MDS Maps for Product Attributes and Market Response: An Application to Scanner Panel Data

Rick L. Andrews • Ajay K. Manrai
Department of Business Administration, University of Delaware, Newark, Delaware 19716
andrewsr@udel.edu • manraia@udel.edu

Abstract
There is theoretical and empirical evidence that consumers have limited cognitive resources and thus cannot maintain direct preferences for each choice alternative on the store shelves. Instead, they likely form their overall preferences for choice alternatives by evaluating the attributes describing each item. Rather than mapping the locations of and preferences for all choice alternatives in a multidimensional space, as is the current practice in marketing research, it is insightful to map the locations of and preferences for the attributes consumers use to evaluate the choice alternatives. The model proposed in this study unifies latent class preference models (choice models or conjoint models) with latent class multidimensional scaling (MDS) analysis. Dimensional restrictions are imposed on latent class preference models such that the locations of attribute levels and market response parameters can be mapped in reduced-dimension spaces. Interactions between attributes can be graphically examined, which is not feasible with the traditional MDS approach. Also, the effects of price reductions and promotions on the locations of attribute levels can be graphically examined. An empirical application with scanner panel data shows the capabilities and limitations of the proposed model. In addition to the managerial insights provided by the model, it is also much more parsimonious than existing methods, and it forecasts holdout choices significantly better. In the empirical application, a model with two-dimensional attribute maps has 50 fewer parameters than the best unrestricted latent class choice model, yet the fit is comparable. The predictive performance of our model is shown to be superior to that of latent class MDS approaches and latent class conjoint approaches.

(brand choice; choice models; marketing mix; scaling methods; segmentation research)
Introduction

Literature on new product development is replete with analytical methods that summarize consumer judgment data to produce insights on optimal attribute configurations, brand positioning, and market segmentation (Urban and Hauser 1993). Recent research on preference analysis has produced latent class conjoint analysis models (e.g., DeSarbo et al. 1992) and latent class choice models (e.g., Kamakura and Russell 1989) that estimate segment-specific preferences for product attributes and/or responses to marketing mix variables. Recent research on latent class MDS has produced models that map brand positions and preferences of segments in a joint space (Böckenholt and Böckenholt 1991, DeSarbo et al. 1991, DeSoete and Heiser 1993, DeSoete and Winsberg 1993, Chintagunta 1994, Wedel and DeSarbo 1996).

This study develops a model that unifies latent class preference models with latent class MDS. The model imposes dimensional restrictions on a latent class choice or conjoint model so that the preferences can be depicted in reduced-dimension spaces. Unlike existing models, the proposed model produces a separate map for each product attribute, including a graphical examination of interactions between attributes, as well as a map showing responses to marketing mix variables. As suggested by Wedel and DeSarbo (1996), the entire family of exponential distributions can form the basis of such models, allowing many possible types of response data (e.g., rating scale measures, choices, pick-any data, purchase frequencies, duration-type data, etc.). We provide an empirical application using scanner panel choice data for stock keeping units (SKUs, cf. Fader and Hardie 1996).

Most, if not all, research in marketing has focused on mapping choice alternatives. From a managerial standpoint, there are advantages to mapping product attributes rather than choice alternatives. First, from a consumer behavior perspective, the attribute level is a more appropriate unit of analysis than the choice alternative. Though consumers choose from a large assortment of SKUs on the store shelves, Fader and Hardie (1996) assert that consumers have limited cognitive resources and thus do not form preferences for each individual SKU. With 50–150 SKUs on the shelves in many product categories, consumers may not be able to maintain well-defined preferences for each one. Instead, they probably form preferences for the attributes describing each item (brand name, size, formulation, etc.) to derive their overall preference for an SKU, which is a much more manageable task. Fader and Hardie cite theoretical justification for this assertion from both economics and psychology and provide empirical justification in their own study. Given the important role of attributes in the formation of preferences, it may be more meaningful to map these attributes rather than the choice alternatives.

Second, the attribute level maps are easier to interpret than maps containing all choice alternatives. Existing MDS approaches map the locations of all choice alternatives in a single space, possibly resulting in a space that is very crowded and difficult to interpret. In a later application involving 57 SKUs of fabric softeners, a map produced by existing MDS methods would show all 57 alternatives in a single space, along with segment-specific preference vectors. Our model maps each attribute in a separate space, resulting in maps that are simpler to interpret. Instead of a single map containing all 57 SKUs, our method produces five maps (one for each attribute—brand name, size, product form, and formulation—and one for the marketing mix). No map contains more than 10 attribute levels and the preference vectors. Graphical examination of an interaction between two attributes does not require additional maps—it can be incorporated into one of the maps of main effects. Given that it is not at all uncommon to find product categories with many alternatives on the store shelves, the proposed method could be applicable in many studies.

Third, the approach is very parsimonious. The existing approach for mapping the locations of and preferences for discrete product attributes, known as simultaneous reparameterization, requires that restrictions be placed on the latent class MDS model. Frequently, these restrictions are rejected by the data (cf. Wedel and DeSarbo 1996), as we later demonstrate empirically. Our model instead imposes restrictions on a latent class preference model. In contrast, we have not found a single instance in which the restrictions required by our model are rejected by the data. The restrictions we impose typically have little effect on the
fit of the model, yet they often reduce parameter requirements significantly. In a later empirical application, we show that a model with two-dimensional maps for all attributes uses 50 fewer parameters than the best unrestricted latent class model, yet fits as well and forecasts better.

Finally, the proposed attribute-level maps provide unique and compelling managerial insights not obvious from plots of part worths. The maps provide all the preference information available from part worths, but they can provide managerial insights beyond those provided by part worths. Attribute levels located in close proximity have very similar preferences and probably need to be differentiated from other levels. Interactions between attributes can be shown graphically. Heterogeneous responses to marketing mix variables such as price, aisle display, and store feature advertising can be shown graphically. We are not aware of any MDS approach that maps marketing mix effects. Later, we extend the proposed model to show how price reductions and promotional activities such as aisle displays and store feature advertisements affect the mapped locations of brand names and other attributes.

This paper first reviews recent developments in latent class preference analysis and MDS that are relevant to the current study. Following the review of literature, the formulation of the proposed model is presented. Finally, we present an application to scanner panel choice data for SKUs of fabric softeners.

**Relevant Literature**

Figure 1 shows the positioning of the proposed model in relation to existing models and serves as an outline for this review of relevant literature.

Conjoint analysis models that simultaneously segment consumers and estimate segment-level conjoint part worths are now quite numerous (DeSarbo et al. 1992, Green and Helsen 1989, Green and Krieger 1991, Hagerty 1985, Kamakura 1988, Kamakura et al. 1994, Ogawa 1987, Wedel and Kistemaker 1989, Wedel and Steenkamp 1989). Of particular interest in this study is the latent class approach of DeSarbo et al. (1992). In their model, the segments and the utility function part worths are estimated simultaneously using a model involving mixtures of multivariate normal distributions.

The simulation study by Vriens et al. (1996) compares nine metric conjoint segmentation methods and concludes that the latent class model and a fuzzy clusterwise regression procedure (Wedel and Steenkamp 1989) generally outperform other procedures with respect to coefficient and segment membership recovery.

Kamakura et al. (1994) develop a similar approach based on logit in which the mixing weights are parameterized as a function of consumer descriptor variables (see also Gupta and Chintagunta 1994 and Wedel and DeSarbo 1996). Fader and Hardie (1996) reparameterize the choice alternative-specific constants of a logit model into part worths in the context of scanner panel data for SKUs of fabric softeners, bridging the gap between logit analysis and conjoint analysis. Our model builds on the contribution of Fader and Hardie by depicting preferences in reduced-dimension joint spaces. Kamakura et al. (1994) and Fader and Hardie (1996) are examples of nonmetric or choice-based conjoint analysis.

In the MDS literature, latent class models have recently been developed to represent locations of choice alternatives and segment preferences in a joint space. DeSarbo et al. (1991) develop a vector model for normally distributed ratings data, while DeSoete and Heiser (1993) formulate an ideal point model for normally distributed data. Vector models are based on the notion that more of a desirable dimension is better, but ideal point models suggest that there is an ideal level of any dimension. Böckenholt and Böckenholt (1991) estimate vector and ideal point models for binary (pick any or pick any/J) data. Chintagunta (1994) develops a vector model for multinomial choice data. DeSarbo et al. (1994) and Wedel and DeSarbo (1996) review the literature on latent class MDS models.

To improve the interpretability of the derived dimensions of a joint space, researchers have used so-called property fitting methods (Böckenholt and Böckenholt 1991, Carroll et al. 1989, Wedel and DeSarbo 1996). For example, the estimated MDS coordinates of choice alternatives can be regressed on a design matrix describing the alternatives on a group of attributes. The resulting weights for attributes can be used to plot attribute vectors in the same joint space as the choice alternatives and preferences.

Others have proposed reparameterizing coordinates...
of choice alternatives into attribute coordinates as part of an integrated MDS analysis (DeSarbo et al. 1982, DeSarbo and Rao 1986, DeSoete and Heiser 1993, DeSoete and Winsberg 1993, Wedel and DeSarbo 1996), which is known as *simultaneous* reparameterization. This procedure involves computing the coordinates of choice alternatives as the product of a design matrix describing the alternatives on conjoint attributes and a matrix of estimated attribute locations. Thus, simultaneous reparameterization imposes constraints on the conventional stochastic mixture MDS model and requires the estimation of attribute coordinates instead of the coordinates of choice alternatives. The end result is a single map that contains all attribute levels and segment preferences (and perhaps choice alternatives if desired). We note that such maps may be crowded and difficult to interpret, as we demonstrate in a later empirical application.

More importantly, the restrictions imposed by the simultaneous reparameterization may not be consistent with the data. The study by Wedel and DeSarbo (1996) found that the restrictions were not consistent with the data and relied on conventional property fitting methods instead. DeSoete and Heiser (1993) demonstrate the reparameterization with synthetic data but not data from an actual application. The simultaneous reparameterization is rejected in our empirical application as well. Though the studies by DeSarbo et al. (1982) and DeSoete and Winsberg (1993) did find that simultaneous reparameterization restrictions are consistent with the data, the technique is sometimes
not an acceptable way of depicting preference information in reduced-dimension spaces. Our proposed model is intended to remedy this problem.

In our model, dimensional restrictions imposed on latent class preference models allow the preferences to be shown graphically in reduced-dimension spaces, as in an MDS model. The result is a separate map for each attribute, with each map containing the locations of the attribute levels and the segment-specific preferences for those attribute level locations. As we demonstrate with simulated data in the next section, an interaction between two attributes can be plotted in a single map. Heterogeneous market responses to price, aisle display, and store feature advertising can also be mapped in a reduced-dimension space. As our empirical analysis demonstrates, these dimensional restrictions can result in a significant reduction in parameters without a corresponding loss of fit. The forecasting performance of our approach is shown to be superior to that of all existing approaches as well.

An Integrated Latent Class Preference-MDS Model

We first present a latent class preference model based on that of DeSarbo et al. (1992) and then impose dimensional restrictions on the model so that the preferences can be graphically depicted in reduced-dimension spaces. Let

\[ i = 1, \ldots, I, \text{ consumers;} \]
\[ j = 1, \ldots, J, \text{ choice alternatives;} \]
\[ k = 1, \ldots, K, \text{ derived segments;} \]
\[ l = 1, \ldots, L, \text{ variables (continuous or discrete) describing the alternatives;} \]
\[ Y_{ij} = \text{ the response to choice alternative } j \text{ by consumer } i; \]
\[ Y_i = \text{ the } J \times 1 \text{ column vector of responses by consumer } i; \]
\[ X_{lj} = \text{ the value of the } l\text{th variable for the } j\text{th alternative;} \]
\[ X = (X_{lj}), \text{ which is } J \times L; \]
\[ \beta_{lk} = \text{ the estimated response to the } l\text{th variable for the } k\text{th segment;} \]
\[ \beta = ((\beta_{lk})) \text{, which is } L \times K; \]
\[ \Sigma_k = \text{ a } J \times J \text{ covariance matrix estimated for segment } k; \text{ and} \]
\[ \Sigma = (\Sigma_1, \Sigma_2, \ldots, \Sigma_K). \]

Restrictions to Reduce Dimensionality of Latent Class Preference Models

The model presented thus far estimates \( K \) sets of segment-specific preferences. The contribution of our model is that it imposes dimensional restrictions on the latent class preference model above so that preferences
may be graphically depicted in reduced-dimension spaces.

We first describe maps for discrete attributes such as brand name and later extend the maps to marketing mix attributes, which may be continuous (e.g., price) or discrete (e.g., store feature). Let
\[ f = 1, \ldots, F \] attributes;
\[ L_f = \text{the number of levels for attribute } f; \]
\[ \beta_k^f = \text{the } L_f \times 1 \text{ vector of preferences for attribute } f, \text{ segment } k; \]
\[ M' = \text{the dimensions of the space into which } \beta_k^f \text{ is to be mapped (to be determined by the data);} \]
\[ A' = \text{the } L_f \times M' \text{ vector of household invariant locations of } \beta_k^f \text{ in the space; and} \]
\[ w_k^f = \text{the } M' \times 1 \text{ vector of importance weights which segment } k \text{ attaches to the } M' \text{ dimensions}. \]

The estimated preference for an attribute \( f \) for segment \( k \), \( \beta_k^f \), can be decomposed into the location of the corresponding attribute levels in multidimensional space, \( A' \), and the segment-specific importance weights consumers attach to those locations, \( w_k^f \) (cf. Chintagunta 1994, Elrod 1988). Specifically, assume that \( \beta_k^f \) is a linear function of the levels’ locations within an \( M' \)-dimensional map such that
\[ \beta_k^f = A'w_k^f. \tag{5} \]

For an attribute with \( L_f \) levels, only \( L_f - 1 \) part worths are identified per segment, so one of the levels must be normalized to zero. Likewise, one attribute level must be normalized to the origin of the map, and hence one row of \( A' \) will contain only zeros (it does not matter which row contains the zeros). This is analogous to arbitrarily setting one element of \( \beta_k^f \) to zero in a dummy variable regression, which is required for identification. Restrictions are also imposed on the importance weights \( w_k^f \) as explained below.

Consider the simple example of an unrestricted five-segment latent class preference model that has only one attribute with four levels. Fifteen part worths (3 for each segment) would be estimated for such a model. The unrestricted model can be cast into the framework of model (5) above. The \( A' \) matrix for the unrestricted latent class model would be \( 4 \times 5 \), reflecting a five-dimensional solution. (For an unrestricted latent class model, the number of dimensions is the same as the number of segments; for our restricted latent class model, the number of dimensions is less than the number of segments). Since one row of \( A' \) contains all zeros for identification, there are 15 location parameters in the matrix. It is important to note that the number of parameters used for an attribute can never be greater than the number required for the unrestricted latent class preference model. In this example, it would not be possible to identify any importance weights since all 15 of the allowable parameters are used in the location matrix. Thus, the segments must have weight vectors \( w_1^f = [1,0,0,0,0], w_2^f = [0,1,0,0,0], w_3^f = [0,0,1,0,0], w_4^f = [0,0,0,1,0], \) and \( w_5^f = [0,0,0,0,1] \). This setup would produce exactly the same solution as an unrestricted five-segment latent class model (DeSarbo et al. 1992).

Now if we impose a two-dimensional map restriction on the attribute, the location matrix \( A' \) is \( 4 \times 2 \) and contains 6 location parameters (one row again contains all zeroes). The dimensional restrictions are imposed such that the two dimensions of the map represent the first two segments’ preferences. The remaining three segments’ preferences are linear combinations of the first two segments’ preferences. The segments would have weight vectors \( w_1^f = [1,0], w_2^f = [0,1], w_3^f = [w_{31}, w_{32}], w_4^f = [w_{41}, w_{42}], \) and \( w_5^f = [w_{51}, w_{52}] \). The first dimension of the map represents the preferences of segment 1, while the second dimension of the map represents the preferences of segment 2. The other segments’ preferences are then linear combinations of the first two segments’ preferences. Rather than estimate more part worths for the other three segments (as in an unrestricted latent class preference model), we simply estimate importance weights that are applied to the first two segments’ part worths. The two-dimensional model contains 12 parameters (6 location parameters and 6 weight parameters) rather than the 15 parameters required for the unrestricted five-segment latent class model, thus requiring 3 fewer parameters for that attribute.

If we impose a one-dimensional map restriction, the \( A' \) matrix is \( 4 \times 1 \) and contains 3 location parameters. The first segment’s preferences are estimated directly, but the other four segments’ preferences are weighted functions of the first segment’s preferences. The one dimension of the map therefore represents the preferences of segment 1. The segments would have weight
vectors \( w_1 = 1, w_2 = w_3 = w_4 = w_5 = w_6 \), for a total of 7 parameters.

Note that three- and four-dimensional maps for this example would not constitute a reduction in parameters. The three-dimensional version would require 15 parameters (9 locations and 6 preference weights), while the four-dimensional version would require 16 (12 locations and 4 preference weights). It is not possible to identify more than the 15 parameters required by the unrestricted latent class preference model. Thus, for a four-level attribute with five preference segments, one would only consider two- and one-dimensional maps.

We have not encountered any data for which three- or higher-dimensional maps were needed; two-dimensional map restrictions are seldom, if ever, rejected by the data. If two-dimensional maps are not rejected by the data, then one has no need to consider higher-dimensional maps. One-dimensional maps impose quite severe restrictions and are often rejected by the data even when two-dimensional restrictions are not rejected. In addition, the managerial insights provided by one-dimensional maps are much less interesting.

For a map to be identified, the number of parameters required for any given attribute must be less than or equal to the number of parameters required for that attribute in the unrestricted latent class model. For example, for a latent class model with five segments, an attribute with four levels would require 5 \((4 - 1) = 15\) parameters. The total number of parameters in \( M_k \) \( w_k \) for all segments \( k \) must therefore be less than or equal to 15. More generally, if \( M' \) is the number of dimensions for factor \( j \), \( K \) is the number of segments, and \( L_j \) is the number of levels, the identifiability restriction can be expressed as:

\[
\text{number of location parameters} + \text{number of preference weight parameters} \leq \text{number of parameters in unrestricted model}
\]

or

\[
M' (L_j - 1) + M' (K - M') \leq KL_j - 1).
\] (6)

If a one-dimensional map is desired (\( M' = 1 \)), any number of segments (\( K \geq 1 \)) can be used with any type of attribute (\( L_j \geq 2 \)). There are no unidentified combinations of \( K \) and \( L_j \) when \( M' = 1 \). If a two-dimensional map is desired, there must be at least two segments (\( K \geq 2 \)) regardless of the number of levels for the attribute. If there are only two attribute levels (\( L_j = 2 \)), only a two-segment map (\( K = 2 \)) is identified; models with more segments (\( K \geq 3 \)) are not identified when \( L_j = 2 \). Thus, two-dimensional maps are identified for \( K \geq 2 \) except for the combination of \( L_j = 2 \) and \( K > 2 \).

Otherwise, there do not appear to be any identification problems with the maps beyond those inherent in latent class models (see McLachlan and Basford 1988, Titterington et al. 1985, Section 3.1). The DeSarbo et al. (1992) article on which our model is based does not suggest that there are any special identification problems with latent class preference models. Given that our model imposes more restrictions on the latent class preference model, we would not expect there to be unidentified parameters. Titterington et al. (1985) give a formal definition of identifiability for finite mixtures, illustrate that nonidentifiability can occur, and derive both necessary and sufficient conditions for identifiability. Of course, for a \( K \)-segment model, there are \( K! \) permutations of the segment labels (e.g., what was called segment 1 on one run may emerge as segment 2, 3, \ldots, \( K \) on other runs). This is of no consequence and does not constitute lack of identifiability (McLachlan and Basford 1988).

In summary, if \( M' \)-dimensional restrictions are imposed on the latent class preference model with \( K \) segments, we estimate \( M' \) unique sets of preferences, with the remaining preferences (\( K - M' \) sets of them) computed as linear combinations of the first \( M' \) sets. The restrictions pose no identification problems except in the very limited circumstances noted above, yet they often save a significant number of parameters. Two-dimensional maps are sufficient in our experience; we have not encountered any data for which higher-dimensional maps were needed. One-dimensional maps are rarely appropriate since they require stringent parameter restrictions; in addition, the managerial insights provided by one-dimensional maps are much less useful. Formal model selection criteria can be used to supplement this heuristic for determining the appropriate number of dimensions, as we discuss in a later section.
Representation of Preference Vectors and Attribute Levels

To graphically depict the map for an attribute, one simply plots the location values in the $A'$ matrix. As with other mapping methods, the interpretation of map dimensions is purely subjective. With a two-dimensional map, however, the horizontal dimension is aligned with segment 1’s preferences, and the vertical dimension represents segment 2’s preferences.

Given the locations, it is straightforward to plot segment-specific preference vectors. The weights in the weight vector $w_k^c$ give the direction of maximum preferences. Assuming a two-dimensional map in which attribute 1 is the horizontal dimension and attribute 2 the vertical dimension, the preference vector for a segment begins at the origin of the map and has a slope equal to the ratio of the second element and the first element of $w_k^c$.

To determine a segment’s preferences for an attribute level, one drops a perpendicular line from the location of the attribute level in product space to that segment’s preference vector. The attribute level farthest along the segment’s preference vector is the one most preferred by that segment. One can extend the vector back through the origin if necessary to determine the preferences for all attribute levels. It is straightforward to show that all attribute levels on a line perpendicular to the vector have equivalent preferences (cf. Green et al. 1989). Multiplying the locations of the attribute levels by the lengths of the preference vectors produces the conventional part worths (cf. Equation 5).

Extension for Marketing Mix Attributes

We can also impose restrictions on the latent class coefficients for marketing mix variables, whether they are continuous or discrete. If we let $\beta_k^c$ be the $L_c \times 1$ vector of marketing mix response coefficients of segment $k$, $\beta_k^c$ can be decomposed into the location of the attributes in $A^c$, an $M^c$-dimensional space, and the segment-specific importance weights consumers attach to those locations, $w_k^c$, as in Equation (5):

$$\beta_k^c = A^c w_k^c.$$  (7)

Unlike the maps for features, it is not necessary that one of the coefficients (corresponding to one row of $A'$) be normalized to the origin. Hence, if the map is two-dimensional, there are two more location parameters requiring estimation compared to the case of a discrete attribute. The constraints on the weight vectors are the same as for discrete attributes. The total number of parameters used in $A^c$ (which is $M^cL_c$) and $w_k^c$ (which is $M^c(K-M^c)$ cannot exceed the number of parameters required for the unrestricted latent class preference model (which is $KL_e$). Like the maps for discrete attributes, we recommend two-dimensional maps for marketing mix coefficients.

One issue that arises is that the scaling of continuous variables can affect their mapped locations. For instance, we can multiply the prices for all brands by 10 (which is legitimate since only the relative prices matter), and the coefficient that results from the estimation will be $1/10$ its previous value. The coordinates of the price coefficient in space will be $1/10$ their previous values as well. To make locations of continuous marketing mix attributes comparable in space, we recommend rescaling them prior to analysis so that they have the same means and standard deviations (standardizing is one option).

Extension for Interactions

The proposed attribute-level maps can also provide useful insights about interactions between attributes that are not obvious from traditional mapping methods. Consider an example in which there is a five-level attribute and also a two-level attribute, resulting in 10 choice alternatives. For example, the attribute with five levels may be brand name, and the attribute with two levels may be size (say, small and large). An interaction between brand and size would suggest that brand preferences differ depending on which size we are considering (and vice versa).

We generated simulated ratings data from 200 consumers (2,000 observations) for this example, including significant interaction effects, such that each part worth was distributed normally. Figure 2A shows a traditional MDS solution to the problem, with 2 segments, in which all choice alternatives are mapped in

\(^1\text{This assumes that higher values of the rating scale mean that respondents have higher preferences for the profile. If not, then the profile evaluations need to be rescaled prior to analysis.}\)

\(^2\text{We stopped at two segments for simplicity. It is possible that more segments could have better explained the data.}\)
the same space. For example, B1S1 is the alternative that has brand level 1 and size level 1. The locations of the alternatives in space reflect the brand name and size main effects as well as the brand-size interaction. It is not straightforward to isolate the main effects or the interaction with the traditional MDS approach since attributes play no role in the analysis.

Our approach can be extended to take into account interactions between factors. In this example, a brand-size interaction means that there would be two brand maps (one for each size). If the preference vectors are restricted to be the same for the two maps, the brand maps for the two size levels can be overlaid on the same graph, as in Figure 2B. For example, B1S1 is the location of brand 1 when the size level is 1, and B1S2 is the location of brand 1 when the size level is 2. This graph clearly shows the main effect of brand name and the interaction between brand and size. The larger the interactions between brand and size, the farther apart the two points for each brand level (e.g., B1S1 and B1S2) will be in the graph. In this example, the true interactions between size and brands 1, 2, 3, and 4 have normal distributions with means 0.1, 0.2, 0.3, and 0.4, respectively. Notice that for each successive pair of brands, the two sizes become farther apart in Figure 2B, as we would expect. That is, B1S1 and B1S2 are much closer together than B4S1 and B4S2 since the true interaction for the former pair is 0.1, and the true interaction for the latter pair is 0.4. If the interaction between brand 1 and size had been zero, B1 would have been located in exactly the same location, except for random error, regardless of whether it was coupled with S1 or S2. In this map, the interaction tells us that brand 4 benefits from being coupled with size 1 more than brand 1 does. This is not obvious from the MDS map in Figure 2A.

The brand name main effect is recovered in the graph as well. Brands 1 through 5 have true part worths of 5, 3, 2, −2, and 0, which is roughly consistent with the scales on both dimensions of the graph. Though the size main effect is not shown in the graph, the model correctly recovers it as well. The true size main effect is 3 for size 1 (with size 2 fixed at zero), and the recovered preferences for size 1 for segments 1 and 2 were estimated as 2.90 and 3.13.

Though the fit of this model is exactly the same as that of the traditional approach in Figure 2A, our approach allows us to graphically isolate the nature of interactions between factors.
the interaction without contaminating it with the size main effect. In applications with more than two factors, the traditional MDS approach of mapping all profiles in the same space would reflect all the main effects and interactions, which could be nearly impossible to disentangle. With our approach, one could selectively include certain interactions if desired, without having all interactions in the model. For example, if there are three factors with three levels each, the MDS plot of all $3^3 = 27$ profiles would reflect all main effects, 2- and 3-way interactions. With our approach, we could allow, say, two of the factors to interact while not allowing any of the other possible interactions. If we did this, two maps would completely characterize the preferences—the one map for the main effect of the noninteracting factor and one map that shows the interaction between the other two factors (similar to Figure 2B). In the absence of a priori expectations about which interactions to examine, a simple ANOVA could be used (at least with ratings data) to determine which interactions are statistically significant before specifying the mapping model.

In this regard, the proposed attribute-level maps have a significant advantage over the traditional MDS approach of mapping all choice alternatives in the same joint space. We can graphically examine as many or as few interactions as we wish with our approach, but the locations of choice alternatives obtained from a traditional MDS are obtained without reference to attributes.

**Estimation Procedure**

Optimization procedures can be used to maximize (4) with respect to the unknown parameters in $\alpha$, $\beta$, and $\Sigma$. Due to significant parameter reductions, the proposed restricted models take less time to estimate than the unrestricted latent class preference model.

There are two unknowns to be determined in estimating the proposed class of models—the dimensionality of the maps and the number of preference segments. In most applications, two-dimensional maps are consistent with the data and provide useful managerial insights. However, a more formal strategy is as follows.

1. Estimate unrestricted latent class preference models with various numbers of segments (say, 3, 4, 5, and 6) until the optimal number of segments is found.

2. Re-estimate these models with the two-dimensional restriction for all attributes. To determine whether the two-dimensional restrictions are consistent with the data, compare the 3-segment unrestricted model to the 3-segment model with two-dimensional restrictions using some criterion that takes into account the number of parameters required, such as AIC, BIC, or CAIC. Make similar comparisons of restricted vs. unrestricted models for the 4- and 5-segment models. We have never encountered data for which the two-dimensional restrictions are rejected.

3. If the two-dimensional restrictions are not rejected, then the models could be estimated again with the one-dimensional map restrictions. The one-dimensional map restrictions are much more severe than the two-dimensional restrictions and hence are more likely to be rejected by the data. In any case, the managerial implications available from one-dimensional maps are more limited, and we do not recommend using them.

4. From the sets of estimated models (unrestricted, 2-dimensional, and possibly 1-dimensional), one can simply pick the one with the lowest value of the information criterion.

It is possible that the optimal number of dimensions varies across attributes. For example, if brand name and size are the two attributes of interest, it is possible that the brand name map should be two-dimensional and the size map should be one-dimensional (or vice-versa). The dimensionality of one map has nothing to do with the dimensionality of another map—they can be determined independently. However, if there are $F$ features, then the researcher would have to estimate $2^F$ models to completely exhaust the combinations of one- and two-dimensional maps for that preference analysis. This is not practical for large data sets because of excess estimation time. It is our philosophy that overspecifying the dimensionality of the maps is less of a problem than underspecifying the dimensionality. In our empirical example, we specify two-dimensional maps for all attributes, though we did estimate a model with all one-dimensional maps as well. Overspecification merely costs extra parameters, but fit and managerial implications may be sacrificed by underspecifying the number of dimensions. In any case, even if
models with all two-dimensional maps are overspecified, they are still much more parsimonious and much less overspecified than the commonly-used unrestricted latent class preference model.

The problem of determining the number of dimensions and segments in latent class multidimensional scaling models is still in need of research. We acknowledge that the violation of regularity conditions that makes likelihood ratio tests invalid for this type of model technically affects such statistics as BIC as well (Titterington et al. 1985). As a result, some researchers have resorted to simulation methods such as bootstrapping for determining the number of dimensions and number of segments (e.g., Böckenholt and Böckenholt 1991, DeSoete and Heiser 1993, DeSoete and Winsberg 1993). However, this is a computationally burdensome procedure when there are many parameters and many observations. Thus, this important problem is not yet solved and is in need of further research.

Application to Choice Data: Fabric Softener SKUs

Fader and Hardie (1996) show how a set of discrete attributes can be used to describe every SKU on the store shelves in a parsimonious manner. In their study, 56 SKUs of fabric softeners sold by stores in the Philadelphia market are described by brand name (10 levels), size (4 levels), product form (4 levels), and formulation (4 levels). Their data set consists of panelists' purchases of actual products in the marketplace, not evaluations of artificial stimuli used in a typical conjoint experiment. The natural and less obtrusive nature of routine grocery shopping makes the analysis of scanner panel data an appealing complement to the standard conjoint experiment (Fader and Hardie 1996).

We apply the proposed preference model to a sample of the IRI fabric softener scanner panel used by Fader and Hardie (the reader is referred to that study for details of the data set, including a detailed description of the attributes of each SKU). Four hundred randomly chosen panelists comprise our sample. Two hundred of these (2,364 purchases) were randomly chosen for model estimation, with the remaining two hundred (2,527 purchases) used for model validation. There are 594 panelists in the entire data set, over two-thirds of whom are used in our study.

Besides the dummy variables for the various attributes, we also include regular price, price cut, aisle display, and store feature variables in the utility functions of the models since marketing mix variables are often important explanatory variables in choice models. We map the coefficients for regular price, price cut, aisle display, and store feature advertising in a common space, as described in Equation (7). Obviously, the scaling of continuous variables such as price can affect the size of the coefficient and hence the coordinates of price in the map. In our data, regular price has an average value of one and very low variance so that its maximum impact on utilities is comparable to that of the discrete variables, and so scaling is not an issue.

Latent structure probabilistic choice models are typically parameterized to be consistent with the proposition that consumers follow a zero-order brand switching pattern (Chintagunta 1994, Dillon et al. 1994, Kamakura and Russell 1989), so we do not include loyalty variables in the specifications. In any case, Fader and Hardie (1996) find high correlations between part worths estimated with loyalty variables present and those estimated without loyalty variables.

Since the SKUs are from a real-world market and not a carefully-designed conjoint experiment, the profiles are correlated, and hence it is not possible to include interactions in the model. The 57 SKUs represented in the data are only a small fraction of the 640 (\(= 10 \times 4 \times 4 \times 4\)) such profiles that would be used in a full factorial design.

The likelihood values and BIC statistics for the unrestricted latent class models and the restricted two-dimensional models are shown in Table 1A. The unrestricted latent class model has 18 dummy variables for the four major attributes (brand name, size, product form, and formulation), in addition to four marketing mix variables (regular price, price cut, aisle display, and store feature). Thus, each additional segment adds 22 parameters plus a weight parameter. The restricted models have two-dimensional maps for all attributes and marketing mix effects. Forecasting statistics for these and several benchmark models are shown in Table 1B.

As Table 1A shows, the two-dimensional restrictions
Table 1  Results for Fabric Softener Data

A. Fit Statistics for Unrestricted Latent Class Model and Proposed 2-Dimensional Model for Fabric Softener Data

<table>
<thead>
<tr>
<th>Model</th>
<th>Log L</th>
<th>P</th>
<th>BIC</th>
<th>Log L</th>
<th>P</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unrestricted Latent Class</td>
<td></td>
<td></td>
<td></td>
<td>Restricted Latent Class 2-Dimensional Restrictions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 segment</td>
<td>−8040</td>
<td>45</td>
<td>16429</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3 segment</td>
<td>−7771</td>
<td>68</td>
<td>16070</td>
<td>−7779</td>
<td>56</td>
<td>15944</td>
</tr>
<tr>
<td>4 segment</td>
<td>−7556</td>
<td>91</td>
<td>15819</td>
<td>−7564</td>
<td>67</td>
<td>15648</td>
</tr>
<tr>
<td>5 segment</td>
<td>−7289</td>
<td>114</td>
<td>15464</td>
<td>−7327</td>
<td>78</td>
<td>15260</td>
</tr>
<tr>
<td>6 segment</td>
<td>−7160</td>
<td>137</td>
<td>15384</td>
<td>−7184</td>
<td>89</td>
<td>15059</td>
</tr>
<tr>
<td>7 segment</td>
<td>−7001</td>
<td>160</td>
<td>15244</td>
<td>−7089</td>
<td>100</td>
<td>14955</td>
</tr>
<tr>
<td>8 segment</td>
<td>−6903</td>
<td>183</td>
<td>15228</td>
<td>−7015</td>
<td>111</td>
<td>14892</td>
</tr>
<tr>
<td>9 segment</td>
<td>−6850</td>
<td>206</td>
<td>15300</td>
<td>−6970</td>
<td>122</td>
<td>14887</td>
</tr>
<tr>
<td>10 segment</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>−6895</td>
<td>133</td>
<td>14822</td>
</tr>
<tr>
<td>11 segment</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>−6885</td>
<td>144</td>
<td>14889</td>
</tr>
</tbody>
</table>

B. Comparison to Benchmark Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Estimation Sample 200 consumers, n = 2,364</th>
<th>Validation Sample 200 consumers, n = 2,527</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>Log L</td>
</tr>
<tr>
<td>I.</td>
<td>60</td>
<td>−8237</td>
</tr>
<tr>
<td>II.</td>
<td>156</td>
<td>−6959</td>
</tr>
<tr>
<td>III.</td>
<td>80</td>
<td>−7472</td>
</tr>
<tr>
<td>IV.</td>
<td>183</td>
<td>−6903</td>
</tr>
<tr>
<td>V.</td>
<td>133</td>
<td>−6895</td>
</tr>
</tbody>
</table>

*Number of parameters.

*BIC = \(-2 \log L + P \ln(n)\), where n is the sample size.

are consistent with the data regardless of the number of segments. For example, for the 3-segment models, the restricted model has 12 fewer parameters but only gives up 8 likelihood points to the unrestricted latent class model. BIC favors the restricted model. For the 8-segment models, the restricted model has 72 fewer parameters but gives up about 112 likelihood points. BIC again favors the restricted model.

For the unrestricted latent class model, 8 segments are optimal according to BIC, while for the restricted 2-dimensional models, 10 segments are optimal.

Table 1B compares the forecasting performance of the unrestricted latent class model (Model IV), the model with 2-dimensional restrictions (Model V), and several other benchmark models. The first model presented in the table is the traditional logit model with SKU-specific constants. With 57 SKUs purchased by panelists\(^4\), requiring 56 constants, the total parameter count is 60 with the four marketing mix variables included in the utility function.

Model II is a latent class version of the basic logit model I, but with restrictions on the dimensionality of the space occupied by the 57 SKUs. This model produces a single 2-dimensional map containing all 57 SKUs, similar in spirit to that by Chintagunta (1994).

\(^4\)Fader and Hardie (1996) include only 56 SKUs in their estimation sample and forecast the market share of the 57th SKU.
The optimal 7-segment model requires 156 parameters and performs much better than the basic logit model I according to all criteria.

We then re-estimate the 7-segment model II with the coordinates of choice alternatives simultaneously reparameterized as attribute coordinates, as described by DeSoete and Winsberg (1993). Since the map is two-dimensional, the matrix containing locations of choice alternatives in model II is $57 \times 2$. In model III, the design matrix that describes the 57 choice alternatives on the four attributes is $57 \times 18$ since 18 dummy variables are required to code four attributes. The matrix of attribute level locations in the two-dimensional space is $18 \times 2$. Thirty-six ($= 18 \times 2$) location parameters are required for the reparameterized model III, compared to 112 ($= 56 \times 2$) for the unrestricted model II\(^5\), resulting in a savings of $112 - 36 = 76$ location parameters over model II. The simultaneous reparameterization is definitively rejected by the data. The BIC value for the reparameterized model is 15,566 compared to 15,130 for the unrestricted 7-segment model II. Also, note that the validation likelihood ($-7,683$) is not as good as that of the unrestricted 7-segment model II ($-7,502$), which also suggests that the simultaneous reparameterization restrictions are inappropriate.

We show the map from the simultaneous reparameterization in Figure 3 to further demonstrate the differences between our approach and the existing approach. Unlike our proposed model, which maps attributes in separate spaces, the existing approach produces a joint space that is quite crowded and difficult to interpret. The locations of the 18 dummy variables used to code the attributes and the seven preference vectors are shown in the map. Much information (sizes, formulations, etc.) is eliminated from the map by necessity.

Model IV estimates part worths instead of SKU-specific constants (cf. Fader and Hardie 1996). The 8-segment model has $8 \times 22 = 176$ parameters plus another 7 parameters for segment weights, for a total of 183. The forecasting performance is significantly better than that of models I–III.

The proposed model with 2-dimensional restrictions (model V) imposes restrictions on the latent class preference model (model IV) such that the preferences are restricted to two-dimensional space. As mentioned earlier, the 10-segment restricted model achieves as good a fit as the unrestricted 8-segment model, but with 50 fewer parameters. Forecasting performance on the validation sample is also impressive—the likelihood value of the restricted model is 109 likelihood points better than that of the unrestricted latent class preference model.\(^6\)

We also estimated the proposed model with one-dimensional maps for all attributes, but the restrictions were not consistent with the data. It is possible that some, but not all, of the maps could be reduced to one dimension. Obviously, we would need to estimate all possible combinations of one- and two-dimensional maps for the four attributes ($2^4 = 16$ models) to fully explore whether the overall model could be improved by allowing some maps to be one-dimensional and

---

\(^5\)Recall that one row of the location matrix must be fixed at the origin in model II; this is not necessary in model III.

\(^6\)For all latent class models II, III, IV, and V, forecasting was accomplished by first assigning consumers to their respective segments using posterior probabilities (cf. Kamakura and Russell 1989).
some to be two-dimensional. Given that this is not practical with large data sets such as this one, we are comfortable risking overspecification of some maps. As mentioned earlier, overspecification does cost extra parameters, but underspecification may cause loss of model fit as well as managerial insights. In any case, the proposed model cannot be nearly as overspecified as the unrestricted latent class preference model.

The Attribute Level Maps
Figure 4 shows the attribute level maps for brand names (A), sizes (B), product forms (C), formulations (D), and the marketing mix (E) produced by the optimal 10-segment model. The lengths of the preference vectors are scaled such that they are proportional to the sizes of the latent segments. For example, a vector that is twice as long as another means that the segment has twice as many consumers as the other segment. Attribute levels that are very close together in the space (e.g., Arm & Hammer, Bounce, and Downy in Figure 4A) have more similar preferences than levels not close together (e.g., Generic and StaPuf). Brands located very close to one another may suffer from lack of differential advantage.

Dropping perpendicular lines from each brand name location in Figure 4A to the vector of the largest segment (segment 1), we see that segment 1’s preference ordering is Snuggle, Cling Free, Final Touch, Downy, Arm & Hammer, Bounce, Toss n’ Soft, StaPuf, Private Label, and the Generic brand. Segment 1 comprises 23 percent of the sample. Segment 2 (13 percent of the sample) prefers StaPuf, Private Label, Toss n’ Soft, Final Touch, Cling Free, Generic, Snuggle, Bounce, Arm & Hammer, and Downy. Despite the very different preferences of segments 1 and 2, note that Arm & Hammer, Bounce, and Downy (all brand names in very close proximity on the map) are close to each other in each segment’s preference ranking.

Segments 3 and 4 have somewhat similar preferences. Segment 3 prefers Generic, StaPuf, and Private Label, while segment 4 prefers StaPuf, Generic, and Private Label. These segments comprise 12% and 11% of the sample, respectively. Segments 5, 6, 7, 8, and 9, comprising 10%, 8%, 8%, 7%, and 4% of the sample, respectively, prefer Final Touch, StaPuf, Final Touch, Downy, and Downy, respectively. The smallest segment, segment 10, strongly prefers Generic to Private Label and StaPuf. This segment, comprising over 3% of the sample, is most likely too small to be actionable.

The size map in Figure 4B shows that segment 1 prefers medium, small, large, and extra large. Segment 2 prefers medium and is virtually indifferent between small, large, and extra large. Similarly, segments 3, 6, 7, and 9 prefer medium as well. Segments 4, 5, and 8 prefer small. Segment 10, the smallest segment, appears to be economy-minded, with strong preferences for the generic brand (Figure 4A) and the extra large size (Figure 4B). Overall, there is less heterogeneity in size preferences than in brand name preferences since more of the preference vectors point in the same direction.

As shown in the map for product forms in Figure 4C, segment 1’s preference ordering is sheets, concentrate, liquid, and refill, while segment 2 is almost indifferent between sheets and refill, with liquid and concentrate being less preferred. Segment 3 prefers sheets, while segment 4 is nearly indifferent between concentrate, liquid, and refill. Consistent with earlier results, the economy segment (segment 10) prefers refill product forms.

The map for formulations in Figure 4D shows that segment 1 prefers regular and is then nearly indifferent between light and staingard. Segments 2, 3, 4, and 10 have fairly similar preferences for formulations, with a slight preference for light over regular. Clearly, there is less heterogeneity in preferences for formulations than for brand names and product forms.

Finally, the map for marketing mix variables in Figure 4E shows that segments 1, 5, 8, 9, and 10 are not very sensitive to marketing activity. Segments 2, 4, 6, 7 are most sensitive to marketing activity, with segments 3 and 5 having intermediate sensitivity. Display and store feature ad activity add to utility for all segments. For example, when we drop a perpendicular line from display to any of the preference vectors, we find that the line intersects the preference vector beyond the origin of the vector in the direction of the

7Final Touch is mostly obscured by the arrow of the preference vector for segment 6.

8The segments are the same in each map—that is, the consumers in segment 1 in map A are the same as the consumers in segment 1 in maps B, C, D, and E.
Figure 4  Fabric Softener Data Results

A. Brand Names: 2-Dimensional Map and 10 Preference Vectors

B. Sizes: 2-Dimensional Map and 10 Preference Vectors

C. Product Forms: 2-Dimensional Map and 10 Preference Vectors

D. Formulations: 2-Dimensional Map and 10 Preference Vectors

E. Marketing Mix: 2-Dimensional Map and 10 Preference Vectors
arrow. Regular price subtracts from utility for all segments. Price cut adds to utility for some segments (e.g., segment 3) but subtracts from utility for others (e.g., segment 1). As we will later see, the price cut effect is actually negligible for all segments.

Information can be integrated across maps to determine which attribute influences each segment the most. Table 2 was computed by first finding the maximum change in utility possible through each variable. For brand name, size, product form, and formulation, this was the difference between the largest and smallest part worth. For aisle display and store feature ad (both 0/1 variables), the maximum difference in utility produced by the variables is simply the value of the projections onto the segment vectors. To compute the maximum change in utility possible through regular price changes and price cuts, we use the maximum values found in the data set, which are 63% and 100%, respectively. For each segment, we then normalize the utility changes across attributes within segment so that the importance weights sum to one. The largest segment (1) is most concerned with product form, while segment 2 is most concerned with formulations and then brand names. With the exception of segment 4, which is most sensitive to aisle display, the remaining segments pay most attention to brand names. It is surprising that three of the four largest segments weight some attribute other than brand name as most important. Given the amount of attention devoted in the literature to the marketing environment relative to product attributes, it is also surprising that the product attributes usually have much more affect on utilities than the marketing environment does.

Combining (i) the preferences for each segment, (ii) the relative attribute importances for each segment, and (iii) the relative sizes of each segment, interesting managerial insights can be obtained. For example, segment 1, which comprises 23% of the sample, is most influenced by the product form attribute (Table 2) and strongly prefers sheets (Figure 4C). Segment 2, comprising 13% of the sample, is most influenced by formulation and slightly prefers light over regular. Segment 4, comprising 11% of the sample, is most influenced by aisle display. Similar analyses can be performed for the other segments as well.

As a basis for comparison, we have plotted the part worths from the unrestricted latent class solution in Figure 5. Eight latent segments were optimal for that model according to BIC. It is our opinion that the attribute-level maps in Figure 4 convey information more readily than the plots of part worths in Figure 5. For example, it would take a considerable amount of inspection to determine that two brands had similar preferences across all segments. Our model conveys

---

**Table 2** Attribute Importances by Segment

<table>
<thead>
<tr>
<th>Segment</th>
<th>Brand</th>
<th>Size</th>
<th>Form</th>
<th>Formula</th>
<th>R. Price</th>
<th>P. Cut</th>
<th>Display</th>
<th>Store Ad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segment 1</td>
<td>0.11</td>
<td>0.10</td>
<td><strong>0.41</strong></td>
<td>0.08</td>
<td>0.00</td>
<td>0.10</td>
<td>0.11</td>
<td>0.08</td>
</tr>
<tr>
<td>Segment 2</td>
<td>0.26</td>
<td>0.05</td>
<td>0.05</td>
<td><strong>0.30</strong></td>
<td>0.05</td>
<td>0.01</td>
<td>0.17</td>
<td>0.11</td>
</tr>
<tr>
<td>Segment 3</td>
<td>0.32</td>
<td>0.15</td>
<td>0.26</td>
<td>0.10</td>
<td>0.03</td>
<td>0.03</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>Segment 4</td>
<td>0.11</td>
<td>0.14</td>
<td>0.15</td>
<td>0.07</td>
<td>0.09</td>
<td>0.06</td>
<td><strong>0.23</strong></td>
<td>0.15</td>
</tr>
<tr>
<td>Segment 5</td>
<td>0.48</td>
<td>0.16</td>
<td>0.04</td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
<td>0.13</td>
<td>0.09</td>
</tr>
<tr>
<td>Segment 6</td>
<td>0.38</td>
<td>0.04</td>
<td>0.12</td>
<td>0.19</td>
<td>0.04</td>
<td>0.04</td>
<td>0.12</td>
<td>0.07</td>
</tr>
<tr>
<td>Segment 7</td>
<td>0.27</td>
<td>0.10</td>
<td>0.16</td>
<td>0.12</td>
<td>0.06</td>
<td>0.06</td>
<td>0.14</td>
<td>0.09</td>
</tr>
<tr>
<td>Segment 8</td>
<td>0.23</td>
<td>0.18</td>
<td>0.22</td>
<td>0.09</td>
<td>0.02</td>
<td>0.06</td>
<td>0.12</td>
<td>0.08</td>
</tr>
<tr>
<td>Segment 9</td>
<td>0.75</td>
<td>0.04</td>
<td>0.03</td>
<td>0.11</td>
<td>0.00</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>Segment 10</td>
<td>0.46</td>
<td>0.09</td>
<td>0.23</td>
<td>0.10</td>
<td>0.00</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
</tr>
</tbody>
</table>

*Note: Since we are computing the maximum effects for each attribute, the largest changes found in the data set for regular price and price cut were used, which are 63% for regular price and 100% for price cut.*
Figure 5  Part Worths from Unrestricted Latent Class Solution

A. 8-Segment Latent Class Solution: Brand Part Worths

B. 8-Segment Latent Class Solution: Size Part Worths

C. 8-Segment Latent Class Solution: Product Form Part Worths

D. 8-Segment Latent Class Solution: Formulation Part Worths
Managerial Implications: The Effects of Marketing Mix Variables on Mapped Attribute Locations

In this section we investigate some unique managerial implications of the maps that cannot be gleaned readily from plots of part worths. Specifically, we suggest an extension of the model that allows an examination of the effects of marketing activity such as price and promotions on the mapped locations of attributes.

For the proposed maps appearing in Figure 4, the vector of 57 SKU utilities for segment \( k \) on purchase occasion \( t \) was computed as

\[
U_k^t = \exp(Brand \cdot A^1 \cdot w_{k1}^t + Size \cdot A^2 \cdot w_{k2}^t + Form \cdot A^3 \cdot w_{k3}^t + Formulation \cdot A^4 \cdot w_{k4}^t + X_t \cdot A^5 \cdot w_{k5}^t),
\]

where the \( A^i \) matrices contain coordinates of attribute levels in the spaces, the \( w_{ki} \) are the preference vectors (cf. Equation (5)), the Brand, Size, Form, and Formulation matrices are design matrices which describe each of the 57 SKUs on each of the attributes\(^{10} \), and \( X_t \) is the matrix which describes each SKU on the price, price cut, aisle display, and feature variables at purchase occasion \( t \). For example, Figure 4A maps the locations in \( A^1 \) and the preference vectors \( w_{k1}^t \).

We can respecify the above model so that the preference vectors \( w_{ki} \) are common to selected attributes. The effect of this respecification is that the attributes having a common set of preference vectors share the same reduced-dimension space. This allows us to study the impact of marketing mix variables on the locations of attribute levels in the derived spaces. If we constrain the above model so that the brand name and marketing mix attributes share a common space, the utilities become

\[
U_k^t = \exp(Brand \cdot A^1 + X_t \cdot A^5) \cdot w_{k1.5}^t \cdot Size \cdot A^2 \cdot w^2_k + Form \cdot A^3 \cdot w^3_k + Formulation \cdot A^4 \cdot w^4_k + X_t \cdot A^5 \cdot w^5_k),
\]

where \( w_{k1.5}^t \) is the preference vector for segment \( k \) shared by the brand name and marketing mix attributes. This constrained model fits very well (LogL = −6,998, BIC = 14,905) compared to the unrestricted latent class models, though not quite as well as the model with separate spaces for brand name and marketing mix variables (Model V in Table 1B).

Figure 6A plots the coordinates in \( A^1 \) and \( A^5 \) and the preference vectors \( w_{k1.5}^t \) for \( k = 1, \ldots, 10 \). First, note that the results from this model are quite consistent with those of the model presented in Figure 4. Comparing Figure 4A to Figure 6A, we see that segment 1 (the largest segment) still prefers Snuggle, Cling Free, and Final Touch (though in slightly different order), and segment 2 still prefers StaPuf. Segment 2 is very sensitive to aisle display and store feature ads (consistent with Figure 4E). The segment is quite price sensitive as well, with the projection of price onto the segment 2 vector having a value of −1.11. The Generic brand name has a large negative effect for segment 2 as well, with a projection on the segment 2 vector of −0.88. Such information can be used to determine how much consumers would have to be compensated with price for switching from one brand name to another. For example, if a segment 2 consumer were to switch from Arm & Hammer (which has a projection value of zero) to Generic, their utility would decrease by 0.88. However, if the regular price of the generic were lower by 0.79 (−1.11 × −0.79 = 0.88), consumers would be indifferent between Arm & Hammer and Generic. This magnitude of price variation is not present in the data, so apparently it is not feasible to compensate segment 2 consumers enough for them to use generics. Given their deal proneness, consumers in this segment apparently look for named brands that are on sale rather than resort to generics.

Using this model, we can graphically describe what happens to the locations of attribute levels when there are price reductions and promotional activity. Figure 6B shows the brand locations at the average price level, assuming no aisle display or store feature advertisements. The locations plotted in the graph are

\[
A^1 + \bar{X}A^5,
\]

and the preference vectors are the same as in Figure 6A. Also shown in the graph is the movement of the...
Figure 6 Effects of Marketing Mix on Brand Name Locations
A. 2-Dimensional Model with Brand Name and Marketing Mix in Common Space

B. Locations of Brand Names at Average Prices, No Promotion; Movement in Bounce Brand with Average Price Cut (23.75%), Aisle Display, Store Feature Ad

Bounce brand name when there is a 23.75% price cut (the average size of the price cut when there was a price cut), a store feature advertisement, and an aisle display for that brand. As in Figure 6A, we see that segment 2 is very sensitive to promotions. For this segment, when Bounce was heavily promoted, the positioning moved from among the least preferred brand names to a location far preferred over the other brands. The promotion makes Bounce most preferred for segment 1 consumers as well, but not for segment 3 consumers, who still prefer StaPuf, Generic, and Private Label. In fact, the promotions improve the positioning of Bounce for every segment.

One could compute a map such as that in Figure 6B for each purchase occasion, depending on the prices and promotional activity in the marketing environment. Thus, this example shows that it is possible to make the proposed maps dynamic to reflect the purchase environment. Similar maps could be developed for the size, product form, and formulation attributes as well. For example, to study the effects of promotions on the positioning of sizes, we could plot

\[ A^2 + \bar{X}A^5 \]

in a joint space with preference vectors constrained to be the same for the size and marketing mix attributes.

Conclusion
A fundamental premise of this paper is that consumers are much more likely to form preferences for attributes of products rather than for each individual product on store shelves. Given limited cognitive resources, it is highly unlikely that consumers form direct preferences for each choice alternative on the store shelves. Mapping preferences for attributes is therefore much more consistent with consumer behavior than the traditional MDS approach of mapping all choice alternatives. In addition to the behavioral basis for mapping attributes, the model is much more parsimonious than an unrestricted latent class preference model, it forecasts holdout choices better, and it provides unique managerial insights not readily obtainable from traditional MDS approaches or plots of part worths.

The parsimony of the model is very important with large data sets such as the fabric softener data used in this study. Compared to the best unrestricted latent class preference model, our restricted model requires 50 fewer parameters yet provides a comparable fit. Equally impressive is the fact that the model forecasts
holdout choices better by 109 likelihood points. Indeed, our model forecasted choices better than MDS approaches that reparameterize choice alternatives as attributes as well.

In addition to parsimony and improved forecasting performance, the proposed model provides some unique managerial insights not readily available from MDS or conjoint approaches. First, the interpretation of the maps is very intuitive. With a glance, we can determine which attribute levels have similar preferences across segments. Attribute levels very close to each other in the space have very similar preferences, and are likely to be prone to more substitution and switching. A focal brand in close proximity to other brands needs some sort of differential advantage over those brands.

Second, the model allows us to graphically examine interactions between attributes. In an example with simulated data, we show how an interaction between, say, brand names and sizes can be shown as separate brand maps for each level of the size factor. We can then see how the preferences for a brand name change when the brand name is coupled with the various size levels. In contrast, a map produced by a traditional MDS approach reflects not only the brand and size main effects but also the brand-size interaction. There is no satisfactory approach for examining interactions within the traditional MDS approach—simultaneous reparameterization has limitations, as shown in this study, and with two-stage property fitting procedures, there is no guarantee that all parameters are optimized.

Third, the effects of changes in the price and promotional environment on positioning and preferences can be studied. In our empirical application, we respecified the maps so that marketing mix effects could be mapped in the same space as the attributes. This makes the maps dynamic since the locations could change for each purchase occasion depending on the prices and promotional activities being used on that occasion. For each segment, we can also study how any combination of price changes, aisle display, or store feature advertising affects the positioning of the focal brand (or other attribute level). Since nearly all scanner panel data contains some information on SKU attributes such as brand name, size, formulation, flavor, etc., this approach could be used widely in applied choice modeling.

The study by Fader and Hardie (1996) calls for a comprehensive research program on the topic of SKU choice. We answer this call with an integrated MDS/preference analysis model that provides information on the positioning of attribute levels and the segment-specific preferences for those positions. The results from the fabric softener category show that product attributes usually have more impact on choices than do marketing environment variables, which suggests the need for more research into the role of attributes in choice. Future research could investigate the possibility of developing ideal point models for attributes. Instead of the “more is better” interpretation associated with vector models, ideal point models posit that there is an ideal level of each dimension of the map. Another possible next step for the model is to parameterize the mixing weights directly as a function of actionable segmentation variables (cf. Gupta and Chintagunta 1994, Kamakura et al. 1994). Another possibility is to incorporate profitability information into the model (e.g., Green and Krieger 1991), leading to an integrated model of optimal product design.

More research is also needed on the technical issues involved in estimating such models. There is need for a thorough Monte Carlo analysis that examines algorithm performance, parameter recovery, model selection, CPU time, etc., as a number of model, data, and error factors are experimentally manipulated.\(^{11}\)

References


———, J. Douglas Carroll, Donald R. Lehmann, and John O’

\(^{11}\)The authors thank the editor, area editor, and two anonymous reviewers for their contributions to this study and Pete Fader for providing the data.


---

This paper was received June 23, 1998, and has been with the authors 5 months for 4 revisions; processed by Wagner Kamakura.