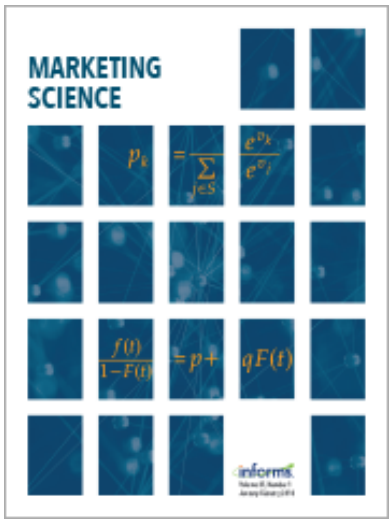


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Spatial Representation of Market Information

Wayne S. DeSarbo, Alexandru M. Degeratu, Michel Wedel, M. Kim Saxton,

To cite this article:
Wayne S. DeSarbo, Alexandru M. Degeratu, Michel Wedel, M. Kim Saxton, (2001) The Spatial Representation of Market Information. Marketing Science 20(4):426-441. https://doi.org/10.1287/mksc.20.4.426.9759

Research Note

The Spatial Representation of Market Information

Wayne S. DeSarbo • Alexandru M. Degeratu • Michel Wedel • M. Kim Saxton
Marketing Department, Smeal College of Business, Pennsylvania State University, University Park, Pennsylvania 16802
McKinsey & Co., 55 East 52nd Street, New York, New York 10022
Department of Marketing Research, Faculty of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands
Eli Lilly & Co., Lilly Corporate Center, Indianapolis, Indiana 46285-4113
desarbows@aol.com • alexandru_degeratu@mckinsey.com • m.wedel@eco.rug.nl • saxton_kim@lilly.cc

Abstract

To be used effectively, market knowledge and information must be structured and represented in ways that are parsimonious and conducive to efficient managerial decision making. This manuscript proposes a new latent structure spatial model for the representation of market information that meets this requirement. When applied to a priori defined (e.g., socioeconomic) segments, our proposed methodology provides a new way to display marketing data parsimoniously via dimension reduction through a factor-analytic specification. In post hoc studies, we simultaneously derive market segments from the data and represent the structure of market information within each of the unobserved, derived groups/segments. We summarize all relevant information concerning derived market segments via a series of maps that prove conducive to the quick and accurate dissemination of customer and competitor market information. The associations between the variables are captured in a reduced space, where each variable is represented by a vector that emanates from the origin and terminates on a hypersphere of unit (the vector length is arbitrary) radius (e.g., a unit circle in a two-dimensional space). The angles between the variable vectors capture the correlation structure in the reduced space. The method is very general and can be utilized to identify latent structures in a wide range of marketing applications. We present an actual commercial marketing application involving the (normalized) prescription shares (of specialists) of ethical drugs to demonstrate the effectiveness of representing market information in this manner and to reveal the advantage of the proposed methodology over a more general finite mixture-based method. The proposed methodology derives three segments that tend to group specialists with respect to the stage of adoption of innovation in this therapeutic category. The specialists in the first group appear to be laggards because they prescribe more of the older class of brands. However, they also have a higher-than-average preference for a newer and somewhat cheaper brand. This suggests that some of the specialists belonging to this segment may be price sensitive, while others may exhibit a slower adoption cycle, replacing the older class with the newer brands, and thus, skip one stage in the cycle of innovation. The specialists in the second segment are heavy users of the newer class of brands but are not particularly fast to adopt the latest brands. Finally, the last segment clearly consists of innovators. Traditionally, pharmaceutical marketers have viewed specialists in one of two extremes—all specialists are the same (i.e., the market has only one segment) or all specialists are very different (i.e., the market consists of 10,000+ segments of one physician each). Not surprisingly, this analysis suggests a more moderate perspective: specialists adopt new products at different rates.

(*Spatial Models; Market Segmentation; Latent Structure Analysis; Maximum Likelihood Estimation; Ethical Drugs*)

1. Introduction

Market information has profound strategic importance for organizations, and its positive effects on business performance have been well documented (e.g., Day 1994, Li and Calantone 1998). To be effective, however, information needs to be transferred efficiently throughout the organization beyond those who acquire it (Day 1994). This can be achieved only if the information is parsimoniously summarized in a format that facilitates rapid interpretation and communication. Graphical representations help to quickly identify patterns in multivariate data and serve to communicate the very essence of marketing research results, capitalizing on human abilities to recognize, process, and remember visual patterns (Spence and Lewandowski 1990). Although insightful, graphical display is fre-

0732-2399/01/2004/0426/$05.00
1526-548X electronic ISSN

quently obtained in a model-free context and therefore does not do justice to the complexity of marketing data. We propose to fill that gap by providing a new model-based technique for the parsimonious representation and graphical display of large, high-dimensional, heterogeneous data sets. This type of data is encountered in a wide range of marketing problems, such as in customer satisfaction, market segmentation, studies of customer value, positioning, and brand evaluation studies. When applied to *a priori derived* (e.g., socioeconomic) groups, our proposed method provides a new way to display marketing data parsimoniously via a spatial map. In post hoc studies, we simultaneously derive such diverse groups from the data (market segments) and represent the structure of market information within each of the unobserved groups/segments. We summarize all relevant information at the derived group/segment level via a series of maps that are conducive to the quick and accurate dissemination of customer and competitor market information across the firm.

Our approach falls into the class of latent structure (MDS or factor) mixture models, which we review in the next section. This proposed methodology can be accurately viewed as a covariance-restricted spatial version (nested) of the general Wolfe (1970) pattern clustering procedure involving finite mixtures of multivariate normal distributions. Unlike several latent structure MDS models (e.g., MULTICLUS; DeSarbo et al. 1990), which are spatial mean-restricted special cases of this same Wolfe general framework, the proposed methodology restricts/reparameterizes the derived covariance matrices, not the estimated centroids. We then describe our approach technically and apply it to an actual commercial data set, investigating physicians' prescribing behavior in a category of ethical pharmaceuticals. We first present the empirical results from the most general Wolfe pattern clustering model with unrestricted mean vectors and covariance matrices. Finally, the results of our covariance-restricted methodology are presented and formally compared to this more general, highly parameterized model. The last section summarizes the conclusions.

2. Latent Structure Spatial Mixture Models

Several existing spatial models provide graphical representations and abound in such areas as competitive market structure analysis (Elrod and Keane 1995) and market segmentation (DeSarbo et al. 1994). However, heterogeneity is a major factor hampering the graphical display of multivariate data that is often encountered in marketing research studies. As has been aptly demonstrated in the segmentation and mixture modeling literature, aggregate sample results frequently mask the true structure underlying marketing data. After the data are classified according to a priori defined groups or segments, many of the above methods can be applied separately to each group. However, splitting the sample a priori according to demographic and lifestyle variables often yields limited discrimination with respect to the variables of key interest depending on the particular variables used (cf., Wedel and Kamakura 1999, Chapter 2). Latent structure mixture models, specifically mixture MDS models (cf., Bockenholt and Bockenholt 1991, DeSarbo et al. 1994, Wedel and DeSarbo 1996) and mixture factor models (cf., Hinton et al. 1997, Tipping and Bishop 1999), were developed for such situations, where a priori market segmentation based on observed variables fails.

In mixture MDS (mMDS) models (DeSarbo et al. 1994), the relationships between brand locations and derived segment vectors or ideal points render information about the conditional centroids of the component mixture distributions. As will be shown, this is quite different from our proposed procedure, which derives separate spaces for each of the derived latent groups/segments and thus allows the brand locations to vary across the groups/segments. This may better serve the needs of graphically representing the structure of high-dimensional data, especially when the user wants to represent the different pattern of interdependencies among variables between the unobserved groups/segments. A second class of mixture-latent variable models comprises (normal) mixtures involving structural equation models. These models are constrained versions of Wolfe's (1970) mixture of multivariate normal distributions. Wolfe

(1970) assumed that the data follow a multivariate normal distribution within each of a number of unobserved groups/segments and attempts to recover the latent segments, as well as the mean vector and covariance matrix for each segment. As in structural equation modeling (Jöreskog 1973), latent class structural equations modeling imposes constraints on the covariance matrix in the unobserved segments, which are dictated by prior theory (e.g., Jedidi et al. 1997, Dolan and van der Maas 1998). Specifically, mixtures of confirmatory factor (mCFA) models are contingent upon available prior theory that dictates the structure of the factor loadings. Although identifying restrictions are necessary for our approach, it does not require a theoretical motivation of the pattern of loadings. In addition, our parameterization of the group-specific covariance matrix is different from that in mCFA models in that we enable a spatial representation of the segment level covariance matrix. Also, per the mCFA literature, the factor structure is not graphically represented, which is a main feature of our approach. Thus, our proposed graphical display of mixtures of multivariate normals has features that have not been proposed heretofore in the literature and may provide a more parsimonious fit over competing latent structure approaches containing more parameters.

3. The Proposed Spatial Methodology

We intend to portray the structure in market segment level data covariance structures via a series of maps. Heterogeneity and multivariate dependencies in the data are captured by specific parameters of a factor model underlying these maps. The covariance structure of the variables is represented by the angles among variable vectors that terminate on the boundaries of a unit hypersphere. Such a representation is comparable to biplot graphical representations in exploratory factor analysis. We assume the existence of segments that are ordered on the basis of their (unknown) sizes or prior/mixing probabilities: $w_{(1)}, \ldots, w_{(S)}$. Then, the covariance structure of the variables within a segment is represented by conveniently depicting them as vectors that terminate at the boundary of a unit hypersphere, analogous to factor analytic specifications. The directions of the variables in the reduced space portray the covariance structure; color intensities can be utilized to identify quickly the variables in the plots. Our proposed methodology can be viewed as a mixture of exploratory factor models that simultaneously derive latent segments, as well as depicting the covariance structure within each segment vis-à-vis a reduced space, factor model.

To formulate the model underlying this spatial representation, we let:

$i = 1, \ldots,$ I observations (e.g., consumers);

$j = 1, \ldots,$ J variables (e.g., brands);

$s = 1, \ldots,$ S market segments (unknown);

$X_i = (x_{ij})$ a vector of variables for the ith observation.

We assume that the X_i are generated from a finite but unknown number of segments, ordered in size from 1 to *S*. Each segment is characterized by its covariance structure and associated mixing proportion w_s. Within each segment, the variables are multivariate-normal distributed:

$$X_{i|s} \sim N(\mu_s, \Sigma_s). \tag{1}$$

Equation (1) represents Wolfe's (1970) classical mixture of multivariate normals—the most general model in this particular context (i.e., no explicit constraints are placed on estimated mean vectors or covariance matrices). We now impose a factor structure for the covariance matrix of each segment (cf., Elrod and Keane 1995), with *M* ortho-normal (i.e., unit-variance and orthogonal) common factors and J-specific factors:

$$\Sigma_S = L_S L_S' + \Phi_S. \tag{2}$$

L_S is a segment-specific J $\times$ M matrix of loadings of the J variables on the M common factors. The matrix $\phi_S = diag(\phi_{js})$ is the J $\times$ J diagonal matrix capturing the contribution of the J specific factors to Σ_S. Without any loss of generality, the matrix of factor loadings can be uniquely expressed as:

$$L_S = D_S C_S, \tag{3}$$

which, in effect, is a transformation in scale, where $D_S = diag(d_{js})$ is a J × J segment-specific diagonal matrix of standard deviations and C_S is a J × M segment-specific matrix of factor loadings for the unit-variance common factors of segment s, F^S. That is, $\Psi_S = C_S C_S'$ is the correlation matrix induced by the common factors with the J × M matrix of scores F^S in segment s. This can be seen by writing the model in Equations (1) and (2) equivalently as $X_{i|s} = \mu_S + L_S F^S + e_S$, with $F_S \sim N(0, I_M)$ and $e_S \sim N(0, \Phi_S)$. We now transform the ortho-normal factors of segment s, F^S as:

$$X^{CS} = C_S F^S, \tag{4}$$

which yields a J-dimensional vector X^{CS} of unit-variance (i.e., standardized). This transforms the model in Equations (1) and (2) into: $X_{i|s} = \mu_S + D_S X^{CS} + e_S$, with $X^{CS} \sim N(0, \Psi_S)$. It can be shown that C_S is, in fact, the matrix of cosines of the angles between X^{CS} and the axes (i.e., common factors F^S) in the reduced M-dimensional space, since $C_S = D_S^{-1} L_S$ and C_S has unit length:

$$C_S = \cos(\angle\{X^{CS}, F^S\}). \tag{5}$$

We now combine Equations (1)–(5) into a single equation:

$$X_{i|s} \sim N[\mu_s, D_s \cos(\angle\{X^{cs}, F^s\}) \cos(\angle\{X^{cs}, F^s\})' D_s + \Phi_s]. \tag{6}$$

Equation (6) completely describes a model in which each segment is characterized by the means μ_S, the common-factor variance D_S, the angles between variables and the directions of the reduced space $\angle\{X^{CS}, F^S\}$, and its specific-factors variance Φ_S. In specific applications, the parsimony of this representation is substantially increased if the equality of means and/or the equality of specific factor variances across segments (i.e., $\mu_S = \mu$, and/or $\Phi_S = \Phi$) is justified by appropriate statistical hypothesis tests. The estimated parameters of the model—in particular the w_S, μ_S, and C_S—form the basis for the graphical representation.

3.1. Estimation

It is well known that factor analytic models are not identified because the solutions remain invariant to rotation and reflection (Elrod and Keane 1995). In addition, the model described by Equation (6) is affected by indeterminacies induced by the constraint that each row vector in the matrix C_S be of unit length. This makes one column in the matrix $\angle\{X^{CS}, F^S\}$ of angles between variables and the directions of the reduced space a combination of the remaining columns. This indeterminacy is removed by estimating only the elements in the first M − 1 columns of $\angle\{X^{CS}, F^S\}$ and using them to compute the elements of the last column (based on the constraint $\Sigma_{m=1}^{M} \cos(\angle X_{jm}^{CS} F_m^S)^2 = 1$, for any $j = 1, \ldots, J$). We remove within-segment invariance to rotations by imposing that the first k of the M factors form an orthogonal base for (i.e., be ortho-normal transformations of) the first k variables, for all integer k between 1 and M. Under these conditions, the first factor must be collinear with the first variable, while the other factor(s) must be orthogonal to it. Because the second variable is contained in the plane determined by the first two factors, the other M − 2 factors must be orthogonal on this variable. In general, because the kth variable is contained in the hyperplane determined by the first k factors, the other M − k factors are constrained to be orthogonal on the kth variable. Because of these constraints and the requirement that rows in C_S be of unit length, in models with two factors M = 2, only (J − 1)(M − 1) elements in $\angle\{X^{CS}, F^S\}$ are estimated. In models with three or more factors M ≥ 3, only (J − 1)(M − 1) − (M − 1)(M − 2)/2 elements in $\angle\{X^{CS}, F^S\}$ are estimated. Finally, we remove the invariance to reflections by arbitrarily constraining the correlations between the kth variable and the kth factor to be nonnegative, for $k = 2, \ldots, M$. The constraints imposed on our model for identification are analogous to the constraints imposed on a confirmatory factor model for identification. However, there does not exist a one-to-one relationship between these sets of constraints.

The model is estimated using maximum likelihood. The conditional likelihood that observation i belongs to segment s is:

$$L_{i|s} = [(2\pi)^J|\Sigma_s|]^{-1/2}\exp\left[-\frac{1}{2}(X_i - \mu_s)'\Sigma_s^{-1}(X_i - \mu_s)\right], \quad (7)$$

where

$$\Sigma_s = D_s\cos(\angle\{X^{cs}, F^s\})\cos(\angle\{X^{cs}, F^s\})'D_s + \Phi_s. \quad (8)$$

The complete log-likelihood function for all observations and all segments becomes

$$\ln L = \sum_{i=1}^{I} \ln\left(\sum_{s=1}^{S} w_s L_{i|s}\right), \quad (9)$$

where w_S are segment weights or mixing proportions. The log-likelihood described by Equations (7)–(9) is maximized numerically under the constraints that the estimated standard deviations are positive, and segment weights are positive and sum to one.

3.2. Model Selection and Evaluation

The standard likelihood ratio test for testing the M-factor model against the M + 1-factor model is invalid in this case. Thus, we use CAIC = $-2\ln L + (\ln I + 1)P$ (Bozdogan 1987) to select both M and S. For each M, we estimate our model for increasing values of S, and deem that value of S appropriate that yields the minimum value of CAIC (Bozdogan 1987). For selection of S, similar problems of the LR test as for the selection of M occur (Aitkin et al. 1981). Because the log-likelihood function of mixtures of multivariate normal distributions is not globally concave, the estimation procedure may yield only a local maximum. Thus, for each S and M, we estimate the model 10 times from random starting points and retain the solution with largest log-likelihood. We select the solution that yields the minimum CAIC. In addition, we can also estimate reduced models with the same number of factors and segments as the retained solution, with equal variances of specific factor variances across segments (i.e., $\Phi_s = \Phi$), and use the CAIC criterion to test if they perform better than the retained solution. We would like to note that whereas the estimation of our model is feasible for many dimensions, clearly interpretation is more difficult than for a lower dimensional representation. In practical applications we can, however (at the cost of fit), always use a two-dimensional representation of the data comparable to what is generally accepted for biplots and correspondence analysis, for example.

4. Application: Ethical Pharmaceuticals

4.1. Background: The Pharmaceutical Industry

The pharmaceutical industry is driven by the need to innovate new molecular compounds. The process of developing a new product and getting it to market takes as long as 8 years and costs in excess of $500 million. At the same time, only 1 in 10 chemical compounds that are chosen for clinical testing is launched into the market. Because of the underlying discovery of a new chemical chain that offers therapeutic possibility, multiple pharmaceutical companies pursue the development of similar or related chemical compounds at the same time. Typically, several classes of ethical drugs, with several related brands within each class, are available for each disease state. Within a particular class of therapies, several (as many as 5 to 10, depending on the disease state) brands may be available that work fairly similarly, yet seem to work better in particular types of patients (for example, overweight patients, older patients, women, men, or patients with certain specific symptoms). Based on their training, experience, types of patients, and pharmaceutical marketing activities, most doctors will write prescriptions for all or most brands in a therapeutic class. At the same time, most doctors will develop certain preferences over time. So, while they write prescriptions for many products, one brand may capture a higher share of the doctor's therapeutic choices than other brands. As a result, medical doctors are likely to exhibit different prescribing patterns, not only because they treat different types of patients but also because they may come to different beliefs regarding the relative performance of the drugs for any given patient type. Thus, an analysis of physician's prescription patterns can provide valuable and timely information about the perceived efficacy of drugs and the emerging structure of competition within a given therapeutic class.

Table 1 Actual Drug Market Shares

Brand	Market Share	Description
Brand A	About 20%	Oldest class of drugs
Brand B	15–17%	Early entrants in new class
Brand C	15–17%	Early entrants in new class
Brand D	15–17%	Early entrants in new class
Brand E	5–7%	Newest brands in new class
Brand F	5–7%	Newest brands in new class
Brand G	5–7%	Newest brands in new class

4.2. A Commercial Pharmaceutical Market Research Study

A major U.S. pharmaceutical manufacturer commissioned a market research study among specialists (n = 258) (such as endocrinologists, obstetricians, psychiatrists, diabetologists, etc.) to understand their prescribing rates for various brands of ethical pharmaceuticals in a particular therapeutic class and how these rates relate to physician demographic and psychographic measures. In general, specialists tend to be more knowledgeable than general or family practice doctors about drugs, their pharmacology, and rationale for how and why the drugs work the way they do. In any given disease state, specialists tend to represent less than 20% (perhaps as little as 5%) of all doctors that prescribe a therapeutic class, but may represent as much as 50% of the prescriptions in the class.

In this study, we focus on the top seven brands that account for more than 85% of the drugs prescribed by the specialists in the sample. The brands are labeled with letters from A to G, to reflect the order of their introduction in the market (e.g., Brand C has been introduced prior to Brand D). Brand A belongs to an older class of drugs, while brand G belongs to the newest class of drugs in this particular product class. Table 1 reports the market shares of these brands in the sample.

We first calculate prescription shares across all the brands mentioned in the survey and focus on these seven brands. The prescription shares are column-standardized prior to analysis. As a basis of comparison, we contrast the results of the proposed methodology with results from the more general Wolfe's (1970) mixture of multivariate normal distributions.

Table 2 Model Selection Heuristics for Wolfe's Mixture Model

Segments	Log-likelihood	Parameters	AIC	CAIC
1	2517.23	35	5104.46	5263.81
2	2368.86	71	4879.72	5202.98
3*	2249.47	107	4712.94	5200.11
4	2207.82	143	4701.50	5352.71
5	2172.75	179	4703.50	5518.48
3a	2530.50	23	5107.00	5211.72
3b	2490.21	44	5068.42	5268.75
3c	2517.23	86	5206.46	5598.02

*Denotes minimum CAIC solution.

Because brand shares are standardized prior to the analysis, segment-specific means capture departures from the average prescription shares in the sample. For example, if the mean for Brand A's standardized market shares in segment s is positive, then Brand A's average market share in segment s is greater than its average market share in the entire sample.

4.3. The General Model: Wolfe's (1970) Pattern Clustering

We first apply Wolfe's (1970) pattern clustering finite mixture approach for $S = 1, \ldots, 5$ segments across this sample of specialists. Table 2 presents the various goodness-of-fit heuristics for this analysis. Based on CAIC, we select the solution with $S = 3$ segments. Table 3 reports the estimated means and covariance matrices for this solution, while Figure 1 displays the estimated segment means.

Specialists in the first and the largest segment (w_1 = 0.45) exhibit above average preferences for Brand A, the older (and cheaper) class of therapy, and substantially lower than average preferences for the newer class of therapies, particularly for the first ones introduced on the market (i.e., Brand B and Brand E). Note that there is a high positive covariance in prescription between Brand G and Brand C, indicating that specialists in this segment preferring C also tend to prescribe G. Specialists in the second segment (w_2 = 0.40) favor the first brands in the newer class of therapy, particularly Brand B, and have the lowest usage of Brand F and Brand G, the newest brands on the market. Brand E and Brand B prescriptions covary

Table 3 Estimates for the $S = 3$ Wolfe Solution

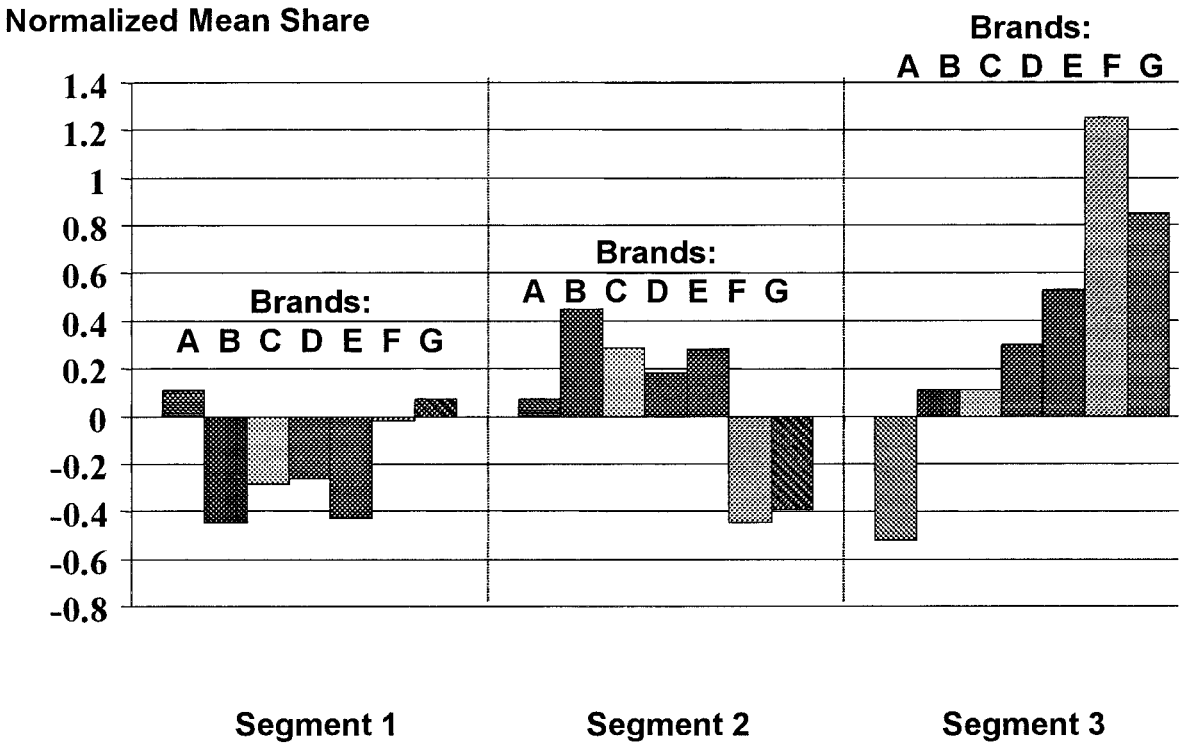
Segment (weight)	Brand A	Brand B	Brand C	Brand D	Brand E	Brand F	Brand G
				Means			
1 (0.45)	0.11	−0.45	−0.29	−0.26	−0.43	−0.02	0.07
2 (0.40)	0.07	0.45	0.29	0.18	0.28	−0.45	−0.39
3 (0.15)	−0.52	0.11	0.11	0.30	0.53	1.25	0.85
				Covariance Matrices			
1	0.66	0.05	0.09	−0.08	0.10	−0.01	0.03
	·	0.31	0.15	0.04	0.05	0.00	0.12
	·	·	0.52	−0.03	0.10	0.04	0.30
	·	·	·	0.39	0.01	0.01	0.00
	·	·	·	·	0.33	0.10	0.24
	·	·	·	·	·	0.51	0.17
	·	·	·	·	·	·	0.88
2	1.41	−0.35	−0.57	−0.32	−0.38	−0.05	−0.02
	·	1.30	−0.18	−0.18	−0.34	−0.07	−0.11
	·	·	1.43	−0.14	0.06	0.10	−0.04
	·	·	·	1.35	−0.22	−0.07	−0.06
	·	·	·	·	1.37	0.00	−0.04
	·	·	·	·	·	0.10	0.12
	·	·	·	·	·	·	0.33
3	0.60	−0.23	0.14	−0.12	−0.05	0.11	−0.37
	·	1.09	−0.25	−0.17	−0.11	−0.09	−0.64
	·	·	0.79	0.14	−0.20	−0.23	−0.19
	·	·	·	1.47	−0.44	0.17	0.17
	·	·	·	·	0.96	−0.34	0.07
	·	·	·	·	·	2.77	−1.15
	·	·	·	·	·	·	2.01

negatively, while Brand A prescriptions covary negatively with prescriptions for some of the newer brands. Finally, the third and smallest segment ($w_3 = 0.15$) has a higher than average preference for Brand F and Brand G, the most recently introduced brands, uses less of the earlier brands in the newer therapeutic class, and has the lowest usage of the older class of drugs. In fact, the usage level in this segment, relative to sample averages, correlates very highly with the novelty of the drugs. It is also important to note the strong negative correlation between Brand F and Brand G.

Thus, the three segments tend to group specialists with respect to the stage of adoption of innovation in this therapeutic category. The specialists in the first group appear to be laggards because they prescribe more of the older class. However, they also have a higher than average preference for Brand G, a newer and somewhat cheaper brand. This suggests that some of the specialists belonging to this segment may be price sensitive, while others may exhibit a slower adoption cycle, replacing the older class with the newer brands and, thus, skip one stage in the cycle of innovation. The specialists in the second segment are heavy users of the newer class but are not particularly fast to adopt the latest brands. Finally, the last segment clearly consists of innovators. Traditionally, pharmaceutical marketers have viewed specialists in one of two extremes: all specialists are the same (i.e., the market has only one segment) or all specialists are very different (i.e., the market consists of 10,000+ segments of one physician each). Not surprisingly, this analysis suggests a more moderate perspective: specialists adopt new products at different rates. Knowing brand preferences, and being able to visually depict them as in Figure 1, is insightful and tells a complete story about how to affect those preferences and which products are likely to be used as substitutes. This secondary information is available in the covariance matrices in Table 3. Unfortunately, marketing managers often have difficulty interpreting these numeric matrices to meaningfully impact their decision making. So, while Wolfe's (1970) model does provide some insight into the market structure, it does not allow managers the full utility of the marketing research data/analyses they have procured. Clearly, as the number of brands here increases, such interpretations of the covariance matrix become increasingly problematic!

We also examine three externally constrained Wolfe (1970) three-segment solutions as displayed near the bottom of Table 2. Solution 3a reflects the various model selection heuristics for the solution where only means were estimated and not covariances (here, all segment covariance matrices were set to the identity matrix). Solution 3b displays the same type of results for the situation where only segment level means and variances were estimated (e.g., the segment level covariance matrices estimated were constrained to be diagonal). Finally, Solution 3c denotes similar types of statistics for the case where all the segment means were fixed at zero and the segment covariance matri-

Figure 1 Graphic Representation of $S = 3$ Wolfe Solution



ces were free to vary. As shown by the corresponding CAIC statistics, the three-segment unconstrained solution presented in detail above dominates all three of these constrained solutions. Thus, it appears that means and covariances are *both* necessary for three segments to adequately describe the structure in this particular application.

4.4. The Proposed Spatial Analysis

The proposed spatial latent structure methodology is nested in Wolfe's (1970) normal mixture model, with segment covariance matrices reparameterized as functions of angles in a reduced space. In an attempt to derive the most parsimonious model, we exploit the flexibility of the proposed method and arrive at the best solution in multiple stages. In the first stage, we run the procedure for an increasing number of dimensions for the reduced space. Based on CAIC, the solutions with one and two dimensions are superior to Wolfe's (1970) solution and, thus, are retained for further analysis (among these, the two-dimensional solution is best). In the second stage, we explore if additional constraints across segments can further increase the parsimony of any of these two solutions. The constraint $\Phi_S = \Phi$ is found to lower CAIC for the two-dimensional solution. Thus, the proposed procedure identified three solutions that are substantially more parsimonious than, and statistically superior to Wolfe's (1970) solution (in terms of CAIC). In the final stage, we impose additional restrictions on each of these solutions to ensure that, in each case, the elements of the resulting spatial representation are statistically significant. Specifically, all the standard deviations along common factors that are not statistically different from zero in the previous stage *and* their corresponding angles are constrained to zero. Table 4 reports the goodness-of-fit measures for all the relevant models during this model selection process. Table 5 reports the coefficient estimates, and Figures 2, 3, and 4 provide a graphical representation of the estimated means and angles of the retained ''best'' solution.

This best solution has two dimensions for the reduced space, equal variances of specific factors across

Table 4 Model Selection Heuristics for Proposed Model with 3 Segments

Dimensions	Constraints across Segments	Constraints on Nonsignif. Std. Devs. within Segments	Parameters	Log-Likelihood	AIC	CAIC
1	No	No	65	2334.21	4798.42	5094.36
2	No	No	83	2296.50	4759.00	5136.90
3	No	No	101	2272.15	4746.30	5206.15
2b	$\Phi_s = \Phi$	No	69	2332.64	4803.28	5117.43
1	No	Yes	45	2345.12	4780.24	4985.13
2	No	Yes	68	2299.29	4734.59	5044.18
2b*	$\Phi_s = \Phi$	Yes	43	2350.97	4787.95	4983.72

*Denotes minimum CAIC solution.

segments $\Phi_S = \Phi$ and many variances along common factors set to zero. It is characterized by only 43 parameters, as opposed to 107 parameters in the solution based on Wolfe's (1970) procedure. Similar to Wolfe's (1970) first segment, the specialists in the first segment ($w_1 = 0.50$) exhibit significantly higher than average prescription shares for Brand A, the older class of therapy, and below average preferences for all other drugs. The common factor structure captures a pattern of significant covariance for three of the seven drugs: Brand C, Brand G, and Brand A, explaining 44.4%, 51.6%, and 34.8%, respectively, of each drug's within-segment variance in standardized market shares. Within this reduced space, the spatial representation of the covariance structure reveals that Brand C and Brand G are prescribed in a very similar manner (the differences between their angles are not statistically significant) and almost orthogonal to Brand A (the angles are greater than, but not statistically different from, 90°). When the reduced-space structure is not statistically significant (i.e., all standard deviations along common factors are not statistically different from zero), the covariance structure in the full seven-dimensional space is orthogonal. Thus in this segment, the only departure from a fully orthogonal structure is that Brand C and Brand G are positively correlated. Based on point estimates, this correlation equals 0.47, and the angle between Brand C and Brand G in the seven-dimensional space equals 62°. Thus, specialists in this segment prescribing Brand C have a relatively strong tendency to also prescribe Brand G, which was introduced to the market at a later time.

The specialists in the second segment ($w_2 = 0.27$) have substantially and significantly higher than average prescription shares for the newer brands in the new class of therapy, Brand F and Brand G, and a significantly lower-than-average preference for Brand B. Interestingly, the common factor structure captures a pattern of covariance only for Brand F and Brand G, which explains 88.7% and 84.2%, respectively, of each drug's market share variance. Because Brand F and Brand G are negatively correlated in the reduced space (corr = −0.57), they are also negatively correlated in the full space (corr = −0.42), where they make an angle of 115 degrees. Thus, relative to segment averages, those specialists prescribing Brand F more tend to prescribe Brand G less, and vice versa. The marketing implications in Segment 2 are even more interesting than for Segment 1. Although both Brands F and G are clearly more preferred by Segment 2, increasing share for either drug will come at the expense of the other brand. Clearly, these two brands compete strongly against each other in this segment. The marketing manager who assumes this segment as merely innovative runs the risk of focusing on the wrong competitors if she or he positions only against the older products. The advantage of our new methodology is that the marketing manager can quickly see the trade-offs occurring between Brands F and G without having to interpret the full covari-

Table 5 Estimates for the S = 3, M = 2 Solution for the Proposed Model

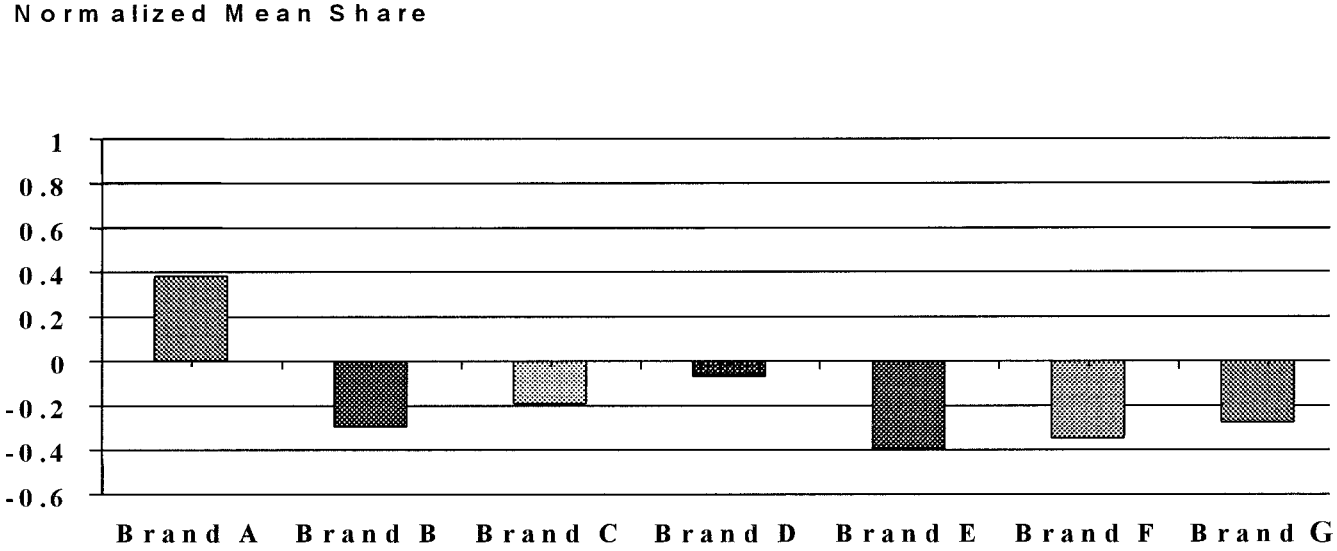
	Segment Weight	1 0.50 (0.05)	2 0.27 (0.04)	3 0.23 (0.03)			
Segment	Brand A	Brand B	Brand C	Brand D	Brand E	Brand F	Brand G
				Means			
1	0.38 (0.14)	−0.30 (0.08)	−0.19 (0.12)	−0.07 (0.12)	−0.40 (0.10)	−0.35 (0.06)	−0.28 (0.09)
2	−0.24 (0.14)	−0.28 (0.08)	−0.14 (0.10)	0.06 (0.16)	0.19 (0.11)	0.90 (0.31)	0.98 (0.20)
3	−0.53 (0.17)	0.96 (0.25)	0.57 (0.26)	0.08 (0.18)	0.63 (0.23)	−0.30 (0.09)	−0.55 (0.08)
				Standard Deviations along Common Factors			
1	0.60 (0.16)	0 (—)	0.58 (0.12)	0 (—)	0 (—)	0 (—)	0.47 (0.11)
2	0 (—)	0 (—)	0 (—)	0 (—)	0 (—)	1.32 (0.16)	1.06 (0.12)
3	0 (—)	1.09 (0.20)	1.16 (0.14)	0 (—)	0.99 (0.25)	0 (—)	0 (—)
				Angles (in Degrees)			
1	0 (—)	—	112.0 (21.0)	—	—	—	105.1 (14.3)
2	—	—	—	—	—	0 (—)	—124.6 (11.7)
3	—	0 (—)	—125.2 (11.1)	—	141.9 (13.8)	—	—
				Standard Deviations of Common Factors (Equal across Segments)			
	0.82 (0.06)	0.66 (0.05)	0.65 (0.06)	1.00 (0.05)	0.77 (0.05)	0.47 (0.04)	0.46 (0.06)
				Resulting Covariance Matrices			
1	1.03	0.00	−0.13	0.00	0.00	0.00	−0.07
	.	0.44	0.00	0.00	0.00	0.00	0.00
	.	.	0.76	0.00	0.00	0.00	0.27
	.	.	.	1.00	0.00	0.00	0.00
	.	.	.	.	0.59	0.00	0.00
	.	.	.	.	.	0.22	0.00
	.	.	.	.	.	.	0.44
2	0.67	0.00	0.00	0.00	0.00	0.00	0.00
	.	0.44	0.00	0.00	0.00	0.00	0.00
	.	.	0.42	0.00	0.00	0.00	0.00
	.	.	.	1.00	0.00	0.00	0.00
	.	.	.	.	0.59	0.00	0.00
	.	.	.	.	.	1.96	−0.79
	.	.	.	.	.	.	1.34
3	0.67	0.00	0.00	0.00	0.00	0.00	0.00
	.	1.64	−0.73	0.00	−0.85	0.00	0.00
	.	.	1.77	0.00	−0.06	0.00	0.00
	.	.	.	1.00	0.00	0.00	0.00
	.	.	.	.	1.57	0.00	0.00
	.	.	.	.	.	0.22	0.00
	.	.	.	.	.	.	0.21

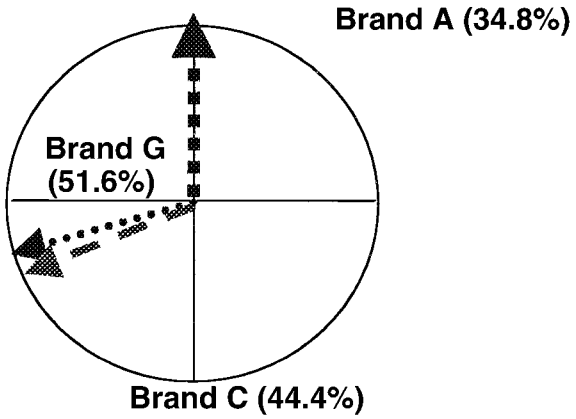
* Numbers in parentheses represent the computed standard errors.

ances matrices in the previous two methods individually.

Finally, the specialists in the last segment (w_3 = 0.23) have substantially and significantly higher-than-average preferences for the earlier brands in the newer class of therapy: Brand B, Brand C, and Brand E. They also exhibit lower-than-average preferences for the newest brands in this class, Brand F and Brand

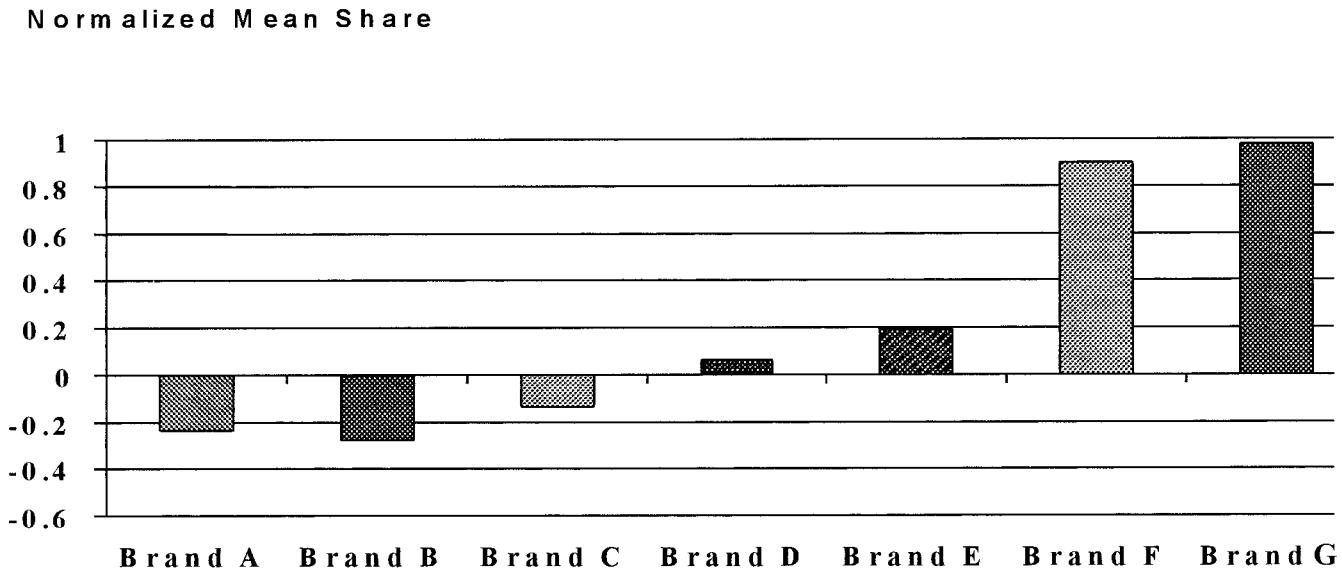
Figure 2 Mean Shares and Derived Common Factor Structures for Segment 1

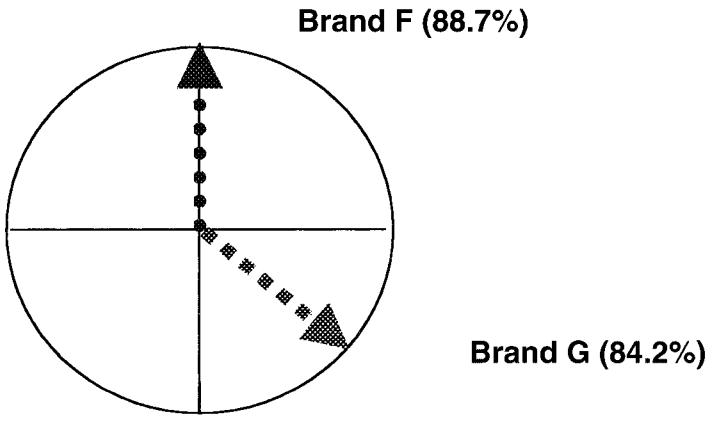




Note: Reported in parentheses is the percent of total variance in brand share explained by the common factors structure.

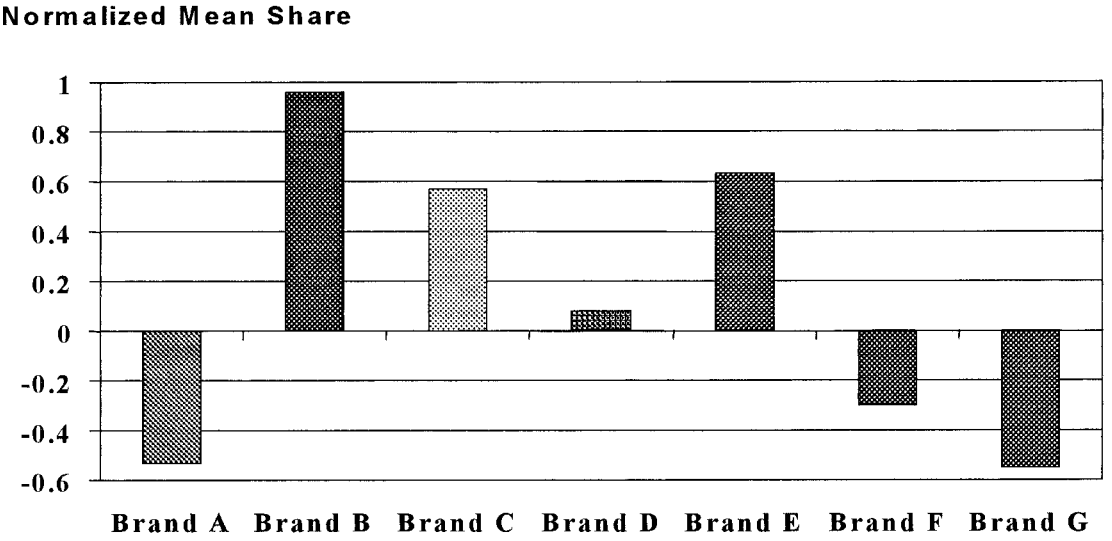
Figure 3 Mean Shares and Derived Common Factors Structures for Segment 2

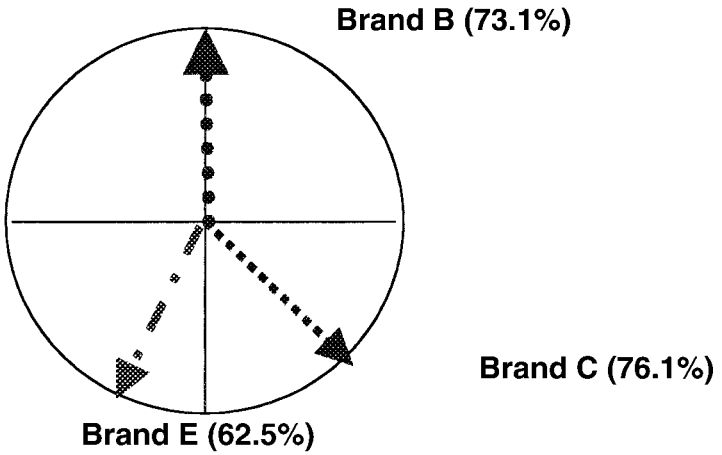




Note: Reported in parentheses is the percent of total variance in brand share explained by the common factors structure.

Figure 4 **Mean Shares and Derived Common Factor Structures for Segment 3**





Note: Reported in parentheses is the percent of total variance in brand share explained by the common factors structure.

G, and for Brand A. The two-factor structure captures a pattern of covariance for Brand B, Brand C, and Brand E and explains 73.1%, 76.2%, and 62.5%, respectively, of each drug's market share variance within this segment. The structure is symmetrical, yielding negative correlations between Brand B and Brand C (corr = −0.57) and between Brand B and Brand E (corr = −0.79), while Brand C and Brand E remain orthogonal (corr = 0.00). This structure indicates that Brand C and Brand E are each substitutes for, and thus compete directly with, Brand B, but they do not compete directly with each other. Reflecting back to Table 1, Brands B, C, and D are the market leaders. The brand manager for Brand B is likely to assume that the key competitors are either the other market leaders, Brands C and D, or the latest entrants, Brands F and G. What the graphs in Figure 4 suggest is that for those physicians who clearly prefer Brand B, not only is one of the other market leaders a key competitor but another product that might not even have been considered is a key competitor: Brand E. Moreover, the newest products are not critical competitors for Brand B.

This retained solution shows some congruence with the Wolfe (1970) solution. Segments 1, 2, and 3 of the retained solution are comparable to Segments 1, 3, and 2 in Wolfe's (1970) solution in terms of order of within-segment normalized prescription shares for brands. The implied covariances among the drug prescriptions found with the proposed procedure also tend to appear in Wolfe's (1970) solution. It is interesting to note that all solutions capture the same negative correlation between Brand F and Brand G in the segment where they exhibit higher-than-average prescription shares. Yet, the solution from our proposed spatial model presented above dominates that of Wolfe (1970), according to the CAIC heuristic, plus makes the results so much easier to communicate in parsimoniously revealing the underlying structure in the data! Rather than having to process cumbersome numerical parameter estimates obtained from full estimated covariance matrices, managers can quickly inspect the spatial representations as shown in Figures 2–4 to assess the structure of mean prescription shares, as well as the covariance of these brand shares by derived market segments.

A posterior analysis is performed utilizing correlation and regression on logit transformed posterior probabilities of membership for each derived market segment to relate the posterior probabilities of the retained solution to various descriptor variables collected during the study. From this analyses, Segment 1 specialists tend to be general specialists, institution-based, high-volume users with large clienteles, they are open to input regarding therapy, and they do not tend to change dosages for refills. Brand A tends to be utilized where costs are a concern, such as in medical institutions. Segment 2 specialists are office-based, private practice MDs. They deal with older patients and are more likely to change dosages. They are innovators, which explains the high Brand F and Brand G usage. Finally, specialists in Segment 3 tend to be child specialists with private practices located in large urban areas. They are low-volume prescribers in this therapeutic category, which suggests that their clientele is smaller, and they adhere to more conservative prescription policies.

5. Discussion

There are a number of important features of the proposed methodology. First, it provides a simple and flexible way to impose within- and between-segment constraints on segment-specific covariance matrices. As a consequence, it is effective in identifying parsimonious solutions to segmentation problems that can be easily interpreted. Second, when the common factor structure has three dimensions or less, the method also provides a parsimonious graphical display of the data. Researchers seeking to apply our procedures are free to choose a two-dimensional representation (at the cost of a reduction in fit) that provides a very parsimonious graphical representation, similar to that used in biplots or correspondence analysis. The multivariate covariance structure induced by common factors is intuitively represented terms of angles between vectors representing the variables (brands). Thus, rather than needing to peruse several tables with parameter estimates, managers can very quickly grasp the structure of the data and the implications for strategy. As a result of using this methodology,

marketing practitioners can quickly identify both customer target segments where their products are/are not preferred, and which products need to be marketed against to improve overall market performance. Moreover, the available graphical displays make it easier to communicate these strategic issues to others in their organization. It is our conjecture that these graphical displays will have a much more profound impact on strategy given the great capacity that humans have for fast processing and accurate recall of graphical information. The proposed graphical display is very general and allows for the representation of a wide range of data used to support strategic decisions including market segmentation, customer satisfaction, competitive positioning, and customer value perception. In addition, in such applications, identifying restrictions may be derived from substantive theory, similar to confirmatory factor analysis.

We believe that our contribution to the literature is mostly technical. The issue of identifying factor models in different latent groups has not been dealt with before in the way we address it. Our method enables managers to quickly grasp the basic structure of a dataset visually. The structure that we identify may have interesting substantive interpretation in many applications. Graphical representation of data hinges strongly upon data reduction techniques. We attempt to reduce the dimensionality of a two-way, two-mode dataset in two directions simultaneously: We reduce the rows of the dataset through the assumption of discrete latent segments and the columns through the assumption of continuous latent variables. Most of the interpretation of the latent structure of marketing data to date has been based on one of those two modes: either a factor model to reduce the variables dimension, or a latent class model to reduce the subjects dimension.

The power of factor models to provide a representation that lends itself to interesting substantive interpretations, based on the covariance pattern in the data, is well established. The correlation structure of the variables is reduced and depicted in a way that allows for easy interpretation. Although a similar interpretation can be obtained from the original segment level correlation matrices derived from Wolfe's (1970) model, we would like to emphasize that the three circular graphs that we provide are much easier to interpret and the results can be communicated much faster through these graphs than through the original correlation matrices. In the case of our drug prescription application, obviously the differences in mean prescription rates between the segments are important for substantive interpretation. Different groups of physicians have different overall tendencies to prescribe drugs. Our model offers the advantage of conveniently depicting the basic structure of the within segment covariance matrix, and in the application displays which drugs, within a segment of physicians, tend to be prescribed together. If a marketer wants to target the particular segment in question, that could give important clues for the bundling and cross-selling strategies of drugs.

References

Aitkin, M., D. Anderson, J. Hinde. 1985. Statistical modeling of data on teaching style (with discussion). *J. Roy. Statist. Soc.* **A144** 419–461.

Bockenholt, Ulf, Ingo Bockenholt. 1991. Constrained latent class analysis: simultaneous classification and scaling of discrete choice data. *Psychometrika* **56** 699–716.

Bozdogan, Hamparsum. 1987. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika* **52** 345–370.

Day, George S. 1994. The capabilities of market driven organizations. *J. Marketing* **47** 79–89.

DeSarbo, Wayne S., D. J. Howard, Kamel Jedidi. 1990. MULTICLUS: A new method for simultaneously performing multidimensional scaling and cluster analysis. *Psychometrika* **56** 121–136.

——, Kamel Jedidi, K. J. Cool, P. Schendel. 1991. Simultaneous multidimensional unfolding and cluster analysis: an investigation of strategic groups. *Marketing Lett.* **2** 129–146.

——, Ajay K. Manrai, Lalita A. Manrai. 1994. Latent class multidimensional scaling: A review of recent developments in the marketing and psychometric literature. R. P. Bagozzi, *Advanced Methods of Marketing Research.* 190–222. Cambridge, MA, Blackwell.

Dolan, Conor V., Han L. J. van der Maas. 1998. Fitting multivariate normal finite mixtures subject to structural equation modeling. *Psychometrika* **63** 227–253.

Elrod, Terry, Michael P. Keane. 1995. A factor-analytic probit model for representing the market structure. *J. Marketing Res.* **32** 1–16.

Hinton, Geoffrey E., Peter Dayan, Michael Revow. 1997. Modeling the manifold of images of handwritten digits. *IEEE Trans. Neural Networks* **8** 65–74.

Jedidi, Kamel, Harsharanjeet S. Jagpal, Wayne S. DeSarbo. 1997. Finite mixture structural equation models for response based

segmentation and unobserved heterogeneity. *Marketing Sci.* **16** 39–59.

Jöreskog, Karl G. 1971. Simultaneous factor analysis in several populations. *Psychometrika* **36** 409–426.

Kamakura, Wagner A., Gary J. Russell. 1989. A probabilistic choice model for market segmentation and elasticity structure. *J. Marketing Res.* **26** 379–390.

Li, Tiger, Roger J. Calantone. 1998. The impact of market knowledge competence on new product advantage: Conceptualization and empirical examination. *J. Marketing* **62** 13–29.

Spence, Ian, Stephan Lewandowski. 1990. Graphical perception. In *Modern Methods of Data Analysis.* John Fox, J. Scott Long, eds. Sage, London. 1–57.

Tipping, Michael E., Christopher M. Bishop. 1999. Mixtures of probabilistic principal component analyzers. *Neural Comput.* **11** 443–482.

Wedel, Michel, Wayne S. DeSarbo. 1996. An exponential family multidimensional scaling mixture methodology. *J. Bus. Econom. Statist.* **144** 447–459.

——, Wagner A. Kamakura. 1999. *Market Segmentation: Conceptual and Methodological Foundations.* Kluwer, Dordrecht. The Netherlands.

Wolfe, John H. 1970. Pattern clustering by multivariate mixtures. *Multivariate Behavioral Res.* **5** 329–350.

This paper was received June 4, 1999, and was with the authors 13 months for 4 revisions; processed by Wagner Kamakura.