lift generated with segment-specific target marketing exceeds

³Note that the profit estimation of the standardized promotion benefits from the assumption of ignoring costs since it ensures a more conservative calculation of the output of our segment-specific approach.

the undifferentiated promotion by up to 128%.

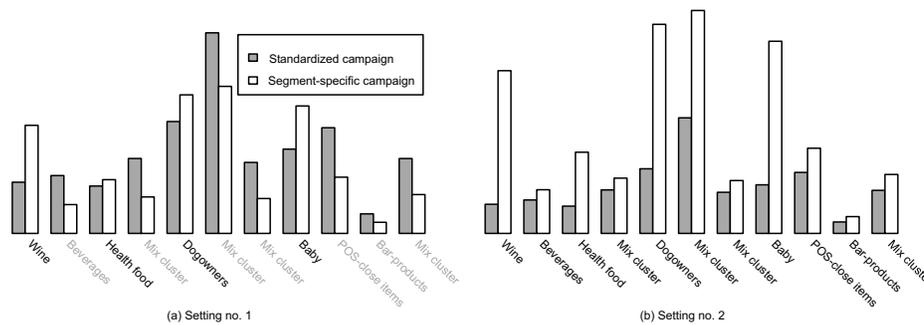


Figure 6: Expected gross-profit growth for (a) setting no. 1 and (b) setting no. 2. The gray bars depict the expected gross-profit gains for the standardized undifferentiated promotion campaign (against profit expectations in the hold-out period assuming stationary conditions); the white bars depict the corresponding values resulting for a segment-specific promotion campaign.

5. Discussion and Future Research

We present and empirically demonstrate the performance of a new approach to support loyalty program managers and direct marketers in customizing their segment-specific target marketing activities on a product category level. Our proposed decision-support framework requires customers' past purchase histories as input data, builds on state-of-art data mining techniques and integrates an optimization procedure which provides the decision maker with a list of candidate product categories for segment-level targeted promotion campaigns. This list is derived such that the included categories maximize the cross-category spillover effects for non-promoted categories.

There are many occasions in which marketing managers can benefit from such itemset recommendations for target marketing purposes. These include but are not limited to designing segment-specific rewards in loyalty programs, cross-merchandising activities, targeted direct mailings, customized supplements in catalogues, and customized promotions. For example, the latter can be delivered both offline directly in the store (e.g., by issuing customized check-out-coupons as provided by Catalina Marketing services) or in online environments by sending targeted emails or during shopping trips in online stores.

We demonstrate the application of the stepwise procedure using transaction data from a real-world loyalty program offered by an anonymous major grocery retailer. In the scope of our

empirical application study we also explored the projected profitability implications of utilizing the derived recommendations for designing segment-specific, category level targeted promotions. A scenario-based simulation study suggests that the adoption of targeted promotions might boost profitability between 15% to 128% relative to an undifferentiated standard campaign and that at least for some segments targeting can be the preferred option even under very conservative assumptions on the effectiveness of segment-specific target marketing actions. Of course, a more thorough evaluation of the relative effectiveness of targeted campaigns derived by utilizing our approach would be desirable. Such an evaluation strategy could entail a series of randomized field experiments. The evaluation framework introduced by Wang et al (2016) offers a promising starting point for endeavors toward this direction, which we leave open for future research.

It also would be beneficial to combine our approach, which is primarily concerned with data compression, with a more predictive approach for modeling customers' multi-category choice decisions (such as the work by Manchanda et al, 1999; Dippold and Hruschka, 2013; Hruschka, 2013). Further promising extensions of our approach could also concentrate on making the segmentation approach dynamic in order to account for changes in customers' purchasing habits. Finally, to accommodate larger numbers of categories or for applications on a sub-category or even item-level the proposed approach needs to be made scalable for very high-dimensional transaction data. Such an attempt requires some type of variable selection or variable weighting; the contributions by Carmone Jr. et al (1999); Brusco and Credit (2001) might be promising candidates to deal with this kind of challenges.

References

- Agrawal R, Srikant R (1994) Fast algorithms for mining association rules. In: Proceedings of the 20th International Conference on Very Large Databases, Santiago, Chile, pp 487–499
- Agrawal R, Imielinski T, Swami A (1993) Mining association rules between sets of items in large databases. In: Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, Washington, DC, pp 207–216
- Ailawadi KL, Bradlow ET, Draganska M, Nijs V, Rooderkerk RP, Sudhir K, Wilbur KC, Zhang J (2010) Empirical models of manufacturer-retailer interaction: A review and agenda for future research. *Marketing Letters* 21(3):273–285
- Aldenderfer MS, Blashfield RK (1984) *Cluster Analysis, Quantitative Applications in the Social Sciences*, vol 44. Sage University Paper, Beverly Hills
- Anderson C (2006) *The Long Tail: how endless choice is creating unlimited demand*. RH Business Books, London
- Basu S, Davidson I, Wagstaff K (2008) *Constrained Clustering: Advances in Algorithms, Theory, and Applications*. Chapman & Hall
- Bawa K, Shoemaker RW (1989) Analyzing incremental sales from a direct mail coupon promotion. *Journal of Marketing* 53(3):66
- Bayardo RJ, Agrawal S (1999) Mining the most interesting rules. In: Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 145–154
- Bell DR, Chiang J, Padmanabhan V (1999) The decomposition of promotional response: An empirical generalization. *Marketing Science* 18(4):504–526
- Berry J (2013) Bulking up: The 2013 colloquy loyalty census – growth and trends in us loyalty program activity. *Colloquy* June
- Bijmolt THA, van Heerde HJ, Pieters RGM (2005) New empirical generalizations on the determinants of price elasticity. *Journal of Marketing Research* 42(2):141–156

- Bock HH (1999) Clustering and neural network approaches. In: Gaul W, Locarek-Junge H (eds) *Classification in the Information Age, Proceedings of the 22nd Annual Conference of the Gesellschaft für Klassifikation e.V.*, Springer, Heidelberg, Germany, pp 42–57
- Bodapati A, Gupta S (2004) A direct approach to predicting discretized response in target marketing. *Journal of Marketing Research* 41(1):73–85
- Bodapati AV (2008) Recommendation systems with purchase data. *Journal of Marketing Research* 45(1):77–93
- Boztug Y, Reutterer T (2008) A combined approach for segment-specific analysis of market basket data. *EJOR European Journal of Operational Research* 187(1):294–312
- Brijs T, Swinnen G, Vanhoof K, Wets G (2004) Building an association rules framework to improve product assortment decisions. *Data Mining and Knowledge Discovery* 8(1):7–23
- Brusco MJ, Cradit DJ (2001) A variable-selection heuristic for k-means clustering. *Psychometrika* 66(2):249–270
- Carmone Jr F, Kara A, Maxwell S (1999) A new model to improve market segment definition by identifying noisy variables. *Journal of Marketing Research* 36(4):501–509
- Chen Y, Hess JD, Wilcox RT, Zhang ZJ (1999) Accounting profits versus marketing profits: A relevant metric for category management. *Marketing Science* 18(3):208–229
- DelVecchio D, Krishnan HS, Smith DC (2007) Cents or percent? the effects of promotion framing on price expectations and choice. *Journal of Marketing* 71(3):158–170
- Dimitriadou E, Dolnicar S, Weingessel A (2002) An examination of indexes for determining the number of clusters in binary data sets. *Psychometrika* 67(1):137–160
- Dippold K, Hruschka H (2013) A model of heterogeneous multicategory choice for market basket analysis. *Review of Marketing Science* 11(1):1–31
- Drèze X, Hoch SJ (1998) Exploiting the installed base using cross-merchandising and category destination programs. *International Journal of Research in Marketing* 15(5):459–471

- Elberse A (2008) Should you invest in the long tail? *Harvard Business Review* 86(7/8):88–96, hBS Centennial Issue
- Ferguson R, Hlavinka K (2007) The COLLOQUY loyalty marketing census: sizing up the us loyalty marketing industry. *Journal of Consumer Marketing* 24(5):313–321
- Gordon AD, Vichi M (1998) Partitions of partitions. *Journal of Classification* 15(2):265–285
- Hahsler M, Hornik K, Reutterer T (2006) Implications of probabilistic data modeling for mining association rules. In: *From Data and Information Analysis to Knowledge Engineering (Proceedings of the 29th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Magdeburg, March 9–11, 2005)*, Springer-Verlag, Heidelberg, *Studies in Classification, Data Analysis, and Knowledge Organization*, pp 598–605
- Hettich S, Hippner H (2001) Assoziationsanalyse. In: Hippner H, Küsters UL, Meyer M, Wilde K (eds) *Handbuch Data Mining im Marketing - Knowledge Discovering in Marketing Databases*, Viewag, Wiesbaden, pp 427–463
- Hornik K (2005) A CLUE for CLUster Ensembles. *Journal of Statistical Software* 14(12):1–25
- Hruschka H (1991) Bestimmung der Kaufverbundenheit mit Hilfe eines probalistischen Messmodells. *Zeitschrift für betriebswirtschaftliche Forschung* 43(5):418–434
- Hruschka H (2013) Comparing small-and large-scale models of multicategory buying behavior. *Journal of Forecasting* 32(5):423–434
- Hui X, Tan PN, Kumar V (2006) Hyperclique pattern discovery. *Data Mining and Knowledge Discovery* 13(2):219–242
- Humby C, Hunt T, Phillips T (2004) *Scoring points: How Tesco is winning customer loyalty*. Kogan Page Publishers
- Kalwani MU, Yim CK (1992) Consumer price and promotion expectations: An experimental study. *Journal of Marketing Research (JMR)* 29(1):90–100
- Kamakura WA (2012) Sequential market basket analysis. *Marketing Letters* 23(3):505–516

- Kaufman L, Rousseeuw PJ (2005) *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley
- Khan R, Lewis M, Singh V (2009) Dynamic customer management and the value of one-to-one marketing. *Marketing Science* 28(6):1063–1079
- Kivetz R, Simonson I (2003) The idiosyncratic fit heuristic: Effort advantage as a determinant of consumer response to loyalty programs. *Journal of Marketing Research* 40(4):454–467
- Kumar V, Shah D (2004) Building and sustaining profitable customer loyalty for the 21st century. *Journal of Retailing* 80(4):317–329
- Kwak K, Duvvuri SD, Russell GJ (2015) An analysis of assortment choice in grocery retailing. *Journal of Retailing* 91(1):19–33
- Leisch F (2006) A toolbox for k-centroids cluster analysis. *Computational Statistics and Data Analysis* 51(2):526–544
- Leisch F, Grün B (2006) Extending standard cluster algorithms to allow for group constraints. In: Rizzi A, Vichi M (eds) *COMPSTAT 2006, Proceedings in Computational Statistics*, Physica-Verlag, Heidelberg, pp 885–892
- Liu Y (2007) The long-term impact of loyalty programs on consumer purchase behavior and loyalty. *Journal of Marketing* 71(4):19–35
- Liu Y, Yang R (2009) Competing loyalty programs: Impact of market saturation, market share, and category expandability. *Journal of Marketing* 73(1):93–108
- Manchanda P, Ansari A, Gupta S (1999) The "shopping basket": a model for multicategory purchase incidence decisions. *Marketing Science* 18(2):95–114
- Mannila H (1997) Methods and problems in data mining. In: Afrati FN, Kolaitis PG (eds) *Database Theory — ICDT '97, 6th International Conference, Delphi, Greece, January 8–10, 1997, Proceedings*, Springer, Lecture Notes in Computer Science, vol 1186, pp 41–55

- McQueen J (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, CA, USA, pp 281–297
- Milligan GW, Cooper MC (1985) An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 50(2):159–179
- Musalem A, Bradlow ET, Raju JS (2008) Who's got the coupon? Estimating consumer preferences and coupon usage from aggregate information. *Journal of Marketing Research* 45(6):715–730
- Omicinski E (2003) Alternative interest measures for mining associations in databases. *IEEE Transactions on Knowledge and Data Engineering* 15(1):57–69
- Osuna I, González J, Capizzani M (2016) Which categories and brands to promote with targeted coupons to reward and to develop customers in supermarkets. *Journal of Retailing*
- Pancras J, Sudhir K (2007) Optimal marketing strategies for a customer data intermediary. *Journal of Marketing Research* 44(4):560–578
- Rand WM (1971) Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association* 66(336):846–850
- Reutterer T, Mild A, Natter M, Taudes A (2006) A dynamic segmentation approach for targeting and customizing direct marketing campaigns. *Journal of Interactive Marketing* 20(3-4):43–57
- Reutterer T, Hahsler M, Hornik K (2007) Data Mining und Marketing am Beispiel der explorativen Warenkorbanalyse. *Marketing: Zeitschrift für Forschung und Praxis* 29(3):163–179
- Rossi PE, McCulloch RE, Allenby GM (1996) The value of purchase history data in target marketing. *Marketing Science* 15(4):321–340
- Rowley J (2005) Building brand webs: Customer relationship management through the tesco clubcard loyalty scheme. *International Journal of Retail & Distribution Management* 33(3):194–206

- Russell GJ, Petersen A (2000) Analysis of cross category dependence in market basket selection. *Journal of Retailing* 76(3):367–392
- Sahni N, Zou D, Chintagunta PK (2014) Effects of targeted promotions: Evidence from field experiments. Available at SSRN 2530290
- Shaffer G, Zhang ZJ (1995) Competitive coupon targeting. *Marketing Science* 14(4):395–416
- Shoemaker RW, Tibrewala V (1985) Relating coupon redemption rates to past purchasing of the brand. *Journal of Advertising Research* 25(5):40–47
- Shugan SM (2005) Brand loyalty programs: are they shams? *Marketing Science* 24(2):185–193
- Singh SS, Jain DC, Krishnan TV (2008) Research note - customer loyalty programs: Are they profitable? *Management Science* 54(6):1205–1211
- Tellis GJ (1988) The price elasticity of selective demand: A meta-analysis of econometric models of sales. *Journal of Marketing Research* 24(4):331–341
- Thorndike RL (1953) Who belongs in the family? *Psychometrika* 18(4):267–276
- Venkatesan R, Farris PW (2012) Measuring and managing returns from retailer-customized coupon campaigns. *Journal of Marketing* 76(1):76–94
- Wagstaff K, Cardie C, Rogers S, Schroedl S (2001) Constrained k -means clustering with background knowledge. In: *Proceedings of the International Conference on Machine Learning (ICML)*, pp 577–584
- Wang Y, Lewis M, Cryder C, Sprigg J (2016) Enduring effects of goal achievement and failure within customer loyalty programs: A large-scale field experiment. *Marketing Science* 10.1287/mksc.2015.0966, URL <http://dx.doi.org/10.1287/mksc.2015.0966>, to appear
- Zaki MJ, Parthasarathy S, Ogihara M, Li W (1997) New algorithms for fast discovery of association rules. In: *KDD*, pp 283–286
- Zhang J, Breugelmans E (2012) The impact of an item-based loyalty program on consumer purchase behavior. *Journal of Marketing Research* 49(1):50–65

Zhang J, Krishnamurthi L (2004) Customizing promotions in online stores. *Marketing Science* 23(4):561–578

Zhang J, Wedel M (2009) The effectiveness of customized promotions in online and offline stores. *Journal of Marketing Research* 46(2):190–206