

A STOCHASTIC MULTIDIMENSIONAL SCALING PROCEDURE FOR THE EMPIRICAL DETERMINATION OF CONVEX INDIFFERENCE CURVES FOR PREFERENCE/CHOICE ANALYSIS

WAYNE S. DESARBO

DEPARTMENTS OF MARKETING AND STATISTICS
SCHOOL OF BUSINESS
THE UNIVERSITY OF MICHIGAN

KAMEL JEDIDI

GRADUATE SCHOOL OF BUSINESS
COLUMBIA UNIVERSITY

JOEL H. STECKEL

DEPARTMENT OF MARKETING
STERN SCHOOL OF BUSINESS
NEW YORK UNIVERSITY

The vast majority of existing multidimensional scaling (MDS) procedures devised for the analysis of paired comparison preference/choice judgments are typically based on either scalar product (i.e., vector) or unfolding (i.e., ideal-point) models. Such methods tend to ignore many of the essential components of microeconomic theory including convex indifference curves, constrained utility maximization, demand functions, et cetera. This paper presents a new stochastic MDS procedure called MICROSCALE that attempts to operationalize many of these traditional microeconomic concepts. First, we briefly review several existing MDS models that operate on paired comparisons data, noting the particular nature of the utility functions implied by each class of models. These utility assumptions are then directly contrasted to those of microeconomic theory. The new maximum likelihood based procedure, MICROSCALE, is presented, as well as the technical details of the estimation procedure. The results of a Monte Carlo analysis investigating the performance of the algorithm as a number of model, data, and error factors are experimentally manipulated are provided. Finally, an illustration in consumer psychology concerning a convenience sample of thirty consumers providing paired comparisons judgments for some fourteen brands of over-the-counter analgesics is discussed.

Key words: multidimensional scaling, paired comparisons data, maximum likelihood estimation, microeconomics, consumer psychology.

1. Introduction

Psychometric procedures developed for the analysis of paired comparisons data typically attempt to utilize spatial models to display the structure in such data. Typically, one of two general classes of models have been utilized to represent such preference/choice data: vector and unfolding models. A vector or scalar products multidimensional scaling model (Slater, 1960; Tucker, 1960) represents subjects as vectors and objects as points in a T-dimensional space. The projection of object points onto subject vector(s) is assumed at least monotone with that subject's preferences. These analyses estimate both the subjects' vector directions and object coordinates in a prescribed

Requests for reprints should be sent to Wayne S. DeSarbo, Departments of Marketing and Statistics, School of Business, University of Michigan, Ann Arbor, MI 48109-1234.

dimensionality. Vector models have the property that preference or utility changes monotonically with all dimensions. That is, these models assume that if a certain amount of a dimension is good, even more must be better. (The iso-utility contours in two dimensions therefore are parallel straight lines perpendicular to a subject's vector.) Work by Bechtel, Tucker, and Chang (1971), Cooper and Nakanishi (1983), Carroll (1980), De Soete and Carroll (1983, 1986), and DeSarbo, Oliver, and De Soete (1986) present alternative probabilistic vector MDS models that operate on such paired comparisons data.

The other major type of psychometric model used to represent such paired comparisons preference/choice data is the ideal-point or unfolding model. We will only discuss Coombs' (1964, pp. 80–119) simple unfolding model. Here, both subjects and objects are represented as points in a T -dimensional space. The points for the subjects represent ideal-points, or optimal sets of dimensional values. The farther a given object point is from a subject's ideal-point, the less utility that object provides to that subject. This notion of relative distance often assumes a Euclidean metric on the space implying that, in two dimensions, iso-utility contours are families of concentric circles centered at a subject's ideal-point. Carroll (1980) demonstrates that the vector model is a special case of this unfolding model where the ideal-point goes off to infinity. The goal in unfolding analysis is to estimate both the subjects' ideal-points and the coordinates for the objects in a prescribed dimensionality in some optimal fashion. Cooper and Nakanishi (1983), Schönemann and Wang (1972), Wang, Schönemann, and Rusk (1975), Zinnes and Griggs (1974), De Soete, Carroll, and DeSarbo (1986), De Soete and Carroll (1986), and DeSarbo, De Soete, and Eliashberg (1986) provide alternative probabilistic MDS unfolding models that operate on paired comparisons data.

Most economists would be quite critical of the nature of the utility functions implied by both these vector and unfolding models. In the vector model, subjects have linear indifference or iso-utility curves that imply a constant rate of substitution (i.e., the amount of one dimension a subject will sacrifice to obtain a unit of a second is constant and independent of the amount of either dimension possessed by the object in question). In simple unfolding models, subjects possess circular (in two dimensions) indifference or iso-utility curves around their ideal point. This makes it possible for two object points: (c_1, c_2) and (d_1, d_2) to be equally preferred by a subject, where $c_1 > d_1$ and $c_2 > d_2$, and thus, the first object dominates the second object on both dimensions. In contrast, basic microeconomic theory assumes that subjects will have convex indifference curves with a decreasing marginal rate of substitution, where the amount of one dimension or characteristic a subject will sacrifice for one unit of a second is negatively related to the amount of the second characteristic already possessed (see Henderson & Quandt, 1984; Lancaster, 1971).

Against this background, this paper develops a new multidimensional scaling procedure incorporating the microeconomic notion of convex subject preferences. The underlying model also incorporates recent developments in multidimensional scaling that allow for a translation of physical object attributes into derived object coordinates (e.g., DeSarbo & Rao, 1984, 1986). These two features described above, convex preferences and translation from physical attributes to perceptual characteristics, enable us to relate our procedure to Lancaster's (1966, 1971) microeconomic theory of consumer demand. Hence, we call the procedure MICROSCALE. The next section briefly reviews basic concepts of utility theory in microeconomics including developments by Lancaster (1966, 1971). MICROSCALE is then described in terms of the model, maximum likelihood estimation, and program options. The results of a modest Monte Carlo analysis are discussed showing somewhat consistent algorithm performance over a number of different model, data, and error specifications. Finally, an application in

consumer psychology is provided where consumer preferences for some 14 brands of over-the-counter analgesics are analyzed by MICROSCALE.

2. Microeconomic Conceptualizations of Preference and Utility

We first briefly review the traditional microeconomic conceptualization of preferences and describe several standard regularity assumptions concerning them. In particular, we discuss the utility function and its relationship to preference orderings. Then, the problem of utility maximization subject to a budget constraint is presented in terms of traditional views on consumer behavior. Finally, we describe the Lancaster (1966, 1971) view of consumer behavior in terms of its differences from the traditional views. For the discussion that follows, we interchange the use of consumer with subject, and goods/services, commodities, or brands with objects.

Traditional microeconomic theory pictures a consumer faced with possible consumption bundles in some set X^* , his consumption set. A consumption/commodity bundle \mathbf{x} is a collection of amounts of all goods and services consumed. If there are N goods and services in the economy, then \mathbf{x} is a vector of real numbers $[x_1, x_2, \dots, x_N]'$, where x_j is the amount of the j -th commodity consumed. Equivalently, \mathbf{x} is a point in Euclidean N space. The following discussion is adapted from Henderson and Quandt (1984), Varian (1984), and Russell and Wilkinson (1979).

Assumption 1: (Consumption Space). The consumption space, X^* , is the non-negative Euclidean N -orthant.

The consumer is assumed to have preferences over the various commodity bundles in X^* . $\mathbf{x} \succcurlyeq \mathbf{y}$ indicates bundle \mathbf{x} is preferred to \mathbf{y} , $\mathbf{x} \sim \mathbf{y}$ indicates bundle \mathbf{x} and \mathbf{y} are equally preferred (consumer is indifferent), and $\mathbf{y} \succcurlyeq \mathbf{x}$ indicates that bundle \mathbf{y} is preferred to \mathbf{x} . Weak preference relationships \succeq and \preceq are similarly defined. We use \leq , \geq and \neq to designate less than or equal to, greater than or equal to, or not equal to, respectively, for specifying order relations between the amounts of all commodities between any pair of bundles.

Assumption 2: (Completeness). For all \mathbf{x} and \mathbf{y} in X^* , either $\mathbf{x} \succeq \mathbf{y}$, or $\mathbf{y} \succeq \mathbf{x}$, or both.

Assumption 3: (Reflexivity). For all \mathbf{x} in X^* , $\mathbf{x} \succeq \mathbf{x}$.

Assumption 4: (Transitivity). For all \mathbf{x} , \mathbf{y} , and \mathbf{z} in X^* , if $\mathbf{x} \succeq \mathbf{y}$, $\mathbf{y} \succeq \mathbf{z}$, then $\mathbf{x} \succeq \mathbf{z}$.

Assumption 5: (Continuity). For all \mathbf{y} in X^* , $\{\mathbf{x}: \mathbf{x} \succeq \mathbf{y}\}$ and $\{\mathbf{x}: \mathbf{x} \preceq \mathbf{y}\}$ are closed sets; or equivalently, $\{\mathbf{x}: \mathbf{x} \succ \mathbf{y}\}$ and $\{\mathbf{x}: \mathbf{x} \ll \mathbf{y}\}$ are open sets.

This fifth assumption is necessary to rule out certain discontinuous behavior; it says that, if (\mathbf{x}^i) is a sequence of consumption bundles that are all preferred to a bundle \mathbf{y} and if this sequence converges to some bundle \mathbf{x}^* , then \mathbf{x}^* is preferred to \mathbf{y} . The most important consequence of continuity is that if \mathbf{y} is strictly preferred to \mathbf{z} and if \mathbf{x} is a bundle that is close enough to \mathbf{y} , then \mathbf{x} must be strictly preferred to \mathbf{z} .

Assumption 6: (Strong Monotonicity). If $\mathbf{x} \succeq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$, then $\mathbf{x} \succ \mathbf{y}$.

This states that more of a bundle is better than less. This is akin to the assumptions made in vector/scalar products MDS models. Note that ideal-point (finite) MDS models violate this assumption.

Assumption 7: (Strict Convexity). Given $\mathbf{x} \neq \mathbf{y}$ and \mathbf{z} in X^* , if $\mathbf{x} \succeq \mathbf{z}$ and $\mathbf{y} \succeq \mathbf{z}$, then $t\mathbf{x} + (1 - t)\mathbf{y} \succ \mathbf{z}$ for all $0 < t < 1$.

Note that a lexicographic preference ordering violates strict convexity (Henderson & Quandt, 1984, p. 13). Thus, such preferences are not truly captured by the classical theory.

In classical economic theory, one often summarizes a consumer's behavior by means of a utility function, that is, a continuous function $u: X^* \rightarrow R$ such that $\mathbf{x} \succeq \mathbf{y}$ if and only if $u(\mathbf{x}) \geq u(\mathbf{y})$. It can be shown that, if the preference ordering is complete, reflexive, transitive, and continuous, then it can be represented by some utility function. A utility function $u(\mathbf{x})$ is often a very convenient way to describe preferences, but it may not have any coherent psychological interpretation (see Cooper, 1973, for distinctions between preference and utility). One relevant property of a utility function is its ordinal character. Thus, if $u(\mathbf{x})$ represents some preferences and $f: R \rightarrow R$ is a monotonic, increasing function, then $f(u(\mathbf{x}))$ will represent exactly the same preferences since $f(u(\mathbf{x})) > f(u(\mathbf{y}))$ if and only if $u(\mathbf{x}) > u(\mathbf{y})$. Utility functions are usually graphically displayed with the help of indifference curves, or the sets of all consumption bundles that a consumer is indifferent between. One can think of indifference curves as being level sets of the utility function. The set of all bundles on or above an indifference curve, $\{\mathbf{x} \in X^* : \mathbf{x} \succeq \mathbf{x}_0\}$, is called an *upper contour set*. A particular level of utility or satisfaction can be derived from a variety of commodity bundles. An indifference curve is thus defined as $\{\mathbf{x} \in X^* | u(\mathbf{x}) = u_0\}$ for some constant u_0 . Note, Assumption 5 guarantees that the derived indifference curves are continuous, while Assumption 7 guarantees that the derived indifference curves are convex to the origin. Convex indifference curves imply that a consumer has diminishing marginal rates of substitution (i.e., the amount of one commodity a consumer will substitute for another and keep the same level of utility is a decreasing function of the second commodity).

In the basic problem of preference maximization, the set of feasible alternatives is just the set of all bundles that the consumer can afford. Let w be the fixed amount of money available to a consumer, and let $\mathbf{p} = (p_1, \dots, p_N)'$ be the vector of prices of goods 1, \dots , N . The set of affordable bundles, the budget set of the consumer, is given by:

$$S = \{\mathbf{x} \in X^* : \mathbf{p}'\mathbf{x} \leq w\},$$

where, as defined, S is a convex polytope. The problem of preference maximization can then be written as:

$$\begin{aligned} \max \quad & u(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{p}'\mathbf{x} \leq w, \\ & \mathbf{x} \in X^* \end{aligned}$$

The solution obtained via Lagrangian maximization is the point of tangency of the indifference curve associated with the highest utility and the budget constraint w . The regularity conditions, Assumptions 2 through 6, and strict convexity, Assumption 7, guarantee the existence and uniqueness of the solution.

Indifference curves form the basis of much model building in economics and psychology and it is surprising that few empirical studies exist that focus on their structure.

The earliest such study was performed by Thurstone (1931) who constructed indifference curves for quantities of pairs of three different goods: hats, pairs of shoes, and overcoats. He constrained the mathematical form of these curves to be hyperbolic (convex) and then fit parameters to observed preference data. The second study concentrated on the consumption of eggs and bacon (Rousseas & Hart, 1951). A useful aspect of the Rousseas and Hart study was that the most preferred combination of eggs and bacon had to be consumed by the subjects, hopefully inducing them to consider the alternatives more carefully. One final study, performed by MacCrimmon and Toda (1969), represents a major departure from the two papers cited above. Rather than assuming convexity, they derived a procedure to experimentally determine the shape of individual indifference curves. They examined tradeoffs between money and pens, and between money and pastries for seven subjects. The authors concluded that subjects had a satiation point for pastries beyond which subjects had no desire for additional pastries. This violated Assumption 6 (Strong Monotonicity) and suggested that money-pastry preferences did not fit the pure economic paradigm. Money-pen preferences were more appropriate. Among the seven subjects, MacCrimmon and Toda found only two examples of nonconvexity. One subject exhibited step-function like indifference curves. Another subject displayed noticeable concavity in one of his outer indifference curves. The authors attributed this latter instance to subject carelessness and concluded that their money-pen results are largely consistent with convexity. To be fair though, we should recognize that the degree of convexity was not great, and that some indifference curves exhibited long linear appearing segments.

Lancaster's Theory of Consumer Demand

Lancaster's (1966, 1971) theory of consumer demand has been an important contribution to the consumer psychology literature. It has formed the basis for much research in consumer psychology (see Haines, 1975; Ladd & Zober, 1977; Pekelman & Sen, 1975; Ratchford, 1975, 1979; Ryans, 1974). It has also been the basis for the per dollar product spaces used in many consumer psychology models (see Hauser, 1988; Hauser & Shugan, 1983; Hauser & Simmie, 1981). The latter two references develop a type of probabilistic vector model for external analysis for first choice data in a framework similar to the Lancaster theory.

Lancaster's model differs from traditional economic presentations of consumer demand in that utility is derived from the characteristics that goods possess and not the goods themselves. As Lancaster (1971) argues:

A steak is seen, in this sense, as something to eat, with many properties that could be agreed on by everyone, and not as something with which to make clothing. Some may not like steak, or may prefer chicken, or may have a religious tabu against eating it, but can still agree that a steak is something with specific properties of its own. Those who do not like steak dislike it because of some of its properties, not because they see an entirely different set of properties from the steak-eaters (p. 6).

Satisfaction is derived from such properties or characteristics.

In a system that consists of N different goods and r different characteristics, consumers observe a vector of goods $\mathbf{x} = (x_1, \dots, x_N)'$ and transform them into a vector of characteristics $\mathbf{c} = (c_1, \dots, c_r)'$ in a linear fashion:

$$c_i = \sum_{j=1}^N b_{ij}x_j, \quad i = 1, \dots, r,$$

or equivalently,

$$\mathbf{c} = \mathbf{B}\mathbf{x},$$

where \mathbf{B} is the matrix of coefficients (b_{ij}). Lancaster calls this the consumption technology matrix (CTM). The j -th column of \mathbf{B} represents the amount of each characteristic possessed by the j -th good. The relationship $\mathbf{c} = \mathbf{B}\mathbf{x}$ implies that a goods vector will be associated with a unique characteristics vector; but not necessarily the converse. The set of affordable combinations of goods, $S = \{\mathbf{x} | \mathbf{p}'\mathbf{x} \leq w\}$, can be mapped into an affordable set of characteristics using the CTM, $K = \{\mathbf{c} | \mathbf{c} = \mathbf{B}\mathbf{x} \text{ and } \mathbf{p}'\mathbf{x} \leq w\}$. K is the image of a linear transformation of a convex set, and consequently, is also convex. A fundamental property of any convex set is that it can be characterized as all convex combinations of its extreme points (Lancaster, 1971; p. 26). Therefore, S is the set of all convex combinations of its $N + 1$ extreme points:

$$\begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \begin{bmatrix} w/p_1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ w/p_2 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ w/p_N \end{bmatrix},$$

where $p_j =$ the price of one unit of the j -th good. K is therefore the image of the mapping of all convex combinations of these points. Since K is convex, it also can be characterized as the convex combinations of its extreme points. Lancaster demonstrates that all extreme points in K are the images of extreme points in S . However, not every extreme point in S need map into one in K .

Lancaster makes most of the traditional assumptions about a consumer's utility function. In particular, the utility function is continuous, monotonic, and produces indifference curves that are convex to the origin. The only difference is that the utility function operates on characteristics rather than goods. The consumer's problem then becomes:

$$\begin{aligned} \text{Max } & u(\mathbf{c}) \\ \text{s.t. } & \mathbf{c} \in K. \end{aligned}$$

Just as in the classical theory, basic optimization principles imply that the optimal \mathbf{c} , \mathbf{c}^* , must lie at the highest indifference curve which intersects K . Since utility functions are monotonic, this must occur on the outer boundary of K . A point \mathbf{c} lies in the outer boundary if there is no point $\mathbf{c}' \in K$ such that $\mathbf{c}' \succ \mathbf{c}$. This set of outer boundary points, E , is called the *efficiency frontier*. The solution to Lancaster's model is the point on the efficiency frontier that is tangent to an indifference curve if such exists. If not, the solution is a vertex of E , each of which is the image of a vertex in S . Lancaster (1971) shows that the combination of goods \mathbf{x}^* that corresponds to \mathbf{c}^* satisfy $\mathbf{p}'\mathbf{x}^* = w$, and can be found through the following linear program:

$$\begin{aligned} \text{Min } & \mathbf{p}'\mathbf{x} \\ \text{s.t. } & \mathbf{B}\mathbf{x} = \mathbf{c}^*, \mathbf{x} \geq 0. \end{aligned}$$

We would be remiss if we did not point out that Lancaster's theory has its critics (Ladd & Zober, 1977). Three implicit assumptions have been questioned. First, Hendler (1975) suggests that the marginal utilities of certain characteristics may be negative. Second, Lucas (1975) and Hendler both suggest that consumption technologies may be nonlinear. Finally, Lucas and Hendler both argue that utility may not be independent

of the characteristics' distribution across categories. For example, a consumer would prefer one combination of food to another with the same carbohydrates, protein, et cetera. If any or all of these criticisms apply, Lancaster's analysis may break down. In particular, the optimal c , may not be the point of tangency between an indifference curve and E , as defined earlier (Ladd & Zober, 1977).

In MICROSCALE, physical attributes are translated into dimensions upon which consumers base their decisions. In Lancaster's model, goods are translated into characteristics upon which consumers base their decisions. Thus, there is a natural correspondence between MICROSCALE and Lancaster's model. Each takes an observable construct, a brand's physical attributes in MICROSCALE and a collection of goods in Lancaster's model, and transforms them into an unobservable one, dimensions in MICROSCALE and characteristics in Lancaster's model, which consumer's ultimately use. Our dimensions are somewhat different from Lancaster's characteristics. As Ratchford (1975) points out, Lancaster asserts that characteristics should be defined in terms of objective measurement and not people's reactions. In contrast, multidimensional scaling procedures (ours included) can focus on a consumer's perceptions or preferences as well as objective characteristics. Ratchford suggests that the vector model option in PREFMAP (Carroll, 1972) may be useful in empirical applications of Lancaster's theory. Such an approach, however, carries a number of limitations with respect to Lancaster's theory. First, as indicated earlier, the indifference curves are linear. Second, rotational indeterminacies in PREFMAP (and other two-way operationalizations of the vector model) inhibit the dividing by price in goods space. Therefore, construction of the efficiency frontier becomes impossible. In contrast, MICROSCALE not only has convex indifference curves, but also produces unique solutions in which all characteristics are positive ratio scales. Additionally, while Lancaster (and microeconomics in general) concerns himself with choice among bundles of different products, we focus on choice among different brands within a single product class. Finally, unlike Lancaster, MICROSCALE takes on a probabilistic form with a just noticeable difference term, similar in form to many economic random utility models of choice (see McFadden, 1976).

3. The MICROSCALE Procedure

The Full MICROSCALE Model

Let:

$i = 1, \dots, I$ consumers;

$j, k = 1, \dots, N$ brands;

$t = 1, \dots, T$ dimensions;

$$\delta_{ijk} = \begin{cases} 1 & \text{if consumer } i \text{ prefers brand } j \text{ to } k; \\ 0 & \text{if consumer } i \text{ is indifferent between brands } j \text{ and } k; \\ -1 & \text{if consumer } i \text{ prefers brand } k \text{ to } j; \end{cases}$$

X_{jt} : the t -th coordinate for brand j , where $X_{jt} > 0$;

U_{ij} : the latent utility derived from the choice of brand j by consumer i as described by a Cobb-Douglas formulation:

$$= a_i \prod_{t=1}^T X_{jt}^{\alpha_{it}} + e_{ij}, \quad (1)$$

where:

a_i : a multiplicative constant for consumer i , $a_i \geq 0$;

α_{it} : the t -th coordinate of the exponent for consumer i with:

$$0 \leq \alpha_{it} \leq 1,$$

$$\sum_t \alpha_{it} = 1,$$

e_{ij} = an error term $\sim N(0, \sigma_i^2)$, where it is assumed that:

$$\text{Cov}(e_{ij}, e_{ik}) = 0, \quad \text{for all } i, j \neq k,$$

$$\text{Cov}(e_{ij}, e_{i'k}) = 0, \quad \text{for all } j, k, i \neq i'.$$

Equation (1) is of a form that produces convex indifference curves. Lancaster (1971) suggests two possible functional forms: the Cobb-Douglas function (1) and the CES (constant elasticity of substitution) production function:

$$U_{ij} = a_i \left[\prod_{t=1}^T \beta_{it} X_{jt}^{-\rho} \right]^{-1/\rho}, \quad (2)$$

where a_i , β_{it} , and ρ are model parameters with $a_i, \rho > 0$, $\sum_{t=1}^T \beta_{it} = 1$, and $0 < \beta_{it} < 1$.

Lancaster cites two limitations of these forms with respect to utility representation. First, they are homogeneous of degree 1, implying that doubling each characteristic or dimension would double the utility. Second, they have unit elasticities of substitution implying that the proportion of total income spent on each characteristic or dimension depends only on the function's parameters and not on consumer's income. We note that these also apply to the linear utility representation in the traditional vector model. However, as Lancaster points out, these limitations are of minimal concern since properties of *aggregate* demand are determined by the distribution of preference parameters (α_{it} and β_{it}) across the population. Furthermore, also following Lancaster, we choose (1) over (2) for reasons of simplicity. Note, in most choice models that have multiplicative forms similar to (1), the error term is multiplicative and lognormal (see Nakanishi & Cooper, 1974). In contrast, we will assume that the error relationship with the utility function is additive and the e_{ij} are iid normal with mean zero and variance σ_i^2 . This is necessary to avoid taking logarithms of negative numbers in the subsequent development.

In MICROSCALE, $P(\text{consumer } i \text{ prefers brand } j \text{ to } k) = P(\delta_{ijk} = 1) = P(U_{ij} > U_{ik} + s_i)$, where s_i is a just-noticeable difference parameter (Thurstone, 1927),

$$\begin{aligned}
 &= P\left(e_{ik} - e_{ij} < a_i \left(\prod_t X_{jt}^{\alpha_{it}} - \prod_t X_{kt}^{\alpha_{it}} \right) - s_i \right), \\
 &= \Phi\left(\frac{a_i \left(\prod_t X_{jt}^{\alpha_{it}} - \prod_t X_{kt}^{\alpha_{it}} \right) - s_i}{(2\sigma_i^2)^{1/2}} \right) \\
 &\stackrel{\text{wlog}}{=} \Phi\left(a_i \left(\prod_t X_{jt}^{\alpha_{it}} - \prod_t X_{kt}^{\alpha_{it}} \right) - s_i \right) = G_{ijk}^{(1)}, \tag{3}
 \end{aligned}$$

where Φ is the standard normal cdf. The denominator vanishes since the variance term can be embedded in the a_i and s_i terms. Similarly,

$$\begin{aligned}
 P(\text{subject } i \text{ prefers brand } k \text{ to } j) &= P(\delta_{ijk} = -1) = P(U_{ik} > U_{ij} + s_i) \\
 &= \Phi\left(a_i \left(\prod_t X_{kt}^{\alpha_{it}} - \prod_t X_{jt}^{\alpha_{it}} \right) - s_i \right) = G_{ijk}^{(2)}. \tag{4}
 \end{aligned}$$

And

$$\begin{aligned}
 P(\text{consumer } i \text{ is indifferent between brands } j \text{ and } k) &= P(\delta_{ijk} = 0) \\
 &= P(|U_{ij} - U_{ik}| < s_i) \\
 &= 1 - G_{ijk}^{(1)} - G_{ijk}^{(2)} = G_{ijk}^{(3)}. \tag{5}
 \end{aligned}$$

Note that we model a consumer specific just-noticeable threshold term, s_i , which is a departure from Lancaster (1966, 1971). MICROSCALE assumes that consumers make choices that follow a random utility model (McFadden, 1976) with thresholds (Krishnan, 1977). Most utility based models of consumer choice assume that the decision makers always select the alternative that provides them with maximum utility. Accordingly, they have no problem recognizing differences in utilities, no matter how small. We take the view espoused by Lioukas (1984) who suggests that consumers may be insensitive to small differences and that certain thresholds have to be exceeded before the consumer notices a difference (the just-noticeable difference). In economics, an analogous notion was proposed by Quandt (1956) who argued that a band of indifference exists around the level of utility derived from a given quantity of a good. Small changes in the quantity will not affect utility. Later, Krishnan (1977) described thresholds as minimum perceivable differences between the utilities of choice alternatives. According to him, an individual would prefer one alternative to another only if the corresponding difference in utilities exceeded that threshold. We adopt this view in MICROSCALE.

Finally, an important special case of the above model arises when data are collected as forced-choice paired comparisons. In that case, s_i is equal to zero and the last term in (5) vanishes. We note that, like Thurstone's Case V unidimensional scaling, MICROSCALE assumes strong stochastic transitivity.

Estimation

Assuming independence over $i, j,$ and k subscripts, the general form of the likelihood function is

$$L = \prod_{i=1}^I \prod_{j < k}^N G_{ijk}^{(1)\delta_{ijk}^{(+)}} G_{ijk}^{(2) - \delta_{ijk}^{(-)}} G_{ijk}^{(3)1 - |\delta_{ijk}|}, \quad (6)$$

and the log likelihood function is:

$$\ln L = \sum_{i=1}^I \sum_{j < k}^N [\delta_{ijk}^{(+)} \ln G_{ijk}^{(1)} - \delta_{ijk}^{(-)} \ln G_{ijk}^{(2)} + (1 - |\delta_{ijk}|) \ln G_{ijk}^{(3)}], \quad (7)$$

where

$$\delta_{ijk}^{(+)} = \max(0, \delta_{ijk}),$$

$$\delta_{ijk}^{(-)} = \min(0, \delta_{ijk}).$$

Given $\Delta = ((\delta_{ijk}))$ and T , MICROSCALE attempts to estimate (in the full model) a_i , s_i , $\mathbf{X} = (X_{jt})$, and $\alpha = (\alpha_{it})$. We use the method of conjugate gradients with automatic restarts (Powell, 1977) for parameter estimation. We impose the following constraints consistent with our earlier development (for all i, j , and t): (i) $a_i \geq 0$, (ii) $X_{jt} \geq 0$, (iii) $0 \leq \alpha_{it} \leq 1$, (iv) $\sum_t \alpha_{it} = 1$, and (v) $s_i \geq 0$. The nonnegativity constraints in (i), (ii), and (v) are implemented by substituting squared quantities for these parameters (e.g., substituting a_i^2 , s_i^2 , X_{jt}^2 for their respective unsquared terms in the likelihood formulation—see Gill, Murray, & Wright, 1981, pp. 268–269). The constraints on α_{it} in (iii) and (iv) above are operationalized by reparameterizing

$$\alpha_{it} = \frac{\theta_{it}^2}{\sum_t \theta_{it}^2}, \quad (8)$$

and estimating θ_{it}^2 . These techniques of constraint implementation still allow us to utilize an unconstrained optimization procedure that simplifies the estimation process considerably. The Appendix contains a technical description of the alternating conditional maximum likelihood (ACML) procedure we developed for the estimation of the MICROSCALE model.

Program Options

MICROSCALE can accommodate a number of different model specifications and options. One can perform either an internal analysis (where the user estimates both α and \mathbf{X}) or an external analysis (where the user can fix one or more sets of coordinates throughout the analysis). The user can also select among a number of methods for generating starting estimates, including a user-defined singular value decomposition (SVD) option for \mathbf{X} , where all values are made positive by adding the absolute value of the smallest X_{jt} to all values of \mathbf{X} plus .001. Also, a variety of specifications are allowed on a_i (a_i given externally, $a_i = a$, $a_i = 1$, estimate general a_i), s_i (s_i given externally, $s_i = s$, $s_i = 0$, estimate general s_i), and α_{it} (α_{it} given, $\alpha_{it} = \alpha_t$, $\alpha_{it} = \frac{1}{T}$, estimate general α_{it}).

Perhaps the most valuable option that ties MICROSCALE directly into the theoretical work of Lancaster concerns the potential for reparameterizing the brand/product coordinates as direct functions of their observed physical attributes. That is:

$$X_{jt} = f_1(H_{j1}, H_{j2}, \dots, H_{jM}), \quad (9)$$

where H_{jm} represents the value of the m -th physical attribute of brand j ($m = 1, \dots, M$). Expression (9) captures the notion that dimensions are a function of physical attributes. Following Lancaster's CTM, we assume that f_1 is linear,

$$X_{jt} = \sum_{m=1}^M H_{jm} \gamma_{mt}, \quad (10)$$

where the γ_{mt} are unknown parameters. As in CANDELINC (Carroll, Pruzansky, & Kruskal, 1980), three-way multivariate conjoint analysis (DeSarbo, Carroll, Lehmann, & O'Shaughnessy, 1982), and GENFOLD2 (DeSarbo & Rao, 1984, 1986), one can use these reparameterization options to examine what impact such features/attributes have on the derived solution. This can often aid in interpreting the resulting solution (see Bentler & Weeks, 1978; Bloxom, 1978; de Leeuw & Heiser, 1980), and can replace the post-analysis property-fitting method often used in an attempt to interpret the results. In addition, as we shall discuss, the imposition of these sets of reparameterizations can provide an effective tool for understanding the nature of preference or choice.

It should be noted that when a linear function replaces a product coordinate, the number of background variables in the linear function cannot exceed the number of entities that exist for those variables. For example, if N brands have M attributes, $N \geq M$ since one can only identify at most NT coordinates (for \mathbf{X}). Thus, in most applications, such reparameterizations actually improve the degrees of freedom of the model by reducing the number of parameters to be estimated. Given $I(N)(N-1)/2$ independent paired comparison judgments, the number of estimated model parameters in the full model ($a_i, s_i, \alpha_{it}, \mathbf{X}$) is:

$$P = NT + I(T+1) - 1.$$

Note, one of the benefits of MICROSCALE is the fact that one obtains an \mathbf{X} configuration that is unique in the positive orthant of a T -dimensional space (there is no rotational indeterminacy as in competing models) up to a scale indeterminacy with the a_i coefficients in the full model (one must set one of the a_i 's equal to 1).

Goodness-of-Fit Indices

A number of goodness-of-fit measures are computed for MICROSCALE.

1. A deviance measure (McCullagh & Nelder, 1983; Nelder & Wedderburn, 1972): $D = -2 \ln L$. Note that one can potentially test nested models within MICROSCALE as the difference between respective deviance measures. One obvious problem with this approach concerns the presence of incidental parameters in the likelihood function (i.e., parameters whose order vary according to the order of Δ , such as the α_{it} 's) as there are no within-subject replications. In such cases, maximum likelihood estimators may not be consistent. Other measures should also be examined.

2. The proportion of correct predictions in Δ . Here, the proportion of times the solution correctly predicts δ_{ijk} is calculated using a simple matching coefficient for the total sample as well as for each subject. In addition, a Spearman's rho correlation coefficient between δ_{ijk} and $\hat{\delta}_{ijk}$, predicted choices, is also computed.

3. Akaike's (1974) information criterion defined as: $-2 \ln L + 2P$. However, as noted by Bozdogan (1987), the AIC statistic tends to overestimate the appropriate dimensionality for certain models tested, emphasizing the need to examine multiple measures.

4. For situations where δ_{ijk} contains no ties, a phi coefficient and point biserial correlation are also computed.

Microeconomic Extensions

Given the above MICROSACLE formulation, a number of interesting analogs to the microeconomic theory of the consumer are possible. For example, one can define a consumer's marginal utility with respect to a particular dimension of a specific brand as:

$$\begin{aligned} MU_i(X_{jt}) &= \frac{\partial \hat{U}_{ij}}{\partial X_{jt}} = \frac{\partial}{\partial X_{jt}} \left[a_i \prod_{t=1}^T X_{jt}^{\alpha_{it}} + e_{ij} \right] \\ &= a_i \alpha_{it} X_{jt}^{\alpha_{it}-1} \prod_{l \neq t} X_{jl}^{\alpha_{il}}, \end{aligned} \quad (11)$$

or in the case of reparameterizing \mathbf{X} in terms of physical attributes:

$$MU_i(H_{jm}) = \frac{\partial \hat{U}_{ij}}{\partial H_{jm}} = a_i \alpha_{it} \gamma_{mt} X^{\alpha_{it}-1} \prod_{s \neq t} X_{js}^{\alpha_{is}}. \quad (12)$$

Of particular interest in each case is that marginal utility is a function of both consumer (i) and brand (j) (and/or its physical attribute (m)) parameters. With such notions of marginal utility, one could determine the marginal rate of substitution (MRS) that, in the case of product coordinates, is defined as the rate at which consumer i 's consumption of dimension v can be reduced, without reducing his overall utility, when his consumption of dimension r is increased at the rate of an additional unit:

$$\text{MRS}_{rv}^{(i)} = \frac{MU_i(X_{jr})}{MU_i(X_{jv})} = \frac{\frac{\partial U_{ij}}{\partial X_{jr}}}{\frac{\partial U_{ij}}{\partial X_{jv}}}. \quad (13)$$

This is also equal to the negative of the slope of the derived indifference curve defined in this region. Given the strict convexity to the origin requirement of the indifference curves estimated by MICROSACLE, the marginal rate of substitution of dimension r for dimension v with respect to a change in the consumption of dimension r , along an indifference curve, is strictly decreasing. This implies the notion of a diminishing marginal rate of substitution, here defined for the dimensions. A similar extension can be derived for the physical attributes as in Lancaster (1966, 1971).

As an analog to vector MDS models (e.g., MDPREF), one can also solve for individual specific preference vectors. For convenience, assume one consumer and two dimensions, where \mathbf{X} is confined in the positive quadrant of a two dimensional space. Given the requirement of strict convexity to the origin of the resulting indifference curves, the direction or vector of increasing utility will necessarily have a 0° - 90° orientation. The question is to find an exact orientation, which will necessarily vary according to the specific orientation of a consumer's indifference curve. Suppose we hold utility constant for a consumer and thus concentrate on one indifference curve. Then, the shortest ray from the origin to a point on that curve will define the direction of increasing preference. If one were to pick an alternative point, one could reorient a ray between the origin and that point in the direction of the point with the shortest distance, and with that original (longer) distance attain a higher level of utility (on a higher indifference curve). A formal proof by contradiction can easily be shown on these grounds. Thus, we want to find the point (r_1, r_2) that lies on some fixed or

constant level of utility U_i^* that is minimum distance or squared distance from the origin, where squared distance here is:

$$d^2 = (r_1 - 0)^2 + (r_2 - 0)^2 = r_1^2 + r_2^2. \quad (14)$$

Now

$$U_i^*(r_1, r_2) = a_i r_1^{\alpha_1} r_2^{1-\alpha_1}, \quad (15)$$

and solving for r_1 , we obtain

$$r_1 = \left(\frac{U_i^*}{a_i r_2^{1-\alpha_1}} \right)^{1/\alpha_1}. \quad (16)$$

Substituting in (14), one obtains

$$d^2 = \left(\frac{U_i^*}{a_i r_2^{1-\alpha_1}} \right)^{2/\alpha_1} + r_2^2. \quad (17)$$

Taking derivatives of (17), setting them equal to zero, and solving for r_2 , gives

$$r_2^* = \frac{U_i^*}{a_i} \left(\frac{1-\alpha_1}{\alpha_1} \right)^{\alpha_1/2}. \quad (18)$$

Substituting this in (16), we solve for r_1^*

$$r_1^* = \left(\frac{U_i^*}{a_i r_2^{*1-\alpha_1}} \right)^{1/\alpha_1} = \frac{U_i^*}{a_i} \left[\frac{\alpha_1}{1-\alpha_1} \right]^{(1-\alpha_1)/2}. \quad (19)$$

Given (r_1^*, r_2^*) for a specific indifference curve, one can now draw a ray from the origin to this point to indicate the direction of increasing utility—or what we call the preference expansion direction (PED). Given the fact that indifference curves for any consumer will be parallel, we can extend this ray to infinity to denote this direction for higher levels of utility (see also Grandmont, 1978, for an alternative vector formulation for the Cobb-Douglas utility function and interpretations relating to the elasticity of substitution). This formulation generalizes to $T > 2$ dimensions in a straightforward manner.

Finally, we wish to discuss the remaining issue of utility maximization and optimal choice in the derived space. As mentioned, MICROSCALE will produce a configuration of brands/products (\mathbf{X}), indifference curves (α_{it}), and we can derive a PED (also, if \mathbf{X} is reparameterized, we can also represent γ_{mt} by direction vectors in the space emanating from the origin). Unlike traditional microeconomic theory, MICROSCALE analyses are typically performed for brands in a designated product class (\mathbf{X} are brands in a T -dimensional space) as opposed to vectors of commodities. As such, there is little sense to impose a budget constraint in MICROSCALE. For example, if the product class is over-the-counter analgesics, one typically assumes all are affordable given the rather small range of price variation. Instead, analogous to Lancaster's efficiency frontier, MICROSCALE can construct a technological feasibility frontier (TFF) defined as the outer boundary of the convex hull of the outer set of brand coordinates. Depending upon the product class and nature of the derived dimensions, this TFF can often be interpreted as a limit to the available technology with respect to the derived dimensions and the present set of brands studied. If, for the over-the-counter analgesics illustration, the dimensions relate directly to ingredients, the TFF could be utilized to define

TABLE 1

Independent Variables Utilized in the MICROSCALE Monte Carlo Analysis

<u>Factor</u>	<u>Designation</u>	<u>Levels</u>	<u>Code</u>
I	Number of Consumers	10	-
		20	X ₁
		30	X ₂
II	Number of Brands	5	-
		10	X ₃
		15	X ₄
III	Number of Dimensions	2	-
		3	X ₅
		4	X ₆
IV	Type of Threshold (s _i)	0	-
		Constant Varying	X ₇ X ₈
V	Amount of Error	$\sigma = 1$	-
		$\sigma = 2$	X ₉
		$\sigma = 3$	X ₁₀
VI	Error Distribution	Normal	-
		Uniform	X ₁₁
		Exponential	X ₁₂
VII	Type of Analysis	External	-
		Internal	X ₁₃
VIII	Starting Solution	Rational (SVD)	-
		Random	X ₁₄

existing limits on the amount of specific ingredients (e.g., the FDA limits the amount of aspirin per tablet that can be sold without a prescription). This notion may be more relevant for external preference analyses where the dimensions are known physical attributes. This convex hull can be easily constructed by splines whose knots are the brand locations of the outermost brands in the space. Given this convex hull approximation, the point of maximal feasible utility for an individual would be the point of intersection of his/her highest utility indifference curve and this TFF.

4. Monte Carlo Analysis

To examine the performance of the procedure, a Monte Carlo analysis was performed as a number of data, model, and error factors were experimentally manipulated.

TABLE 2

Dependent Measures Utilized in the MICROSACLE Monte Carlo Analysis

A. Computational Effort:

Y_1 : Number of iterations required for convergence.

B. δ_{ijk} Goodness-of-Fit:

Y_2 : A matching coefficient of $(\delta_{ijk}, \hat{\delta}_{ijk})$;

Y_3 : Spearman rho correlation of $(\delta_{ijk}, \hat{\delta}_{ijk})$.

C. Parameter Recovery for each of X_j , α_j , and ζ_j :

Y_4, Y_5, Y_6 : Root mean square error (3 measures-- X_j, α_j, ζ_j);

Y_7, Y_8, Y_9 : Sum-of-squares accounted-for (3 measures-- X_j, α_j, ζ_j).

Table 1 presents the list of some eight independent variables and their respective levels that were utilized in the Monte Carlo analysis. The dependent measures were defined around three major constructs: (a) computational effort, (b) overall goodness-of-fit with δ_{ijk} , and (c) parameter recovery. Table 2 provides a list of the multiple dependent measures used here.

Given the $3^{6 \cdot 2^2}$ structure of the independent variable set, 2,916 trials would be required for a full factorial design with no replications. As in DeSarbo (1982) and DeSarbo and Carroll (1985), a main effects only design was utilized via an asymmetric fractional factorial design (see Addelman, 1962) to provide a feasible computer experiment. While the assumption of main effects only is clearly a limitation, the resulting 27 trials from the design provide a modest test of the procedure and its performance in recovering several synthetic data sets with known structure. Note, in all trials, $\sigma_i = a_i = 1$ was fixed. For each of the 27 trials, synthetic data were generated according to the first six factors in Table 1. Values of the multiple dependent measures were collected for each trial and regression/ANOVA analyses of these Monte Carlo results were performed. Note, for those dependent measures ranging between 0 and 1, the arcsin transformation was applied to the square root of that measure—a step suggested in Snedecor and Cochran (1981, p. 196) to stabilize the resulting variance. The significant results of these analyses are summarized below.

Concerning computational effort and the first dependent measure (Y_1), the only consistent finding is that internal analyses (X_{13}) tend to significantly increase the amount of computation involved in estimation, as opposed to external analyses where X is fixed.

Concerning overall goodness-of-fit and the second pair of dependent measures (Y_2 and Y_3), we observe consistent findings with respect to the two higher values of $\sigma(X_9$ and $X_{10})$ significantly reducing overall goodness-of-fit. Also, the uniform error mis-

specification appears to significantly increase both measures here. The matching coefficient is significantly lower for larger sets of brands (X_3 and X_4), constant threshold specifications (X_7), and external analyses (X_{13}).

Concerning parameter recovery and the pair of measures concerning \mathbf{X} (Y_4 and Y_7), consistent results are obtained with respect to the fact that external analyses (X_{13}) result in better recovery of \mathbf{X} (since \mathbf{X} is fixed) and that somewhat poorer recovery occurs in larger dimensionality (X_5 and X_6) where more parameters are estimated. Surprisingly, using the rational start procedure (X_{14}) tends to result in poorer sums-of-squares accounted-for-values. With respect to α recovery, the only consistent finding across the two measures (Y_5 and Y_8) is the fact that larger stimulus (X_3 and X_4) sets (i.e., more data) tends to result in significantly better α recovery. For Y_5 , specifying larger dimensionality (X_5 and X_6) detracts significantly from α recovery, while the uniform error misspecification surprisingly improves such parameter recovery. For Y_8 , holding \mathbf{X} fixed in external analyses (X_{13}) improves the RMS measure. Finally, with respect to s recovery and Y_6 and Y_9 , the only consistent finding is that it appears harder to recover the true s when s is allowed to vary over subjects (X_8), and thus, when there are more parameters to estimate.

Overall, with some exceptions, these Monte Carlo analyses appear to give a somewhat consistent description of the performance of MICROSACLE. All else held constant, MICROSACLE tends to require more computational effort for internal analyses; data recovery suffers as σ is larger; and, parameter recovery tends to improve as the data set size increases (especially with respect to J) and there are smaller numbers of parameters to estimate—a common finding in nonlinear model estimation. Perhaps the most unexplainable finding concerns the role of the uniform and exponential error misspecification that one would initially hypothesize to have deleterious effects across all these dependent measures, but yet has scattered incidence of positive benefits in some cases. In addition, the rational start had relatively no effect on all but one measure, and that effect was negative (Y_4).

Obviously, there are limitations to such a modest study. Barring restrictions on computer time, it would have been preferable to perform a full factorial experiment with replications to measure possible interaction effects that may confound main effects only estimation. Such a study would have been prohibitive from a cost/time perspective. In addition, it would have been beneficial to specify additional levels per factor (especially concerning the error misspecification factor) and perhaps additional factors (e.g., control parameters for the algorithm).

5. A Consumer Psychology Illustration

Study Design

A sample of $I = 30$ undergraduate students at a major northeastern business school were asked to take part in a small study designed to measure preferences for various brands of existing over-the-counter (OTC) analgesic pain relievers. These respondents were initially questioned as to the brand(s) they currently use (as well as frequency of use) and their personal motivations for why they chose such brand(s) (e.g., ingredients, price, availability, etc.). They were then presented fourteen existing OTC analgesic brands: Advil (ADV), Anacin (ANA), Anacin-3 (ANA3), Ascriptin (ASC), Bayer (BAY), Bufferin (BUF), Cope (COP), CVS Buffered Aspirin (a generic) (CVS), Datril (DAT), Excedrin (EXE), Nuprin (NUP), Panadol (PAN), Tylenol (TYL), and Vanquish (VAN).¹ The letters in parentheses are plotting codes that will be used throughout this paper.

¹ The products listed are registered trademarks of their respective companies.

TABLE 3
OTC Analgesic Brand Descriptions

Plot Code	Brand	Mg. of Aspirin 1	Mg. of Aceta- minaphon 2	Mg. of Ibuprofen 3	Mg. of Caffeine 4	Mg. of Buffered Compounds 5	Price in U.S. Dollars 6	Maximum Recommended Dosage 7	Market Share 8
ADV	Advil	0	0	200	0	0	6.99	6	3.7
ANA	Anacin	400	0	0	32	0	3.97	10	10.3
ANA3	Anacin-3	0	500	0	0	0	5.29	8	6.8
ASC	Ascriptin	325	0	0	0	150	3.29	12	2.6
BAY	Bayer	325	0	0	0	0	2.69	12	20.6
BUF	Bufferin	324	0	0	0	100	3.89	10	8.8
COP	Cope	421	0	0	32	75	5.31	8	0.4
CVS	CVS	500	0	0	0	100	1.99	12	0.3
DAT	Datril	0	500	0	0	0	5.75	8	2.1
EXE	Excedrin	250	250	0	65	0	4.99	8	12.0
NUP	Nuprin	0	0	200	0	0	7.59	6	3.4
PAN	Panadol	0	500	0	0	0	4.99	8	3.2
TYL	Tylenol	0	325	0	0	0	3.69	8	22.7
VAN	Vanquish	227	194	0	33	75	4.99	12	1.1

Correlations

Feature

1	1.000								
2	-.653	1.000							
3	.442	-.323	1.000						
4	.354	-.063	-.239	1.000					
5	.623	-.458	-.288	-.065	1.000				
6	-.673	.199	.716	.061	-.447	1.000			
7	.704	-.302	-.611	-.019	.624	-.828	1.000		
8	-.074	.038	-.203	.013	-.424	-.356	.051	1.000	

Initially, they were presented colored photographs of each brand and its packaging, together with price per 100 tablets, ingredients, package claims, and manufacturer (see DeSarbo & Carroll, 1985). Table 3 presents summaries of selected portions of the descriptions for each of the brands. Each subject/consumer was requested to read this information and return to it anytime during the experiment if he/she so wished. After a period of time, they were asked to make paired comparison preference judgments for all possible 91 pairs of brands. They were told that they had to choose one from each pair (i.e., no ties were allowed). The presented pairs were randomized for each subject. DeSarbo, De Soete, Carroll, and Ramaswamy (1988) have utilized these data in fitting

TABLE 4

Goodness-of-Fit Indices for Analgesics Data

T	Model DF	ln L	MC	PBC	PHI
1	13	-1494.71	.713	.510	.425
2	55	-1148.70	.808	.674	.615
3	98	-1037.77	.828	.717	.655
4	141	-875.54	.855	.769	.710

Correlations Between \bar{X} and Brand Descriptors

<u>Descriptor</u>	<u>Dimension</u>	
	I	II
1	.474	-.806
2	-.127	.156
3	-.541	.849
4	-.072	-.328
5	.067	-.585
6	-.856	.749
7	.558	-.813
8	.597	.030

various stochastic ultrametric tree models designed for such paired comparisons. DeSarbo, De Soete, and Jedidi (1986) have also reported analyses (two-dimensional solutions) utilizing the DeSarbo, Oliver, and De Soete (1986) and DeSarbo, De Soete, and Eliashberg (1987) vector and unfolding models for these paired comparisons data using reparameterization options for the brand coordinates.

MICROSCALE Analysis

A MICROSCALE analysis was performed on this resulting $30 \times 14 \times 14$ array in $T = 1$ to 4 dimensions with $s = 0$, given the absence of ties in δ_{ijk} , and $\mathbf{a} = \mathbf{1}$ to reduce the number of estimated parameters. Given the excessive collinearity between many of the brand attributes as shown in Table 3, we decided to perform non-reparameterized analyses. Table 4 provides a number of goodness-of-fit indices, that appear to suggest

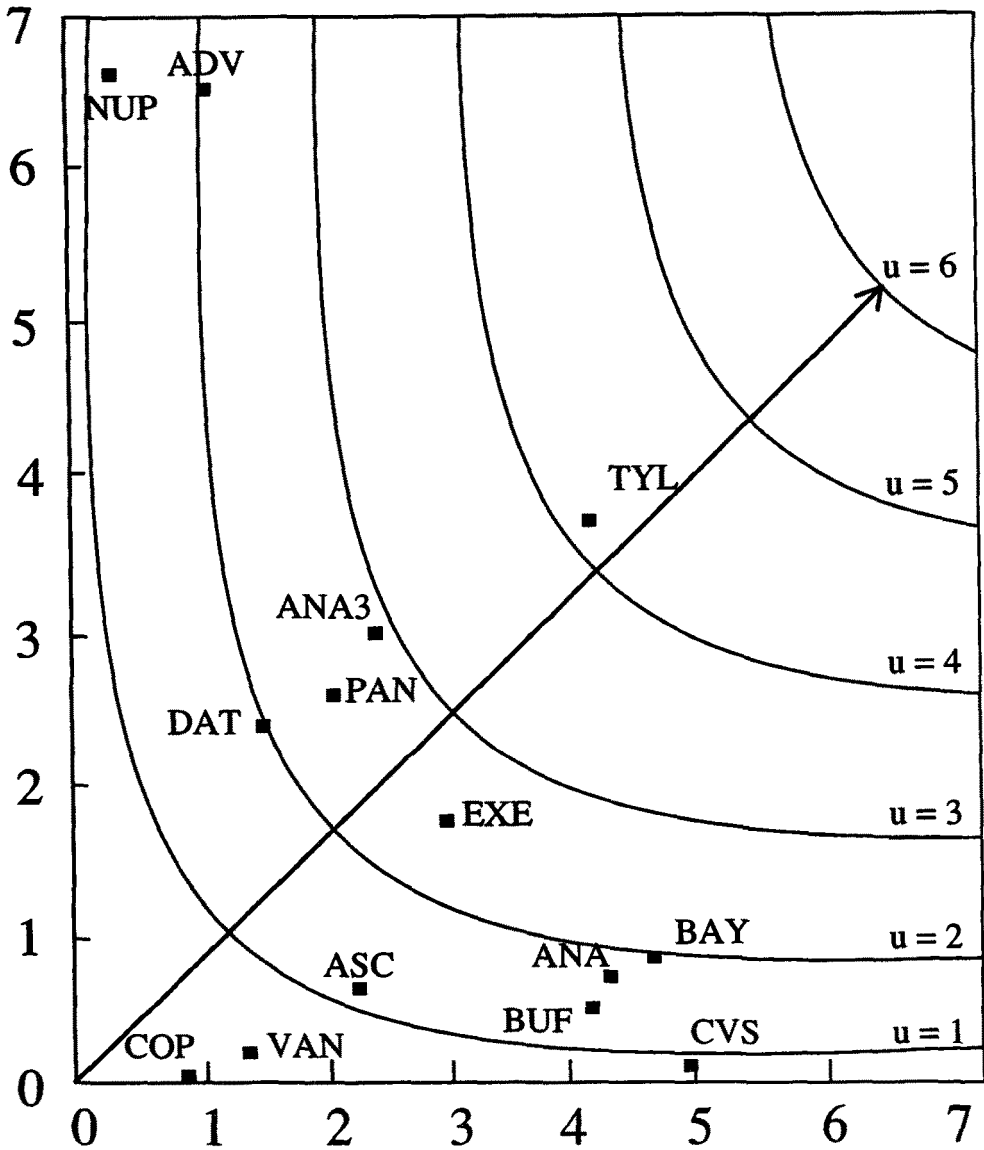


FIGURE 1
Student one's MICROSCALE joint space representation for the analgesics data.

the $T = 2$ dimensional solution is most parsimonious in describing the structure in the δ_{ijk} (AIC = 2407.16).

Figures 1 and 2 present the resulting joint space two-dimensional plots for two selected students. Common to all these figures is the notion that the first dimension (horizontal) appears to separate the major, highly advertised, high market-share brands from the remaining brands. Of particular interest here is the fact that CVS, a local generic brand popular in the area of the study, is also located in this group of higher market share/advertised brands (Anacin, Bayer, Bufferin, Tylenol), reflecting the particular environment under which the study was conducted. Also, this dimension is also highly correlated with price ($-.856$) indicating that the brands toward the positive end of this dimension (mostly aspirin-based) are much cheaper in cost than the ibuprofen

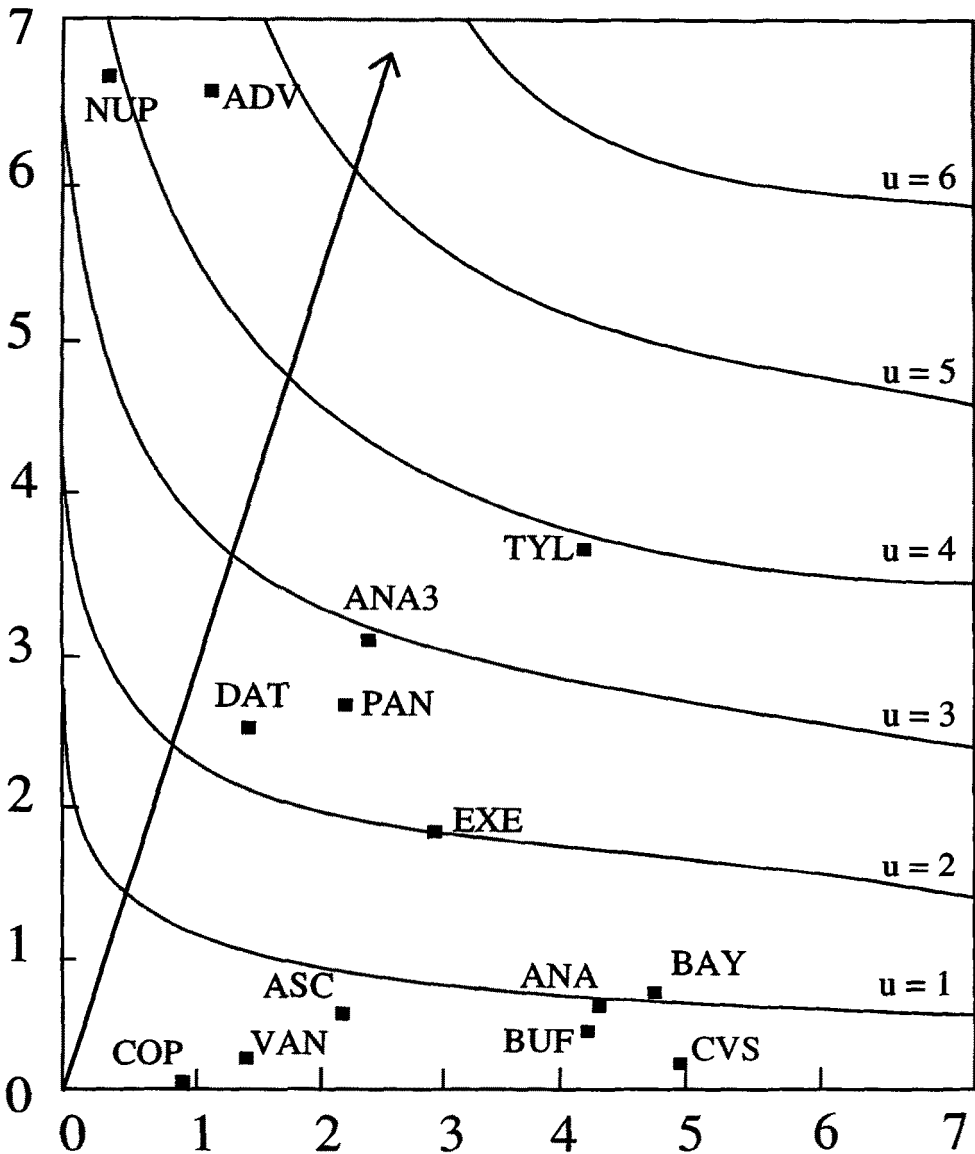


FIGURE 2

Student two's MICROSCALE joint space representation for the analgesics data.

and specialty aspirin brands located near the origin. The second dimension (vertical) clearly separates the ibuprofen brands from the aspirin-based brands, with the acetaminophen brands near the middle of the dimension. Note how Excedrin is located below the acetaminophen brands near the aspirin brands given the fact that it is comprised of both ingredients. Thus, the two dimensions appear to relate rather directly to the type of analgesic ingredient, cost, and popularity as demonstrated by the correlations between X and the features presented in Table 4.

Each figure shows the PED for each of the two students. We do not construct the TFF given the nature of the derived dimensions. Student one (Figure 1) displays symmetric looking indifference curves. For this student, the aspirin substitutes appear to be more desirable, especially Tylenol that has highest predicted utility. The PED for this student points near the direction of Tylenol. On the other hand, the second student

(Figure 2) displays rather skewed indifference curves. For this student, the ibuprofen brands, Nuprin and Advil, have highest utility. The PED for this student is directed just to the southeast of these two brands. Thus, MICROSCALE can reveal individual differences in conceptualization of preference/choice/utility—a feature missing in the wandering vector or ideal point models (De Soete & Carroll, 1983; De Soete, Carroll, & DeSarbo, 1986).

One can also compute the predicted utility values \hat{U}_{ij} for each of the 14 brands by subject. A noticeable congruence between \hat{U}_{ij} , the brand utility means, and the actual market shares ($r = +0.78$) is observed, where the major, heavily advertised brands such as Tylenol, Advil, Anacin, Anacin-3, Bayer, Bufferin, Excedrin, and Panadol are predicted on average to have higher overall utility.

Finally, given the fact that MICROSCALE is prone to locally optimal solutions, we performed ten computer analyses for the two dimensional solution with different random starts (SVD default option for X) and examined statistical summaries for these ten analyses, including means and standard deviations. With the possible exception of one run, all of the solutions rendered relatively similar goodness-of-fit measures (as well as parameter values for X and α). Thus, there is some evidence of somewhat stable solutions with MICROSCALE.

The DeSarbo, Oliver, and De Soete (1986) Vector Model Comparative Results

Given some similarity in utility assumptions between MICROSCALE and the vector model (i.e., the more the better), we decided to compare the above MICROSCALE two-dimensional solution with a two-dimensional solution obtained from applying the DeSarbo et al. (1986) vector model on the same data using similar program options. The resulting two-dimensional solution with $\sigma_{ijk} = 1$ and no reparameterization option, produced a log likelihood value (the overall mathematical structure of this likelihood and that of MICROSCALE are equivalent in the case of no ties) of -1152.44 (AIC = 2472.88), a phi coefficient = 0.637, a matching coefficient of 0.819, and a point biserial correlation of 0.629. Thus, this alternative model is not as good a representation as MICROSCALE according to the AIC statistic. Two of the other goodness-of-fit statistics are higher because many more free parameters are estimated in this vector model as compared to MICROSCALE.

Table 5 presents the correlations of this solution with the eight brand attributes reported in Table 3. Note, unlike MICROSCALE that produces a unique brand space, these loadings (interpretation) will change according to the specific type of rotation used in this vector model to interpret the results! As with most two-way, bilinear models, there is an indeterminacy of the resulting coordinates to any non-singular transformation. The first dimension appears to relate to market share. The second dimension contrasts the aspirin versus ibuprophen brands. The former also tend to be cheaper and involve higher maximum dosage recommendations. This interpretation is apparent by inspection of the resulting joint space representation shown in Figure 3. One of the advantages of this type of procedure over MICROSCALE is the ability to portray all subjects in one joint space plot versus having to inspect 30 of them. As shown, preference tends to be heavily in favor of the more popular, high market-share/highly advertised brands (the first dimension). However, there is substantial variation in preference concerning the second dimension—ingredients.

But how similar are the results from this vector model with those obtained from MICROSCALE? We performed canonical correlation analyses for the two resulting brand coordinate spaces as an “approximate” configuration matching procedure and obtained canonical correlations of 0.966 and 0.883. Given the uniqueness properties of

TABLE 5

The DeSarbo et al. (1986) Solution Correlations Between
the Derived Brand Coordinates and Attributes

Descriptor	Dimension	
	I	II
1	.313	-.701
2	-.370	.238
3	.129	.759
4	.278	-.301
5	.482	-.476
6	.320	.813
7	.048	-.792
8	-.748	-.208

the MICROSCALE solution, these canonical correlations would necessarily be overstated somewhat. Using the MICROSCALE dimensions as dependent variables and the brand coordinates for the vector model as independent variables in multiple regression analyses, the two resulting adjusted R^2 values were 0.882 and 0.867 indicating somewhat high consistency between the two solutions. In addition, measures of association between the utility scores were computed for each of the two models across the 30 subjects (average correlation across subjects = 0.94). With the possible exception of one or two subjects, there appears to be substantial congruence between these derived utility scores across this sample.

Thus, the MICROSCALE procedure provides a slightly better fit to this data according to an AIC criteria. Its primary advantage is the unique brand coordinate space derived where the user need not worry about selection of an appropriate rotation. Yet, both solutions are reasonably congruent with respect to the basic structure of utility derived from this same data set. Note, there are differences concerning policy implications for individual utility maximization. In the DeSarbo et al. (1986) vector model, to maximize a subject's utility, a brand would be positioned in the space (within a technologically feasible area) where its projection on a subject's vector would be highest. In MICROSCALE, this utility maximizing brand location for a given subject would be positioned at the intersection of the subject's highest indifference curve and the TFF. Future experimental research should focus upon the specific conditions under which a particular utility representation is most appropriate.

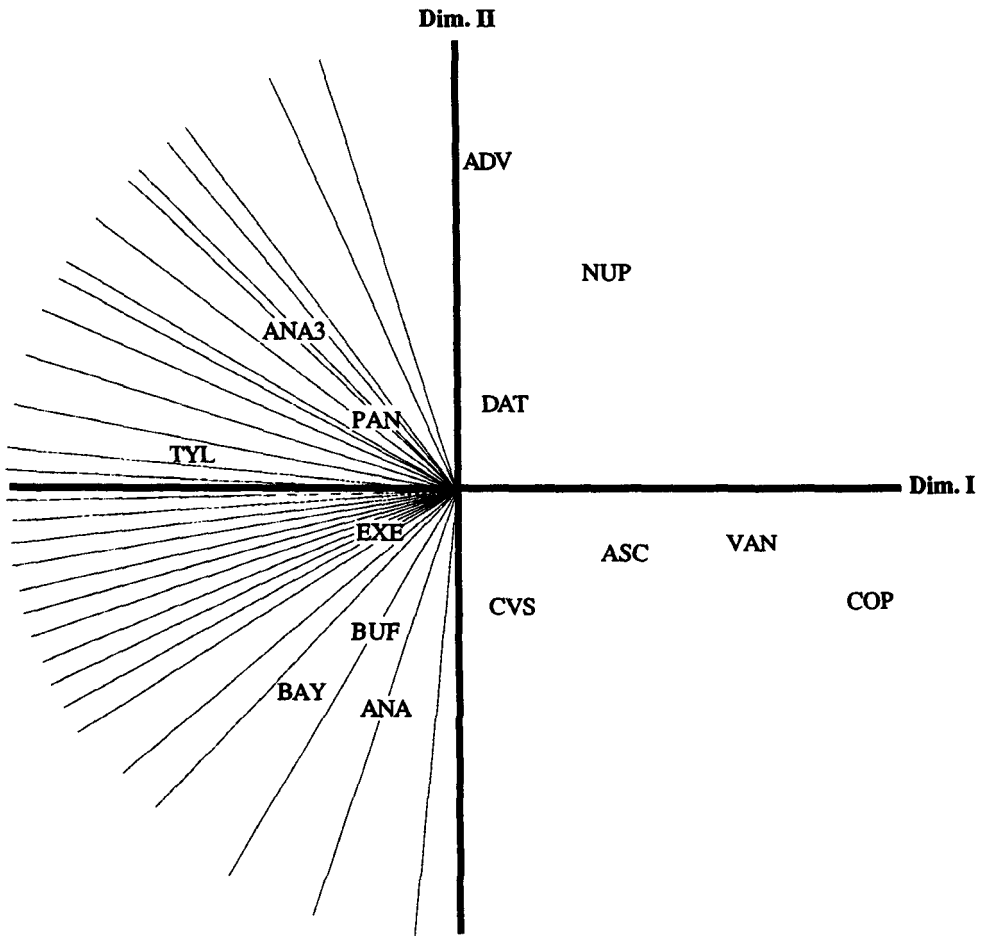


FIGURE 3
DeSarbo et al. (1986) vector model joint space for the analgesics data.

6. Conclusion

We have described the complete MICROSCALE procedure including data requirements, model structure and properties, the ACML algorithm for parameter estimation, a modest Monte Carlo analysis documenting algorithm performance, and a consumer psychology application contrasting MICROSCALE with the results of a recently developed vector model. While these initial results look promising, further research is required to more fully investigate the MICROSCALE procedure. As mentioned in section 4, more ambitious Monte Carlo experimental designs are required allowing for the estimation of interaction effects and investigating the frequency of locally optimal solution problems. The sensitivity of the procedure to empirical/synthetic data that do not satisfy strong stochastic transitivity should also be thoroughly tested. Extensions of MICROSCALE to accommodate incomplete paired comparisons designs should be investigated given the difficulty in collecting such $N(N - 1)/2$ responses per subject. Finally, further applications are desired using empirical data from a number of different disciplines to examine the appropriateness of this convex indifference curve specification, contrast the results of MICROSCALE with the results of more traditional vector and unfolding models, investigate respective policy recommendations, and so forth.

Appendix: The MICROSCALE ACML Procedure

Algorithm:

MICROSCALE uses a conjugate gradient method with automatic restarts (see Powell, 1977). The estimation method is an alternating conditional maximum likelihood procedure (ACML) where the log of the likelihood function is maximized with respect to a specific set of parameters (i.e., the object/brand coordinates, the consumers' α parameters, the multipliers, a_i 's, and the thresholds s_i 's) holding the other sets constant. Estimation alternates across all the parameter sets until a convergence tolerance is satisfied.

Phase I: Input Options

The user must specify the number of consumers (I), the number of brands (J), the number of dimensions (T), the type of analysis (internal or external), the option for the multiplier a_i (varying by consumer, constant, or set equal to one), the option for the threshold (varying by consumer, constant, or set equal to zero), and, the file name containing the input data. This file should include the brand coordinates if an external analysis is to be performed or if a rational start is available. In addition, it should also contain **H** if the reparameterization options are requested.

Phase II: Starting Estimates

Options exist for the user to specify a random start, a rational start for **X** (a singular value decomposition, SVD, analysis on the aggregated preference data obtained by summing over the rows of the complete pairwise comparison data (upper and lower diagonal) as suggested in Carroll (1980), with subsequent translation of origin to satisfy the positivity constraint of the brand coordinates) or, a given start.

Phase III: Estimate X, α , a, and s

We need to impose nonnegative constraints on the brand coordinates ($X_{jt} \geq 0$), the subject parameters ($\alpha_{it} \geq 0$ and $\sum_{t=1}^T \alpha_{it} = 1$), the multipliers ($a_i \geq 0$), and the threshold parameters ($s_i \geq 0$). To accomplish this, we estimate the corresponding squared entities:

$$X_{jt}^2, \frac{\theta_{it}^2}{\sum_{t=1}^T \theta_{it}^2}, a_i^2, \text{ and } s_i^2$$

as suggested by Gill, Murry, and Wright (1981, p. 268–269), and as described in section three of this paper.

Partial Derivatives

Let

$$U_{ij} = \prod_{t=1}^T (X_{jt}^2)^{(\theta_{it}^2 / \sum \theta_{it}^2)},$$

$$UD_{ijk} = a_i^2 (U_{ij} - U_{ik}),$$

$$\Phi(\cdot) = \Phi(UD_{ijk} - s_i^2),$$

$$\Phi(*) = \Phi(-UD_{ijk} - s_i^2).$$

Then

$$\begin{aligned} \frac{\partial \ln L}{\partial a_i} &= \sum_{j < k} \sum \delta_{ijk}^+ \frac{1}{\Phi(\cdot)} \phi(\cdot) [2a_i(U_{ij} - U_{ik})] \\ &\quad - \delta_{ijk}^- \frac{1}{\Phi(*)} \phi(*) [-2a_i(U_{ij} - U_{ik})] \\ &\quad + \frac{(1 - |\delta_{ijk}|)}{(1 - \Phi(\cdot) - \Phi(*))} (\phi(*) - \phi(\cdot)) [2a_i(U_{ij} - U_{ik})]; \end{aligned} \quad (\text{A1})$$

$$\begin{aligned} \frac{\partial \ln L}{\partial s_i} &= \sum_{j < k} \sum \delta_{ijk}^+ \frac{1}{\Phi(\cdot)} \phi(\cdot) [-2s_i] \\ &\quad - \delta_{ijk}^- \frac{1}{\Phi(*)} \phi(*) [-2s_i] \\ &\quad + \frac{(1 - |\delta_{ijk}|)}{(1 - \Phi(\cdot) - \Phi(*))} (\phi(\cdot) + \phi(*)) [2s_i]; \end{aligned} \quad (\text{A2})$$

$$\begin{aligned} \frac{\partial \ln L}{\partial X_{j'}} &= \sum_i \sum_{(j,k)} \delta_{ijk}^+ \frac{\phi(\cdot)}{\Phi(\cdot)} \left[\frac{a_i^2 U_{ij}}{X_{j'}} \left(\frac{2\theta_{i'}^2}{\sum_t \theta_{it}^2} \right) \right] \\ &\quad - \delta_{ijk}^- \frac{\phi(*)}{\Phi(*)} \left[\frac{-a_i^2 U_{ij}}{X_{j'}} \left(\frac{2\theta_{i'}^2}{\sum_t \theta_{it}^2} \right) \right] \\ &\quad + \frac{(1 - |\delta_{ijk}|)}{(1 - \Phi(\cdot) - \Phi(*))} (\phi(*) - \phi(\cdot)) \left[\frac{a_i^2 U_{ij}}{X_{j'}} \left(\frac{2\theta_{i'}^2}{\sum_t \theta_{it}^2} \right) \right]; \end{aligned} \quad (\text{A3})$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \theta_{i'}} &= \sum_{j < k} \sum \delta_{ijk}^+ \frac{\phi(\cdot)}{\Phi(\cdot)} \left[a_i^2 \left(\frac{\partial U_{ij}}{\partial \theta_{i'}} - \frac{\partial U_{ik}}{\partial \theta_{i'}} \right) \right] \\ &\quad - \delta_{ijk}^- \frac{\phi(*)}{\Phi(*)} \left[-a_i^2 \left(\frac{\partial U_{ij}}{\partial \theta_{i'}} - \frac{\partial U_{ik}}{\partial \theta_{i'}} \right) \right] \\ &\quad + \frac{(1 - |\delta_{ijk}|)}{(1 - \Phi(\cdot) - \Phi(*))} (\phi(*) + \phi(\cdot)) \left[a_i^2 \left(\frac{\partial U_{ij}}{\partial \theta_{i'}} - \frac{\partial U_{ik}}{\partial \theta_{i'}} \right) \right], \end{aligned} \quad (\text{A4})$$

where

$$\frac{\partial U_{ij}}{\partial \theta_{it'}} = [-\log(U_{ij}) + \log(X_{jt'}^2)] \left(\frac{2\theta_{it'}}{\sum_t \theta_{it'}^2} \right) U_{ij}.$$

Note, for the case where $a_i = a$, and $s_i = s$, and $\theta_{it} = \theta_t$, for all i , we need to sum over all i 's

$$\frac{\partial \ln L}{\partial a} = \sum_i \frac{\partial \ln L}{\partial a_i},$$

$$\frac{\partial \ln L}{\partial s} = \sum_i \frac{\partial \ln L}{\partial s_i},$$

and

$$\frac{\partial \ln L}{\partial \theta_t} = \sum_i \frac{\partial \ln L}{\partial \theta_{it}}.$$

The Reparameterized Case

$$\begin{aligned} \frac{\partial \ln L}{\partial \gamma_{mt'}} &= \sum_i \sum_{j < k} \delta_{ijk}^+ \frac{\phi(\cdot)}{\Phi(\cdot)} \left[a_i^2 \left(\frac{\partial U_{ij}}{\partial \gamma_{mt'}} - \frac{\partial U_{ik}}{\partial \gamma_{mt'}} \right) \right] \\ &\quad - \delta_{ijk}^- \frac{\phi(*)}{\Phi(*)} \left[-a_i^2 \left(\frac{\partial U_{ij}}{\partial \gamma_{mt'}} - \frac{\partial U_{ik}}{\partial \gamma_{mt'}} \right) \right] \\ &\quad + \frac{(1 - |\delta_{ijk}|)}{(1 - \Phi(\cdot) - \phi(*))} (\Phi(*) + \phi(\cdot)) \left[a_i^2 \left(\frac{\partial U_{ij}}{\partial \gamma_{mt'}} - \frac{\partial U_{ik}}{\partial \gamma_{mt'}} \right) \right], \quad (A5) \end{aligned}$$

where

$$\frac{\partial U_{ij}}{\partial \gamma_{mt'}} = U_{ij} \left(\frac{H_{jm}}{X_{jt'}} \right) \left(\frac{2\theta_{it'}^2}{\sum_t \theta_{it'}^2} \right),$$

H_{jm} = the value of feature m for brand j ,

$$X_{jt'} = \sum_{m=1}^M H_{jm} \gamma_{mt'}.$$

At the start, $\alpha_{it}^2 = 1$, $a_i^2 = 1$, $s_i^2 = 0$, and X_{jt} 's are randomly generated from a uniform distribution if no rational (or given) starting solution is provided for an internal analysis. In the case of external analysis, the methodology solves for whatever parameter set(s) are not fixed/given. Holding α , a , and s constant, the algorithm uses the conjugate gradient method with automatic restarts (Powell, 1977) to estimate X (assuming an internal analysis). Then it estimates α , holding a , s , and X constant. Finally, a and s are estimated in this same manner. The algorithm repeats these major iterations until a convergence tolerance is satisfied.

Phase V: Output

The following goodness-of-fit measures are computed: (a) the likelihood function—an AIC (Akaike, 1974) statistic is also calculated; however, concerns have been expressed by Bozdogan (1987) on its tendency to over-estimate the true dimensionality, (b) a matching coefficient (MC), (c) Spearman's correlation (SCOR), (d) Phi coefficient (PHI), and (e) point biserial correlation (PBC). The last measure can be computed when there are no ties allowed in $\delta_{ijk}(s_i = 0)$. One also obtains the history of iterations, the brand coordinates (\mathbf{X}), the subject coordinates ($\boldsymbol{\alpha}$), the \hat{U}_{ij} scores, and the indifference curve plots by subject including the TFF and PEP.

References

- Addelman, S. (1962). Orthogonal main effect plans for asymmetrical factorial experiments. *Technometrics*, 4, 21–46.
- Akaike, H. (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control*, AC-19, 716–723.
- Bechtel, G. G., Tucker, L. R. & Chang, W. (1971). A scalar product model for the multidimensional scaling of choice. *Psychometrika*, 36, 369–388.
- Bentler, P. M., & Weeks, D. G. (1978). Restricted multidimensional scaling models. *Journal of Mathematical Psychology*, 17, 138–151.
- Bloxom, B. (1978). Constrained multidimensional scaling in N spaces. *Psychometrika*, 43, 397–408.
- Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52, 345–370.
- Carroll, J. D. (1972). Individual differences and multidimensional scaling. In R. N. Shepard, A. K. Romney & S. Nerlove (Eds.), *Multidimensional scaling: Theory and application in the behavioral sciences. Vol. I: Theory* (pp. 105–155). New York: Academic Press.
- Carroll, J. D. (1980). Models and methods for multidimensional analysis of preferential choice (or other dominance data). In E. D. Lantermann & H. Feger (Eds.), *Similarity and choice* (pp. 234–289). Bern: Hans Huber.
- Carroll, J. D., Pruzansky, S., & Kruskal, J. B. (1980). CANDELINC: A general approach to multidimensional analysis of many-way arrays with linear constraints on parameters. *Psychometrika*, 45, 3–24.
- Coombs, C. H. (1964). *A theory of data*. New York: Wiley.
- Cooper, L. G., (1973). A multivariate investigation of preferences. *Multivariate Behavioral Research*, 8, 253–272.
- Cooper, L. G., & Nakanishi, M. (1983). Two logit models for external analysis of preference. *Psychometrika*, 48, 607–620.
- de Leeuw, J., & Heiser, W. (1980). Multidimensional scaling with restrictions on the configuration. In P. R. Krishnaiah (Ed.), *Multivariate analysis V* (pp. 285–316). Amsterdam: North-Holland.
- DeSarbo, W. S. (1982). GENNCLUS: New models for general nonhierarchical clustering analysis. *Psychometrika*, 47, 449–475.
- DeSarbo, W. S., & Carroll, J. D. (1985). Three-way metric unfolding via alternating weighted least squares. *Psychometrika*, 50, 275–300.
- DeSarbo, W. S., Carroll, J. D., Lehmann, D. R., & O'Shaughnessy, J. (1982). Three-way multivariate conjoint analysis. *Marketing Science*, 1, 323–350.
- DeSarbo, W. S., De Soete, G., Carroll, J. D., & Ramaswamy, V. (1988). A new stochastic ultrametric tree unfolding methodology for assessing competitive market structure and deriving market segments. *Applied Stochastic Models and Data Analysis*, 4, 185–204.
- DeSarbo, W. S., De Soete, G., & Eliashberg, J. (1986). A new stochastic multidimensional unfolding model for the investigation of paired comparisons in consumer preference/choice data. *Journal of Economic Psychology*, 8, 357–384.
- DeSarbo, W. S., De Soete, G., & Jedidi, K. (1987). Probabilistic multidimensional scaling models for analyzing consumer choice behavior. *Communication and Cognition*, 20, 93–116.
- DeSarbo, W. S., Oliver, R. L., & De Soete, G. (1986). A probabilistic multidimensional scaling vector model. *Applied Psychological Measurement*, 10, 79–98.
- DeSarbo, W. S., & Rao, V. R. (1984). GENFOLD2: A set of models and algorithms for the GENERAL unFOLDing analysis of preference/dominance data. *Journal of Classification*, 1, 147–186.
- DeSarbo, W. S., & Rao, V. R. (1986). A constrained unfolding model for product positioning. *Market Science*, 5, 1–19.

- De Soete, G., & Carroll, J. D. (1983). A maximum likelihood method for fitting the wandering vector model. *Psychometrika*, *48*, 553–566.
- De Soete, G., & Carroll, J. D. (1986). Probabilistic multidimensional choice models for representing paired comparisons data. In E. Diday et al. (Eds.), *Data analysis and informatics IV* (pp. 485–497). Amsterdam: North-Holland.
- De Soete, G., Carroll, J. D., & DeSarbo, W. S. (1986). The wandering ideal point model: A probabilistic multidimensional unfolding model for paired comparisons data. *Journal of Mathematical Psychology*, *30*, 28–41.
- Gill, P. E., Murray, W., & Wright, M. H. (1981). *Practical optimization*, New York: Academic Press.
- Grandmount, J. M. (1978). Intermediate preferences and the majority rule. *Econometrica*, *46*, 317–330.
- Haines, G. (1975). Commentary on Ratchford, 'the new economic theory of consumer behavior'. *Journal of Consumer Research*, *2*, 77–78.
- Hauser, J. R. (1988). Competitive price and positioning strategies. *Marketing Science*, *7*, 76–91.
- Hauser, J. R., & Shugan, S. (1983). Defensive marketing strategies. *Marketing Science*, *2*, 319–360.
- Hauser, J. R., & Simmie, P. (1981). Profit maximizing perceptual positions: An integrated theory for the selection of product features and price. *Management Science*, *27*, 33–56.
- Henderson, J. M., & Quandt, R. E. (1984). *Microeconomic theory* (3rd ed.). New York: McGraw-Hill.
- Hendler, R. (1975). Lancaster's new approach to consumer demand and its limitations. *American Economic Review*, *64*, 194–199.
- Krishnan, K. S. (1977). Incorporating thresholds of indifference in probabilistic choice models. *Management Science*, *23*, 1224–1233.
- Ladd, G. W., & Zober, M. (1977). A model of consumer reaction to product characteristics. *Journal of Consumer Research*, *4*, 89–101.
- Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy*, *84*, 132–157.
- Lancaster, K. (1971). *Consumer demand: A new approach*. New York: Columbia University Press.
- Lioukas, S. K. (1984). Thresholds and transitivity in stochastic consumer choice: A multinomial logit analysis. *Management Science*, *30*, 110–122.
- Lucas, R. E. (1975). Hedonic price functions. *Economic Inquiry*, *13*, 157–178.
- MacCrimmon, K. R., & Toda, M. (1969). The experimental determination of indifference curves. *Review of Economic Studies*, *36*, 432–451.
- McCullagh, P., & Nelder, J. A. (1983). *Generalized linear models*. New York: Chapman & Hall.
- McFadden, D. (1976). Quantal choice analysis: A survey. *Annals of Economic and Social Measurement*, *5*, 363–390.
- Nakanishi, M., & Cooper, L. G. (1974). Parameter estimation for a multiplicative competitive interaction model—A least squares approach. *Journal of Marketing Research*, *11*, 303–311.
- Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, Series A*, *135*, 370–384.
- Pekelman, D., & Sen, S. (1975, August). *A Lancasterian approach to multiattribute marketing models*. Paper presented at the American Marketing Association Meeting, Rochester, NY.
- Powell, M. J. D. (1977). Restart procedures for the conjugate gradient method. *Mathematical Programming*, *12*, 241–254.
- Quandt, R. E. (1956). A probabilistic theory of consumer behavior. *Quarterly Journal of Economics*, *70*, 507–536.
- Ratchford, B. T. (1975). The new economic theory of consumer behavior: An interpretive essay. *Journal of Consumer Research*, *2*, 65–75.
- Ratchford, B. T. (1979). Operationalizing economic models of demand for product characteristics. *Journal of Consumer Research*, *6*, 76–87.
- Rousseas, S. W., & Hart, A. G. (1951). Experimental verification of a composite indifference map. *Journal of Political Economy*, *59*, 288–318.
- Russell, R. R., & Wilkinson, M. (1979). *Microeconomics: A synthesis of modern and neoclassical theory*. New York: Wiley & Sons.
- Ryans, A. B. (1974). Estimating consumer preferences for a new durable brand on an established product class. *Journal of Marketing Research*, *11*, 434–443.
- Schönemann, P. H., & Wang, M. M. (1972). An individual difference model for the multidimensional analysis of preference data. *Psychometrika*, *37*, 275–305.
- Slater, P. (1960). The analysis of personal preference. *British Journal of Statistical Psychology*, *13*, 119–135.
- Snedecor, G. W., & Cochran, W. G. (1981). *Statistical methods* (7th ed.). Ames: Iowa State University Press.
- Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, *34*, 273–286.
- Thurstone, L. L. (1931). The indifference function. *Journal of Social Psychology*, *2*, 139–167.

- Tucker, L. R. (1960). Intra-individual and inter-individual multidimensionality. In H. Gulliksen & S. Messick (Eds.), *Psychological scaling: Theory and applications* (pp. 155–167). New York: Wiley.
- Varian, H. R. (1984). *Microeconomic analysis*. New York: W. W. Norton.
- Wang, M. M., Schönemann, P. H., & Rusk, J. G. (1975). A conjugate gradient algorithm for the multidimensional analysis of preference data. *Multivariate Behavioral Research*, 10, 45–99.
- Zinnes, J. L., & Griggs, R. A. (1974). Probabilistic multidimensional unfolding analysis. *Psychometrika*, 39, 327–350.

Manuscript received 6/6/89

Final version received 4/28/90