Constructing MDS Joint Spaces from Binary Choice Data: A Multidimensional Unfolding Threshold Model for Marketing Research
Author(s): Wayne S. DeSarbo and Donna L. Hoffman
Source: *Journal of Marketing Research*, Vol. 24, No. 1 (Feb., 1987), pp. 40-54
Published by: American Marketing Association
Stable URL: http://www.jstor.org/stable/3151752
Accessed: 01/09/2009 16:25

WAYNE S. DeSARBO and DONNA L. HOFFMAN*

The authors present a new multidimensional unfolding methodology that can analyze various types of individual choice data. The model represents choice data, defined by dichotomous variables that indicate whether a particular brand was chosen or not, in terms of a joint space of consumers and brands. Explicit treatment of marketing and subject background variables is allowed through optional model reparameterizations of consumers and brands. Together with the joint space representation of both consumers and brands, these optional reparameterizations can provide information on appropriate market segmentation bases and respective product positioning strategies. The authors apply this spatial choice model to data on consumer (intended) choices for 12 residential communications devices and demonstrate how the results can be used for optimal positioning decisions.

Constructing MDS Joint Spaces from Binary Choice Data: A Multidimensional Unfolding Threshold Model for Marketing Research

Multidimensional scaling (MDS) has been an area of prolific marketing research over the past two decades (cf. Green 1975; Cooper 1983). It has been used to investigate problems in market segmentation (Johnson 1971), positioning (DeSarbo and Rao 1986; Wind 1982), competitive market structure (Day, Shocker, and Srivastava 1979; Shocker and Stewart 1983; Srivastava, Alpert, and Shocker 1984), consumer preferences/perceptions (DeSarbo and Carroll 1985; Green and Carmone 1970; Green and Rao 1972), etc. However, as noted by Holbrook, Moore, and Winer (1982), in certain cases traditional MDS procedures are inappropriate or impractical. One such case is the analysis of binary (0,1) data, which may arise, for example, when marketing researchers collect actual or intended choice data from a number of actual or potential consumers evaluating a set of competitive brands. Traditional metric MDS procedures for analyzing such two-way (e.g., subjects × brands) choice data, such as MDPREF (Carroll 1972), PREFMAP (Carroll 1972), and GENFOLD2 (DeSarbo and Rao 1984, 1986), would attempt to solve for row (subject) and column (brand) coordinates whose distances or scalar products would "best match" the binary input data. Many of these procedures thus optimize a loss function comparing the actual binary data with predicted continuous values. As a result, some of the predicted values may be outside the 0, 1 range. Because the predicted values are not constrained to be probabilities, it is also not clear how one interprets these values. Many of these procedures could be modified so that the predicted values would be in the 0, 1 range, but the imposition of such constraints would entail the use of very cumbersome estimation algorithms and severe problems with locally optimum solutions. Nonmetric MDS procedures such as KYST (Kruskal, Young, and Seery 1973) and ALSCAL (Takane, Young, and de Leeuw 1977) applied to such two-way binary data could also encounter problems. There would be a large number of tied ranks for each subject, and it is not clear how such ties would affect the resulting solutions.

We develop a multidimensional unfolding threshold

*Wayne S. DeSarbo is the Howard and Judy Berkowitz Associate Professor of Marketing, Wharton School, University of Pennsylvania. Donna L. Hoffman is Assistant Professor of Marketing, Graduate School of Business, Columbia University.

Professor Hoffman gratefully acknowledges support from the Faculty Research Fund of the Graduate School of Business, Columbia University. The authors thank the editor and three anonymous *JMR* reviewers for their insightful comments that helped improve the article.

Journal of Marketing Research
Vol. XXIV (February 1987), 40–54

methodology purposely designed to accommodate the spatial analysis of binary data. After reviewing the recent MDS literature on procedures for analyzing such binary data, we describe the proposed MDS unfolding threshold model and its many features. We illustrate the model with intention-to-buy binary choice data for potential consumers choosing among actual competing brands of a residential communication device. The results are extended to show how optimal positioning/repositioning strategies can be developed. Finally, we discuss other potential applications and directions for future research.

REVIEW OF MDS PROCEDURES FOR BINARY DATA ANALYSIS

Literature Review

Several multidimensional scaling (MDS) methods have been suggested in the psychometric literature to accommodate the spatial analysis of two-way "binary" choice data. Torgerson (1958) presented deterministic conjunctive and disjunctive models of choice and corresponding spatial methods that operate on binary choice data. He also generalized Thurstone's (1927, 1929) method of similar reactions in obtaining a matrix of "directed distances," which is converted into scalar products and then factor analyzed.

Most nonmetric multidimensional scaling methods (e.g., Kruskal 1964) and nonmetric factor analysis (Kruskal and Shepard 1974) estimate a joint space of subjects and objects in which choices would, for nonmetric unfolding MDS procedures, correspond to short distances between object and ideal points whereas nonchoices would correspond to long distances. For nonmetric MDS vector methods (e.g., PREFMAP2–Model 4) and nonmetric factor analyses, choices would correspond to larger scalar products. In either case, one would find a large number of ties in the data, and it is not clear what effect they would have on subsequent solutions.

Binary factor analysis methods (Bartholomew 1980; Christoffersson 1975; Muthen 1978, 1981) extend classical factor analysis to the analysis of binary data. Maximum likelihood estimation of the factor analysis model for dichotomous variables initially was treated by Bock and Lieberman (1970) for the special case of only one factor. Christoffersson (1975) generalized this model to multiple factors and presented a simpler approach to estimation using response information that considers only the first-order marginal distributions. Muthen (1978) extended Christoffersson's approach by proposing a generalized least squares estimator that is computationally more efficient. Both Christofferson and Muthen propose the use of a latent unobservable variable ("response strength") and threshold values in a stochastic framework to derive factor loadings and an error covariance matrix. Bartholomew (1980) presents a latent-structure-like approach for the analysis of categorical data using a logit response function.

Correspondence analysis (Benzécri et al. 1973) is perhaps the most popular type of scaling procedure for analyzing such data. The method typically handles aggregate choice data in the form of a frequency matrix and derives a joint space of row and column objects based on an eigenstructure analysis of the normalized frequency matrix. It also can be used to analyze the raw binary data. Greenacre (1984) and Lebart, Morineau, and Warwick (1984) discuss extensions of the procedure to map external information (e.g., brand features, subject demographics) via "supplementary points analysis." Hoffman and Franke (1986) discuss the various uses of correspondence analysis and provide several marketing applications. Other "optimal scaling" approaches such as dual scaling methods (Nishisato 1980) and homogeneity analysis (de Leeuw 1973; Gifi 1981a,b; Heiser 1981) have been shown (Tenenhaus and Young 1985) to be equivalent to correspondence analysis. In fact, Levine's (1979) centroid method also can be viewed as a special case of correspondence analysis where a different type of normalization is employed. Green and DeSarbo (1980) and Holbrook, Moore, and Winer (1980, 1982) present applications of Levine's model in marketing.

Takane (1983) describes a method for the analysis of "pick any/n" data. Such data are viewed as a special type of successive categories data in which there are only two response categories. Here, each item (stimulus), rather than each category of an item, may be represented as a point, and the subjects are assumed to choose (or not to choose) the item according to its closeness to their respective ideal points. Takane develops an EM algorithm (cf. Dempster, Laird, and Rubin 1977) to maximize a marginal likelihood to estimate the joint space. His model is somewhat similar to a special case of one of the various unfolding models developed here.

Table 1 is a comparison of these eight particular approaches (we discuss the ninth approach shortly) to the analysis of binary choice data. The columns of Table 1 correspond to the following set of features.

1. Whether or not the data are assumed to have been generated from some stochastic process or choice rule. Within a few general approaches, there are deterministic and stochastic methods.
2. What type of spatial representation is rendered for subjects and/or products. Typically, factor analytic methods involve an analysis or decomposition of some form of scalar products matrix, which implies the projection of points onto vectors. Other methods, such as correspondence analysis and many forms of optimal scaling procedures, have been shown (Gifi 1981a,b; Heiser 1981) to be related to special forms of unfolding analysis whereby both subjects and products are represented by points. Finally, unfolding methods represent subjects as ideal points and products as points.
3. How the appropriate dimensionality is selected. This tends to be related to the deterministic/stochastic assumption. In deterministic models, *ad hoc* rules involving scree plots, interpretation, and parsimony are used to identify the appropriate dimensionality. In many stochastic approaches

Table
COMPARISON OF VARIOUS PSYCHOMETRIC APPROACHES

	Model	Type of representation	Dimensionality identification
Torgerson's (1958) methods	Deterministic or stochastic	Points	Ad hoc
Nonmetric MDS (Kruskal 1964)	Deterministic	Ideal points, vectors	Ad hoc
Nonmetric factor analysis (Kruskal and Shepard 1974)	Deterministic	Vectors	Ad hoc
Binary factor analysis (Christoffersson 1975)	Stochastic	Vectors	χ^2 test
Correspondence analysis (Benzécri et al. 1973)	Deterministic	Points	Ad hoc and χ^2 test if contingency table
Optimal scaling approaches (Nishisato 1980)	Deterministic or stochastic	Points	Ad hoc and χ^2 tests
Levine's (1979) centroid method	Deterministic	Points	Ad hoc
Takane's (1983) method	Stochastic	Ideal points	Ad hoc and χ^2 tests
Multidimensional unfolding threshold model	Stochastic	Ideal points	χ^2 tests

using maximum likelihood methods, asymptotic statistical tests (e.g., χ^2 tests) can be implemented in addition to these other methods to establish the appropriate dimensionality.

4. The type of data the model can accommodate. Some of the models discussed can handle any type of binary data (e.g., pick $1/n$, pick any$/n$, pick k/n, etc.). Some versions, such as Takane's method, typically are applied to pick any$/n$ data.
5. How the effect of marketing mix decisions and/or product features is reflected in the model. Nearly all of the techniques in Table 1 allow the use of correlation or regression property fitting procedures after an analysis is completed to interpret the dimensions extracted for the product/brands. For these cases, the user investigates the relationships between these features and/or marketing mix variables (e.g., price) and the extracted dimensions. This procedure not only aids intepretation, but also examines the impact of features and/or marketing mix decision variables such as price on each of the derived dimensions. One problem is that if the dimensions or coordinates for products/brands are not fully explainable in terms of these features/marketing mix variables, such analyses may prove fruitless. Many general MDS procedures (e.g., GENFOLD2 by DeSarbo and Rao 1984, 1986; DeSarbo, Oliver, and De Soete 1986) allow for brand (subject) coordinates to be an exact user-specified function of designated features (background variables) and/or marketing mix variables. This option allows for the simultaneous estimation of the joint space and "impact" coefficients relating these features and variables to the brand (subject) locations. There is no error component relating brand (subject) coordinates to features (background variables). (Only one optimal-scaling-like procedure entitled OVERALS—Gifi 1981a—can be easily modified to allow for such reparameterizations.) Given this option, once a function is specified for such a reparameterization (discussed subsequently) and coefficients estimated, the marketer can directly link such features/marketing mix decisions to locations in the derived space. The estimated coefficients would indicate the importance or impact of the feature/marketing mix variable on a dimension and aid in interpretation (cf. Bentler and Weeks 1978; Bloxom 1978; de Leeuw and Heiser 1980; Noma and Johnson 1977).
6. How the effect of subject background information (e.g., demographic, psychographic, etc.) is reflected in the model. The preceding discussion of property fitting versus reparameterization also applies here. Reparameterization options here allow for subject coordinates to be specified functions of selected demographic, psychographic, and/or other relevant individual background data. As before, reparameterization coefficients are estimated that indicate the importance or impact of a particular background variable on a particular dimension. This aids in interpreting the obtained solution in terms of subject characteristics, which is useful for segmentation and classification.
7. How the user can make predictions for subjects and/or products/brands not utilized in the analysis. Often, it is desirable to examine the position that a new subject/product would have in an MDS space derived from other subjects/products being tested. Where property fitting techniques such as regression are employed in 5 and 6, one can obtain predicted locations (with regression model error) for these subjects and products. Where product and/or subject reparameterization (product/subject coordinates are some specified function of designated variables) is allowed, one can obtain predictions (without property fitting error) on subsequent coordinates.
8. What type of utility theory underlying choice is suggested by the model. Basically, nearly all of the techniques described in Table 1 are data analytic procedures that provide for a particular type of spatial representation to the rows and/or columns of the input data matrix. Items selected by a subject would appear close to the subject point in an unfolding type model, or project very positively on a subject vector in a scalar products model. Only one model is based on any underlying choice (utility) theory positing how choices are made, where the error distribution can be interpretable as a distribution of underlying utility or preferences among consumers (cf. Chow 1983).
9. Whether there are alternative models and if the "best" one can be selected. Most of the alternative methods listed in Table 1 allow for the fitting of one basic model. With

1
TO THE ANALYSIS OF BINARY CHOICE DATA

Type of data	Marketing mix	Subject information	Predictions	Theory of choice	Nested models & tests
Binary	Property fit	Property fit	Property fit methods	Yes	None
Binary	Property fit	Property fit	Property fit methods	None	None
Binary	Property fit	Property fit	Property fit methods	None	None
Binary	Property fit	Property fit	Property fit methods	None	None
Binary, contingency table frequencies	Supplementary points analysis	Supplementary points analysis	Supplementary points	None	None
Binary, contingency table frequencies	Property fit or reparameterize	Property fit or reparameterize	Property fit methods or via reparameterization	None	Nested models
Pick any	Property fit	Property fit	Property fit methods	None	None
Pick any, pick any/n	Property fit	Property fit	Property fit methods	Yes	None
Binary	Reparameterize	Reparameterize	Via the reparameterizations	Yes	Nested models and tests

modification, one could allow for the reparameterization of subject and/or product coordinates via relevant background variables in one or two of these methods. However, it is not clear how the user can select a "best" model when options are provided to fit a number of related alternative models.

Research Objectives

Our goal is to develop a new MDS-based choice model for the estimation of a joint space of subjects and objects from binary choice data appropriate for marketing problems involving competitive market structure, positioning/repositioning, and market segmentation. For the analysis of such binary choice data, a new spatial choice model is proposed that eliminates some of the limitations in the current psychometric (and econometric) approaches.

In terms of the features and benefits in Table 1, our spatial choice model can be described as (1) a stochastic model (2) utilizing an unfolding representation, (3) allowing for an asymptotic statistical test for identifying the appropriate dimensionality and for nested models testing, (4) accommodating virtually any type of binary data, but especially pick any and pick any/n, (5) having reparameterization options for subject and/or product coordinates that allow for the prediction of subjects and/or products not used in the analysis, and (6) based on utility theory via an indirect utility function (cf. McFadden 1976) that posits choice as a function of a latent variable and an error function. We believe it is an improvement over most current procedures in that its structure is more related to the actual choice process (rather than merely another data analytic procedure), the reparameterization options allow for the direct and simultaneous investigation of features, marketing mix, and psychographic as well as other variables, and, as we demonstrate, the proposed method allows for validation in a simple, straightforward manner. Finally, the new method can be augmented easily with postanalytic optimization routines to examine issues in optimal positioning, as we demonstrate.

MULTIDIMENSIONAL UNFOLDING THRESHOLD METHOD

The Model(s)

Let:

$t = 1 \ldots T$ dimensions (extracted in an MDS context),
$i = 1 \ldots I$ respondents,
$j = 1 \ldots J$ products,
$k = 1 \ldots K$ product features and/or marketing mix variables (e.g., price, brand name, physical features objectively measured, etc.),
$l = 1 \ldots L$ respondent characteristics (e.g. demographic and/or psychographic background variables),
y_{ij} = 1 if respondent i chooses product j, 0 otherwise;
x_{jk} = the value of feature k for product j,
z_{il} = the value of characteristic l for respondent i,
p_{ij} = the probability respondent i chooses product j,
a_{it} = the t^{th} coordinate for respondent i,
w_{it} = the salience of dimension t for respondent i,
c_i = an additive constant for respondent i, and
b_{jt} = the t^{th} coordinate for product j.

We define a latent, unobservable "disutility" variable D_{ij} such that

$$D_{ij} = \sum_{t=1}^{T} w_{it}(a_{it} - b_{jt})^2 + c_i + \mu_{ij} \quad (1)$$

where μ_{ij} is a stochastic error component. Equation 1 denotes that D_{ij}, respondent i's latent disutility for product j, can be represented by a weighted unfolding model (Carroll 1972) involving respondent i's ideal point ($\mathbf{A}_i$), product j's coordinates ($\mathbf{B}_j$), the respondent's importance weights for each dimension ($\mathbf{W}_i$), and an additive constant (c_i). If $w_{it} = 1$ for all i and t, the simple unweighted unfolding model results. Now, if $D_{ij} \leq d_i^*$ (some individual threshold value), a choice is made for product j; that is, one observes $y_{ij} = 1$. If $D_{ij} > d_i^*$, no choice is made for product j and $y_{ij} = 0$.

Therefore, if the individual threshold value d_i^* is *greater than or equal* to the value of the unobservable latent variable, the respondent will choose the product in question; if the threshold value is *less* than the value of the unobservable latent variable, the respondent will not choose the product. In terms of the ideal point representation, product j is selected by respondent i when it is not very far from his or her ideal notion of a product. Therefore,

$$(2)\quad P(y_{ij} = 1) = P(D_{ij} \leq d_i^*)$$
$$= P\left[\sum_t w_{it}(a_{it} - b_{jt})^2 + c_i + \mu_{ij} \leq d_i^*\right]$$
$$= P\left[\mu_{ij} \leq -\sum_t w_{it}(a_{it} - b_{jt})^2 - c_i^*\right]$$
$$= P(\mu_{ij} \leq -f_{ij}),$$

where:

$$c_i^* = c_i - d_i^* \text{ and}$$

$$f_{ij} = \sum_t w_{it}(a_{it} - b_{jt})^2 + c_i^*.$$

Similarly,

$$(3)\quad P(y_{ij} = 0) = P(D_{ij} > d_i^*)$$
$$= 1 - P(\mu_{ij} \leq -f_{ij}).$$

If we assume the model holds with no error, c_i^* is an estimate of respondent i's (negative) threshold value. The size of the negative of the estimated additive constant is the *magnitude* of the choice threshold, indicating the sensitivity of a person's choice process.

The general form of the likelihood function is

$$(4)\quad L = \prod_{y_{ij}=0} P(y_{ij} = 0) \prod_{y_{ij}=1} P(y_{ij} = 1).$$

If we assume $y_{ij} \sim$ binomial $(1, p_{ij})$ with independence across respondents and products, the choice process of respondent i choosing product j is an independent "coin toss" with probability of choice given by p_{ij}. The implications of this assumption are discussed subsequently. The effects of violations of this independence assumption were examined (DeSarbo and Hoffman 1986) and the model was found to be robust. We assume that μ_{ij} has a logistic distribution function (see McFadden 1976 for a discussion of this assumption in logit type models). Then

$$(5)\quad P(y_{ij} = 1) = \frac{1}{1 + e^{f_{ij}}} = p_{ij}$$

and

$$(6)\quad P(y_{ij} = 0) = \frac{e^{f_{ij}}}{1 + e^{f_{ij}}} = 1 - p_{ij}$$

so that the spatial model can be expressed as (taking the ln of the ratio of equations 6 and 5)

$$(7)\quad \ln\left[\frac{1 - p_{ij}}{p_{ij}}\right] = \sum_t w_{it}(a_{it} - b_{jt})^2 + c_i^* = f_{ij},$$

which is a logistic function (cf. Chow 1983), where the proximity of a product to a respondent's ideal point indicates some degree of the magnitude of the probability of choice. The "closer" (in terms of smaller weighted distance) a product point is to an ideal point, the higher the probability of that respondent choosing/buying that product.

The log likelihood function, expressed in terms of the model parameters, can be obtained by substituting equations 5 and 6 in equation 4 and taking logs.

$$(8)\quad \theta = \ln L = \sum_i \sum_j [(1 - y_{ij})f_{ij} - \ln(1 + e^{f_{ij}})]$$

Estimates of the a_{it}'s, w_{it}'s, c_i^*'s, and b_{jt}'s, given $\mathbf{Y} = ((y_{ij}))$ and T, are obtained by maximum likelihood methods. For the computational details of the conjugate gradient algorithm, see DeSarbo and Hoffman (1986).

Reparameterized Models

The model we have defined in equation 7 can be generalized to incorporate additional data in the form of product attributes and/or respondent background variables. The coordinates for products or respondents are reparameterized as linear functions of background variables (e.g., product features and respondent demographics). If product attribute data are available, $\mathbf{B} = ((b_{jt}))$ can be reparameterized as

$$(9)\quad \mathbf{B} = \mathbf{X}\boldsymbol{\gamma}$$
$$= \sum_k x_{jk}\gamma_{kt}$$

where x_{jk} is the value of feature k for product j and γ_{kt} is the impact of feature k on dimension t. The location of a product is modeled as a direct function of its respective features, as in CANDELINC (Carroll, Pruzansky, and Kruskal 1980), three-way multivariate conjoint analysis (DeSarbo et al. 1982), and GENFOLD2 (DeSarbo and Rao 1984, 1986). Thus, the x_{jk} are objectively quantified features that are related to subjective attributes.

If respondent background data are available, $\mathbf{A} = ((a_{it}))$ also can be reparameterized as

$$(10)\quad \mathbf{A} = \mathbf{Z}\boldsymbol{\alpha}$$
$$= \sum_l z_{il}\alpha_{lt}$$

where z_{il} is the value of characteristic l for individual i

and α_{lt} is the impact of characteristic l on dimension t. If both product attribute and respondent background data are available, both individual and product coordinates can be reparameterized. These reparameterizations can assist in dimensional interpretation and also provide an effective tool for product positioning and market segmentation analyses.

Model Implications

Unlike the conditional logit (McFadden 1976) and conditional probit (Hausman and Wise 1978), our model need *not* have the constraint that $\Sigma_j p_{ij} = 1$ because the product choices are modeled as being independent of each other. Therefore the model can accommodate those choice situations in which, say, complementary or multiple-purchased products/brands are bought with correspondingly high probabilities. For example, large families may consistently buy the same two or more brands of breakfast cereals on the same purchase occasion. This independence assumption is why the model is best suited to the analysis of pick any/n or pick any data, which place no "structural dependencies" on the number or format of the choices and nonchoices of a respondent. The respondent is assumed to choose any number from 0 to n of the objects tested. Thus, these types of data best match the independence assumption of the proposed unfolding threshold model. In some sense, this independence assumption is fairly restrictive. For many types of product choice scenarios (e.g., heavy appliances) in which only one unit typically is purchased, the resulting data may not be amenable to analysis by such a methodology. However, we have clearly demonstrated the robustness of the proposed method in analyzing synthetic data that severely violate these independence assumptions for rows *and* columns simultaneously (DeSarbo and Hoffman 1986). Though one could modify the proposed procedure to accommodate pick 1/n choice data by constraining the predicted probabilities to sum to one and be non-negative, such a modification would imply the use of a constrained optimization/estimation procedure that would be very cumbersome.

The spatial choice methodology enables the user to estimate either a simple unfolding (Coombs 1964) or weighted unfolding (Carroll 1972, 1980) model. If the weighted unfolding model is estimated, the user may allow the weights to be constrained to be non-negative given the concerns raised by Srinivasan and Shocker (1973) and Davison (1976) about interpreting negative weights. Note that the simple unfolding model is a special case of the weighted unfolding model where $w_{it} = 1$ for all i and t. The simple unfolding model assumes that a given distance on a dimension makes as much difference to one individual as another. Here, the "isochoice" contours are circles ($T = 2$), spheres ($T = 3$), or hyperspheres ($T \geq 4$). The area within the isochoice contour for an individual denotes the area of the space where brands would be chosen. In the weighted unfolding model, individuals are permitted to weight the dimensions differentially, where the w_{it}'s represent the saliences of each dimension for each respondent. In this case, the isochoice contours are ellipses ($T = 2$), ellipsoids ($T = 3$), or hyperellipsoids ($T \geq 4$).

This method gives the user great flexibility. In addition to specifying the type of unfolding model and reparameterization options, the user may specify the type of analysis. An internal analysis may be performed by estimating **A** and/or **B** directly. For an external analysis, the user must supply coordinates values for products or respondents. That is, in external analyses, one set of coordinates (typically **B**) is obtained through another analysis (e.g., MDS of similarities) and is fixed in the multidimensional unfolding threshold model.

Goodness-of-Fit Measures

Three goodness-of-fit measures are available for evaluating the spatial choice model selected by the user.

The deviance measure (McCullagh and Nelder 1983; Nelder and Wedderburn 1972).

$$D = -2\left[\sum_i \sum_j y_{ij} \ln(\hat{p}_{ij}) + (1 - y_{ij}) \ln(1 - \hat{p}_{ij})\right] = -2\theta \quad (11)$$

where $\hat{p}_{ij}$, the estimated probability, is expressed as in equation 5 using estimated values for a_{it}, w_{it}, and b_{jt}. Nested models are tested as the difference between respective deviance measures, which is asymptotically distributed as χ^2 with degrees of freedom equal to the difference in model degrees of freedom. One obvious question in the use of maximum likelihood procedures here is the validity of the statistical properties of estimators and associated tests. This question is of particular importance in relation to the spatial choice model presented because of the presence of incidental parameters (e.g., $\mathbf{A}_i$, $\mathbf{B}_j$, c_i^*) whose numbers vary according to the order of **Y**. According to Andersen (1980), maximum likelihood estimators in such cases may not be consistent. He suggests the use of conditional maximum likelihood estimators in such cases. Our major concern stems from our use of the χ^2 test for various nested models. A small Monté Carlo analysis on a synthetic data set (DeSarbo and Hoffman 1986) has shown that, for this data set, the χ^2 test appears to be robust to such issues. However, more testing is required. This asymptotic test can be used to test competing nested model specifications (e.g., weighted vs. unweighted models) as well as for testing the appropriate dimensionality. Note that nested models can be tested via this χ^2 test only if one of the two models being compared (i.e., the simpler model) is the "true" underlying model; otherwise, the resulting difference in deviance statistics follows a noncentral χ^2 distribution. This requirement may prove troublesome in tests for dimensionality, for example, where one rarely knows the "true model." We therefore provide other goodness-of-fit measures (r_{pb}^2 and the sum of squares statistic) to use in addition to the χ^2 test to verify the "correct" solution. Also, as with all other MDS solutions, users should weigh the interpretability of the resulting solution for model selection.

Sums of squares.

$$SSQ = \Sigma\Sigma[y_{ij} - \hat{p}_{ij}]^2$$

Point biserial correlation between Y and $\hat{P}$.

$$r_{pb}(\mathbf{Y},\hat{\mathbf{P}}), \text{ where } \hat{\mathbf{P}} = ((\hat{p}_{ij}))$$

The last two measures also are provided for row and column objects so that goodness of fit for each respondent and product can be examined. The overall $r^2_{pb}(\mathbf{Y},\hat{\mathbf{P}})$ can be regarded as a variance-accounted-for statistic for a particular solution.

AN ILLUSTRATION: PICK ANY/n RESIDENTIAL COMMUNICATIONS DEVICES

Study Description

A marketing research project was sponsored by a major communications firm to gain understanding of intended choices for various brands of a residential communication device and their features, as well as to find potential markets for new offerings. The study involved a total of 499 personal interviews at four geographically dispersed high traffic shopping malls. Potential respondents were screened on the basis of information on family income, size of family, and head of household's age, education, and occupation. Eligible respondents were escorted to a completely enclosed interviewing area where 12 actual products/brands were exhibited. Respondents were read the description of each brand by the interviewer, who then demonstrated how each brand was to be used. Respondents were asked whether they would be buying each of the 12 devices within the next six months, assuming the other 11 devices were not available (0 = would not buy; 1 = would buy). Because of the proprietary nature of the research, the 12 brands are identified by the letters A through L. The exact product class cannot be revealed, but is one in which multiple brand purchases are common.

The Data

To present a *detailed* description of the results, the various analyses were performed on a subsample of 30 subjects selected randomly from the larger sample. Our objective is to illustrate what different methods can and cannot do, rather than make explicit generalizations to the universe about the results of these analyses.

In addition to the pick any/*n* choice data, the sponsoring firm provided seven descriptor variables for the 12 brands which they hypothesized to be the most important features to consumers. The selection was based on prior consumer testing and research. This information was conveyed to the respondents in the interviews. The seven descriptor variables for the 12 brands (**X**) were calling feature (1 = send and receive calls, 0 = receive calls only), repertory dialing (1 = yes, 0 = no), range (in feet), speakerphone capability (1 = yes, 0 = no), price, physical setup (1 = stand-up set, 0 = lie-down set), and style (1 = military, resembles a walkie-talkie, 0 = cradle).

The seven variables used in **X** were basically all the features for which measurements were readily available. Such psychophysical variables as color, smoothness, and weight were not available for use, but could conceivably be important in determining preference or intention to buy. Also, though other marketing mix variables besides price could have been used in **X** (e.g., promotions), such information was not available. We label the seven characteristics M through S. Note that DeSarbo and Rao (1984, 1986) have reported analyses of preference data (measured on an 11-point scale) collected in the same study. Here, however, background data on the subjects were not available for use.

Correspondence Analysis Results

Because of the suitability of correspondence analysis for such data and its growing use in marketing research (cf. Hoffman and Franke 1986), and the fact that Tenenhaus and Young (1985) have demonstrated it is equivalent to many of the alternative procedures listed in Table 1 under certain general conditions, we performed a correspondence analysis on the 30 × 12 binary choice indicator matrix of respondents and residential communications devices using a program written in the SAS MATRIX language (SAS Institute 1982).

The associated cumulative proportions of inertia, variance-accounted-for-like measures, for the first four principal axes are .24, .38, .50, and .62. In terms of the residential communications devices, the first dimension is characterized by products J, C, I, G, and K and accounts for 24% of the variance in the choices made by the respondents. All these devices have the ability to send and receive calls, and all but G have the repertory dialing feature. These devices tend to be higher priced and resemble a walkie-talkie in style (except for K). In addition, J and I have speakerphone capabilities. No other devices in the set have this option. These characteristics of the devices comprise the features available and so we tentatively label this a "features" dimension.

Devices A, D, F, and B define a second dimension of the space. A and B both have the ability to receive calls only and have no repertory dialing capabilities. Both A and B are inexpensive. D and F have the ability to send and receive calls and also have repertory dialing. They are in the upper price range. Of these four devices, only A has a lie-down set and a cradle style. That brands A and B are relatively far from most respondents indicates they were not chosen very often. We loosely label this an "appearance" or "style" dimension.

The third axis is defined by devices E and C, which both load positively, and by D, which loads negatively. This dimension is rather difficult to interpret, but we note that E and C are inexpensive and D is expensive, and that E and C have a lie-down set and D has a stand-up set. Thus, this dimension could be considered to represent "price." The remaining dimensions are very difficult to interpret.

The respondent points are reasonably well dispersed throughout the space. In fact, several respondents are in positions that would correspond to choosing devices with few features and a low price, but there are no devices in the region.

Note that in correspondence analysis we may *not* interpret distances *between* points in different sets. That is, the distance between a respondent point and a device point has no meaning in a correspondence analysis. Between-set comparisons are restricted to the interpretation we have just discussed. Carroll, Green, and Schaffer (1986) propose a different normalization of the coordinates from correspondence analysis that renders the between-set distances interpretable. Nevertheless, caution must be used in examining between-set relations, for they are not similar to those obtained via a multidimensional unfolding analysis.

Multidimensional Unfolding Threshold Model Results

An analysis was performed in 1 through 4 dimensions with the reparameterization option of restricting $\mathbf{B} = \mathbf{X}\gamma$, where $\mathbf{X}$ is the matrix of brand features standardized by column to zero mean and unit variance given the different measurement scales in which each feature was measured. Again, individual differences on the subjects via demographics ($\mathbf{Z}$) were not available for these analyses. Table 2 gives the goodness-of-fit measures for each of the four solutions. The χ^2 test for deviance difference scores indicates that three dimensions is the appropriate dimensionality. Given the small sample size and concerns about incidental parameters, one should pay attention to the other two goodness-of-fit statistics; they also indicate that the three-dimensional solution is most parsimonious. Here, the three-dimensional solution accounts for more than 67% of the variance in $\mathbf{Y}$ as indicated by r_{pb}^2.

Table 3 reports the correlations between the seven feature variables and the three dimensions of $\mathbf{B}$ rotated to its principal axes. From an inspection of these correlations, one can begin to interpret the solution space for the brands ($\mathbf{B}$). Dimension I is dominated by the physical setup and style variables. To load highly positive on dimension I, brands have a stand-up design and resemble a walkie-talkie. We therefore label this a "style" or "appearance" dimension.

Dimension II is highly correlated with calling feature, repertory dialing, speakerphone capability, and price. To load highly positive on this dimension, brands must have send and receive capability, repertory dialing, speakerphone options, and a higher price. We label this a "feature" dimension.

Finally, dimension III is highly correlated with range and somewhat correlated with price. To load highly positive on this dimension, brands transmit/receive over long ranges and are typically higher priced. We label this a "power" dimension.

The three one-dimensional joint space plots ($\mathbf{A}$,$\mathbf{B}$) are shown in Figure 1, depicting the interrelationship between the 12 brands and the 30 ideal points. One striking feature in these plots is that the solution is not of the "degenerate" type that seems to plague many unfolding methods (cf. DeSarbo and Rao 1984). That is, one does not see instances of wide separation between the two sets of points. Rather, they are well intermixed, and their interrelationship tells us something about these intention to buy judgments. Unlike the correspondence analysis solution, the interrelationships between these two sets of points *are* meaningful.

In the plot for the style dimension (I), we see that brands between K and J tend to be most preferred because that range is where most ideal points lie. Note how brand A is isolated on this dimension away from the locus of ideal points, as are brands G, D, and F. Also notice the concentration of ideal points between the origin (0) and −1 where no brands (A–L) are located.

In the plot for the features dimension (II), there seems to be a wider dispersion of ideal points all along the axis. Notice the mixing of the various brands with these ideal points. Brands A and B load negatively on this dimension but appear to have some ideal points near them, though brand B appears to be more isolated.

In the plot for the power dimension (III), ideal points appear to be concentrated in a much smaller range; ideal points are concentrated between E and I. Here, brands K, F, G, D, J, A, and B appear to be most preferred.

Table 2
GOODNESS-OF-FIT MEASURES FOR RESIDENTIAL COMMUNICATIONS DATA

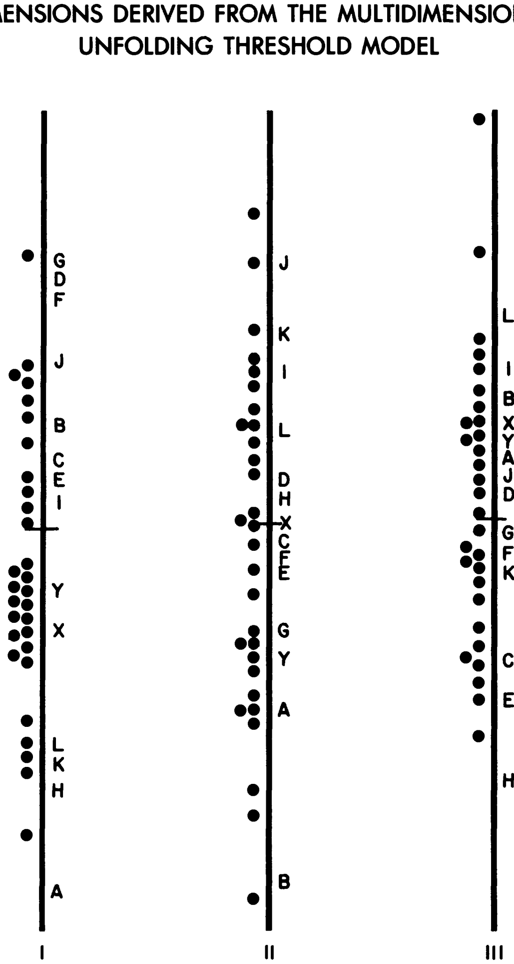
T	Model d.f.	Sum of squares	r_{pb}	r_{pb}^2	θ	Deviance	Deviance difference
1	66	56.41	.588	.346	−167.03	334.06	—
2	101	40.78	.726	.527	−123.23	246.46	87.60[a]
3	135	28.54	.819	.671	−90.74	181.48	64.98[a]
4	168	24.06	.853	.727	−78.73	157.46	24.02

[a] $p \leq .01$.

Table 3
CORRELATIONS BETWEEN B AND X

Feature variable	Dimension I	II	III
M	.313	.807	−.331
N	.249	.736	.067
0	−.113	.364	.731
P	.201	.542	.391
Q	.394	.764	.415
R	.775	−.239	.211
S	.928	−.083	.016

Figure 1
ONE-DIMENSIONAL JOINT SPACES FOR THE THREE DIMENSIONS DERIVED FROM THE MULTIDIMENSIONAL UNFOLDING THRESHOLD MODEL



We also examined the correlations between the three dimensions for the brand coordinates (**B**) and ideal points (**A**). Though these correlations will vary according to the particular rotation used, our solution indicates that no dimension (in **A**, **B**, or γ) shares more than 15% variance with any other dimension. This finding appears to indicate that we are obtaining three somewhat "independent" views of the structure in the choice data.

Table 4 lists the additive constants (c_i^*) and the squared distance (d_{ij}^2) calculations between each ideal point and each brand. Recall that according to our spatial choice model, whenever $d_{ij}^2 < -c_i^*$, the model predicts an intended buy; otherwise, no intended buy. Thus, the negatives of these additive constants can be viewed as threshold values. In comparing them with the computed squared distances in Table 4, one can make predictions about the brands selected by each subject. For example, the negative of c_i^* is 2.385 for subject 1 and brands H, E, and C are predicted to be within this threshold squared distance. Thus, according to the model predictions, subject 1 should intend to buy these three brands. In examining the raw data, one notes that he does intend to buy brands H and E, but not C. As Table 4 shows, the squared distance for brand C is very close to $-c_i^*$. Recall that the size of the negative of the additive constant is the magnitude of the choice threshold by subject indicating the sensitivity of a person's choice process to distance from his/her ideal point. With this simple unweighted unfolding model in three dimensions, one can draw spherical threshold patterns around each subject's ideal point. The volume within that individual's sphere denotes the volume of the space where brands are predicted to be chosen.

Finally, to compare the three-dimensional solutions derived from the correspondence analysis and the multidimensional unfolding threshold model, we used canonical correlation as an "approximate" configuration-matching procedure in comparing subject coordinates, brand coordinates, and the joint space (concatenating the subject and brand coordinates), given the rotational indeterminacies of both models.[1] These analyses demonstrated that only one dimension was somewhat common in each analysis. Major differences were seen with respect to the remaining two dimensions. Hence, the two procedures do in fact render different solutions. (We return to further comparisons in validation between these two models.)

Further Analyses

Because of potential problems associated with possible locally optimum solutions, we reran the three-dimensional solution in our spatial choice model from a different random start. The algorithm converged in 29 iterations, producing a sum of squares of 28.686, point biserial correlation of .819, and a deviance of 183.165. The resulting canonical correlations between this solution and the one reported before are .995, .982, and .971 for **A**; .999, .988, .964 for **B**; and .999, .998, .881 for γ. Hence, there is substantial congruence between the two solutions.

Another analysis was performed in three dimensions on a different randomly selected group of 30 subjects. The algorithm converged in 33 iterations, producing a sum of squares of 25.906, point biserial correlation of .839, and deviance of 169.464. The canonical correlations between this solution and the one reported before are .999, .977, and .636 for **B** and .999, .988, and .458

[1]Canonical correlation is used as a method of configuration matching throughout the article. It is labeled as rendering an "approximate" congruence between the various solutions because of the type of indeterminacies exhibited by the various two-way methods. For example, distances (from the multidimensional unfolding threshold model) are invariant to centering of the joint space or orthogonally rotating **A** and **B** by the same transformation matrix **T**. A more appropriate configuration-matching procedure might be the Cliff-Match procedure (Cliff 1966) after centering the two joint spaces. Because of its unavailability, we chose to use canonical correlation, which should not severely distort the true relationship.

Table 4
ADDITIVE CONSTANTS AND SQUARED DISTANCES

	Brand												Additive constants
Subject	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	$C_{i.}$
1	6.690	10.448	2.359	9.337	2.077	7.564	8.335	1.107	10.097	12.097	4.790	9.090	−2.385
2	11.581	14.397	0.926	4.838	1.113	4.470	6.336	2.688	4.170	4.494	2.404	6.014	−5.185
3	8.466	5.792	2.527	2.854	2.814	2.788	3.102	7.405	2.557	4.903	6.162	4.316	−2.595
4	1.906	5.643	9.275	15.814	8.874	13.684	12.736	6.622	15.295	21.439	10.990	10.526	−7.393
5	3.149	3.389	8.917	13.947	8.544	11.896	10.565	8.446	14.677	20.365	12.687	11.208	−6.397
6	10.431	5.618	0.823	3.918	0.653	2.479	2.536	5.219	7.691	8.136	8.220	10.356	−2.479
7	15.724	9.446	3.842	1.121	4.431	1.958	2.698	12.471	1.198	1.861	8.786	6.026	−2.002
8	8.685	8.720	3.887	3.786	4.379	4.332	5.143	8.043	1.190	4.152	4.894	2.309	−2.216
9	25.527	18.617	4.486	1.592	5.202	2.809	4.806	14.715	2.679	0.159	10.205	10.644	−1.688
10	5.437	6.180	3.425	5.136	3.663	4.922	5.264	5.890	3.356	6.985	4.904	3.085	−5.371
11	2.834	3.612	7.163	10.252	7.114	9.176	8.410	7.892	9.102	14.679	9.266	6.446	−7.472
12	1.477	9.450	7.153	14.851	6.870	13.072	13.214	2.777	12.546	18.105	6.086	7.539	−4.005
13	9.775	12.121	2.226	3.788	2.665	4.177	5.739	5.087	1.225	3.042	2.455	2.601	−3.279
14	3.464	12.348	4.817	9.996	5.006	9.633	10.927	2.284	5.298	9.872	1.498	2.128	−2.143
15	6.651	4.393	3.829	4.435	4.055	4.183	4.090	7.972	3.841	7.239	7.172	4.459	−5.037
16	7.256	5.452	4.009	10.145	3.496	7.679	7.112	4.779	13.911	16.155	10.583	13.448	−6.125
17	14.020	8.388	0.142	1.898	0.175	1.068	1.738	6.192	5.444	4.613	7.527	9.905	−0.516
18	8.540	17.850	7.992	9.755	8.711	10.985	13.061	7.638	2.053	6.625	2.159	0.172	−4.236
19	5.112	3.099	4.074	6.171	4.079	5.257	4.778	6.999	6.520	10.273	8.111	6.205	−7.516
20	8.042	12.263	1.463	5.551	1.648	5.174	6.725	2.116	3.748	5.529	1.704	4.029	−5.948
21	18.107	15.496	0.809	2.607	1.069	2.457	4.329	6.412	4.397	2.525	5.782	9.582	−2.931
22	16.264	13.920	3.167	1.833	3.816	2.849	4.512	9.958	0.482	0.667	5.712	4.906	−7.899
23	9.293	5.751	13.622	11.129	14.095	11.617	10.510	18.979	8.249	14.642	15.766	7.329	−11.617
24	12.555	25.680	9.394	12.549	10.188	13.999	17.180	7.436	3.614	7.216	1.321	1.805	−8.398
25	3.268	5.695	4.276	7.596	4.344	6.871	6.956	4.845	5.775	10.241	5.254	3.950	−4.584
26	12.617	13.013	1.165	2.756	1.543	2.963	4.665	5.152	1.858	2.212	3.222	4.769	−8.131
27	9.850	13.264	1.547	4.501	1.867	4.533	6.292	3.364	2.434	3.785	1.802	3.627	−2.789
28	16.793	10.437	0.249	1.742	0.316	1.015	1.952	7.053	5.773	4.027	8.180	11.262	−7.655
29	9.483	3.872	7.710	5.692	8.086	5.890	5.176	14.277	5.104	9.182	12.157	6.366	−5.271
30	19.675	16.476	2.367	1.763	2.928	2.514	4.531	9.581	1.940	0.483	6.277	7.711	−4.345

for γ. There is somewhat less congruence here than that reported with a different random start, though the results are still very positive.

Comparative Validation

To compare correspondence analysis thoroughly with the proposed method, we ran both procedures on a subset of the data and used the derived solutions to make choice predictions on that part of the choice data held out of the analysis. More specifically, we randomly selected two of the 12 brands to hold out of the analysis (brands G and K). We eliminated the two respective columns from the input choice data and performed a correspondence analysis and multidimensional unfolding threshold analysis on the resulting 30 × 10 matrix. After obtaining a solution by each method, we used the feature/attribute data in **B** for the two brands to make predictions on where their locations in the derived spaces would be. Then, for each model, we constructed predicted choices for all 30 subjects for the two brands and compared these predictions with the actual data for the two brands. These particular validations are possible because these feature variables were available for all 12 brands. Similar validations would have been possible for the remaining 499 − 30 = 469 consumers had we had "relevant" background data (**Z**) for these consumers and utilized one of the reparameterized models that estimated α. We next describe these comparative validation analyses for brands G and K utilizing the two procedures.

Correspondence analysis was performed on the 30 × 10 binary choice data matrix. Products G and K were deleted from the analysis. The corresponding proportions of inertia on the first three principal axes were .239, .165, and .160. Thus, these three axes account for 56.4% of the weighted variation in the data. We also used canonical correlation to compare this solution with the previously reported one performed on all 12 brands. The canonical correlations are .997, .989, and .940 for subject coordinates and .999, .998, and .981 for brands (for the same 10 common brands), indicating substantial congruence between the two solutions.

The 10 × 7 design matrix of product features was column standardized to zero mean and unit variance and regressed on the product coordinates from the three-dimensional correspondence analysis. The regression coefficients from the analysis were applied to the standardized features on products G and K to obtain point locations for these two products in the space of the original 10 products.

We then "reconstituted" the binary choice data matrix according to the relation (Greenacre 1984)

$$\hat{\mathbf{P}} = \mathbf{r}\mathbf{c}' + \mathbf{D}_r\mathbf{F}\mathbf{D}_u^{-1}\mathbf{G}'\mathbf{D}_c \quad (12)$$

where $\hat{\mathbf{P}}$ is the normalized data matrix to be approximated, $\mathbf{r}$ is the 30×1 vector of scaled row (subject) marginals, $\mathbf{c}$ is the 12×1 vector of scaled column (product) marginals, $\mathbf{D}_r = \text{diag}(\mathbf{r})$, $\mathbf{D}_c = \text{diag}(\mathbf{c})$, $\mathbf{D}_u$ is the 3×3 diagonal matrix of singular values, $\mathbf{F}$ is the 30×3 matrix of subject coordinates from this three-dimensional correspondence analysis solution, and $\mathbf{G}$ is the 12×3 matrix of product coordinates (the 10 obtained directly from this analysis and the two predicted).

Because products G and K were not used in the correspondence analysis, their marginals are zero. Zero marginals would necessarily yield predictions of zero. To get around this problem, we tried two approaches. Initially, we substituted the marginals for $\mathbf{c}$ that we *would have obtained* if all 12 products had been in the original analysis. Thus, in a sense we were using information that one typically would *not* have in real applications. This obviously changed the marginals of the 10 products that were in the analysis, but the change was very minor and enabled us to get predictions for G and K that we could not obtain otherwise. Once $\hat{\mathbf{P}}$ was obtained, we compared it with $\mathbf{P}$, the original normalized binary choice data matrix, by counting the number of correct matches. According to our procedure, 53.3% of the choices for product G and 73.3% of the choices for product K were predicted correctly.

Given that we were "cheating" in that one rarely would have such marginals, we then alternatively substituted (for $\mathbf{c}$) averages of the marginals for the 10 brands in the analysis. With this procedure, we predicted 43.3% of the choices for product G and 70.0% of the choices for product K. As one might expect, these predictions are less accurate than when the actual marginals are used. However, the latter figures are more representative of actual validation prediction accuracy because users rarely know such information *a priori*.

Multidimensional unfolding threshold model results. A three-dimensional analysis was run on the 30×10 choice data with the reparameterization option for the brand coordinates. The resulting sum of squares was 17.267, with $r_{pb} = .875$ and $-2\theta = 117.18$. Three canonical correlations analyses were performed to compare the $\mathbf{A}$, $\mathbf{B}$ (only the relevant rows in the first analysis), and $\boldsymbol{\gamma}$ that were estimated separately with the full data set ($\mathbf{Y}$) and this 30×10 subset. For comparing the ideal points ($\mathbf{A}$), the canonical correlations are .966, .955, and .921; for $\mathbf{B}$, .986, .820, and .740; and for $\boldsymbol{\gamma}$, .998, .994, and .836. These statistics indicate congruence between these two solutions, though not as strong as in the correspondence analysis. By multiplying the standardized attribute data for the two holdout brands G and K by the impact coefficients estimated in this secondary analysis, one obtains predicted locations in the derived three-dimensional space for these two brands. By then calculating squared distances between these two new brand locations and the resulting set of 30 ideal points and comparing these squared distances with the newly estimated threshold coefficient c_i^* for each subject, one obtains predicted choices for each of the 30 subjects for each of the two brands. In comparing these predictions with the actual choices for these two brands, we are able to predict 21 of 30 choices/nonchoices for brand G (70%) and 24 of 30 for brand K (80%). This represents a substantial increase over prediction accuracy by guessing (50%) as well as the results obtained by use of correspondence analysis (in both cases using actual and average marginals). Thus, the multidimensional unfolding threshold model appears to have validated a small holdout sample of choice data better than would chance or the correspondence analysis.

DISCUSSION

Competitive Market Structure and Segmentation

The analysis from the simple unfolding option of our spatial choice model for the residential communication device choice data provides some interesting results for marketers. The analysis affords valuable insight about the nature of the competitive structure of this particular market. Figure 1 shows a parsimonious decomposition of the relevant dimensions along which the various brands compete, and the respective location of each brand on each dimension. For example, brands A and B compete in terms of being located near one another on dimensions II (features) and III (power), but are far away from each other on dimension I (style). Thus, brands may have different "substitutes" or competitors along different dimensions.

The analyses also provide information on market segmentation. Figure 1 is a graphic summary of the bases of consumer choice for the set of 12 brands. Subjects with similar ideal points share similar bases underlying their respective choices and can be segmented or classified according to such behavior-related criteria. For example, one could cluster subjects' ideal points for these three dimensions to examine what submarkets appear—what types of subjects have similar choice bases on these three dimensions. Assuming the sample is representative of a target market(s), one could address subsequent clusters or subgroups differently via promotion depending on the respective clusters' characteristics or choice bases.

Optimal Positioning/Repositioning

Assuming this small sample is representative of some target audience, one has the bases for performing some interesting product positioning/repositioning analyses. Here is where the $\boldsymbol{\gamma}$ coefficients come into play. Because the brand locations are explicit functions of the seven feature variables, one can relate changes in these feature variables directly to changes of brand coordinates in the space. That is, because $\mathbf{B}_j = \mathbf{X}_j\boldsymbol{\gamma}$, where $\mathbf{B}_j$ is the vector of brand j coordinates in T dimensions and $\mathbf{X}_j$ is the vector of feature variables for brand j, if one were to, say, redesign brand j to have $\mathbf{X}_j^*$ as design variables, one could get brand j's new position in the space as $\mathbf{B}_j^* = \mathbf{X}_j^*\gamma$. For example, brands A and B are the less popular brands in the study. Suppose the product manager for brand A decides to redesign the product via changes in some subset of the seven variables mentioned before. Given new product design specifications, the manager could obtain predictions not only for the new brand's location, but also for how many customers in the sample will intend to buy it. More specifically, let us examine what effects a change in the design of brand A will have on predicted demand and location. Brand A can be described as being able to receive only and having no repertory dialing, a range of 300 feet, no speakerphone option, a price of \$119.95, and a lie-down, cradle style. Its three-dimensional representation is (−2.616, −1.455, .486) and the model predicts seven people will intend to buy it (actual = eight) as seen from Table 4. Suppose the brand manager adds the capability of sending and receiving calls, the ability to transmit 1000 feet, a speakerphone option, and a military style at an increased price of \$225. On the basis of the independence assumption, the model would predict that 20 of the 30 consumers would state they intend to buy the revised product, and the repositioned brand location for this new brand A* would be (.236, .314, .873). Clearly this estimate of demand is optimistic given the threshold assumption of the model that a consumer will intend to buy *all* brands within c_i^* radius of his/her ideal point. Suppose we assume the consumer will only buy *one* unit—the one closest to his/her ideal point. The original brand A would be predicted to have 5 of the 30 and repositioned product A* would be predicted to have 10 of the 30 who intend to buy only that unit. Thus, this analysis can be conducted with a variety of different assumptions about the buying behavior and competitive structure of the market.

Let us take one step further. Suppose the brand manager is interested in optimally positioning a new product or repositioning a current one spatially in such an analysis. That is, what should the design of the brand be, in terms of these seven features, to maximize expected profit? Suppose we assume the (hypothetical) costs of the various levels of these features shown in Table 5. The nine levels of price are not costs, but feasible prices. Note that all levels of each feature are within those of the brands used in the original analysis. We also assume an auxiliary fixed cost of \$20 per unit for other materials (e.g., the casing). Again, these costs are hypothetical. Though the analysis can be performed under a variety of competitive assumptions, assume the consumer actually will buy only one device—the one which, in comparison with the other 12, is closest to his or her ideal point. Thus, the problem can be stated as

$$\max_{\mathbf{X}_j} \Pi = (P_j - C_j)q_j \tag{13}$$

where:

Π = net profit,

P_j = the price for brand j,

C_j = the costs for brand j as a function of the specific levels selected for the features in Table 5, and

q_j = the number of intended purchasers predicted from the model as a function of $\mathbf{X}_j$.

Because one has both continuous and discrete variables with a nonlinear objective function, some mixed-integer nonlinear programming algorithm would be necessary to find the globally optimum result (e.g., Dakin's 1965 modification of Land and Doig's 1960 branch and bound procedure would be appropriate). To simplify the problem, we discretize the price variable to include nine levels and the range variable to include 10 levels.

A combinatorial optimization routine (cf. DeSarbo and Rao 1986) was programmed to maximize Π and find the corresponding set of design variables. It is a modification of the Lin and Kernighan (1973) procedure. The routine generates several locally optimum solutions of optimal bundles of features. The procedure sequentially alters the discrete levels in $\mathbf{X}_j$ until no further improvement is obtained in the maximand in equation 13. This procedure was run for 50 trials, generating a number of locally optimum solutions. The optimum product, *from the producer's standpoint,* generates an expected profit of \$2390.18 for these 30 consumers, with 13 of the 30 (predicted) buying it. The optimal features, if we assume

Table 5
HYPOTHETICAL FEATURE COSTS PER UNIT

Feature	*Levels*	*Costs (\$)*
1. Calling features	Receive only	8.88
	Send and receive	19.18
2. Repertory dialing	None	0
	Yes	15.75
3. Range (feet)	50	7.12
	150	7.12
	250	7.12
	350	7.12
	450	7.12
	550	12.93
	650	12.93
	750	12.93
	850	12.93
	1000	12.93
4. Speakerphone option	None	0
	Yes	12.15
5. Price (\$)	100	—
	125	—
	150	—
	175	—
	200	—
	225	—
	250	—
	275	—
	300	—
6. Setup	Lie down	6.24
	Stand up	9.77
7. Style	Cradle	7.63
	Military	11.75

the cost structure in Table 5, are receive only, no repertory dialing, range 50 feet, speakerphone option, price $250, lie-down setup, and military style. Obviously, there are more popular offerings from a *consumer's standpoint* (e.g., "high" features, lowest price) but, for the hypothetical cost structure in Table 5, this redesign renders maximum expected profit. This solution was verified by a (more computationally expensive) complete enumeration procedure as the globally optimum solution. The coordinates of this optimal brand are (−.866, .021, .588), shown as X in the plots in Figure 1.

Similarly, this analysis can be performed with different assumptions about the buying behavior of consumers. If we assume the independence of multiple purchases, as specified by the spatial choice model, we can perform a similar maximization of profit. The only computation that changes is the calculation for q_j. The combinatorial optimization procedure was altered to reflect this assumption change, and 50 locally optimum solutions were generated that locally maximized Π. The global maximum, as verified by complete enumeration, rendered Π = $5526.24 with 24 of the 30 consumers predicted to intend to buy the brand. The features were receive only, repertory dialing, range 50 feet, no speakerphone, price $275, lie-down setup, and military style. Again, this is the optimum product from the manufacturer's perspective. The coordinates of this brand are (−.528, −.593, .427), shown as Y in Figure 1.

Though these policy simulations and optimizations also could be attempted with many of the other procedures in Table 1, several problems would arise. First, without exact reparameterizations on brands, predicted locations for them in a derived space via regression will be produced with error that could render misleading results. Second, as demonstrated by the validation comparison with correspondence analysis, it is not clear how one can validate the results or make choice predictions on a holdout sample or set of brands/consumers not included in the analysis. Many of these procedures, especially those mentioned to be related to correspondence analysis, require additional information rarely available at the time such predictions are requested or needed.

Limitations

We use our small example as an illustration of what can be done with the new method, not as a basis for drawing substantive recommendations for this market. Clearly, a sample size of 30 is too small to justify clear-cut policy/strategy. The method, however, is *not* limited to merely 30 people and 12 products.[2] A similar analysis with a vector spatial choice model (DeSarbo, Keramidas, and Clark 1985) was performed on 350 subjects and 12 brands on a Cray-I machine. Unfortunately, the resulting joint space plots became very messy. We used 30 subjects here to facilitate the analysis and graphically display all the detailed analyses.

Second, this method, like all others, is not a substitute for good theory and common sense. It does not tell the user which background variables (for subjects) or features (for brands) should be specified in **X** and **Z**. This choice is a matter of expert judgment, previous research findings, and other factors. In addition, concept testing of the designated "optimal brand(s)" would be desirable to verify demand for the item(s).

Finally, the present version of the model can be extended in the postanalytic optimization phase to accommodate multiple brand positionings and dynamic competitive reactions. The first author is currently working on accommodating competitor reactions and a dynamic marketplace within such spatial models via spatial equilibrium analysis.

Future Research

The spatial choice model presented can be used in a variety of ways to analyze binary data. In addition to analyzing pick any/n or pick any data as described in the application, the spatial choice model can provide a spatial representation of binary profile data, for example, products having or not having certain attributes or benefits. The resulting space can be used for product positioning, as well as obtaining insight into competitive market structure. Another application would involve recording items purchased in a grocery shopping trip by different subjects to investigate the issue of complementary goods. The method also could be used to study sociological networks (e.g., opinion leaders) where respondents denote choices for friends out of particular social groups. Another application would be in investigating potential job choices. Clearly, many marketing and behavioral applications involve binary choice data.

Several interesting research directions are suggested by the development of this procedure for spatially representing binary data. One extension would be to model the log of the odds as a vector model instead of a distance model. The ideal point model in expression 3 may not be appropriate for many applications where subjects have "the more the better" utility functions for the various underlying dimensions/bases of choice. Here, one could use the vector model and provide a different spatial representation. Such work is currently underway by DeSarbo, Keramidas, and Clark (1985). Further Monté Carlo work must be performed on all models under a diversity of conditions to test the algorithm for local minimum or other potential problems, such as other types of violations to the independence assumption. The statistical properties of the estimates must be evaluated via bootstrapping and using second derivative information. Also, the small-sample properties of the χ^2 test must be

[2]The current version was written in APL and is executing on a VAX 8600 machine. Depending on volume of usage, an analysis of data sets as large as I = 100, j = 20 is computationally feasible. Current efforts involve converting the program to FORTRAN (which should allow for at least double the data set size) and putting it on a diskette for microcomputer usage.

examined more rigorously, as should the consistency of the estimated parameters. Finally, work is needed to generalize the method to handle three-way choice arrays where, for example, choices are obtained from subjects for products purchased/chosen over time or for different situations, or where one collects brand × attribute × consumer data as is done in many types of marketing studies. Jedidi (1987) currently is investigating these extensions of the method. In this context, DeSarbo et al. (1985) have extended the basic unfolding model to accommodate the analysis of three-way asymmetric proximity (binary) data and apply it to the spatial analysis of word associations.

REFERENCES

Andersen, Erling B. (1980), *Discrete Statistical Models with Social Science Applications*. New York: North-Holland Publishing Company.

Bartholomew, David J. (1980), "Factor Analysis for Categorical Data," *Journal of the Royal Statistical Society*, B42, 293–321.

Bentler, Peter M. and David G. Weeks (1978), "Restricted Multidimensional Scaling Models," *Journal of Mathematical Psychology*, 17 (April), 138–51.

Benzécri, Jean P., et al. (1973), *L'Analyse des Données. Vol. 1, La Taxinomie; Vol. 2, L'Analyse des Correspondances*. Paris: Dunod.

Bloxom, Bruce (1978), "Constrained Multidimensional Scaling in n Spaces," *Psychometrika*, 43 (September), 397–408.

Bock, R. Darrell and Marcus Lieberman (1970), "Fitting a Response Model for n Dichotomously Scored Items," *Psychometrika*, 35 (June) 179–97.

Carroll, J. Douglas (1972), "Individual Differences and Multidimensional Scaling," in *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences. Vol. 1, Theory*, R. N. Shepard, A. K. Romney, and S. B. Nerlove, eds. New York: Seminar Press, Inc., 105–55.

——— (1980), "Models and Methods for Multidimensional Analysis of Preferential Choice (or Other Dominance Data)," in *Similarity and Choice*, E. D. Lantermann and H. Feger, eds. Vienna: Hans Huber Publishers, 234–89.

———, Paul Green, and Catherine M. Schaffer (1986), "Interpoint Distance Comparisons in Correspondence Analysis," *Journal of Marketing Research*, 23 (August), 271–80.

———, Sandra Pruzansky, and Joseph B. Kruskal (1980), "CANDELINC: A General Approach to Multidimensional Analysis of Many-Way Arrays with Linear Constraints on Parameters," *Psychometrika*, 45 (March), 3–24.

Chow, Gregory C. (1983), *Econometrics*. New York: McGraw-Hill, Inc.

Christoffersson, Anders (1975), "Factor Analysis of Dichotomized Variables," *Psychometrika*, 40 (March), 5–32.

Cliff, Norman (1966), "Orthogonal Rotation to Congruence," *Psychometrika*, 31 (March), 33–42.

Coombs, Clyde H. (1964), *A Theory of Data*. New York: John Wiley & Sons, Inc.

Cooper, Lee (1983), "A Review of Multidimensional Scaling in Marketing Research," *Applied Psychological Measurement*, 7 (Fall), 427–50.

Dakin, R. J. (1965), "A Tree Search Algorithm for Mixed Integer Programming Problems," *Computer Journal*, 8, 250–55.

Davison, Mark L. (1976), "Fitting and Testing Carroll's Weighted Unfolding Model for Preferences," *Psychometrika*, 41 (June), 233–47.

Day, George S., Allan D. Shocker, and Rajendra K. Srivastava (1979), "Consumer-Oriented Approaches to Identifying Product Markets," *Journal of Marketing*, 43 (Fall), 8–20.

de Leeuw, Jan (1973), "Canonical Analysis of Categorical Data," unpublished doctoral dissertation, Psychological Institute, University of Leiden, The Netherlands.

——— and Willem Heiser (1980), "Multidimensional Scaling with Restrictions on the Configuration," in *Multivariate Analysis*, Vol. 5, P. R. Krishnaiah, ed. New York: North-Holland Publishing Company, 501–22.

Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin (1977), "Maximum Likelihood Estimation from Incomplete Data Via the E. M. Algorithm," *Journal of the Royal Statistical Society*, B39, 1–38.

DeSarbo, Wayne S. and J. Douglas Carroll (1985), "Three-Way Metric Unfolding via Weighted Alternating Least Squares," *Psychometrika*, 50 (December), 275–300.

———, ———, Donald R. Lehmann, and John O'Shaughnessy (1982), "Three-Way Multivariate Conjoint Analysis," *Marketing Science*, 1 (Fall), 323–50.

——— and Donna L. Hoffman (1986), "A New Unfolding Threshold Model for the Spatial Representation of Binary Choice Data," *Applied Psychological Measurement*, forthcoming.

———, Elaine M. Keramidas, and Linda A. Clark (1985), "A New Multidimensional Scaling Methodology for the Spatial Representation of Binary Choice Data, unpublished paper, University of Pennsylvania.

———, Donald Lehman, Sunil Gupta, Morris Holbrook, and William Havlena (1985), "A Three-Way Unfolding Methodology for Asymmetric Binary Proximity Data," working paper, University of Pennsylvania.

———, Richard L. Oliver, and Geert De Soete (1986), "A New Probabilistic MDS Vector Model," *Applied Psychological Measurement*, 10 (March), 79–98.

——— and Vithala R. Rao (1984), "GENFOLD2: A Set of Models and Algorithms for the GENeral unFOLDing Analysis of Preference/Dominance Data," *Journal of Classification*, 1 (Winter), 147–86.

——— and ——— (1986), "A Constrained Unfolding Model for Product Positioning Analysis," *Marketing Science*, 5 (Winter), 1–19.

Gifi, Albert (1981a), *Nonlinear Multivariate Analysis*, Department of Data Theory, University of Leiden, The Netherlands.

——— (1981b), "Homogeneity Analysis," unpublished paper, Department of Data Theory, University of Leiden, The Netherlands.

Green, Paul E. (1975), "Marketing Applications of MDS: Assessment and Outlook," *Journal of Marketing*, 39 (January), 24–31.

——— and Frank J. Carmone (1970), *Multidimensional Scaling and Related Techniques in Marketing Analysis*. Boston: Allyn & Bacon.

——— and Wayne S. DeSarbo (1980), "Two Models for Representing Unrestricted Choice Data," *Association for Consumer Research Proceedings*, 309–12.

——— and Vithala R. Rao (1972), *Applied Multidimensional*

Scaling. New York: Holt, Rinehart and Winston.

Greenacre, Michael J. (1984), *Theory and Application of Correspondence Analysis*. London: Academic Press, Inc.

Hausman, John A. and David A. Wise (1978), "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependencies and Heterogeneous Preferences," *Econometrika*, 46 (March), 403–26.

Heiser, Willem J. (1981), *Unfolding Analysis of Proximity Data*, Department of Data Theory, University of Leiden, The Netherlands.

Hoffman, Donna L. and George Franke (1986), "Correspondence Analysis: Graphical Representation of Categorical Data in Marketing Research," *Journal of Marketing Research*, 23 (August), 213–27.

Holbrook, Morris B., William L. Moore, and Russell S. Winer (1980), "Using 'Pick-Any' Data to Represent Competitive Positions," *TIMS/ORSA Marketing Measurement Proceedings*. Providence, RI: Institute of Management Science, 129–34.

———, ———, and ——— (1982), "Constructing Joint Spaces from 'Pick-Any' Data: A New Tool for Consumer Analysis," *Journal of Consumer Research*, 9 (June), 99–105.

Jedidi, Kamel (1987), "A Three-Way Stochastic MDS Choice Methodology," dissertation proposal, University of Pennsylvania, Philadelphia.

Johnson, Richard M. (1971), "Market Segmentation: A Strategic Management Tool," *Journal of Marketing Research*, 8 (February), 13–18.

Kruskal, Joseph B. (1964), "Multidimensional Scaling by Optimizing Goodness-of-Fit to a Nonmetric Hypothesis," *Psychometrika*, 29 (March), 1–28.

——— and Roger N. Shepard (1974), "A Nonmetric Variety of Linear Factor Analyses," *Psychometrika*, 39 (June), 123–57.

———, Forrest W. Young, and Judith B. Seery (1973), "How to Use KYST—a Very Flexible Program to Do Multidimensional Scaling and Unfolding," working paper, Bell Laboratories, Murray Hill, NJ.

Land, Ailsa H. and Alison G. Doig (1960), "An Automatic Method of Solving Discrete Programming Problems," *Econometrika*, 28 (July), 297–520.

Lebart, Ludovic, Alain Morineau, and Kenneth M. Warwick (1984), *Multivariate Descriptive Statistical Analysis: Correspondence Analysis & Related Techniques for Large Matrices*. New York: John Wiley & Sons, Inc.

Levine, Joel H. (1979), "Joint-Space Analysis of 'Pick Any' Data: Analysis of Choices from an Unconstrained Set of Alternatives," *Psychometrika*, 44 (March), 85–92.

Lin, Shen and Brian Kernighan (1973), "An Effective Heuristic Algorithm for the Traveling Salesman Problem," *Operations Research*, 21 (March–April), 489–516.

McCullagh, Peter and John A. Nelder (1983), *Generalized Linear Models*. New York: Chapman and Hall.

McFadden, Daniel (1976), "Quantal Choice Analysis: A Survey," *Annals of Economic and Social Measurement*, 5 (Fall), 363–90.

Muthen, Bengt (1978), "Contributions to Factor Analysis of Dichotomous Variables," *Psychometrika* 43 (December), 551–60.

——— (1981), "Factor Analysis of Dichotomous Variables: American Attitudes Toward Abortion," in *Factor Analysis and Measurement in Sociological Research*, Jackson and Borgatta, eds. Beverly Hills, CA.: Sage Press Publications, Inc., 114–36.

Nelder, John A. and R. W. M. Wedderburn (1972), "Generalized Linear Models," *Journal of the Royal Statistical Society*, A135, 370–84.

Nishisato, Shizuhiko (1980), *Analysis of Categorical Data: Dual Scaling and Its Applications*. Toronto: University of Toronto Press.

Noma, Elliot and J. Johnson (1977), *Constraining Nonmetric Multidimensional Scaling Configurations*, Technical Report 60, Human Performance Center, University of Michigan.

SAS Institute, Inc. (1982), *SAS User's Guide: Statistics*, 1982 ed. Cary, NC: SAS Institute, Inc.

Shocker, Allan D. and David W. Stewart (1983), "Strategic Marketing Decision Making and Perceptual Mapping," in *Advances and Practices of Marketing Science—1983 Proceedings*, F. S. Zufyden, ed. Providence, RI: Institute of Management Science, 224–39.

Srinivasan, Seenu and Allan D. Shocker (1973), "Linear Programming Techniques for Multidimensional Analysis of Preferences," *Psychometrika*, 38 (September), 337–69.

Srivastava, Rajendra K., Mark I. Alpert, and Allan D. Shocker (1984), "A Customer Oriented Approach for Determining Market Structures," *Journal of Marketing*, 48 (Spring), 32–45.

Takane, Yoshio (1983), "Choice Model Analysis of the 'Pick Any/n' Type of Binary Data," handout for talk at European Psychometric and Classification Meetings (July), Jouy-en-Josas, France.

———, Forrest W. Young, and Jan de Leeuw (1977), "Nonmetric Individual Differences Multidimensional Scaling: An Alternating Least Squares Method with Optimal Scaling Features," *Psychometrika*, 42 (March), 7–67.

Tenenhaus, Michael and Forrest W. Young (1985), "An Analysis and Synthesis of Multiple Correspondence Analysis, Optimal Scaling, Dual Scaling, Homogeneity Analysis, and Other Methods for Quantifying Categorical, Multivariate Data," *Psychometrika*, 50 (March), 91–119.

Thurstone, Louis L. (1927), "A Law of Comparative Judgment," *Psychological Review*, 34 (July), 273–86.

——— (1929), "Theory of Attitude Measurement," *Psychological Review*, 36 (May), 222–41.

Torgerson, Warren S. (1958), *Theory and Methods of Scaling*. New York: John Wiley & Sons, Inc.

Wind, Yoram (1982), *Product Policy: Concepts, Methods, and Strategy*. Reading, MA: Addison-Wesley Publishing Company.