Title
CHOICE-THEORETIC MDS BY PAIRWISE EXPLOSION OF RANK DATA

Authors
Nakanishi, Masao
Cooper, Lee

Publication Date
1989-08-01

CHOICE-THEORETIC MDS BY PAIRWISE EXPLOSION OF RANK DATA

Masao Nakanishi

and

Lee G. Cooper

August 1989

Marketing Studies Center

John E. Anderson Graduate School of Management, UCLA

Working Paper No. 188

Working Paper No. 188 August, 1989

Choice-Theoretic MDS By Pairwise Explosion Of Rank Data

Masao Nakanishi [1]

Lee G. Cooper [2]

Center for Marketing Studies

Anderson Graduate School of Management

University of California, Los Angeles

[1]School of Business Administration, Kwansei Gakuin University, Nishinomiya-shi JAPAN 662.
[2]Anderson Graduate School of Management, UCLA, Los Angeles, CA 90024-1481.

Choice-Theoretic MDS
By Pairwise Explosion Of Rank Data

Abstract

A new method is developed for estimating nonmetric multidimensional-scaling solutions from pairwise explosion of rank-order data. The class of MDS methods discussed are adaptations of random-utility choice models to the perceptions of dissimilarities and are therefore called *choice-theoretic* MDS models. We propose simple alternatives to maximum-likelihood estimation of the parameters of these models. The estimation techniques are applied to both simulated data and to a study of perceptions of brands in the Japanese beer market.

1 Introduction

In recent years we have seen the development of multidimensional-scaling techniques which incorporate probabilistic assumptions regarding the human perception of similarities/dissimilarities (MacKay and Zinnes 1981, 1986; Zinnes and MacKay 1983). Although their usefulness is largely dependent on the validity of the assumptions on the probabilistic distribution of perceptual "errors," those techniques in general have several advantages over the traditional techniques in that they avoid both the solution indeterminacy and degeneration problems which often plague MONANOVA-based techniques and that their results may be analyzed statistically. On the other hand the use of probabilistic MDS models is hampered by the difficulties associated with estimation. A maximum-likelihood method is often used for estimation, but it sometimes takes an inordinate amount of computing time to reach a solution. We propose a simple and quick estimation technique which can be applied to a series of MDS models, which we call *choice-theoretic* MDS models.

1.1 Choice-Theoretic MDS Models

In the last two decades a large number of consumer brand/product-choice models have been developed and tested. Among those surviving the test of time, the so-called random-utility models have earned a secure place in the toolbox of consumer researchers. Choice-theoretic MDS models are straight adaptations of random-utility models to the perception of dissimilarities.

The basic assumptions of choice-theoretic MDS models are deceptively simple. First, it is assumed that the subject's perception (i.e., internal response) of dissimilarity between a pair of objects, s_{ij}, is a function of the underlying distance between the objects, d_{ij} , in the individual's perceptual space and an error term, ε_{ij}. That is,

$$s_{ij} = f(d_{ij}) + \varepsilon_{ij} \tag{1}$$

where $f(\cdot)$ is a monotone transformation of distance. Katahira [1986, 1989] for example suggested logarithmic, identity, and exponential transformations for f. A simple linear transformation of d_{ij} is equally permissible (Cooper 1972). And, of course, much of the modern interest in MDS stems from the development of analyses based on general monotone transformations (Kruskal 1964a and b). Second, the error term, ε_{ij}, may take many alternative distribution forms. For example, one may assume that the ε_{ij}'s have a common normal distribution (which we will call a PROBIT model later) or a double-exponential distribution (as in LOGMAP below). There may be other distributions which are just as reasonable as normal or double-exponential, although what is or is not reasonable depends on the interpretation of the error term, as is discussed below. By combining various specifications for f and ε_{ij}, a whole family of models may be developed from the simple structure of perception (1). One of the early choice-theoretic MDS models is LOGMAP (Katahira 1986, 1989). In LOGMAP the internal perception of dissimilarity between two objects, s_{ij}, is given by

$$s_{ij} = \log(d_{ij}) + \varepsilon_{ij} \tag{2}$$

This model assumes that human perception of dissimilarity follows the Weber-Fechner law: as the *true* distance, d_{ij}, becomes longer, any incremental distance *felt* by the respondent, Δs_{ij}, becomes progressively less.[3]

[3]More precisely, (2) assumes that Δs_{ij} is approximately proportional to the inverse of d_{ij}.

LOGMAP's rationale for introducing the error term in (2) is that, even at the within-individual level, the perceived dissimilarities may not always be an exact function of inter-object distances. From time to time s_{ij} may change subtly and deviate from $\log(d_{ij})$. One may assume that the deviations, the ε_{ij}'s, at any given point of time are a sample from a multivariate probabilistic distribution and independent samples are taken over time, that is, the ε_{ij}'s at different time points are generated by a stochastic process. Given this stochastic-process interpretation of ε_{ij}, Chapman and Staelin [1982] showed that the likelihood of a given ranking, $\{j_1, j_2, \ldots, j_m\}$ (where j_r is the index of the object ranked r^{th} highest), is given by,

$$\begin{aligned} Prob(j_1, j_2, \ldots, j_m) &= Prob(j_1 \mid \{j_1, j_2, \ldots, j_m\}) \\ &\quad \cdot Prob(j_2 \mid \{j_2, j_3, \ldots, j_m\}) \\ &\quad \cdot Prob(j_3 \mid \{j_3, j_4, \ldots, j_m\}) \\ &\quad \ldots\ldots \\ &\quad \cdot Prob(j_{m-1} \mid \{j_{m-1}, j_m\}). \end{aligned} \tag{3}$$

The expression, $Prob(j_i \mid \{j_i, j_{i+1}, \ldots, j_m\})$, gives the conditional probability that j_i is selected from the set $\{j_i, j_{i+1}, \ldots, j_m\}$.

The above likelihood formulation is known as the *explosion rule* and forms the basis of the estimation method for LOGMAP. In order to compute the likelihood (3) one needs the probabilistic specification of the perceptual-error term, ε_{ij}. In particular, let each ε_{ij} have an independent and identical double-exponential (type-I extreme-value) distribution of the following type.

$$Prob(\varepsilon_{ij} \leq u) = \exp\{-\exp(-\beta u)\} \tag{4}$$

The parameter β is related to the individual's discriminatory ability. Since the variance of (4) is $\pi^2/6\beta^2$, a large value of β implies a high degree of dissimilarity discrimination. A very well-known result in random-utility theory (McFadden 1974) states that, under this double-exponential assumption, the probability that object

j ($j \neq i$) is judged to be the most dissimilar to object i among a set of m objects $\{d_{i1}, d_{i2}, \ldots, d_{im}\}$, is given by

$$Prob(d_{ij} = \max\{d_{i1}, d_{i2}, \ldots, d_{im}\}) = d_{ij}^{\beta} / \sum_{k=1}^{m} d_{ik}^{\beta} \tag{5}$$

If we combine the above expression with (3), we have

$$\begin{aligned} Prob(j_1, j_2, \ldots, j_m) &= (d_{ij_1}^{\beta} / \sum_{\{C_1\}}^{m} d_{ik}^{\beta}) \cdot (d_{ij_2}^{\beta} / \sum_{\{C_2\}}^{m} d_{ik}^{\beta}) \\ &\quad \cdot (d_{ij_3}^{\beta} / \sum_{\{C_3\}}^{m} d_{ik}^{\beta}) \cdot (d_{ij_4}^{\beta} / \sum_{\{C_4\}}^{m} d_{ik}^{\beta}) \\ &\quad \cdots \cdot (d_{ij_{m-1}}^{\beta} / \sum_{\{C_{m-1}\}}^{m} d_{ik}^{\beta}) \\ &= \prod_{s=1}^{m-1} (d_{ij_s}^{\beta} / \sum_{\{C_s\}}^{m} d_{ik}^{\beta}) \end{aligned} \tag{6}$$

where $\{C_s\}$ indicates the set of subscripts $\{j_s, j_{s+1}, \ldots, j_{m-1}, j_m\}$ over which summation is performed. As we show in the next section, Katahira [1986, 1989] exploits this likelihood function to estimate the coordinates of objects in the perceptual space. He showed that the maximum-likelihood approach was adaptable to a wide variety of data-collection techniques. Furukawa [1987] also developed an individual-differences model based on LOGMAP which is similar to INDSCAL (Carroll and Chang 1970).

2 Pivot Ordering and Maximum-Likelihood Estimation

Consider the following data-collection procedure. Suppose that each of n respondents compares m objects and judges their dissimilarities in the following manner. For each object i ($i = 1, 2, \ldots, m$) the respondent is asked to look at the remaining $(m-1)$ objects and rank them according to the dissimilarity from object i. This produces m rankings of $m-1$ objects. This procedure is called "pivot ordering"

(Katahira 1986, 1989), the pivot being the fixed object against which the dissimilarities of others are judged.[4] Note that in this procedure the distance between each object pair (i, j) is judged twice, introducing some degree of stability in the estimates.

LOGMAP recovers the object configurations by a maximum-likelihood method based on (5). The pivot-ordering procedure yields m rankings of $m-1$ objects per respondent, and, since each respondent may give different rankings, the logarithm of (5) and its derivatives will have to be computed for each individual separately. Then the overall log-likelihood (for the entire sample) and its derivatives are computed by summing individual likelihoods and derivatives over all individuals in the sample and over m rankings per individual. Because it is extremely time-consuming to compute the second-order derivatives (i.e., the Hessian matrix) for this overall log-likelihood function, maximization is usually achieved by some first-order method (e.g., the method of steepest descent). The problem with first-order maximization methods is that the step size for improving estimates at each iteration may not be optimal, and many unnecessary steps may be taken before reaching the local maxima. Osawa [1988] reports that in a case where the number of objects is six and the number of dimensions is two, the iterations took two minutes on the average on an IBM 3090-200 before a satisfactory degree of convergence was achieved. Table 1 gives the simulation results from Osawa.

(Insert Table 1 about here.)

The maximum-likelihood estimates are apparently unbiased and accurate, judging from the standard deviations, but they are costly in terms of computing resources. To compound the problem, the maximum-likelihood method requires that the number of dimensions of the perceptual space must be specified beforehand, making it necessary in some cases to compute solutions for several different numbers

[4]This data-collection procedure is also known in the literature as the method of conditional rank orders and the method of rotating anchor points.

of dimensions (and possibly attempt to interpret each solution) before a reasonable configuration is selected.

3 Pairwise-Explosion Method

We propose another estimation method which utilizes the *pairwise explosion* of rankings. At this point it may be emphasized that our view as to the nature of the error term in (1) is distinctly different from that of Katahira. If the data are taken from a group of respondents, it is more reasonable for one to expect some degree of variability in the values of s_{ij} among individuals. One may assume that the group mean of s_{ij} is $f(d_{ij})$ and the deviations from the group mean, the ε_{ij}'s, are distributed over the individuals within the group according to some multivariate distribution. Because we feel that between-individual variations are relatively greater than within-individual variations, we will adopt this *cross-sectional* rationale in this paper and consider the analysis of rank data obtained from a group of individuals. In this situation, one needs to derive the unconditional probability that object j is judged the most dissimilar to object i by a group member. We note that, in the case of LOGMAP, the unconditional probability expression happens to be the same as (5), and Osawa [1988] used this expression in LOGMAP to analyze group data cited above.

We capitalize the fact that a ranking of m objects may be expressed as $m(m-1)/2$ paired comparisons (cf., Cooper and Nakanishi 1983). Table 2(a) gives the result of pairwise explosion of a ranking obtained from a single respondent.

(Insert Table 2 about here.)

In this $m \times m$ matrix, a value of 1 is given to element (i, j) if object i is ranked higher than j; otherwise, a value of 0 is given.[5] If the rankings from n individuals in a sample were pairwise exploded and aggregated, we would obtain a matrix shown

[5] Such data are normally called *row-dominance data*.

in Table 2(b). The n_{ij} in Table 2(b) gives the number of individuals who gave a higher rank to object i than object j. Of course $n_{ij} + n_{ji} = n$.

When the pivot-ordering procedure is used, one will have m tables such as Table 2(b), one for each pivot element. Let us write $n_{(h)ij}$ for pivot element h. In the case of LOGMAP, it may be assumed that $n_{(h)ij}$ is binomially distributed with parameters n and $\pi_{(h)ij} = d^{\beta}_{hi}/(d^{\beta}_{hi} + d^{\beta}_{hj})$. Under this assumption the relative frequency with which i is judged farther from h than j is from h can be expressed as:

$$n_{(h)ij}/n_{(h)ji} \approx \pi_{(h)ij}/\pi_{(h)ji} = d^{\beta}_{hi}/d^{\beta}_{hj}.$$

The approximation improves as the sample size n increases. If we let the usual logit transformation applied to $n_{(h)ij}$ be $p_{(h)ij}$,

$$p_{(h)ij} = \log(n_{(h)ij}/n_{(h)ji}) \approx \beta(\log d_{hi} - \log d_{hj}).$$

By averaging $p_{(h)ij}$ with respect to subscripts j and h, we obtain

$$p_{(h)i\cdot} \approx \beta(\log d_{hi} - \log \tilde{d}_{h\cdot})$$

$$p_{(\cdot)i\cdot} \approx \beta(\log \tilde{d}_{\cdot i} - \log \tilde{d}_{\cdot\cdot})$$

where $\tilde{d}_{h\cdot}$ is the geometric mean of d_{hi} over $j = 1, 2, \ldots, m$ ($d_{hh} = 0$), and $\tilde{d}_{\cdot\cdot}$ is the overall geometric mean of d_{hi}. Noting that $p_{(\cdot)i\cdot} = p_{(\cdot)\cdot i}$ and letting $p_{(\cdot)h\cdot}$ be the value of $p_{(\cdot)i\cdot}$ for $i = h$, we have

$$\delta_{hi} = p_{(h)i\cdot} + p_{(\cdot)h\cdot} \approx \beta(\log d_{hi} - \log \tilde{d}_{\cdot\cdot}) \approx \beta \log(d_{hi}/\tilde{d}_{\cdot\cdot}).$$

Since the $(d_{hi}/\tilde{d}_{\cdot\cdot})$ term in the right-hand side is merely d_{hi} which is *normalized* by the overall (geometric) mean $\tilde{d}_{\cdot\cdot}$, no generality is lost by assuming that $\tilde{d}_{\cdot\cdot} = 1$. Hence we have an important result.

$$\delta_{hi} \approx \beta \log d_{hi} \quad (h = 1, 2, \ldots, m; \; i = h + 1, \ldots, m). \tag{7}$$

We will make the full use of the last expression.

Let X_{it} be the t^{th} coordinate of object i. The distances can be converted into scalar products using the formula developed by Tucker and reported in Torgerson [1958, p. 258]. If we assume that $\sum_{i=1}^{m} X_{it} = 0$ (that is, object coordinates are measured from the centroid), then we obtain the scalar product b_{hi} by double-centering the matrix of squared distances:

$$\begin{aligned} b_{hi} &= d_{hi}^2 - \sum_{h=1}^{m} d_{hi}^2/m - \sum_{i=1}^{m} d_{hi}^2/m + \sum_{h=1}^{m}\sum_{i=1}^{m} d_{hi}^2/m^2 \\ &= -2\sum_{t=1}^{T} X_{ht}X_{it}. \end{aligned}$$

We have an estimate of d_{hi}^2 in $\exp(2\delta_{hi}/\beta)$, and therefore the matrix of estimated values of $\{b_{hi}\}$ may be obtained by double-centering the matrix of $\{(-0.5)\exp(2\delta_{hi}/\beta)\}$. The object coordinates $\{X_{it}\}$ $(i = 1, 2, \ldots, m;\ \ t = 1, 2, \ldots, T)$ are recoverable by applying the Eckart-Young [1936] decomposition to the matrix of $\{b_{hi}\}$.[6] It is necessary to determine the number of dimensions in the perceptual space, but we may use a logic similar to the one we use to select the number of factors in factor analysis by judging the relative size of eigenvalues. Each of the retained eigenvectors is multiplied by the square root of the corresponding eigenvalue to obtain object coordinates.

So far nothing is said about the remaining parameter β, but the following iterative process gives an estimate very quickly.

1. Set an arbitrary β-value initially and find object coordinates $\{X_{it}\}$.
2. Compute inter-object distances $\{d_{hi}\}$ from the estimated coordinates.
3. Since $\delta_{hi} \approx \beta \log d_{hi}$, let δ_{hi} be the dependent variable and $\log d_{hi}$ be the independent variable, and find β by a least-squares method.

[6] Because the $\{\delta_{hi}\}$ matrix is not symmetric, we have several options. We can symmetrize by averaging the matrix with its own transpose, or we can investigate the asymmetric structure as in Cooper [1988]. We choose to symmetrize in the current study and defer investigation of the asymmetries for future efforts.

Repeat steps (1) through (3) until the value of β converges. Convergence is very fast — four-digit stability can be obtained in less than 10 iterations.

4 Generalization of the Pairwise-Explosion Method

Before we turn to the simulation results to investigate the properties of the proposed estimation method, we would like to point out that this method is not limited to LOGMAP (which we call LOGIT1 hereafter). For example, a variant of LOGIT1 is given by a slight modification of (1).

$$s_{ij} = d_{ij} + \varepsilon_{ij} \tag{8}$$

This model (to be called LOGIT2 hereafter) is the case where transformation f in (1) is the identity transformation. If we assume that ε_{ij}'s are identically and independently distributed as the double-exponential distribution (4), we have

$$Prob(d_{ij} = \max\{d_{i1}, d_{i2}, \ldots, d_{im}\}) = \exp(\beta d_{ij}) / \sum_{k=1}^{m} \exp(\beta d_{ik}).$$

For the pairwise-exploded matrices $\{n_{(h)ij}\}$ $(h = 1, 2, \ldots, m)$, a manipulation analogous to the LOGIT1 (LOGMAP) case gives the logit transformation of $n_{(h)ij}$ by

$$p_{(h)ij} \approx \beta(d_{hi} - d_{hj}),$$

with an end result that

$$\delta_{hi} = p_{(h)i\cdot} - p_{(\cdot)h\cdot} \approx \beta(d_{hi} - \bar{d}_{\cdot\cdot})$$

where $\bar{d}_{\cdot\cdot}$ in this case is the overall *arithmetic* mean of the d_{hi}'s.

The estimation of object coordinates $\{X_{it}\}$ $(i = 1, 2, \ldots, m; \quad t = 1, 2, \ldots, T)$ and parameter β for LOGIT2 proceeds as follows.

1. Choose initial values of β and $\bar{d}_{\cdot\cdot}$, and estimate squared distances as

$$\hat{d}^2_{hi} = (\delta_{hi}/\beta + \bar{d}_{\cdot\cdot})^2.$$

   Double-center the $\{\hat{d}^2_{hi}\}$ matrix to obtain the $\{b_{hi}\}$ matrix.

2. Find $\{X_{it}\}$ by the Eckart-Young decomposition of the $\{b_{hi}\}$ matrix.

3. Compute $\{d_{hi}\}$ matrix (normalized such that the overall geometric mean is 1) and then compute $\bar{d}_{..}$.

4. Based on the relationship, $\delta_{hi} \approx \beta(d_{hi} - \bar{d}_{..})$, re-estimate β by a least-squares method.

Repeat steps (1) through (4) until the values of both β and $\bar{d}_{..}$ converge. In our simulation the convergence was fast and it took less than 10 iterations before four-digit stability was achieved.

Next we consider the case in which the ε_{ij}'s in (8) are jointly normally distributed. For this model (which we call PROBIT2), the f in (1) is the identity transformation. We assume that each ε_{ij} is identically and independently distributed as a normal distribution with mean 0 and variance σ^2. Under this assumption, $n_{(h)ij}$ is approximately normally distributed with mean $(d_{hi} - d_{hj})$ and variance $2\sigma^2$ for a reasonably large sample size, n. Hence, if we take the inverse normal (or *probit*) transformation of $n_{(h)ij}/n$,[7] then we have an estimate of $(d_{hi} - d_{hj})/\sqrt{2}\sigma$. If we let $probit(\cdot)$ indicate the probit transformation and

$$p_{(h)ij} = probit(n_{(h)ij}/n) \approx (d_{hi} - d_{hj})/\sqrt{2}\sigma,$$

by the analogy to the LOGIT2 case we have

$$\delta_{hi} = p_{(h)i\cdot} - p_{(\cdot)h\cdot} \approx (d_{hi} - \bar{d}_{..})/\sqrt{2}\sigma$$

where $\bar{d}_{..}$ is the overall *arithmetic* mean of the d_{hi}'s.

The last expression suggests the following steps for computing the coordinates of objects $\{X_{it}\}$ and parameter σ for PROBIT2.

1. Choose initial values of σ and $\bar{d}_{..}$, and estimate squared distances as

$$\hat{d}^2_{hi} = (\sqrt{2}\sigma\delta_{hi} + \bar{d}_{..})^2.$$

[7] The PROBIT function in SAS is convenient to use in this case.

Double-center the $\{\hat{d}_{hi}^2\}$ matrix to obtain the $\{b_{hi}\}$ matrix.

2. Find object coordinates $\{X_{it}\}$ by the Eckart-Young decomposition of the $\{b_{hi}\}$ matrix.

3. Compute $\{d_{hi}\}$ matrix (normalized such that the overall geometric mean is 1) and then compute $\bar{d}_{..}$.

4. Based on the relationship, $\delta_{hi} \approx (d_{hi} - \bar{d}_{..})/\sqrt{2}\sigma$, re-estimate σ by a least-squares method.

Repeat steps (1) through (4) until the values of both σ and $\bar{d}_{..}$ converge. The convergence was very fast as before — four-digit stability requiring less than 10 iterations.

Another variant of the above PROBIT2 model may be derived if we choose the logarithmic instead of the identity transformation in (1) and assume that the error term is normally distributed. The estimation procedure for this model (which we call PROBIT1) follows that of LOGIT1 (LOGMAP) very closely, except that the probit transformation is used in lieu of the logit transformation. It may be pointed out that those PROBIT models are not easily estimable by the maximum-likelihood method, because conditional probabilities included in the likelihood function (3) (i.e., the *explosion rule*) must be evaluated by numerical integration. In this age of super-computers the maximum-likelihood estimation is a possibility, but our method is far more efficient in the use of computing resources, so much so that it is implementable on a personal computer.

We have shown in this section that the pairwise-explosion method is applicable to PROBIT, as well as LOGIT, models, but it has much wider applications. Indeed this method is applicable to a wide variety of other models, if

1. the probability that an object pair (h, i) is judged more dissimilar than another pair (h, j) $(j \neq i)$, $Prob(s_{hi} > s_{hj})$, may be expressed as the function of

underlying distances d_{hi} and d_{hj}, and

2. more importantly, a suitable inverse transformation exists to the function.

The double-exponential (i.e., LOGIT models) and normal (i.e., PROBIT models) distributions for the perceptual error term ε_{ij} satisfy those requirements, but there are other distributional forms which satisfy those requirements. For example, if we assumed a uniform distribution for ε_{ij}, it would be a relatively simple matter to find an explicit function for $Prob(s_{hi} > s_{hj})$ in terms of d_{hi} and d_{hj}, although one might be hard-pressed to find a psychological argument for such an assumption. In the following parts of this paper we will deal only with LOGIT and PROBIT models, simply because we consider them more important models of dissimilarity perception.

5 Simulation Studies

In order to assess the usefulness of the proposed estimation method (i.e., the pairwise-explosion method), we performed some simulation studies. Our purpose was two-fold: to investigate the statistical properties of the method and to make a comparative study of alternative models of dissimilarity perception. We compared the following four models:

$$\begin{array}{lrcl} \text{LOGIT1:} & s_{ij} & = & \log(d_{ij}) + \varepsilon_{ij} \\ \text{LOGIT2:} & s_{ij} & = & d_{ij} + \varepsilon_{ij} \\ \text{PROBIT1:} & s_{ij} & = & \log(d_{ij}) + \varepsilon_{ij} \\ \text{PROBIT2:} & s_{ij} & = & d_{ij} + \varepsilon_{ij} \end{array}$$

LOGIT models assume a double-exponential distribution (with parameter β), and PROBIT models a normal distribution (with parameter σ), for ε_{ij}.

The object configuration used in the simulation is given in Figure 1.

(Insert Figure 1 about here.)

The geometric mean of the inter-object distances is adjusted to 1. Also the centroid is set at the origin (that is, the mean for the coordinates on each dimension is equal to 0). For each simulated individual we generated a series of random s_{ij}'s for each of m pivot elements using the above specifications, and converted the s_{ij}'s into rankings. Individual rankings were then exploded into paired comparisons, and individual paired comparisons were summed over all individuals to obtain m matrices in the form of Table 2(b). Object coordinates $\{X_{it}\}$ and other parameters (e.g., β or σ) were estimated from those matrices. The same process was repeated 50 times to obtain the means and standard deviations of the estimates. We were concerned here with how well the estimated configuration matched the original and also with how well parameters β and σ were estimated. The σ-value was chosen so that the variance of the normal distribution was the same as that of the double-exponential distribution. Large values of the β's or small values of the σ's would imply that individuals are capable of discriminating small dissimilarities among the objects.

We may add that the estimated configuration from the proposed approach is rotationally indeterminate. This is not a problem for most practical applications because usually one is only interested in recovering the relative locations of objects in the perceptual space, but in simulation studies this property adds another source of discrepancies from the original configuration. We dealt with this problem by finding the orthogonal Procrustes rotation (Cliff 1966, Schönemann 1966) which minimized the sum of squared discrepancies between the estimated and original configurations.

(Insert Tables 3 – 6 about here.)

Tables 3 – 6 summarize the simulation results. The second column of each table gives the average of estimated coordinates for each object. The third column gives the standard deviations of estimated coordinates over 50 trials. The recovery of the original configuration seems reasonably good. In general there is no indication of systematic biases in the estimates for object coordinates. The LOGIT2 model gives a better set of coordinates than the LOGIT1 model, but this was expected. By applying the logarithmic transformation to d_{ij}, the distances between objects were reduced, and thus the variance of the random-error term, ε_{ij}, was effectively increased. For the same reason the PROBIT2 model gives better estimates than the PROBIT1 model. Turning to the estimates of the variance parameters for ε_{ij}, the estimates of β for LOGIT models are slightly overestimated, while those of σ for PROBIT models are estimated without apparent biases.

Though the simulation results show that the pairwise-explosion method estimates object coordinates of the original configuration reasonably well, we have no external criteria to judge its efficiency. Fortunately Table 1 gives the maximum-likelihood estimates for the LOGIT1 model from the Osawa study. The configuration used by Osawa in his simulation study is not the same as the one used in this study, but the number of objects (6), the simulated number of subjects (50) and the number of trials (50) are the same as our study, making it possible to check the efficiency of our method against the maximum-likelihood method. The standard deviations for object coordinates in Table 1 are roughly comparable with those in Table 3, but the configuration used by Osawa were not normalized such that the geometric mean of the inter-object distances was unity. When one takes this fact into account, the object coordinates estimated by the pairwise-explosion method have variances roughly 1.5 times greater than those of the maximum-likelihood estimates.[8] The efficiency for the estimates obtained by the pairwise-explosion method will improve with the increase in the sample size (n), however. It is probably not necessary to double the sample size for the pairwise-explosion method for one to obtain esti-

[8]The average of standard deviations is 0.0321 for Table 1 and 0.0279 for Table 3. Since the geometric mean of inter-object distances for the Osawa's configuration is 1.431, one should compare 0.0224 (=0.0321/1.431) against 0.0279. Thus the maximum-likelihood estimates of object coordinates have standard deviations approximately 20% less (0.0224/0.0279 = 0.803) than those for the pairwise-explosion method.

mates of object coordinates roughly comparable in precision with those obtained by the maximum-likelihood estimates. We believe that the precision achieved by the pairwise-explosion method is acceptable in many practical applications.

We also tried to estimate object coordinates and parameters for wrong models in the hope of discovering the robustness of the proposed method. We estimated the X_{it}'s and β for LOGIT1 from the data for LOGIT2, PROBIT1, and PROBIT2, those for LOGIT2 from the data for LOGIT1, PROBIT1, and PROBIT2, etc. In this manner we hoped to find how well we were able to recover the original configuration by the pairwise-explosion method, even if we happened to use a wrong model. We were particularly interested in the comparisons between LOGIT and PROBIT models for we suspected that a double-exponential distribution and a normal distribution were to some extent interchangeable as a perceptual-error distribution. To facilitate comparisons, we used one set of random numbers to generate data for both LOGIT models and another set for both PROBIT models. Put differently, the simulated data for LOGIT models were independent from those for PROBIT models. Under this procedure the differences between the LOGIT1 and LOGIT2 models were mainly due to the difference in the specification of $f(d_{ij})$ in (1), that is, whether one used the logarithmic or identity transformation. The same may be said of the differences between PROBIT1 and PROBIT2 models.

(Insert Table 7 about here.)

Table 7 summarizes the analyses for various combinations of models and data. Only the components of mean squared errors (biases squared and variances) for estimated object coordinates and the estimates of β and σ are reported. The following are major observations from the table.

1. The main diagonal entries of the table give the cases where the model and the data match. (The MSE values for those entries were already reported in Table 3 – 6.) Those cases are characterized by small biases.

2. When the models use the same transformation of distances (i.e., the logarithmic transformation for LOGIT1 and PROBIT1 and the identity transformation for LOGIT2 and PROBIT2), they tend to yield similar results despite of the difference in the assumption about the distribution of ε_{ij}.

3. PROBIT models have consistently smaller variances, and consequently smaller mean squared errors, than LOGIT models. The PROBIT2 model in particular gives the smallest sum of variances for every data.

The above observations confirmed our suspicion that the normal and the double-exponential distributions may be used interchangeably in the specification of the error distribution in (1). Our simulation study used one set of random numbers for LOGIT models and another set for PROBIT models, so the similarities of the performance between LOGIT and PROBIT models could not be attributed to this source. Why then were models based on two different distributions indistinguishable? A reflection reminded us that we simulated the pivot-ordering procedure in which a sample of five random numbers from the perceptual-error distribution was drawn for each pivot object and converted to one ranking of dissimilarity. This size of the perceptual-error sample, and not the sample size for respondents, was the critical factor. It seems obvious to us now that one cannot expect to distinguish between two similarly shaped distributions with a sample size as small as this (five). This situation will not be improved much if the number of objects to be compared per pivot object is increased to 10 or even to 25, or if the number of respondents is increased. Since one cannot indefinitely increase the number of objects to be compared per pivot element, we might as well accept our inability to distinguish between LOGIT and PROBIT models with the pivot-ordering procedure. It remains to be seen if other data-collection procedures do better in distinguishing between the two types of models, but we doubt that even the straight ranking or paired-comparison procedure (see below) is capable of doing so.

That the PROBIT models had consistently smaller mean squared errors than the LOGIT models is due to the properties of the probit and logit transformations. It is known that the logit transformation results in a larger variance than the probit (=inverse normal) transformation, if n_{ij} or $(n - n_{ij})$ is near zero. Moreover, note that the biases are mainly caused by the misspecification in the transformation for distances ($f(d_{ij})$ in (1)) and not by the misspecification in the error distribution. Our simulation results suggest that we are perhaps better off with PROBIT models with the pairwise-explosion method when we are uncertain of the correct form of underlying perceptual-error distribution.

6 Illustration

The simulation studies show a good ability to recover known configurations from data simulated with error, but we also need to demonstrate that the methods can be meaningfully applied to real data. The LOGIT1 model will be illustrated using an example from a study of the Japanese beer market.

6.1 Background

Asahi Beer which was a weak third among four major breweries in Japan in 1987, came up with a new product called *dry beer*, which used a new type of yeast and, by letting fermentation go on longer than before, had a drier (that is, less sweet) taste. This new product sold so well that in 1988 Asahi achieved an 8 percentage-point market-share gain (12% to 20%) in a single year. Other breweries began producing dry beer, but Asahi, being the first to enter the sub-market, has an overwhelming advantage. Asahi Super Dry (which is Asahi's brand for dry beer) currently dominates the dry-beer portion (approximately one third) of the beer market. Against Asahi's competitive advantage with dry beer, Kirin (the top beer producer with a current share of about 50 percent) is emphasizing the traditional qualities of its lager beer, and Suntory (No. 4 in this market with an approximate share of 10 percent) is pushing its new product, "malts 100% beer." The beer industry is in a state of complete turmoil and has been the object of the frequent comments by mass media.

6.2 Study Design

A team of students at Kwansei Gakuin University decided to study the beer industry and tried the pairwise explosion technique to get perceptual maps. They selected the following five brands of beer for the study: Asahi Super Dry (AD), Kirin Dry (KD), Suntory Dry (SD), Kirin Lager (KL), Suntory Malts 100% (SM). The students selected a convenience sample of 80 male beer drinkers of various age groups, and split them randomly into two groups. The first group of respondents (= the "NOT SHOWN" group) was handed tin-foil covered cans of five brands of beer and, after tasting the contents, was asked to perform the *pivot-ordering* dissimilarity-judgment task. Each brand was taken in turn and became the *pivot*. The respondents were asked to rank-order four other brands according to their dissimilarity from the pivot brand. The second group (= the "SHOWN" group), after tasting from unwrapped beer cans, was asked to perform the same dissimilarity-judgment task. The resulting rank data were analyzed by the pairwise-explosion technique to obtain (group) perceptual maps of five beer brands. In addition, straight preference rankings of five brands supplied by the respondents were pairwise exploded and (group) preference scales for brands were computed from the resulting explosion matrices by Thurstone's [1927] Law of Comparative Judgment.

6.3 Solutions

The logit model (LOGMAP or LOGIT1) was used to analyze this set of data. One solution was obtained for each split-run group. Additionally, the Thurstone preference-scale scores were regressed on the brand coordinates and the regression coefficients are taken as the preference (directional) vectors in the perceptual space.

(Insert Tables 8 a and b, and Figures 2 and 3 about here.)

Two coordinates are given in Table 8 (a) and 8 (b), and plotted in Figures 2 and 3. Both solutions were rotated in such a way that the plane containing Asahi Super Dry (AD) , Kirin Lager (KL) and Kirin Dry (KD) is aligned parallel to the first and second axis. Only the two Suntory brands are set away from this plane. The general shape of brand locations, especially those of KL, AD and KD, are not much different between two solutions, but there are subtle differences.

1. The discrimination parameter, β (note that the standard deviation of the perceptual-error distribution is equal to $\pi/6\beta$), are quite different between two solutions. The value of β is smaller for the NOT SHOWN group than for the SHOWN group (1.30 against 2.33), as expected. If we consider the difference in the β-values, the map for the SHOWN group should be drawn nearly twice as large as that for the NOT SHOWN group. Obviously the SHOWN group has a greater degree of discrimination among brands.

2. The positions of SD and SM are different between two solutions relative to the KL-AD-KD triangle. Compared with the SHOWN group, SD and SM shift toward northeast on the KL-AD-KD plane for the NOT SHOWN group. This indicates that the NOT SHOWN group perceives SM as more different from KL, AD and KD than the SHOWN group does.

3. When the preference vector is plotted on the same map, we find a marked difference. The preference vector for the SHOWN group clearly points to AD, the current favorite, but the one for the NOT SHOWN group points to SM. This is a striking contrast to the relative stability of perceptual maps between two experimental conditions (SHOWN vs. NOT SHOWN). The weights for the first two axes are not different between two maps, but the weight for the third (Suntory) axis is negative for the SHOWN group, but positive for the

NOT SHOWN group. While SD may not have much chance, the fact that SM is clearly distinct from and preferred to other brands in a blind-tasting condition indicates that it may gain acceptance in the market with adequate marketing support.

6.4 Discussion of the Illustration

Despite (or perhaps because of) its tremendous current popularity, Asahi Super Dry has drawn some critical comments from beer connoisseurs (e.g., "Dry beer is not real beer." or "Its raw taste is acceptable only to those who are not real beer drinkers." etc.). This experiment to some extent showed that the popularity of AD is based on factors other than taste. Perhaps it is the newness of AD compared with traditional lager beer that created such popularity. Suntory with the new "Malts 100%" beer has an opportunity to enlarge its share by emphasizing its newness.

6.5 Simulation of Error Distributions

To obtain approximate standard deviations for the scale values we took the coordinates and parameters from the beer illustration and generated pivot-ordering data for 42 simulated subjects. This was done 100 times and the standard deviations from the sample means were computed after Procrustes rotation. For the group NOT SHOWN the beer labels the standard deviations of scale values ranged from a maximum of .1114 to a minimum of .0597 with an average over the 15 coordinated of .0842. For the group SHOWN the beer labels the standard deviations were smaller, as expected, with a maximum of .0569, a minimum of .0334, and an average of .0452.

For the discrimination parameters the means (and standard deviations) were also computed. For the group NOT SHOWN the beer labels the mean was 1.3570 (.3357). For the group SHOWN the beer labels the mean was 2.4486 (.3996) – exhibiting the higher discriminability we expect from this group.

The standard errors corresponding to these standard deviations would be reduced by a factor of 10, underscoring that these solutions seem to have quite acceptable stability.

6.6 Probit Solutions

The PROBIT1 model was applied to the original ranking and the resulting scale values were nearly identitcal. The logit and probit solutions for the group NOT SHOWN the beer labels differed by only .0002 on an average (root mean-square difference), while the scale values for the group SHOWN the labels differed by .0013 on the average.

The discrimination parameters for the probit model were .8721 and .4919 for the NOT SHOWN and SHOWN groups, respectively. Again reflecting the increase in discriminability expected to be associated with being given the brand name as additional information.

7 Extensions of the Pairwise-Explosion Method

We have so far investigated the properties of the proposed estimation method based on the pivot-ordering procedure. This method is, however, adaptable to other data-collection procedures as well. For example, we may use the straight ranking procedure in which the respondent is presented with all $m(m-1)/2$ object pairs and is asked to rank them in the order of dissimilarity. The pairwise-explosion method may be applied to this ranking to yield an aggregate matrix similar to Table 2(b). With a slight change in computation of squared distances, the d_{hi}^2's, the object coordinates and additional parameters are recoverable by the proposed procedure. Of course one may also use the straight paired-comparison procedure in which every object pair is compared against all other pairs in terms of dissimilarity, but, as the number of objects increases, such a data-collection scheme soon becomes too demanding of the respondent's time and effort.

With a large sample size ($n > 400$, say) it may not be necessary for each respondent to rank all objects. After all, the pairwise-explosion method needs only the proportions of respondents who judged one pair of objects as more dissimilar than another pair. While it is desirable that the proportions are estimated as accurately as possible, there are many alternative ways to estimate the proportions. Even the popular *pick-any* scheme (cf. Levine 1979; Holbrook, Moore and Winer 1982) yields data from which the required proportions can be computed. Katahira [1986, 1989] suggested that each respondent ranked only those object pairs which he or she could remember (i.e., the *evoked* set) in the pivot-ordering procedure. He also suggested a selected pivot-ordering procedure in which one may ask each respondent to rank objects for a limited number ($< m$) of pivot elements, the selection of which is randomized for each respondent, or strategically selected as in various missing-data designs for data collection. Another data-collection procedure may be a variation of a full paired-comparison task in which each respondent is asked to compare a small number of randomly chosen object pairs. Any of the procedures discussed here will produce usable data for computing logits or probits. Once the logits or probits are computed, the remaining estimation process will be the same. The investigation of various simplified data-collection procedures as to their efficiency in terms of cost and precision is an important research topic in the future.

8 Summary and Conclusion

In this paper we proposed a quick method for estimating object coordinates in the perceptual space and parameters of discriminatory capacity from a set of ordinal dissimilarities. We applied the choice-theoretic MDS models to data sets obtained from a group of individuals. Our simulation study showed that, when used in conjunction with the pivot-ordering procedure of data collection, the proposed estimation method – pairwise explosion – recovers the original configuration reasonably well. In the case of LOGIT1 (LOGMAP) model, the proposed method yielded estimates which were comparable to those obtained by the maximum-likelihood method, except that the variances of estimates by the proposed method were somewhat greater than those of maximum-likelihood estimates. We could not compare our PROBIT model results with maximum-likelihood estimates, due to the difficulty associated with the maximum-likelihood estimation, but PROBIT models performed slightly better than LOGIT models in terms of variances and mean squared errors of estimates.

We also found that the method was robust in a sense that the recovered configurations were little affected by the model we assumed, so long as the variances of the assumed perceptual-error distributions are equal. The main reason to this robustness is that the double-exponential and normal distributions have a very similar distributional form, and, with the error-sampling scheme achieved by the pivot-ordering procedure, it is impossible to distinguish between LOGIT and PROBIT models. This suggests that in practical applications those two models may be used interchangeably, but, since the probit transformation is less likely to create large variances, PROBIT models should be used when the correct error distribution is not known.

The ability to interpret and understand the marketing implications of maps based on as few as five brands, leads us to believe that there will other product domains in which this style of analysis will be useful.

References

Carroll, J. D. and J. J. Chang [1970], "Analysis of Individual Differences in Multidimensional Scaling Via an N-Way Generalization of 'Eckart-Young' Decomposition," *Psychometrika*, 35, 283-320.

Chapman, R. G., and R. Staelin [1982], "Exploiting Rank Ordered Choice Set Data within the Stochastic Utility Model," *Journal of Marketing Research*, 19, 288-301.

Cliff, N. [1966], "Orthogonal Rotation to Congruence," *Psychometrika*, 31, 31-42.

Cooper, L. G. [1972], "A New Solution to the Additive Constant Problem in Metric Multidimensional Scaling," *Psychometrika*, 37, 311-22.

Cooper, L. G. [1988], "Competitive Maps: The Structure Underlying Asymmetric Cross Elasticities," *Management Science*, 34, 707-23.

Cooper, L. G., and M. Nakanishi [1983], "Two Logit Models for External Analysis of Preference," *Psychometrika*, 48 (December), 607-20.

Eckart, C. and G. Young [1936], "The Approximation of One Matrix by Another of Lower Rank," *Psychometrika*, 1, 211-18.

Furukawa, I. [1987], "An Individual Differences Model for LOGMAP," *Maketingu Saiensu*, 27 (June), 41-53.

Holbrook, M. B., W. L. Moore and R. S. Winer [1982], "Constructing Joint Spaces from Pick-Any Data: A New Tool for Consumer Analysis," *Journal of Consumer Research*, 9, 99-105.

Katahira, H. [1986], "Constructing a Perceptual Map by Rank Order Logit Analysis: a More Efficient Method of Nonmetric MDS," in K. Moler and M. Paltschik (eds.), *Contemporary Research in Marketing, Vol. 2, (Proceedings of the XV^{th} Annual Conference of the European Marketing Academy)*, Helsinki, 791-808.

Katahira, H. [1986], "Perceptual Mapping Using Ordered Logit Analysis," *Marketing Science*, forthcoming.

Kruskal, J. B. [1964a], "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," *Psychometrika*, 29, 1-27.

Kruskal, J. B. [1964b], "Nonmetric Multidimensional Scaling: A Numerical Method," *Psychometrika*, 29, 115-28.

Levine, J. H. [1979], "Joint-Space Analysis of Pick-Any Data: Analysis of Choices from an Unconstrained Set of Alternatives," *Psychometrika*, 44, 85-92.

MacKay, D. B. and J. L. Zinnes [1981], "Probabilistic Scaling of Spatial Distance Judgments," *Geographical Analysis*, 13, 21-37.

MacKay, D. B. and J. L. Zinnes [1986], "A Probabilistic Model for the Multidimensional Scaling of Proximity and Preference Data," *Marketing Science*, 5, 325-44.

McFadden, D. [1974], "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka (ed.), *Frontiers in Econometrics*, New York: Academic Press.

Osawa, Y. [1988], "Parameter Estimation and Computation of Standard Errors for PVLOGMAP," *Osaka University Economic Review*, Forthcoming.

Suppes, P., and J. L. Zinnes [1963], "Basic Measurement Theory" in R. D. Luce, R. R. Bush and E. Galanter (eds.), *Handbook of Mathematical Psychology, I*, New York: John Wiley and Sons.

Schönemann, P. [1966], "A Generalized Solution to the Orthogonal Procrustes Problem," *Psychometrika*, 31, 1-10.

Thurstone, L. L. "A Law of Comparative Judgment," *Psychological Review*, 34, 273-86.

Torgerson, W. S. [1958], *Theory and Methods of Scaling*, New York: John Wiley and Sons.

Zinnes, J. L. and D. B. MacKay [1983], "Probabilistic Multidimensional Scaling: Complete and Incomplete Data," *Psychometrika*, 48, 27-48.

Table 1: Osawa's Simulation Results – LOGMAP

(Number of Simulated Subjects = 50; Number of Trials = 50)

Object Coordinates	True Values	Estimated Means	Standard Deviations	Deviation Rate (Bias/Std. Dev.)
X1	1.000	0.999	0.032	–0.030
Y1	0.000	—	—	—
X2	0.500	0.502	0.032	0.046
Y2	0.866	0.862	0.033	–0.135
X3	–0.500	–0.503	0.039	–0.076
Y3	0.866	0.862	0.028	0.092
X4	–1.000	–1.004	0.031	–0.125
Y4	0.000	–0.002	0.033	–0.053
X5	–0.500	–0.494	0.035	0.178
Y5	–0.866	–0.864	0.026	0.080
X6	0.500	0.500	0.031	0.002
Y6	–0.866	–0.864	0.033	0.050
β	5.000	4.620	0.228	1.669
The log-likelihood		–965.911	23.232	
CPU Seconds on IBM3090-200		123.955	35.201	

Table 2: Numerical Example of Pair-Wise Explosion

(a) Individual Data

		Objects				
Objects	Ranking	1	2	3	4	5
1	1	–	1	1	1	1
2	3	0	–	0	1	1
3	2	0	1	–	1	1
4	5	0	0	0	–	0
5	4	0	0	0	1	–

(b) Group Data (n respondents)

	Objects				
Objects	1	2	3	4	5
1	–	n_{12}	n_{13}	n_{14}	n_{15}
2	n_{21}	–	n_{23}	n_{24}	n_{25}
3	n_{31}	n_{32}	–	n_{34}	n_{35}
4	n_{41}	n_{42}	n_{43}	–	n_{45}
5	n_{51}	n_{52}	n_{53}	n_{54}	–

$n_{ij} + n_{ji} = n$

Table 3: Simulation Results – LOGIT1

(Number of Simulated Subjects = 50; Number of Trials = 50)

Object Coordinates	True Values	Estimated Means	Standard Deviations	MSE
X1	0.0000	–0.0036	0.0134	0.000193
Y1	0.1312	0.1305	0.0190	0.000361
X2	0.6820	0.6830	0.0322	0.001038
Y2	0.5250	0.5254	0.0263	0.000692
X3	0.6820	0.6756	0.0342	0.001211
Y3	–0.2625	–0.2584	0.0287	0.000841
X4	0.0000	0.0046	0.0302	0.000933
Y4	–0.6562	–0.6617	0.0331	0.001126
X5	–0.6820	–0.6830	0.0355	0.001261
Y5	–0.2625	–0.2642	0.0256	0.000658
X6	–0.6820	–0.6765	0.0279	0.000809
Y6	0.5250	0.5284	0.0288	0.000841
β	5.0000	5.1740	0.2796	0.108452

The sum of MSE's for coordinates over all objects = 0.009959

Table 4: Simulation Results – LOGIT2

(Number of Simulated Subjects = 50; Number of Trials = 50)

Object Coordinates	True Values	Estimated Means	Standard Deviations	MSE
X1	0.0000	–0.0037	0.0170	0.000303
Y1	0.1312	0.1327	0.0184	0.000341
X2	0.6820	0.6805	0.0237	0.000564
Y2	0.5250	0.5240	0.0206	0.000425
X3	0.6820	0.6767	0.0256	0.000683
Y3	–0.2625	–0.2573	0.0259	0.000698
X4	0.0000	0.0046	0.0244	0.000617
Y4	–0.6562	–0.6683	0.0274	0.000897
X5	–0.6820	–0.6799	0.0276	0.000766
Y5	–0.2625	–0.2588	0.0204	0.000430
X6	–0.6820	–0.6782	0.0207	0.000443
Y6	0.5250	0.5276	0.0226	0.000518
β	5.0000	5.1514	0.2565	0.088714

The sum of MSE's for coordinates over all objects = 0.006682

Table 5: Simulation Results – PROBIT1

(Number of Simulated Subjects = 50; Number of Trials = 50)

Object Coordinates	True Values	Estimated Means	Standard Deviations	MSE
X1	0.0000	0.0020	0.0142	0.000206
Y1	0.1312	0.1356	0.0233	0.000562
X2	0.6820	0.6877	0.0350	0.001257
Y2	0.5250	0.5243	0.0303	0.000919
X3	0.6820	0.6819	0.0339	0.001149
Y3	–0.2625	–0.2609	0.0318	0.001014
X4	0.0000	–0.0027	0.0278	0.000780
Y4	–0.6562	–0.6563	0.0326	0.001063
X5	–0.6820	–0.6841	0.0348	0.001215
Y5	–0.2625	–0.2574	0.0312	0.000999
X6	–0.6820	–0.6848	0.0336	0.001137
Y6	0.5250	0.5147	0.0303	0.001024
σ	0.2565	0.2526	0.0118	0.000154

The sum of MSE's for coordinates over all objects = 0.011332

Table 6: Simulation Results – PROBIT2

(Number of Simulated Subjects = 50; Number of Trials = 50)

Object Coordinates	True Values	Estimated Means	Standard Deviations	MSE
X1	0.0000	0.0014	0.0173	0.000301
Y1	0.1312	0.1357	0.0240	0.000596
X2	0.6820	0.6836	0.0235	0.000555
Y2	0.5250	0.5233	0.0244	0.000598
X3	0.6820	0.6811	0.0223	0.000498
Y3	−0.2625	−0.2576	0.0220	0.000508
X4	0.0000	−0.0007	0.0224	0.000502
Y4	−0.6562	−0.6669	0.0217	0.000585
X5	−0.6820	−0.6807	0.0233	0.000545
Y5	−0.2625	−0.2514	0.0245	0.000723
X6	−0.6820	−0.6847	0.0224	0.000509
Y6	0.5250	0.5170	0.0245	0.000664
σ	0.2565	0.2536	0.0085	0.000081

The sum of MSE's for coordinates over all objects = 0.006596

Table 7: Simulation Results – Overall

(Number of Simulated Subjects = 50; Number of Trials = 50)

	Simulation Data			
Models: Statistics	LOGIT1	LOGIT2	PROBIT1	PROBIT2
LOGIT1 Model:				
$\sum \text{Bias}^2$	0.000170	0.003059	0.000358	0.004104
$\sum \text{Var}$	0.009789	0.008550	0.014346	0.010951
$\sum \text{MSE}$	0.009959	0.011609	0.014703	0.015056
β	5.1740	5.5421	5.1126	5.5616
LOGIT2 Model:				
$\sum \text{Bias}^2$	0.002875	0.000279	0.002330	0.000241
$\sum \text{Var}$	0.007750	0.006403	0.010078	0.007792
$\sum \text{MSE}$	0.010625	0.006682	0.012407	0.008034
β	4.7676	5.1514	4.7358	5.1854
PROBIT1 Model:				
$\sum \text{Bias}^2$	0.000314	0.002322	0.000209	0.002440
$\sum \text{Var}$	0.007748	0.006799	0.011123	0.008427
$\sum \text{MSE}$	0.008062	0.009121	0.011332	0.010867
σ	0.2498	0.2357	0.2526	0.2351
PROBIT2 Model:				
$\sum \text{Bias}^2$	0.003430	0.000448	0.002926	0.000364
$\sum \text{Var}$	0.006222	0.005202	0.008058	0.006232
$\sum \text{MSE}$	0.009652	0.005650	0.010984	0.006596
σ	0.2716	0.2545	0.2737	0.2536

Table 8: Perceptual Maps for Japanese Beers

NOT SHOWN Group

Brand		Coordinates Axis 1	Axis 2	Axis 3	Preference Scores
Suntory Dry	SD	–0.2228	–0.3419	0.4714	–0.5276
Kirin Dry	KD	–0.2629	0.5645	–0.2633	–0.6341
Suntory Malts 100%	SM	0.4215	0.4075	0.3183	1.4874
Asahi Super Dry	AD	0.5078	–0.3151	–0.2633	0.3759
Kirin Lager	KL	–0.4437	–0.3151	–0.2633	–0.7019
Preference Vector		1.7591	0.5008	0.6420	
Discrimimation Parameter (β) 1.3001					

SHOWN Group

Brand		Coordinates Axis 1	Axis 2	Axis 3	Preference Scores
Suntory Dry	SD	–0.3934	–0.4130	0.3563	–1.6287
Kirin Dry	KD	–0.2279	0.7223	–0.2826	0.4002
Suntory Malts 100%	SM	0.2547	0.1599	0.4916	0.0614
Asahi Super Dry	AD	0.6137	–0.2346	–0.2826	2.0355
Kirin Lager	KL	–0.2471	–0.2346	–0.2826	–0.8685
Preference Vector		2.8229	0.8020	–1.1966	
Discrimimation Parameter (β) 2.3332					

Figure 1. Object Configuration for Simulation

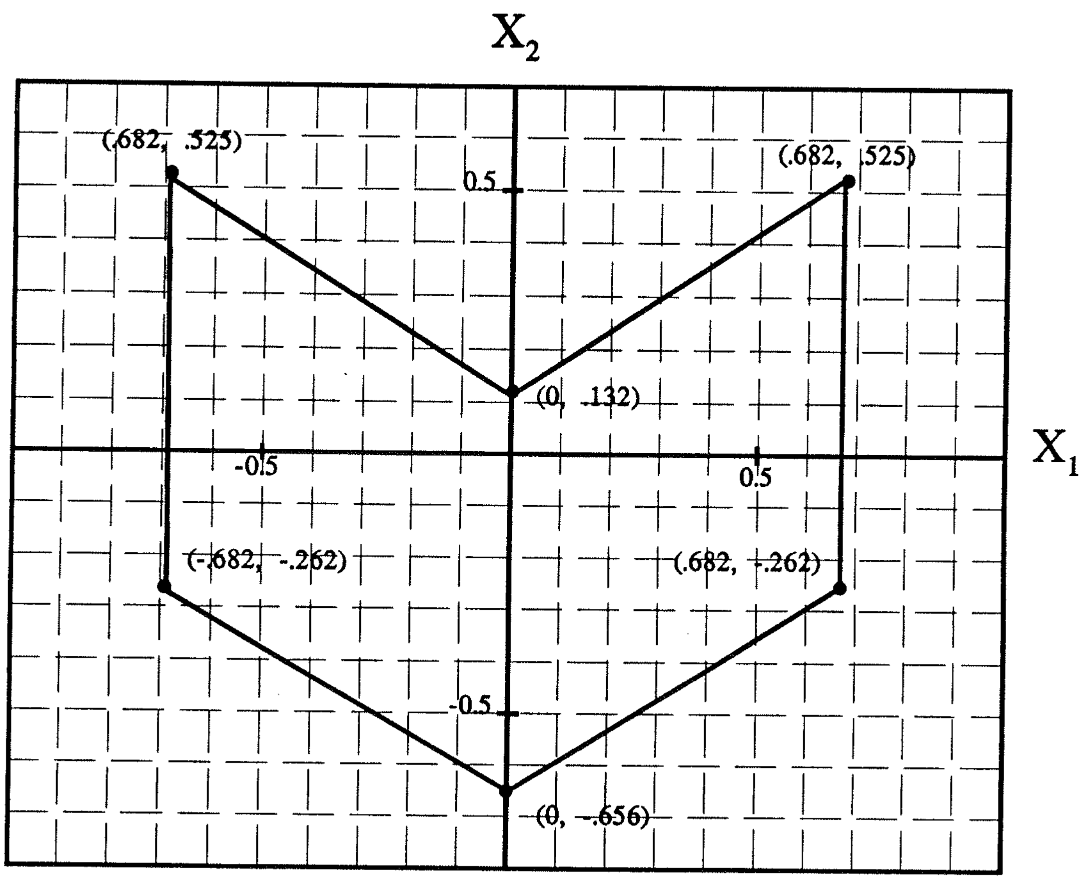

Figure 2. Map for Group NOT SHOWN the Beer Labels.

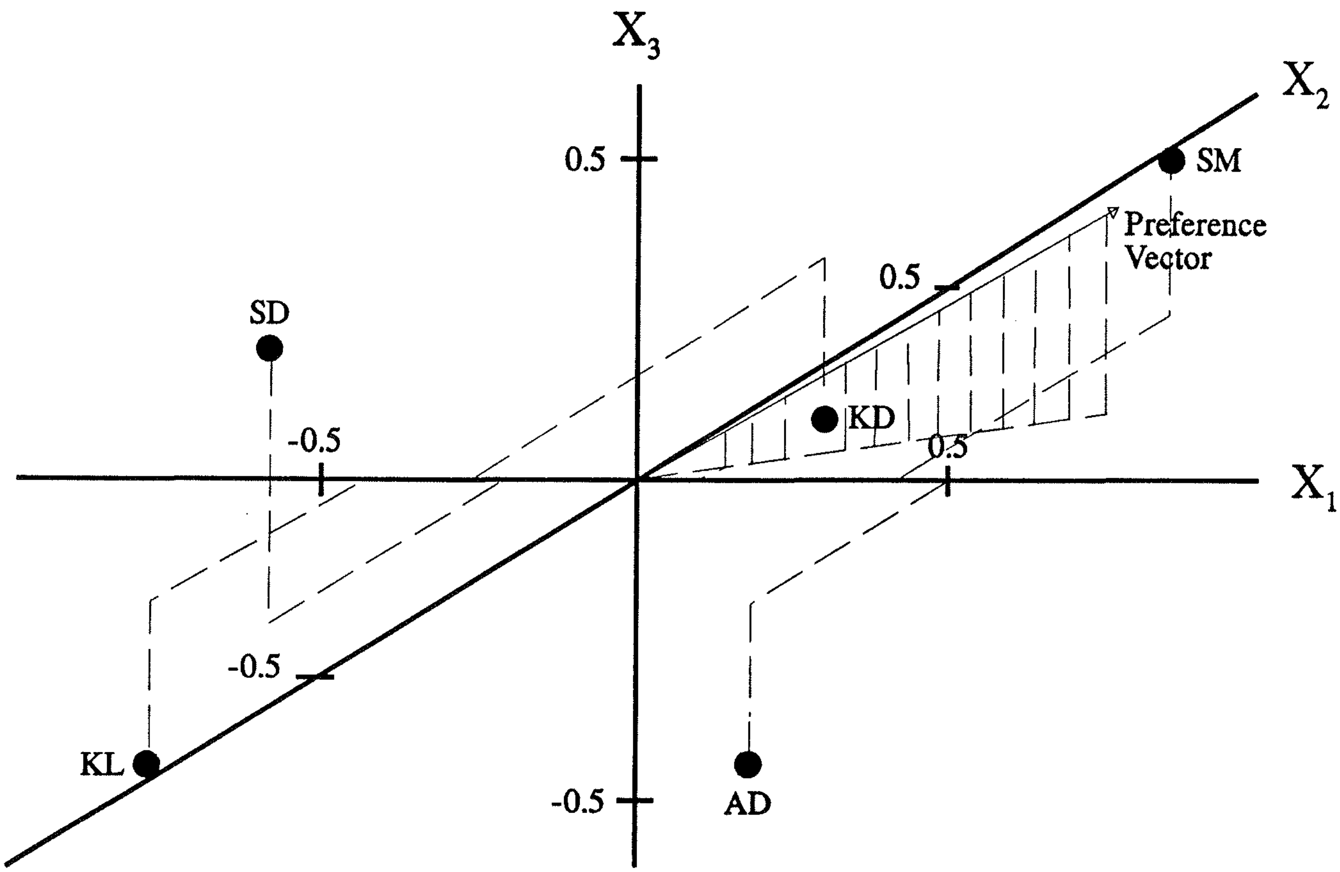

Figure 3. Map for Group SHOWN the Beer Labels.