A New Spatial Classification Methodology for Simultaneous Segmentation, Targeting, and Positioning (STP Analysis) for Marketing Research

Chapter · January 2009

3 authors:

Wayne S. Desarbo
Pennsylvania State University

Simon Blanchard
Georgetown University

Selin Atalay
Frankfurt School of Finance & Management

Review of Marketing Research

A New Spatial Classification Methodology for Simultaneous Segmentation, Targeting, and Positioning (STP Analysis) for Marketing Research

Wayne S. DeSarboSimon J. BlanchardA. Selin Atalay

CHAPTER 4

A NEW SPATIAL CLASSIFICATION METHODOLOGY FOR SIMULTANEOUS SEGMENTATION, TARGETING, AND POSITIONING (STP ANALYSIS) FOR MARKETING RESEARCH

WAYNE S. DESARBO, SIMON J. BLANCHARD, AND A. SELIN ATALAY

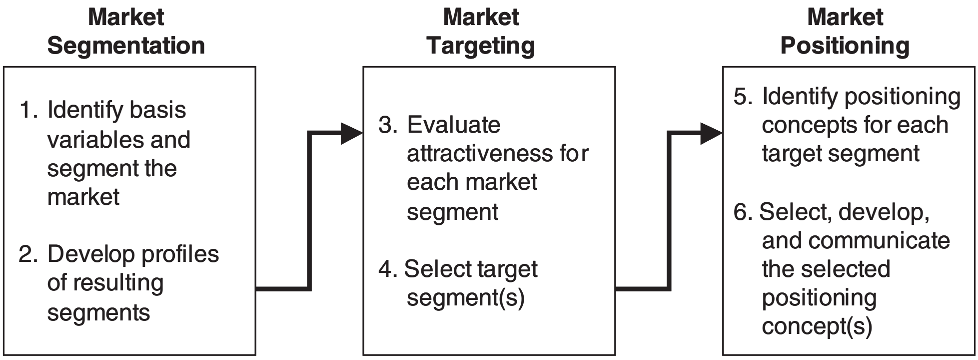
Abstract

The Segmentation-Targeting-Positioning (STP) process is the foundation of all marketing strategy. This chapter presents a new constrained clusterwise multidimensional unfolding procedure for performing STP that simultaneously identifies consumer segments, derives a joint space of brand coordinates and segment-level ideal points, and creates a link between specified product attributes and brand locations in the derived joint space. This latter feature permits a variety of policy simulations by brand(s), as well as subsequent positioning optimization and targeting. We first begin with a brief review of the STP framework and optimal product positioning literature. The technical details of the proposed procedure are then presented, as well as a description of the various types of simulations and subsequent optimization that can be performed. An application is provided concerning consumers' intentions to buy various competitive brands of portable telephones. The results of the proposed methodology are then compared to a naïve sequential application of multidimensional unfolding, clustering, and correlation/regression analyses with this same communication devices data. Finally, directions for future research are given.

Introduction

Competitive market structure analysis is an essential ingredient of the strategic market planning process (cf. Day 1984; Myers 1996; Wind and Robertson 1983). DeSarbo, Manrai, and Manrai (1994) describe the primary task of competitive market structure analyses as deriving a spatial configuration of products/brands/services in a designated product/service class on the basis of some competitive relationships between these products/brands/services (see also Fraser and Bradford 1984; Lattin and McAlister 1985). As mentioned in Reutterer and Natter (2000), to provide managers with more meaningful decision-oriented information needed for the evaluation of actual brands' positions with respect to competitors, competitive market structure analyses have been enhanced with positioning models that incorporate actual consumers' purchase behavior/intentions, background characteristics, marketing mix, etc. (cf. Cooper 1988; DeSarbo and Rao 1986; Green and Krieger 1989; Moore and Winer 1987). Thus, contemporary approaches for the empirical modeling of competitive market structure now take into account both consumer heterogeneity (e.g., market segmentation) and the competitive relationship of products/brands/services (e.g.,

Figure 4.1 **The STP Process**



Source: Modified from Kotler (1997).

positioning). Such analyses are integrated into the more encompassing Segmentation-Targeting-Positioning (STP) approach (Kotler 1997; Lilien and Rangaswamy 2004).

Figure 4.1, modified from Kotler (1997), illustrates the various steps or stages of this STP concept concerning implementation. Here, each of the three stages is depicted as discrete steps performed in a sequential manner. In stage I, *segmentation*, basis variables defined for segmentation (e.g., customer needs, wants, benefits sought, preference, intention to buy, usage situations, etc.) are collected depending on the specific application. Specification of these basis variables is important to ensure that the derived market segments are distinct in some managerially important manner (e.g., with respect to behavior). These data are then input into some multivariate procedure (e.g., cluster analysis) to form market segments. Myers (1996) and Wedel and Kamakura (2000) provide reviews of the various methodologies available for use in market segmentation, as well as their pros and cons. These derived market segments are then typically identified using profile variables that aid the firm in understanding how to serve these customers (e.g., shopping patterns, geographic location, spending patterns, marketing mix sensitivity, etc.), as well as how to communicate to these customer segments (e.g., demographics, media usage, psychographics, etc.). It is important that the derived market segments exhibit the various traits for effective segmentation (cf. DeSarbo and DeSarbo 2002; Kotler, 1997; Wedel and Kamakura, 2000): differential behavior, membership identification, reachability, feasibility, substantiality, profitability, responsiveness, stability, actionability, and predictability.

In stage II, *targeting*, one evaluates the financial attractiveness of each derived market segment in terms of demand, costs, and competitive advantage. Based on this financial evaluation, one or more target segments are selected as "targets" based on the profit potential of such segments and their fit with the firm's corporate goals and strategy. The "optimal" level of resources is then determined to allocate to these targeted market segments. As a final phase of this intermediate targeting stage, customers and prospects are often identified in these targeted segments. Kotler and Keller (2006) list five different patterns of target market selection, including single segment concentration (specialize in one segment), selective specialization (concentrate in S segments), product specialization (produce product variant in one major product class across all segments), market specialization (serve the complete needs in a specified market), and full market coverage (full coverage of all market segments and products).

Finally, in stage III, *positioning*, the marketer identifies a positioning concept for the firm's products/services that attracts targeted customers. Kotler and Keller (2006) define positioning as "the act of designing the company's offering and image to occupy a distinctive place in the mind of the target market" (p. 310). In such a framework where a firm targets one or more groups or market segments with its offerings, positioning then becomes a segment-specific concept. As noted by Wind (1980), most firms produce multiple products/services, and the positioning decision of any given product cannot ignore the place the product occupies in the firm's product line as perceived by the consumer. The product offering of a firm should lead to an optimal mix of product positioning by market segments, that is, the product positioning of any given product should not be designed and evaluated in isolation from the positioning of the firm's other products and the market segments at which they are aimed (p. 70). In addition, the focus of a product line positioning should not be limited to the firm's own products but rather should take explicitly into account the product lines of its competitors as well.

Empirical modeling approaches in the STP area have taken typically two forms: (1) sequential use of multidimensional scaling and cluster analysis, and (2) parametric finite mixture or latent class multidimensional scaling models. The more traditional approaches have employed the sequential use of multidimensional scaling (MDS) and cluster analysis to first spatially portray the relationship between competitive brands and consumers, and then segment the resulting consumer locations to form market segments. As mentioned in DeSarbo, Grewal, and Scott (2008), a number of methodological problems are associated with such a naïve approach. One is that each type of analysis (multidimensional scaling and cluster analysis) typically optimizes different loss functions. (In fact, many forms of cluster analysis optimize nothing.) As such, different aspects of the data are often ignored by the disjointed application of such sequential procedures. Two, there are many types of multidimensional scaling and cluster analysis procedures, as documented in the vast psychometric and classification literature, where each type of procedure can render different results (cf. DeSarbo, Manrai, and Manrai 1994). In addition, there is little a priori theory to suggest which methodological selections within these alternatives are most appropriate for a given marketing application. And the various combinations of types of multidimensional scaling and cluster analyses typically render different results as well. Finally, as noted by Holman (1972), the Euclidean distance metric utilized in many forms of multidimensional scaling is not congruent with the ultrametric distance formulation metric utilized in many forms of (hierarchical) cluster analysis.

Efforts to resolve such methodological problems associated with this naïve sequential administration of multidimensional scaling and cluster analyses have resulted in the evolution of parametric finite mixture or latent class multidimensional scaling models (cf. DeSarbo, Manrai, and Manrai 1994; Wedel and DeSarbo 1996). In particular, latent class multidimensional scaling models for the analysis of preference/dominance data have been proposed by a number of different authors over the past fifteen years employing either scalar products/vector (Slater 1960; Tucker 1960) or unfolding (Coombs 1964) representations of the structure in two-way preference/dominance data. In such latent class multidimensional scaling models, vectors or ideal points of derived segments are estimated instead of separate parameters for every individual consumer. Thus, the number of parameters is significantly reduced relative to individual-level models. Latent class multidimensional scaling models are traditionally estimated using the method of maximum likelihood (E-M algorithms [cf. Dempster, Laird, and Rubin 1977] are typically employed). For example, DeSarbo, Howard, and Jedidi (1991) developed a latent class multidimensional scaling vector model (MULTICLUS) for normally distributed data. DeSarbo, Jedidi, Cool, and Schendel (1990) extended this latent class multidimensional scaling model to a weighted ideal point model. De Soete and Heiser (1993) and De Soete and Winsberg (1993), respectively, extended the MULTICLUS latent class multidimensional scaling model by accommodating linear restrictions on the stimulus coordinates. Böckenholt and Böckenholt (1991) developed simple and weighted ideal-point latent class multidimensional scaling models for binary data. DeSarbo, Ramaswamy, and Lenk (1993) developed a vector latent class multidimensional scaling model for Dirichlet-distributed data, and Chintagunta (1994) developed a latent class multidimensional scaling vector model for multinomial data. Wedel and DeSarbo (1996) extended the entire family of exponential distributions to such latent class multidimensional scaling models for two-way preference/dominance data.

As noted in DeSarbo, Grewal, and Scott (2008), such latent class multidimensional scaling models also have a number of limitations associated with them. One, they are parametric models that require the assumption of specific distributions. Oftentimes, continuous support distributions are utilized in the finite mixture and are applied to traditional discrete response scales (e.g., semantic differential, Likert, etc.), which cannot possibly obey such distributional assumptions. As such, violations of such distributional assumptions may invalidate the use of the procedure. Two, such latent class multidimensional scaling procedures require the underlying finite mixture to be identified (cf. McLachlan and Peel 2000), which is often problematic when the underlying distributions deviate from the exponential family. When multivariate normal distributions are employed, identification problems can arise in the estimation of separate full covariance matrices by derived market segment. Three, most of the latent class multidimensional scaling procedures are highly nonlinear in nature and require very intensive computation. Such procedures typically utilize an E-M approach (cf. McLachlan and Krishnan 1997) or gradient-based estimation procedures, which may take hours of computation time for a complete analysis to be performed. Four, at best, only locally optimum solutions are typically reached, and the analyses have to be repeated several times for each value of the dimensionality and number of groups. Five, the available heuristics employing various information criteria typically result in different solutions being selected. For example, the BIC and CAIC are considered as more conservative measures resulting in the selection of fewer dimensions and groups in contrast to AIC and MAIC, which are considered to be more liberal criterion (cf. Wedel and Kamakura 2000). As discussed in Wedel and Kamakura (2000), there are still other heuristics utilized for model selection for such finite mixture models (e.g., ICOMP, NEC, etc.), which are equally plausible but might also result in different solutions being selected. Finally, the underlying framework assumed by these latent class multidimensional scaling procedures involves a partitioning of the sample space, although the estimated posterior probabilities of membership often result in fuzzy probabilities of segment membership, which could be difficult to interpret or justify in applications requiring partitions. In fact, pronounced fuzziness of the resulting posterior probabilities of segment membership may be indicative of poor separation of the conditional support functions' centroids.

Our goal in this research is to contribute to current STP research by devising a new deterministic, clusterwise multidimensional unfolding procedure for STP that overcomes some of the above-mentioned limitations of the current procedures. Our goal is to *simultaneously* derive a single joint space where derived "segments" are also determined and represented by ideal points, and brands via coordinate points, and their interrelationship in the space denotes some aspect of the structure in the input preference/dominance data. In addition, our approach (like the GENFOLD2 multidimensional unfolding procedure of DeSarbo and Rao [1986]) explicitly relates brand attributes to brand locations in the derived joint space. The proposed approach does not require distributional assumptions, such as latent class multidimensional scaling procedures, and it provides a concise spatial representation for the analysis of preference/dominance data, as will be illustrated in our application to buying intentions for portable telephone communication devices. In the proposed model, no finite mixture distribution identification is required, unlike its latent class multidimensional scaling counterparts. We generalize the DeSarbo, Grewal, and Scott (2008) clusterwise vector model to the unfolding case (see also DeSarbo, Atalay, LeBaron, and Blanchard 2008) with reparametrization of the derived brand space. The estimation procedure developed is relatively fast and efficient, and it converges in a matter of minutes on a PC (although potential locally optimum solutions are possible here also). Finally, our proposed procedure accommodates both overlapping segments (cf. Arabie et al. 1981), as well as nonoverlapping segments. These two segmentation schemes account for consumer heterogeneity within the data. While nonoverlapping segment structures exemplify the separation in preference between segments, overlapping segments acknowledge that individuals may have multiple needs or goals and therefore be members of multiple segments (cf. DeSarbo, Atalay, LeBaron, and Blanchard 2008).

In this chapter, we propose a method that simultaneously identifies segments of consumers and relates product characteristics to the derived product space. The segment-level ideal point representation allows for a parsimonious representation of preferences, and the relation between product characteristics and the attribute space eases the understanding of the brand space and facilitates the generation of actionable new products. In the following section, we review the major trends in optimal product positioning research in the area of STP analysis. Then, we present the clusterwise multidimensional unfolding model (with reparametrization of the attributes) in an optimal product positioning framework for STP. Following this, a comparison with a traditional sequential procedure is presented (MDS and cluster analysis) in terms of an application based on communication devices research data initially presented in DeSarbo and Rao (1986). Finally, we discuss several avenues for future research.

Optimal Positioning in STP

Most recent methods of optimal product positioning operate under the assumption that competition between brands in a designated product/service class can be adequately represented in a joint space map. The preferences of consumers are often represented in the derived joint space map as ideal points that denote consumers' ideal brands: the brands that would maximize their utility and that are most likely to be chosen if they were to be offered in the market. Such maps also help illustrate the relationships between the brands and their attributes. As the most basic models assume that an individual's preference is inversely related to the distance between the brands and his/her ideal point, the closer an ideal point is to a brand in the derived space, the greater is the individual's preference for that brand.

The two most common methods for obtaining such cognitive spaces are multidimensional unfolding (MDU) and conjoint analysis (CA). Baier and Gaul (1999) offer comprehensive lists of the studies in both MDU and CA in the context of optimal product positioning and design. These methods generally focus on either product positioning or product design, which have consistently been confused in the literature (see Kaul and Rao [1995] for a detailed review). Product positioning models use abstract perceptual dimensions (e.g., quality, prestige, beauty, etc.), while product design models use more specific product characteristics or features (e.g., color, price, size). Whereas product design models often miss the impact of the marketing mix variables, the products generated from positioning models are often difficult to execute in practice. The need of integration between derived dimensions and attributes has been stressed in several models (Baier and Gaul 1999; DeSarbo and Rao 1986; Hauser and Simmie 1981; Michalek, Feinberg, and Papalambros 2005; Shocker and Srinivasan 1974). This is critical because specific marketing strategies and actions must be determined to obtain desired or optimal positions (Eliashberg and Manrai 1992).

A model for generating the choice behaviors of the consumers must also be specified. The first generation of optimal positioning models assumed a deterministic model of choice in which consumers would choose the brand that they most prefer with certainty (Albers 1979; Albers and Brockhoff 1977; Gavish, Horsky, and Srikanth 1983; Zufryden 1979). In the context of ideal point models, this suggests that consumers always choose the brand that is positioned closest to their ideal points. The implicit assumption here was that these methods were to be used for choice behaviors about infrequently chosen objects and durables for which consumers have been assumed to select the best option with certainty. These models also assume that the consumers have strongly held preferences and render no consideration to situational (e.g., different uses, environmental changes) or contextual (e.g., variety-seeking tendencies, mood) factors that also may impact choice.

Other approaches followed a probabilistic choice rule where products have a nonzero probability of being chosen depending on their distance to the consumer ideal points (e.g., Bachem and Simon 1981; Gruca and Klemz 2003; Pessemier et al. 1971; Shocker and Srinivasan 1974). Generally, the inverse of the distance between a product and an ideal point is taken as a proxy for the probability of that product's being chosen. These probabilistic models have been found, in simulations, to provide somewhat better solutions than the single choice deterministic models (Shocker and Srinivasan 1979; Sudharshan, May, and Shocker 1987).

Note that this difference between deterministic and probabilistic models generating choice is not to be confused with that between nonstochastic and stochastic models. The latter implies that the model parameters such as the location of the ideal points and the brand positions are random variables drawn from pre-specified distributions (e.g., Baier and Gaul 1999). Both stochastic and nonstochastic methods are generally subject to Luce's axiom (Luce 1959), which relies on often difficult assumptions to justify in the product positioning context (Batsell and Polking 1985; Bordley 2003; Meyer and Johnson 1995).

Early optimization models maximized market share for the new brand based on predicted single choices (Albers 1979; Albers and Brockhoff 1977; Shocker and Srinivasan 1974). These models did not explicitly consider price/cost information (Hauser and Simmie 1981; Schmalensee and Thisse 1988). Several authors followed with profit formulations (Bachem and Simon 1981; Choi, DeSarbo, and Harker 1990; Green, Carroll, and Goldberg 1981; Green and Krieger 1985; DeSarbo and Hoffman 1987). Others improved realism by adding budget constraints (Thakur et al. 2000), by incorporating competitive reactions (Choi, DeSarbo, and Harker 1990), by adding variable costs (Bachem and Simon 1981; Gavish, Horsky, and Srikanth 1983), or by implementing a more complete cost structure that involves variable development costs (Bordley 2003). Even with these developments, all these models are highly dependent on the quality of the brand space and preference representations that they obtain in the first step, especially since degenerate solutions haunt most MDU procedures (cf. DeSarbo and Rao 1986). All such optimal product approaches are also limited by the difficulties associated with the collection of cost information (Green, Carroll, and Goldberg 1981).

In addition to these modifications, these optimal positioning models also differ in a number of other ways. It was argued that because of self-interest, individuals are likely to be more familiar with products they more highly prefer. DeSarbo, Ramaswamy, and Chatterjee (1995) examine the effect of differential familiarity in MDS models. Models should also ideally consider that the chosen product will vary depending on the situation and on individual factors; it will be chosen from a reduced set of options that depends partly on the type of usage that is intended. To account for this, some authors used a reduced consideration set of the k-closest options to their ideal points (Gavish, Horsky, and Srikanth 1983; Gruca and Klemz 2003; Sudharshan, May, and Shocker 1987).

However, the assumption that some products can have a zero probability of being chosen seems excessively strong (Pessemier et al. 1971; Schmalensee and Thisse 1988).

These optimal positioning models also differ in their computational complexity. The initial models of Albers et al. (1977, 1979) are limited by the number of ideal points. As such, many authors suggested that grouping consumers into market segments could lead to reduced computational complexity (McBride and Zufryden 1988; Schmalensee and Thisse 1988). Zufryden (1982) proposed grouping individuals based on their utilities and on their usage intentions, whereas Shugan and Balachandran (1977) suggested forming segments based on the preference orders a priori. The way that the segments are created is especially important, as it should be closely related to the choice behavior at hand. Two-step models, where segmentation is obtained after or before the ideal points and the brand space are obtained, are less efficient than methods that simultaneously cluster individuals based on their ideal points and their perceived brand structure (see DeSarbo, Manrai, and Manrai 1994).

Additionally, most of the optimal positioning methods follow a two-step structure (the brand space and ideal points are first identified *before* any segmentation or the optimal product positioning stage). The latter is advantageous because the positioning step does not depend on the type of data that is initially used to obtain the ideal points (Shocker and Srinivasan 1979). This has resulted in the use of a variety of algorithms and input data, such as paired comparisons and rank-order preferences, among others. However, once the brand positions and the ideal points are obtained, data considerations about the attributes and characteristics become increasingly important. Some attributes are inherently discrete, such as the presence of the air-conditioning option in a vehicle. Others, such as price, miles per gallon, and quality, represent continuous information.

The models for optimal positioning have been limited by the type of data that they can deal with, partly because of the difficulties in using various types in optimization routines, but also because of the difficulties in getting the costs associated with the various attributes and characteristics. To optimize profit functions, it is important that product characteristics and elements of the marketing mix be quantifiable in terms of costs. This can be difficult with some attributes, which lead some to the discretization of continuous variables at levels for which product costs are known. This also simplified the optimization, although it precluded traditional (continuous) gradient procedures from being used in subsequent optimization.

Even when the attribute/characteristics issues are resolved, not all combinations of characteristics may be feasible, and there is a need for models that can incorporate constraints on feasibility (Shocker and Srinivasan 1979). As previously mentioned, it is essential that attribute-based propositions can be translated into a feasible product space that can be successfully produced by the firm (Gavish, Horsky, and Srikanth 1983). Finally, the variants of the optimization problems are often difficult to solve for. The gradient procedures (when the attributes are continuous), given the nonconvex nature of the problem, are subject to numerous local optima. The mixed-integer nonlinear problems are NP-hard. Optimization procedures that have been used to improve on performance and computational times include genetic algorithms (Gruca and Klemz 2003), branch and bound (Hansen et al. 1998), dynamic programming (Kohli and Krishnamurti 1987), and nested partition methods (Shi, Olafsson, and Chen 2001).

Given the above variants regarding optimal product positioning in STP problems, we propose a clusterwise multidimensional unfolding procedure that simultaneously identifies segments of consumers that have similar ideal products, and that reparameterizes the brand space to be a linear function of actionable product characteristics. The segment ideal points, the obtained brand space, and reparametrization coefficients can then be used to generate and select the best new product to introduce in the market. The clusterwise multidimensional unfolding procedure that we propose is a direct generalization of GENFOLD2 (DeSarbo and Rao 1986) to the clusterwise case. It presents several advantages over previous methods. As mentioned, it is important to select segments of consumers who follow a meaningful behavioral/choice pattern. By integrating the segmentation stage into the model, the model optimizes the real underlying objective to find the segments that have the most similar preference structure. It is also independent of all parametric assumptions. Finally, it provides a simple segment-level ideal point structure that will be computationally easier at the product optimization stage as the computation is greatly influenced by the number of ideal points.

The optimization step that we propose is an extension of Shocker and Srinivasan's (1974) "distance as probability" model for discretrized product characteristics, with a profit function and with segment-level ideal points. The brand reparametrization feature permits the link between derived brand coordinates and brand attributes, as it is important to suggest optimal brands that will be realizable, because the dimensions (brand coordinates) are obtained as a linear function of the brand attributes. Using this reparametrization, one now has a way not only to better interpret the dimensions but also to examine the effect in the derived joint space of altering attribute values for repositioning existing brands or positioning new brands. Finally, the optimal product-positioning step identifies and selects, through a probabilistic nonstochastic model, the product to introduce that will maximize a company's profit function. Because of the reduction of the number of ideal points in the first step, it is possible to proceed to the complete enumeration of the realizable products and to select the globally optimal new product to introduce in the market.

The Proposed Clusterwise Unfolding Model

Clusterwise models are common in the marketing and psychometric literatures. These methods simultaneously group individuals into segments to represent sample heterogeneity and derive segment-level parameters for the particular model being estimated. Several clusterwise formulations have been proposed for the regression problem (DeSarbo, Oliver, and Rangaswamy 1989; Spath 1979, 1981, 1982; Wedel and Kistemaker 1989). Here, we develop a clusterwise multidimensional model with brand reparametrization options as in GENFOLD2 (DeSarbo and Rao 1986).

Let:

$i = 1, \ldots, N$ consumers;
$j = 1, \ldots, J$ brands;
$k = 1, \ldots, K$ brand attributes ($K < J$);
$s = 1, \ldots, S$ market segments (unknown);
$r = 1, \ldots, R$ dimensions (unknown);
Δ_{ij} = the dispreference for brand j given by consumer i.

We model the observed data as:

$$\Delta_{ij} = a\sum_{s=1}^{S} P_{is}\sum_{r=1}^{R}(X_{jr} - Y_{sr})^2 + b + \varepsilon_{ij}\,, \tag{1}$$

where:

X_{jr} = the rth coordinate for brand j;
Y_{sr} = the rth coordinate of the ideal point for market segment s;
Z_{jk} = the value of the jth brand on the kth product characteristics;
b = an additive constant;
a = a multiplicative constant;
P_{is} = 1 if consumer i is classified into market segment s,
0 else;

where:

$P_{is} \in \{0,1\}$,

$\sum_{s=1}^{S} P_{is} = 1 \ \forall_i$ (for nonoverlapping segments),

or

$0 < \sum_{s=1}^{S} P_{is} \leq S$ (for overlapping segments);

ε_{ij} = error (deterministic).

As in DeSarbo and Rao (1986), we have an option to reparameterize the brand coordinates as a linear function of specified product characteristics:

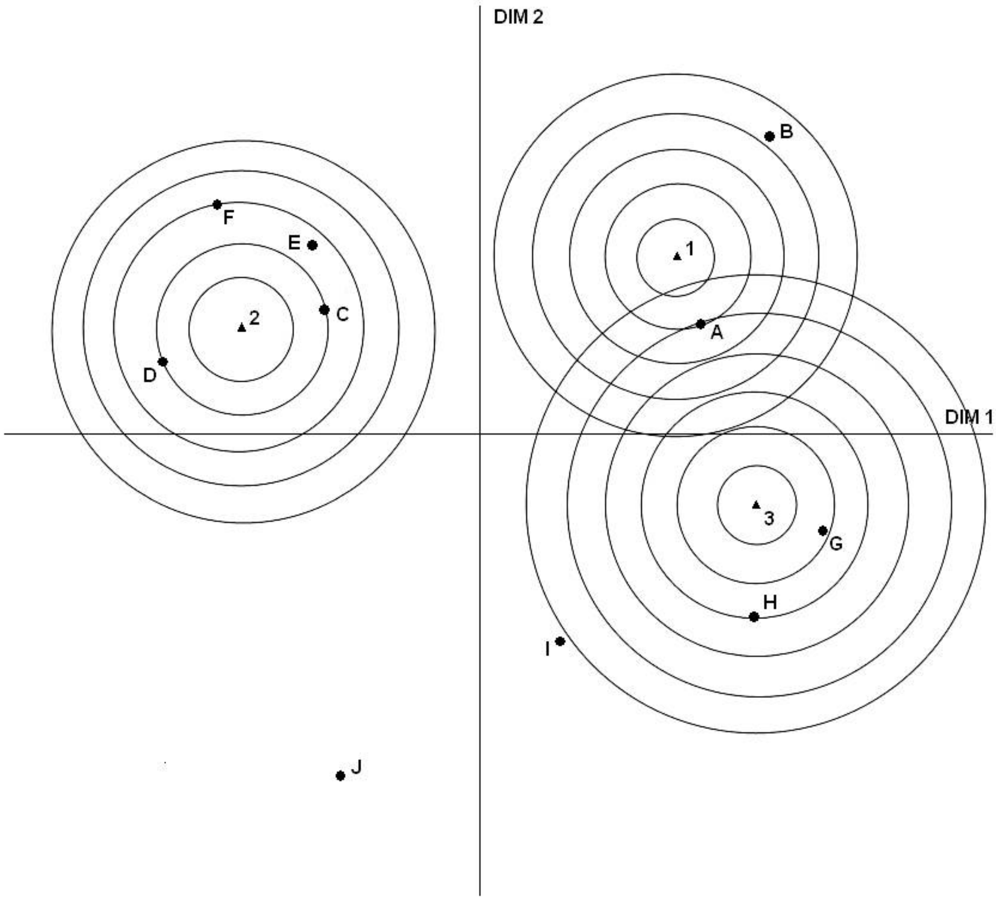
$$X_{jr} = \sum_{k=1}^{K} Z_{jk} \alpha_{kr} \tag{2}$$

or

$$\underline{X} = \underline{Z\alpha} . \tag{3}$$

The proposed model then provides a spatial representation of the brands (as perceived by all consumers) and the location of each of the segment's ideal points. Figure 4.2 is a hypothetical two-dimensional illustration that includes ten brands (labeled A through J) and three segment-level ideal points (labeled 1, 2, and 3). Each brand is positioned in the space in which increasing distances between points implies higher dissimilarity. For example, brands B and J are at opposite ends of the space and hence are implicitly perceived as dissimilar by consumers. For a given segment, an ideal point represents the coordinates of the hypothetical product for which preference would be the greatest among consumers belonging to that particular segment. For segment 1, one can see that A is the most preferred brand: it is the closest in the space. Brand B is the second closest and brand J is the least preferred. For segment 2, brands C and D are equally preferred. This is exemplified by the iso-preference contours (circles that indicate coordinates of equal distances/preference to an ideal point) by which C and D are of equal distance to the ideal point of segment 2, although their coordinates in the space are quite different.

Figure 4.2 **Hypothetical Illustration of the Methodology**



Estimation

Given $\underline{\Delta} = ((\Delta_{ij}))$ and values of S and R, we wish to estimate a, b, $\underline{P} = ((P_{is}))$, $\underline{X} = ((X_{jr}))$, and $\underline{Y} = ((Y_{sr}))$ so as to minimize:

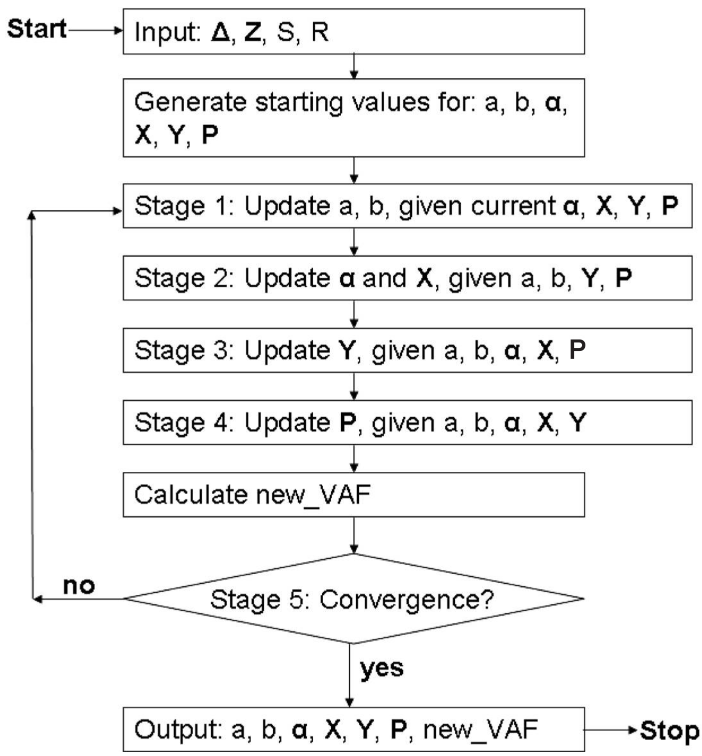
$$\Phi = \sum_{i=1}^{N} \sum_{j=1}^{J} \sum_{t=1}^{T} \left[\Delta_{ij} - \hat{\Delta}_{ij} \right]^2, \tag{4}$$

where:

$$\hat{\Delta}_{ij} = \sum_{s=1}^{S} \hat{P}_{is} \sum_{r=1}^{R} \hat{a}(\hat{X}_{jr} - \hat{Y}_{sr})^2 + \hat{b}. \tag{5}$$

We devise an alternating least-squares iterative estimation procedure involving five stages, all of which conditionally optimize expression (4). Given the nonlinear nature of the underlying

Figure 4.3 **Flowchart of the ALS Algorithm**



model and objective function, convergence to globally optimum solutions is not guaranteed as with all other MDU approaches. These estimation steps are depicted in Figure 4.3 above and described below:

Stage 1: Estimate a and b

We first define:

$$\hat{\Delta}^*_{ij} = \sum_{s=1}^{S} P_{is} \sum_{r=1}^{R} (X_{jr} - Y_{sr})^2, \tag{6}$$

where:

$\underline{L}$ = an NJ × 1 vector containing $\hat{\Delta}^*_{ij}$
$\underline{K}$ = ($\underline{1}$, $\underline{L}$),
$\underline{1}'$ = (1,1,...,1), an NJ × 1 vector of ones,
M = a NJ × 1 vector containing Δ_{ij}.

It is now straightforward to show the following formulation:

$$\begin{pmatrix} \hat{b} \\ \hat{a} \end{pmatrix} = (\underline{K}' \ \underline{K})^{-1} \underline{K}' \ \underline{M}, \tag{7}$$

can be gainfully employed to obtain current estimates of the additive and multiplicative constants, given fixed current values of $\underline{P}$, $\underline{X}$, and $\underline{Y}$. Note, the multiplicative constant $a > 0$ is not identifiable, and its square root is multiplied into $\underline{X}$ and $\underline{Y}$, with a then set equal to 1.00.

Stage 2: Estimate $\underline{X}$ ($\underline{\alpha}$)

Now, holding fixed current values of a, b, $\underline{P}$, and $\underline{Y}$, we estimate $\underline{\alpha}$ via a conjugate gradient procedure (see Nocedal and Wright 2006). Here, the respective partial derivatives of Φ with respect to α_{kr} are:

$$\frac{\partial \Phi}{\partial \alpha_{kr}} = -4 \sum_{i=1}^{N} \sum_{j=1}^{J} (\Delta_{ij} - \hat{\Delta}_{ij}) \left[\sum_{s} P_{is} a (X_{jr} - Y_{sr}) Z_{jk} \right]. \tag{8}$$

Minimization of Φ occurs until a number of user preset minor iterations (MIT) are exhausted or there is failure to improve the objective function. Once $\underline{\alpha}$ is estimated, $\underline{X}$ is then determined via expression (3).

Stage 3: Estimate $\underline{Y}$

Now, holding fixed current values of a, b, $\underline{P}$,and $\underline{X}$, we estimate $\underline{Y}$ via a conjugate gradient procedure as well. Here, the respective partial derivatives with respect to Y_{sr} are:

$$\frac{\partial \Phi}{\partial Y_{sr}} = +4 \sum_{i=1}^{N} \sum_{j=1}^{J} (\Delta_{ij} - \hat{\Delta}_{ij}) \left[P_{is} a (X_{jr} - Y_{sr}) \right]. \tag{9}$$

Here too, minimization of Φ occurs until a number of user preset MIT are exhausted or there is failure to improve the objective function.

Stage 4: Estimate $\underline{P}$

For estimating the membership matrix $\underline{P}$, one notes that the error sums of squares (SSE) objective function minimized in expression (4) is calculated as a sum of squared discrepancies over brands (j), and subjects (i). One can thus rewrite:

$$\Phi = \sum_{i=1}^{N} \Phi_i, \tag{10}$$

where:

$$\Phi_i = \sum_{j=1}^{J} (\Delta_{ij} - \sum_{s-1}^{S} P_{is} \sum_{r=1}^{R} a (X_{jr} - Y_{sr})^2 - b)^2 ; \tag{11}$$

that is, for a given subject i, (4) is separable with respect to that mode of the array. That is, $\underline{P}_i$ and $\underline{\Delta}_i$ are the only two entities indexed by i that affect Φ in expression (4), and thus one could perform the subsequent optimization by subject in minimizing expression (11) for each consumer i. Using this separable property, the proposed procedure performs a complete enumeration over $2^s - 1$ (ignoring the $\underline{0}$ solution) options for the overlapping cluster case (S options for the non-overlapping case) to obtain a globally optimum solution for $\underline{P}$ *conditioned* on current values of $\underline{X}$, $\underline{Y}$, a, and b.

Stage 5: Convergence

Thus, the proposed alternating least-squares estimation procedure iteratively cycles through (7), (8), (9), and the N complete enumerations to minimize the error sums of squares in (4). Stages 2 and 3 involving the conjugate gradient procedures conditionally provide at least locally optimum estimates of the associated set of parameters ($\underline{X}$, $\underline{Y}$) estimated *conditional* on the remaining sets of parameters held at their current values. Stages 1 and 4 provide conditionally global estimates of (a, b) and $\underline{P}$, holding all other parameters fixed at their current values. This process obviously does not guarantee a final globally optimum result, and thus multiple runs (we recommend 10) are required for each solution in order to detect the best one. Convergence is reached when subsequent improvement in the objective function reaches a critical level. Given that the magnitude and scale of the objective function in expression (4) is contingent on such factors as the size of the data array (I, J), the scale of the preference ratings collected, and so on, two normalized goodness-of-fit measures (both ranging between 0 and 1) are also utilized to assess convergence, model selection (i.e., selecting values of R and S), and overall fit via a variance accounted for measure (VAF):

$$VAF = 1 - \frac{\sum_i \sum_j (\Delta_{ij} - \hat{\Delta}_{ij})^2}{\sum_i \sum_j (\Delta_{ij} - \overline{\Delta})^2} \tag{12}$$

for interval scale data, and a sums-of-squares accounted for measure (SSAF):

$$SSAF = 1 - \frac{\sum_i \sum_j (\Delta_{ij} - \hat{\Delta}_{ij})^2}{\sum_i \sum_j \Delta_{ij}^2} \tag{13}$$

for ratio scale data (cf. DeSarbo and Carroll 1985), where:

$$\hat{\Delta}_{ij} = \sum_{s=1}^{S} \hat{P}_{is} \sum_{r=1}^{R} \hat{a}(\hat{X}_{jr} - \hat{Y}_{sr}) + \hat{b} \tag{14}$$

and

$$\overline{\Delta} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{J} \Delta_{ij}}{NJ} \; . \tag{15}$$

Finally, $S \geq R$ is a necessary identification restriction required (Wedel and DeSarbo 1996).

The method that we propose is a direct generalization of GENFOLD2 (DeSarbo and Rao 1986) to the clusterwise case, where the brand space and coordinates are defined as a linear function of the available products' characteristics. It also relates to the work of DeSarbo et al. (2008) concerning their estimation of multiple ideal points with survey data. It is also related to the work of Baier and Gaul (1999), who proposed a parametric model to simultaneously identify segments and ideal points for the optimal product positioning problem. They then proposed to find the optimal new product using a gradient procedure. The method that we propose compares favorably to that of Baier and Gaul (1999). First, we do not make any parametric assumption regarding the distributions of the segment-level ideal points. Also, their method is limited to product characteristics that have a continuous level of measurement and to paired comparison data. The model that we propose is also related to that of Andrews and Manrai (1999) who proposed a latent class vector model that can incorporate reparametrization of a reduced attribute space. Their method maps each attribute separately, focusing rather on explaining the differences between the various attribute levels and the segment-level weights associated with them. A proposed extension also permits the analysis of marketing mix changes on the perceptions of attribute levels at the segment level. However, if multiple changes in marketing mix are advisable, the space needs to be recomputed by restricting the preference vectors to be common for the selected attributes. Moreover, while this model is relevant in terms of reparametrization, its vector formulation makes it more difficult to use in an optimal product positioning context as it possesses a scalar products type of utility function that assumes "the more the better," and optimal brands would necessarily be positioned toward infinity in selected directions, as they would necessarily project higher on consumer/segment vectors.

The Proposed Optimal Brand Positioning Module

Consider index G, which represents the optimal brand to be introduced in the market. From the proposed clusterwise multidimensional unfolding analysis, the following are obtained:

Y_{sr} is the ideal point coordinate for derived segment s on dimension r.
X_{jr} is the coordinate of brand j on dimension r.
α_{kr} is the impact coefficient of attribute k on dimension r.
N_s is the number of expected consumers in segment s.

In case of overlapping clusters, proportions of assignment are used to avoid counting individuals in multiple clusters multiple times. For example, if an individual is a member of both cluster 1 and 2, 0.5 will be added to N_1 and 0.5 will be added to N_2. Now, let ϕ_s be the sum of the inverse distances between a brand and a segment ideal point, for all brands that are assumed to have stable positions in the market:

$$\phi_s = \sum_{j=1}^{J}\left(1/\sum_{r=1}^{R}\left(Y_{sr} - Z_{jr} \times \alpha_{jr}\right)^2\right), \forall s = 1...S. \tag{16}$$

Then let:

$$G_{kl} = \begin{cases} 1, \text{ if the proposed product G has attribute level } 1(1 = 1...L_k) \text{ on attribute k.} \\ 0, \text{ otherwise.} \end{cases}$$

where the attribute levels $1 = 1 \ldots L_k$ are chosen to be the levels that attribute k can take in a new product. Also, let V_{kl} be the value of the level that can be used along with α_{kr} to represent a newly constructed product in the brand space. Finally, let C_{Glk} be the benefit from having option G level 1 on attribute k. Then the expected profit π_G obtained from selling one proposed product G, based on the G_{kl} attribute levels, is given by:

$$\pi_G = \sum_{k=1}^{K} \sum_{l=1}^{L_k} C_{Glk} G_{kl} . \tag{17}$$

Both the costs of using the chosen attributes and the price sold are part of the optimization process. These costs are assumed to be known a priori. The sales price is one of the attributes and is coded as a positive number, whereas the costs of using different attribute levels are coded as negative numbers. Now, let

$$d_{Gs} = 1/\left(\sum_{r=1}^{R} \left(Y_{sr} - \sum_{k=1}^{K} \sum_{l=1}^{L_k} G_{kl} \times V_{kl} \times \alpha_{kr} \right)^2 \right), \forall s = 1 \ldots S \tag{18}$$

be the inverse of the squared Euclidian distance between the proposed product G and segment s. The main objective function becomes:

$$Max_G \sum_{s=1}^{S} \pi_G N_s \left(\frac{\phi_s + d_{Gs}}{d_{Gs}} \right), \tag{19}$$

subject to:

$$\sum_{l=1}^{L_k} G_{kl} = 1, \forall k = 1 \ldots K, \tag{20}$$

where the group of constraints requires that there must be one and only one chosen level 1 for every attribute k. Other constraints can also be added to comply with production and feasibility restrictions. The inverse distance measure in an "us/(us + them)" format permits the assignment of a probability that the product is being chosen. The closer a product is to an ideal point, the higher the probability that individuals within the segment will choose the product. It is assumed that the choice of individuals will not be affected by the choices of other individuals. We then assume that the error in the measure of distances is not correlated in any way across different groups of products. It is assumed there is no latent construct that explains covariation between two products that are perceived as equally distant from an ideal point (i.e., an unobserved feature common to the products).

Application: Portable Telephones

Study Description

A major U.S. communications firm sponsored a research project concerning demand for various portable telephone communication devices. They conducted 499 personal interviews distributed across four large American shopping malls to measure preferences for 12 existing products/brands that were displayed (cf. DeSarbo and Rao 1986). After the respondents were first screened for specific conditions related to family income, household size, head of household's age, educa-

Table 4.1

Portable Telephone Brand Attributes

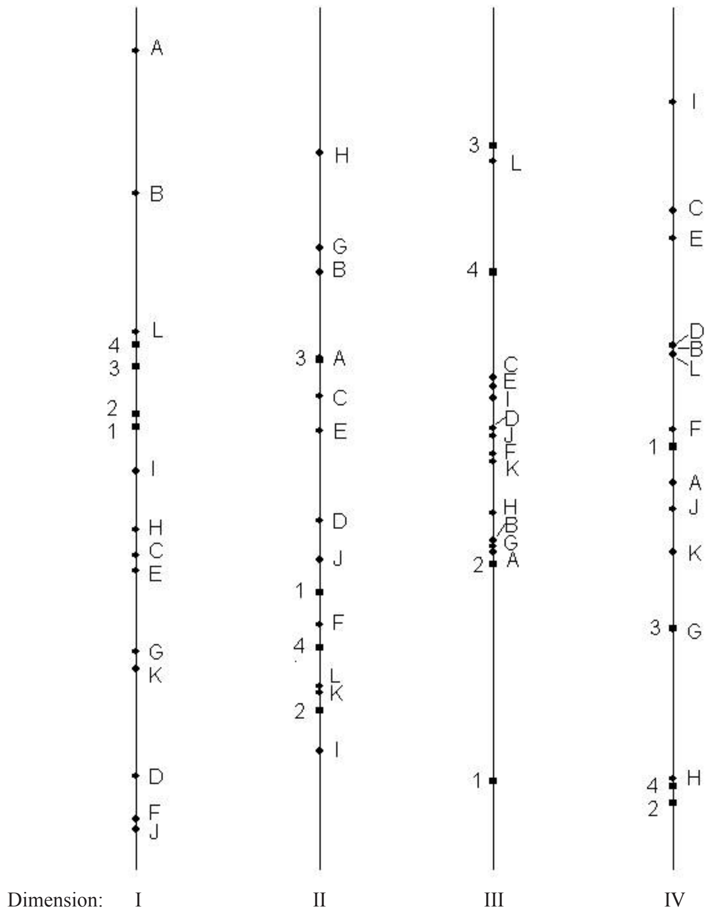
	Attribute Variable								
Brand	A1	A2	A3	A4	A5	A6	A7	A8	A9
A	1	0	0	0	0	300	120	0	0
B	1	0	0	0	0	300	120	1	1
C	1	1	1	1	0	300	220	0	1
D	1	1	1	1	0	300	300	1	1
E	1	1	1	1	0	300	200	0	1
F	1	1	1	1	0	300	240	1	1
G	1	1	0	0	0	300	220	1	0
H	1	1	0	0	0	50	100	0	0
I	1	1	1	3	1	1,000	300	0	0
J	1	1	1	1	1	300	300	1	1
K	1	1	1	1	0	300	250	0	0
L	1	1	1	12	0	1,000	300	0	0

A1 = intercept term
A2 = 0 if receive only, 1 if send and receive
A3 = 0 if no repertory dialing, 1 if repertory dialing
A4 = number of phone numbers in memory
A5 = 0 if no speakerphone option, 1 if speakerphone option
A6 = range in feet
A7 = price in dollars
A8 = 0 lie-down style, 1 if stand-up style
A9 = 0 if cradle style, 1 if walkie-talkie style

tion, and occupation, they were taken to a separate area where the 12 actual products/brands were shown. The interviewer read to the respondents a description of the products/brands and their attributes, and then conducted a demonstration of how each worked. The respondents were then asked to indicate their preference for each of the 12 brands on a 0-to-10 intention- to-buy scale. The same data were used by DeSarbo and Rao (1986) and DeSarbo and Hoffman (1987). Because the data is proprietary, the brands are referred to by letters as brands A to L. We reverse scaled the data to reflect dispreference given the distance model underlying multidimensional unfolding.

The research firm presented the respondents information about eight key product characteristics, which had been selected through prior research and interviews. They are the only characteristics available to us about the product/brands. The characteristics are: whether the phone can send and receive calls (vs. receive only), whether it has repertory dialing, number of phone numbers that can be stored in the phone's memory, whether the phone has a speakerphone option, its wireless transmission range, its price, its style (lie-down versus stand-up), and physical design (cradle style versus walkie-talkie). The attribute values for each of the twelve product/brands are presented in Table 4.1. The firm also provided cost figures for various levels for each of the possible product characteristics. The combination of those levels forms the feasible product space for the new product generation.

Figure 4.4 **Derived Joint Space: Dimensions Obtained from the Proposed Methodology**



Analysis by the Proposed Clusterwise MDU Model

The clusterwise multidimensional unfolding model with brand reparametrization was run for $s = 1 \ldots 5$ segments and $r = 1 \ldots 5$ dimensions, keeping the best of 10 runs for each (s, r) solution set. Table 4.2 displays the various goodness-of-fit values obtained for each of the solutions for both the overlapping and nonoverlapping cases. By inspecting the derived configurations and by considering the VAF statistics, the $S = 4$ segment, $R = 4$ dimensional solution with nonoverlapping segments seems to be the most parsimonious solution given the rate of successive increases in VAF for increasing values of numbers of dimensions and segments, and the fact that overlapping segments add very little additional explanatory power over nonoverlapping segments in this particular application. The resulting goodness-of-fit values for this solution was SSE: 3550.02, VAF: 0.407. The derived joint space is presented in Figure 4.4 by dimension (brands are represented by the letters A–L, and segment-level ideal points are labeled by the integers 1–4 in the figures), and the correlation between the derived brand coordinates and the brand attributes are displayed in Table 4.3.

Table 4.2

Goodness-of-Fit Statistics

		Nonoverlapping		Overlapping	
Clusters	Dimensions	SSE	VAF	SSE	VAF
1	1	5,721.12	0.045	—	—
2	1	5,510.86	0.080	5,472.80	0.086
3	1	5,445.73	0.091	5,391.97	0.099
4	1	5,424.47	0.094	5,346.60	0.107
5	1	5,412.37	0.096	5,296.91	0.115
2	2	5,172.34	0.136	4,925.51	0.177
3	2	5,062.72	0.157	4,687.24	0.217
4	2	4,919.29	0.178	4,580.63	0.235
5	2	4,839.28	0.192	4,459.60	0.255
3	3	4,475.27	0.253	4,254.65	0.289
4	3	4,300.04	0.282	4,087.90	0.317
5	3	4,214.49	0.296	3,831.97	0.360
4	4	3,550.02	0.407	3,513.69	0.411
5	4	3,446.48	0.424	3,404.70	0.429
5	5	3,444.45	0.425	3,227.96	0.444

Table 4.3

Correlation of Brand Attributes with Dimensions: Proposed Methodology

	DIM1	DIM2	DIM3	DIM4
A2	–0.79	–0.38	0.47	0.00
A3	–0.55	–0.82	0.69	0.56
A4	0.19	–0.52	0.90	0.27
A5	–0.22	–0.43	0.09	0.27
A6	0.24	–0.68	0.70	0.58
A7	–0.56	–0.81	0.62	0.41
A8	–0.42	0.13	–0.35	–0.17
A9	–0.37	0.04	0.01	0.38

Dimension 1 seems to identify the traditional (classic) phones. This dimension is negatively correlated with the send and receive capacities, repertory dialing, speakerphone, and stand-up and walkie-talkie style. It clearly discriminates between the brands that have the send and receive (versus the receive only) feature. This also correlates with the associated characteristics that are available only to phones that have the send and receive calls option. DeSarbo and Hoffman (1987) also found a clear distinction between the brands that send and receive versus those that only receive calls. Dimension 2, with option I at one end and option H at the other, discriminates

Table 4.4

Average Brand Preferences by Segment—Proposed Methodology

	Segment 1	Segment 2	Segment 3	Segment 4	Overall Mean
A	5.40	1.85	4.77	2.15	3.14
B	5.02	1.54	4.49	1.59	2.73
C	5.69	2.16	6.13	2.36	3.56
D	5.21	1.88	4.87	1.78	3.00
E	5.07	2.01	5.18	2.00	3.13
F	3.84	2.52	2.01	2.18	2.62
G	5.32	3.16	5.21	2.22	3.62
H	3.34	2.25	4.86	1.69	2.68
I	4.21	1.73	3.06	1.87	2.51
J	4.55	3.12	3.25	1.99	3.08
K	6.14	8.00	3.18	6.27	6.34
L	0.75	1.66	6.96	7.66	4.18

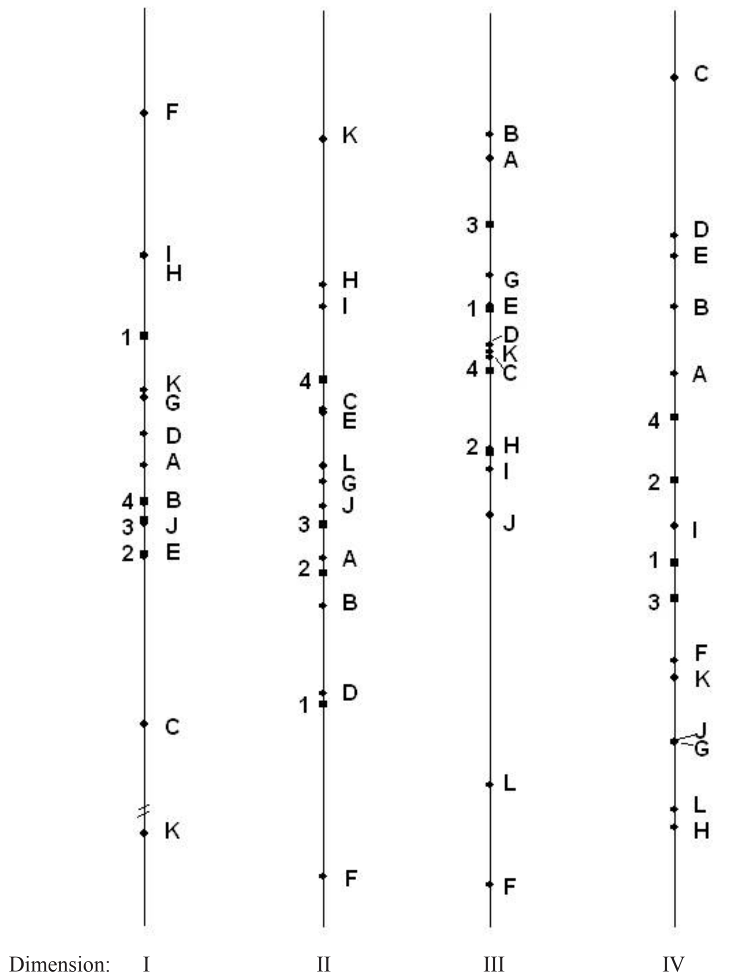
between options mainly on the basis of price. Dimension 3 identifies the luxury features and also includes range and the number of phone numbers storable in memory. Moreover, it correlates with the lie-down style akin to more traditional telephones. There is a clear distinction on this dimension between L and the other brands (L has a 1,000-foot range, 12 numbers storable in memory, and lie-down style). At the other end of this dimension lie brands A and G with a standard 300-foot range, stand-up style, and no numbers storable in memory as it does not have repertory dialing. Finally, Dimension 4 seems to distinguish those brands of portable phones that have a walkie-talkie functioning and longer range from the rest.

As shown in Figure 4.4 and Table 4.4, respondents in all four segments clearly prefer devices that can both send and receive. Segment 1 prefers brand options that are full-featured and not at the maximal price. They prefer options that do not involve the more luxurious features such as longer ranges and high number of options, as shown on Dimension 3. Segment 2 has a clear preference for option K, a cradle-style, lie-down, full-featured, traditional style phone. Segment 3 has a preference for option L, a cradle-style, lie-down phone with the longest range and the maximum number of storable numbers. However, it also prefers option C, which is more affordable with slightly fewer features. Segment 4 has a preference for cradle-style lie-down phones, with full features except for speakerphone. Whereas Segments 2 and 4 are very definite in their preferences, members of Segment 1 and 3 show less variation in their preferences throughout the range of products.

ALSCAL + K-Means Traditional Analysis

To compare with the results of our proposed clusterwise multidimensional unfolding space, a multistep analysis was performed using traditional methods. First, an ALSCAL (Takane, Young, and de Leeuw 1977) MDU joint space configuration in four dimensions was obtained using all the respondents. Respondent preferences were clustered into four nonoverlapping segments using a K-Means algorithm (Shugan and Balachandran 1977; Zufryden 1982). Then, ideal point

Figure 4.5 **Derived Joint Space: Dimensions Obtained from ALSCAL + K-Means**



coordinates were averaged per segment to obtain the segment ideal points. The resulting $R = 4$, $S = 4$ joint space is presented in Figure 4.5. Note, the corresponding goodness-of-fit values for this solution were SSE: 5871.94 and VAF: 0.019, which is clearly dominated by our proposed methodology (SSE: 3550.02, VAF: 0.407). The correlations between the four ALSCAL dimensions of brand coordinates and brand attributes are shown in Table 4.5. In this solution, dimension 1 is closely related to the number of phone numbers storable in memory and range. Dimension 2 discriminates between the classic style phones (lie-down and cradle style) versus those that have either stand up or walkie-talkie style. Dimension 3 relates to the basic phone features: send and receive, repertory dialing, and numbers in memory. Finally, dimension 4 relates to the walkie-talkie feature. Table 4.6 presents the average preference ratings by derived market segment. Segment 1 prefers phones K and L; segment 2 has an overwhelming preference for K; segment 3 prefers phones C and E; and segment 4 prefers K.

Table 4.5

Correlation of Brand Attributes with Dimensions: ALSCAL + K-Means

	DIM1	DIM2	DIM3	DIM4
A2	–0.01	0.25	–0.57	–0.32
A3	–0.21	0.02	–0.54	0.12
A4	–0.79	0.07	–0.58	–0.32
A5	0.16	0.17	–0.15	–0.21
A6	–0.48	0.13	–0.39	–0.16
A7	–0.25	–0.04	–0.52	–0.15
A8	0.29	–0.70	–0.05	–0.08
A9	0.06	–0.57	–0.02	0.56

Table 4.6

Average Brand Preferences by Clusters—ALSCAL + K-Means

	Segment 1	Segment 2	Segment 3	Segment 4	Overall Mean
A	2.220	2.399	5.914	4.323	3.139
B	1.709	2.211	5.334	3.801	2.734
C	2.439	2.654	6.862	5.128	3.561
D	1.930	2.278	5.961	4.322	3.001
E	1.995	2.324	6.076	4.855	3.126
F	2.233	2.284	1.675	5.952	2.624
G	2.616	3.415	4.936	5.305	3.615
H	2.032	2.291	3.683	4.330	2.677
I	1.801	2.205	3.518	4.072	2.507
J	2.026	3.339	4.044	4.058	3.084
K	6.017	8.273	4.910	3.419	6.344
L	7.990	1.352	4.322	1.614	4.175

Comparative Predictive Validation

To test the predictability of the two approaches, both models were reestimated using the product/brands B and E as holdout. Then, using the attributes, predictions were made as to locations for these two brands in the joint space. In the proposed clusterwise MDU model, the impact coefficients were utilized for this purpose given the reparametrization of the brands. In the traditional model using ALSCAL + K-Means, an additional regression step was utilized to estimate these reparametrization coefficients, and then predicted locations were generated. From the predicted locations for these two holdout brands, dispreferences were generated for each brand and compared to the actual holdout dispreferences. The VAF of the clusterwise model for these two brands was 0.445 compared to only 0.030 for the traditional ALSCAL + K-Means approach (SSE of 3082.93

Table 4.7

Comparative Predictive Validation (B and E Brands)

	Proposed Methodology	ALSCAL + K-Means
SSE	3,082.930	5,385.690
VAF	0.445	0.030
Correlation: $\hat{B}$ and cluster dispreferences	0.997	–0.087
Correlation: $\hat{E}$ and cluster dispreferences	0.789	0.498
Correlation: $\hat{B}$, $\hat{E}$ and cluster dispreferences	0.912	0.203

versus 5385.69). The separate correlations between the predicted coordinates for B and E and the original dispreferences were estimated for both models and are presented in Table 4.7. The results clearly indicate the superiority of our proposed clusterwise method.

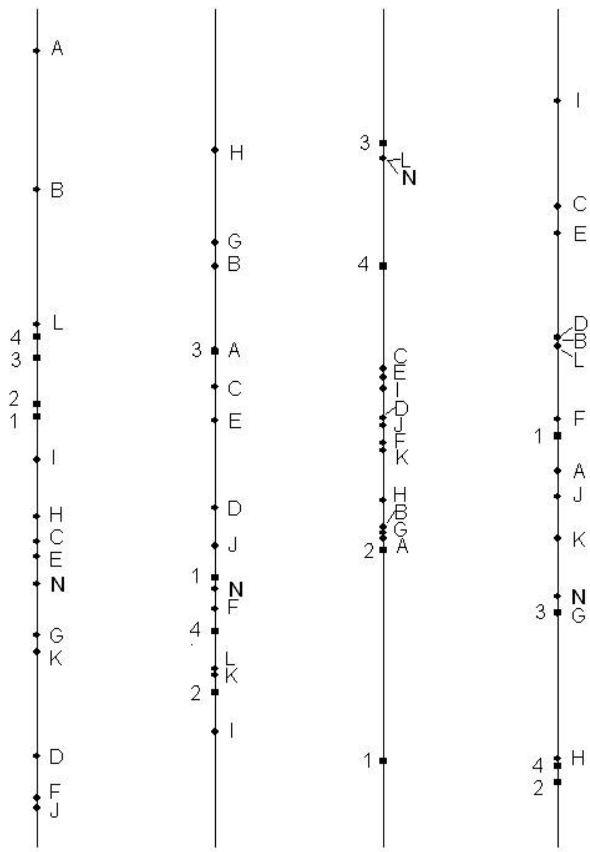
Optimal Portable Telephone Positioning

The solution of the proposed clusterwise multidimensional unfolding model with brand reparametrization was then used as the input for the optimal product positioning model. The communications firm also provided the cost information for varying levels of the product characteristics. A discretized search space was used to enumerate the complete set of possible products that the firm could create. In addition, two more restrictions were put in place reflecting technological constraints at the time. First, a device could have only the option for a greater amount of numbers in memory if it had repertory dialing. Second, it could have only repertory dialing if it had the capacity to both send and receive calls. The whole positioning step was programmed in MATLAB, using approximately 50 lines of code, and ran in less than 3 seconds. The procedure produced 5,210 feasible products. The optimal product devised can be described as follows: send and receive capabilities, repertory dialing with 12 storable numbers, a 750-foot range, is stand-up, and cradle-style. It would be priced at $300. It is represented in the joint space in Figure 4.6 as brand “N”, and the distance and profit information by derived segment is presented in Table 4.8 on pages 98–99.

It can be seen from Figure 4.6 that the new product is positioned to get most of segment 3 and segment 4’s market shares. Both segments 3 and 4 desire phones with full features, longer range, and a high number of storable numbers. The new product offers them exactly that (750 feet, 12 numbers). It also offers less than the current full-featured options (1,000 feet). Members of segment 2 still apparently prefer option K, which is cheaper. However, members of segment 1 seemingly do not develop a clear preference for the new brand. Perhaps it is the high price and the extra features (longer range and numbers) in which they do not see much value.

This positioning results in a higher probability of purchase for members of segments 3 and 4, with 17.3 percent and 18.5 percent predicted choice probability (as in Equation 19) and a reasonable share for the members of segment 2. Moreover, the new product would, in the single-choice model, be the most preferred by two out of the four segments with an estimated profit larger than all other available brands. Thus, one can see how the proposed complete methodology performs the STP process by simultaneously deriving market segments, target-

Figure 4.6 **Derived Joint Space with New Product N from the Proposed Methodology**



ing some subset of these derived market segments, and identifying positions in the derived joint space for optimal new brands.

Discussion

Summary

In this chapter, we have proposed a new clusterwise MDU methodology for the analysis of preference data in which the brand coordinates are a linear function of product characteristics. The proposed method *simultaneously* identifies segments of consumers and relates product characteristics to the derived product space. Moreover, a model for optimal product positioning is proposed to accompany this methodology. A technical description of the models and of estimation procedures has been provided. The optimal product positioning and design literatures have stressed the importance of targeting the proper consumer segments and have focused on reducing the computation difficulties related to the complex formulations of such a problem. The proposed

Table 4.8

Optimal Portable Telephone Brand Positioning Results

		Segment 1 (n = 109)			Segment 2 (n = 156)		
	Unit Profit	Distance	Probability	Expected Profit	Distance	Probability	Expected Profit
A	90.13	17.48	0.087	856.53	21.11	0.068	950.76
B	82.48	15.72	0.097	871.53	24.42	0.058	752.02
C	139.17	19.90	0.077	1161.45	25.11	0.057	1234.49
D	231.32	19.53	0.078	1966.87	21.11	0.068	2440.15
E	134.85	18.24	0.084	1228.12	22.43	0.064	1338.71
F	171.32	19.44	0.078	1463.88	17.99	0.079	2120.78
G	176.3	17.23	0.088	1699.42	19.02	0.075	2064.17
H	59.83	24.05	0.063	413.14	21.56	0.066	618.06
I	221.01	19.19	0.079	1912.58	22.69	0.063	2169.36
J	219.17	21.18	0.072	1719.10	17.76	0.080	2748.83
K	188.97	13.78	0.111	2276.99	7.77	0.184	5417.67
L	231.02	34.59	0.044	1109.19	22.75	0.063	2261.42
New	227.49	36.61	0.042	1032.09	18.95	0.075	2673.80

	Segment 3 (n = 71)			Segment 4 (n = 163)			
	Distance	Probability	Expected Profit	Distance	Probability	Expected Profit	Total Profit
A	21.63	0.059	376.28	21.73	0.059	861.31	3044.88
B	19.08	0.067	390.32	24.89	0.051	688.08	2701.95
C	14.10	0.090	891.14	21.57	0.059	1339.68	4626.77
D	22.78	0.056	916.85	23.36	0.055	2056.05	7379.93
E	14.15	0.090	860.59	19.79	0.064	1414.72	4842.16
F	27.82	0.046	556.01	22.93	0.056	1551.77	5692.44
G	19.93	0.064	798.73	24.23	0.053	1510.97	6073.30
H	17.03	0.075	317.14	23.36	0.055	531.82	1880.16
I	27.48	0.046	726.26	22.20	0.057	2067.83	6876.03
J	24.47	0.052	808.78	21.44	0.059	2123.05	7399.75
K	22.10	0.058	772.05	12.41	0.103	3162.39	11629.11
L	10.15	0.125	2054.29	8.81	0.145	5447.83	10872.73
New	7.33	0.173	2801.18	6.87	0.185	6874.51	13381.58

clusterwise model obtained much better predictive validation than the traditional multistep clustering, multidimensional unfolding, and regression procedures.

The use of the reparametrization of the brand space has also permitted better interpretation of the dimensions and provided the generation of the new optimal product(s) to position in the existing market. This was exemplified using data from a marketing research study concerning intention to buy some 12 brands of existing portable telephone communication devices. The procedure derived four dimensions and four market segments.

The first derived dimension marked the extent to which the proposed devices resemble traditional phones, with an emphasis on the send and receive capacities, storable numbers, and traditional style. It clearly distinguished those portable phones that have the capabilities to both send and receive calls from those that can only receive calls. This was an important attribute, as all derived segments preferred phones that could both send and receive calls. The second dimension was related to price/cost, and the third dimension was related to the more luxurious attributes, including the higher ranges and the number of storable numbers that come along with repertory dialing. Finally, the fourth dimension was related to style (cradle versus walkie-talkie) and range.

We found that the members of segment 1 prefer a product that is full-featured but not necessarily with longer range and a high number of storable numbers. Members of segments 3 and 4 both prefer full-featured phones, but members of segment 4 prefer the phones that have longer range and more storable numbers. However, members of segments 1 and 3 seem to have less strongly defined preferences, whereas they are much more strongly defined for members of segments 2 and 4.

Through our optimal brand positioning analysis, a brand that would maximize a profit function was identified. It was primarily targeted to the members of segments 3 and 4 for whom it clearly became the most preferred option. It would also appeal somewhat to the members of segment 2. Moreover, because of this phone's expensive price and its more luxurious features, members of segment 1 would still likely prefer their original favorites. This led to a probabilistic profit estimate that significantly surpassed that of other product/brands available.

Directions for Future Research

There exist a number of rich possibilities for future research in this area. First, models in the product positioning literature have been deficient when it comes to the behavioral assumptions of the underlying choice model. There is a need for models that can account for the perceived similarity of products. Krumhansl's density-distance model (1978) has been used by DeSarbo and Manrai (1992) and, if extended to the clusterwise case, could be used to generate new brand positions and account for the perceptual similarities between the brands in the marketplace. There is also a need to adapt or derive procedures to create optimal products in the case where product characteristics are a combination of discrete and continuous variables. Our procedure could easily enumerate the complete list of feasible products that could be introduced into the market, partly because of the discretization of the attributes that was imposed by the nature of the cost information that was provided. However, there is a need to find a procedure that can accommodate such different attribute and cost structures. Simulated optimizations can be performed for multiple brands with multiple manufacturers. It would thus be interesting to perform such optimizations for a specific manufacturer interested in producing more than one new brand, and/or attempting to minimize cannibalization with existing brand(s). Finally, generalizing the proposed methodology to the analysis of nonmetric and three-way data would prove advantageous in extending the range of applications for this approach.

References

Albers, S. 1979. “An Extended Algorithm for Optimal Product Positioning.” *European Journal of Operational Research* 3 (May), 222–31.

Albers, S., and K. Brockhoff. 1977. “A Procedure for New Product Positioning in an Attribute Space.” *European Journal of Operational Research* 1 (July), 230–38.

Andrews, K.S., and A.K. Manrai. 1999. “MDS Maps for Product Attributes and Market Response: An Application to Scanner Panel Data.” *Marketing Science* 18 (4), 584–604.

Arabie, P., J.D. Carroll, W.S. DeSarbo, and Y. Wind. 1981. “Overlapping Clustering: A New Method for Product Positioning.” *Journal of Marketing Research* 18 (August), 310–17.

Bachem, A., and H. Simon. 1981. “A Product Positioning Model with Costs and Prices.” *European Journal of Operational Research* 7 (August), 362–70.

Baier, D., and W. Gaul. 1999. “Optimal Product Positioning Based on Paired Comparison Data.” *Journal of Econometrics* 89 (November), 365–92.

Batsell, R.R., and J.C. Polking. 1985. “A New Class of Market Share Models.” *Marketing Science* 4 (3), 177–98.

Böckenholt, U., and I. Böckenholt. 1991. “Constrained Latent Class Analysis: Simultaneous Classification and Scaling of Discrete Choice Data.” *Psychometrika* 56 (December), 699–716.

Bordley, R. 2003. “Determining the Appropriate Depth and Breath of a Firm’s Product Portfolio.” *Journal of Marketing Research* 40 (February), 39–53.

Chintagunta, P.K. 1994. “Heterogeneous Logit Model Implications for Brand Positioning.” *Journal of Marketing Research* 31 (May), 304–11.

Choi, S. C., W.S. DeSarbo, and P.T. Harker. 1990. “Product Positioning Under Price Competition.” *Management Science* 36 (February), 175–99.

Coombs, C.H. 1964. *A Theory of Data.* New York: John Wiley & Sons.

Cooper, L.G. 1988. “Competitive Maps: The Structure Underlying Asymmetric Cross Elasticities.” *Management Science* 34 (June), 707–23.

Day, G.S. 1984. *Strategic Market Planning: The Pursuit of Competitive Advantage.* St. Paul, MN: West Publishing.

Dempster, A.P., M.N. Laird, and D.B. Rubin. 1977. “Maximum Likelihood Estimation from Incomplete Data via the E-M Algorithm.” *Journal of the Royal Statistical Society, Series B* 39 (1), 1–38.

DeSarbo, W.S., A.S. Atalay, D. LeBaron, and S. J. Blanchard. 2008. “Estimating Multiple Consumer Segment Ideal Points from Context Dependent Survey Preference Data.” *Journal of Consumer Research,* 35 (June), 142–53.

DeSarbo, W.S., and J.D. Carroll. 1985. “Three-Way Metric Unfolding via Alternating Weighted Least-Squares.” *Psychometrika* 50 (September), 275–300.

DeSarbo, W.S., and C.F. DeSarbo. 2002. “A Generalized Normative Segmentation Methodology Employing Conjoint Analysis.” In C*onjoint Analysis: Methods and Applications*, pp. 473–504, ed. A. Gustafsson, F. Herrmann, and F. Huber. 3rd ed. Berlin: Springer-Verlag.

DeSarbo, W.S., R. Grewal, and C. Scott. 2008. “A Clusterwise Bilinear Multidimensional Scaling Methodology for Simultaneous Segmentation and Positioning Analyses.” *Journal of Marketing Research*, 45 (3), 280–92.

DeSarbo, W.S., and D.L. Hoffman. 1987. “Constructing MDS Joint Spaces from Binary Choice Data: A Multidimensional Unfolding Threshold Model for Marketing Research.” *Journal of Marketing Research* 24 (February), 40–54.

DeSarbo, W.S., D.J. Howard, and K. Jedidi. 1991. “MULTICLUS: A New Method for Simultaneously Performing Multidimensional Scaling and Cluster Analysis.” *Psychometrika* 56 (March), 121–36.

DeSarbo, W.S., K. Jedidi, K. Cool, and D. Schendel. 1990. “Simultaneous Multidimensional Unfolding and Cluster Analysis: An Investigation of Strategic Groups.” *Marketing Letters* 2, 129–46.

DeSarbo, W.S., and A.K. Manrai. 1992. “A New Multidimensional Scaling Methodology for the Analysis of Asymmetric Proximity Data in Marketing Research.” *Marketing Science* 11 (1), 1–20.

DeSarbo, W.S., A.K. Manrai, and L.A. Manrai. 1994. “Latent Class Multidimensional Scaling Approaches: A Review of the Recent Developments in the Marketing and Psychometric Literature.” In *Advanced Methods of Marketing Research*, ed. R.P. Bagozzi. Oxford, UK: Blackwell.

DeSarbo, W.S., R.L. Oliver, and A. Rangaswamy. 1989. “A Simulated Annealing Methodology for Clusterwise Linear Regression.” *Psychometrika* 54 (September), 707–36.

DeSarbo, W.S., V. Ramaswamy, and R. Chatterjee. 1995. "Analyzing Constant-Sum Multiple Criterion Data: A Segment-Level Approach." *Journal of Marketing Research* 32 (May), 222–32.

DeSarbo, W.S., V. Ramaswamy, and P. Lenk.1993. "A Latent Class Procedure for the Structural Analysis of Two-Way Compositional Data." *Journal of Classification* 10 (December), 159–93.

DeSarbo, W.S., and V.R. Rao. 1986. "GENFOLD2: A Constrained Unfolding Model for Product Positioning." *Marketing Science* 5 (1), 1–19.

De Soete, G., and W.J. Heiser. 1993. "A Latent Class Unfolding Model for Analyzing Single Stimulus Preference Ratings." *Psychometrika* 58 (December), 545–65.

De Soete, G., and S. Winsberg. 1993. "A Latent Class Vector Model for Preference Ratings." *Journal of Classification* 10 (December), 192–218.

Eliashberg, J., and A.K. Manrai. 1992. "Optimal Positioning of New Product-Concepts: Some Analytical Implications and Empirical Results." *European Journal of Operational Research* 63 (December), 376–97.

Fraser, C., and John W. Bradford. 1984. "Competitive Market Structure Analysis: A Reply." *Journal of Consumer Research* 11 (December), 842–47.

Gavish, B., D. Horsky, and K. Srikanth. 1983. "An Approach to the Optimal Positioning of a New Product." *Management Science* 29 (February), 1277–97.

Green, P.E., J.D. Carroll, and S.M. Goldberg. 1981. "A General Approach to Product Design Optimization Via Conjoint Analysis." *Journal of Marketing* 45 (Summer), 17–37.

Green, P.E., and A.M. Krieger. 1985. "Models and Heuristics for Product Line Selection." *Marketing Science* 4 (1), 1–19.

———. 1989. "Recent Contributions to Optimal Product Positioning and Buyer Segmentation." *European Journal of Operational Research* 41 (July), 127–41.

Gruca, T.S., and B.R. Klemz. 2003. "Optimal New Product Positioning: A Genetic Algorithm Approach." *European Journal of Operational Research* 146 (May), 621–33.

Hansen, P., B. Jaumard, C. Meyer, and J.-F. Thisse. 1998. "New Algorithms for Product Positioning." *European Journal of Operational Research* 104 (January), 154–74.

Hauser, J.R., and P. Simmie. 1981. "Profit Maximizing Perceptual Positions: An Integrated Theory for the Selection of Product Features and Price." *Management Science* 27 (January), 33–56.

Holman, E.W. 1972. "The Relation Between Hierarchical and Euclidean Models for Psychological Distances." *Psychometrika* 37 (December), 417–23.

Kaul, A., and V.R. Rao (1995), "Research for Product Positioning and Design Decisions: An Integrative Review." *International Journal of Research in Marketing* 12 (November), 293–320.

Kohli, R., and R. Krishnamurti. 1987. "A Heuristic Approach to Product Design." *Management Science* 33 (December), 1523–33.

Kotler, P. 1997. *Marketing Management.* 7th ed. Englewood Cliffs, NJ: Prentice Hall.

Kotler, P., and K.L. Keller. 2006. *Marketing Management.* 12th ed. Englewood Cliffs, NJ: Prentice Hall.

Krumhansl, C. 1978. "Concerning the Applicability of Geometric Models to Similarity Data: The Interrelationship Between Similarity and Spatial Density." *Psychological Review* 85 (5), 445–63.

Lattin, J.M., and L. McAlister. 1985. "Using a Variety-Seeking Model to Identify Substitute and Complementary Relationships Among Competing Products." *Journal of Marketing Research* 22 (August), 330–39.

Lilien, G.L., and A. Rangaswamy. 2004. *Marketing Engineering: Computer Assisted Marketing Analysis and Planning.* Boston, MA: Addison Wesley.

Luce, R.D. 1959. *Individual Choice Behavior: A Theoritical Analysis.* New York: John Wiley & Sons.

McBride, R.D., and F.S. Zufryden. 1988. "An Integer Programming Approach to the Optimal Product Line Selection Problem." *Marketing Science* 7 (2), 126–40.

McLachlan, G.J, and T. Krishnan. 1997. *The EM Algorithm and Extensions.* New York: John Wiley & Sons.

McLachlan, G.J., and D. Peel. 2000. *Finite Mixture Models.* New York: John Wiley & Sons.

Meyer, R., and E.J. Johnson. 1995. "Empirical Generalizations in the Modeling of Consumer Choice." *Marketing Science* 14 (3), 180–89.

Michalek, J.J., F.M. Feinberg, and P.Y. Papalambros. 2005. "Linking Marketing and Engineering Product Design Decisions via Analytical Target Cascading." *Journal of Product Innovation Management* 22 (January), 42–62.

Moore, W.L., and R.S. Winer. 1987. "A Panel-Data Based Method for Merging Joint Space and Market Response Function Estimation." *Marketing Science* 6 (1), 25–42.

Myers, J.H. 1996. *Segmentation and Positioning for Strategic Marketing Decisions.* Chicago: AMA Publishing.

Nocedal, J., and S. Wright. 2006. *Numerical Optimization.* 2nd ed. New York: Springer.

Pessemier, E.A., P. Burger, R. Teach, and D. Tigert. 1971. "Using Laboratory Preference Scales to Predict Consumer Brand Preferences." *Management Science* 17 (February), 371–85.

Reutterer T., and M. Natter. 2000, "Segmentation Based Competitive Analysis with MULTICLUS and Topology Representing Networks." *Computers and Operations Research*, 27, 1212–27.

Schmalensee, R., and J.-F. Thisse. 1988. "Perceptual Maps and the Optimal Location of New Products: An Integrative Essay." *International Journal of Research in Marketing* 5 (4), 225–49.

Shi, L., S. Olafsson, and Q. Chen. 2001. "An Optimization Framework for Product Design." *Management Science* 47 (December), 1681–92.

Shocker, A.D., and V. Srinivasan. 1974. "A Consumer-Based Methodology for the Identification of New Product Ideas." *Management Science* 20 (February), 921–38.

———. 1979. "Multiattribute Approaches for Product Concept Evaluation and Generation: A Critical Review." *Journal of Marketing Research* 16 (May), 159–80.

Shugan, S.M., and V. Balachandran. 1977. "A Mathematical Programming Model for Optimal Product Line Structuring." *Discussion Paper 265*. Evanston, IL: Northwestern University, The Center for Mathematical Studies in Economics and Management Science.

Slater, P. 1960. "The Analysis of Personal Preferences." *British Journal of Statistical Psychology* 30, 119–35.

Spath, H. 1979. "Algorithm 39 Clusterwise Linear Regression." *Computing* 22 (December), 367–73.

———. 1981. "Correction to Algorithm 39 Clusterwise Linear Regression." *Computing* 26 (September), 275.

———. 1982. "Fast Algorithm for Clusterwise Linear Regression." *Computing* 29 (June), 175–81.

Sudharshan, D., J.H. May, and A.D. Shocker. 1987. "A Simulation Comparison of Methods for New Product Location." *Marketing Science* 6 (2), 182–201.

Takane, Y., F.W. Young, and J. de Leeuw. 1977. "Nonmetric Individual Differences Multidimensional Scaling: An Alternating Least Squares Method with Optimal Scaling Features." *Psychometrika* 42 (March), 8–67.

Thakur, L.S., S.K. Nair, K.-W. Wen, and P. Tarasewich. 2000. "A New Model and Solution Method for Product Line Design with Pricing." *Journal of the Operational Research Society* 51 (January), 90–101.

Tucker, L.R. 1960. "Intra-Individual and Inter-Individual Multidimensionality." In *Psychological Scaling*, ed. H. Gulliksen and S. Messick. New York: John Wiley & Sons.

Wedel, M., and W.S. DeSarbo. 1996. "An Exponential-Family Multidimensional Scaling Mixture Methodology." *Journal of Business & Economic Statistics* 14 (October), 447–59.

Wedel, M., and W.A. Kamakura. 2000. *Market Segmentation: Conceptual and Methodological Foundations.* Boston: Kluwer.

Wedel, M., and C. Kistemaker. 1989. "Consumer Benefit Segmentation Using CLusterwise Linear Regression." *International Journal of Research in Marketing* 6 (1), 45–59.

Wind, Y. 1980, *Product Policy.* Boston: Addison Wesley Publishing.

Wind, Y., and T.S. Robertson. 1983. "Marketing Strategy: New Directions for Theory and Research." *Journal of Marketing* 47 (Spring), 12–25.

———. 1982. "Product Line Optimization by Integer Programming." In *Annual Meeting of ORSA/TIMS,* San Diego, CA.

Zufryden, F.S. 1979. "ZIPMAP—A Zero-One Integer Programming Model for Market Segmentation and Product Positioning." *Journal of the Operational Research Society* 30 (1), 63–70.