DISCUSSION PAPER NO. 319

THE RELATIONSHIP BETWEEN ATTRIBUTES, BRAND PREFERENCE, AND CHOICE: A STOCHASTIC VIEW

Jean-Marie Blin[1] and Joe A. Dodson[2]

February, 1978
(revised)

[1]Northwestern University

[2]N.W. Ayer ABH International

We wish to acknowledge our appreciation for helpful comments and programming assistance by Jerry Wharton. We also thank Hillel Einhorn, Arthur Geoffrion, John Hauser, Morgan Jones and Alan Shocker, who made constructive comments and suggestions at various stages in the preparation of this paper.

The Relationship Between Attributes, Brand Preference, and Choice: A Stochastic View

Introduction

Two separate research traditions have arisen in marketing over the last decade in the study of consumer behavior. Each has tended to focus its attention primarily on one of the links in the brand attribute → brand preference → brand choice chain. One approach which has emerged from a number of disciplines, most notably economics [15] and social psychology [19] has focused attention on the multiattribute nature of the goods chosen by consumers. The one model which has received the most attention in the marketing literature is the expectancy-value model adapted from a formulation proposed by Fishbein [13]. Much work has gone into attribute identification, interaction, and measurement [14]. In spite of the refinements in measurement techniques and the greater reliability of the data bases generated, the predictive power of this type of model has not fulfilled the early expectations of researchers [3]. A substantial, unexplained variance still remains. More recently explanations suggesting situational variables unrecorded in the data base (thus unknown to the researcher) have been offered [8].

An alternative approach to the study of consumer behavior has focused its attention on observed consumer choice. Marketing has long recognized the brand switching phenomenon [12]. In fact, such recognition forms the basis for market penetration strategies and new brand introductions. A variety of stochastic models of brand choice (e.g., Markov or linear learning), has been presented in the marketing literature [16]. Most frequently the need for a stochastic model is attributed to the complexity of the choice process and the inability to take into account all of the factors which influence choice such as promotions, deals, availability, etc.

Whether behavior is fundamentally stochastic, as Bass [3] suggests, or only appears stochastic because of our inability to measure all of the relevant influences adequately, is a debate which likely will never be resolved. However, to progress in the study of consumer behavior, these two schools of thought must recognize and acknowledge the contribution of each toward developing a theory of consumer choice. Stochastic models of brand choice in general have not provided a rationale for their probabilistic basis. Much of the behavioral research has ignored the brand switching phenomenon or chosen to attribute it to situational changes in consumer preference.

In the next section an explanation for this divergence is discussed. This discussion suggests an alternative research strategy embodied in a multi-attribute model of brand choice, developed by Blin and Dodson [10]. This model when fitted against empirical data compares favorably with three alternative models formulated to predict observed consumer choice.

Preference vs. Choice

One explanation for the lack of cross fertilization between multiattribute research and stochastic modeling is that each school of thought has set up differential objectives and measures of success. Multiattribute research in marketing has focused on the evaluative aspects of choice. A majority of the articles reviewed by Wilkie and Pessemier [21] utilized stated preference as the criterion variable. While these studies have provided reasonably good predictions of the stated preferences of consumers, they have failed to provide good predictions of individual brand choice. Actual choice rather than stated preference has been the dependent variable in stochastic models of buyer behavior. The existence of consumer panel diary data prompted the use of stochastic models to represent the purchasing process. Thus to bridge the gap between the two schools of thought the relationships between stated preference and actual choice behavior must be evaluated.

Some evidence on the relationship between preference and choice already exists. Bass, Pessemier, and Lehmann [4] measured the stated preferences of 264 experimental subjects for 8 soft drink brands. They also observed the actual choice behavior over a three week period. Even while stated preferences and attitudes were substantially stable during the experiment, choices of subjects were constantly changing. Using stated preference as a basis for predicting choice resulted in only 55% correct predictions. Axelrod [1], in a comparison of ten measures of stated preference, found that only 44% of those subjects indicating preference for a brand in response to a lottery question purchased that brand during the following three weeks.

As noted above, some researchers have proposed a situational explanation for this lack of agreement [8]. However, with this approach the researcher is led on an extended search for "all" the influences of behavior. Unless he is able to specify the extent and degree of influence of these essentially "random" situational elements, the researcher will never be able to evaluate the viability of a predictive model since it is likely that there will always be some unexplained variance in individual behavior.

An alternative research strategy is proposed in this paper. The authors present a theory of choice which is linked directly to the consumer's attitude structure. Within this structure brand switching can be represented as an integral and rational element of consumer behavior without having to appeal to "situationalism". The mathematical details of the model have been elaborated elsewhere [10] and will only be summarized here. The objective here is firstly to present a multiattribute model which directly links a consumer's attitude structure to his observed choice behavior, and secondly to evaluate the ability of the model to predict actual choice behavior. The model presented in the next section, is later compared to alternative models of brand choice on the basis of preference prediction and choice prediction.

The Model

In this section a model for determining an individual's aggregate preference ordering on a set of stimuli is presented. The model provides insight into the causes of brand switching and also provides a basis for predicting the frequency of choice. Since the analysis does not involve any comparisons across individuals, the procedure can be carried out for each individual separately.

Definition and Notation

We adopt the following notation:

$B = \{b_1, b_2, \ldots, b_i, \ldots, b_m\}$ is the brand set from which the consumer chooses.

$A = \{a_1, a_2, \ldots, a_k, \ldots, a_n\}$ is the attribute set, taken as given. P is a $(m \times m)$ permutation matrix representing a preference ranking (aggregate or for some attribute). In general, a permutation matrix P has a one in each row and column and zeroes everywhere else. For instance, if $B = \{b_1, b_2, b_3\}$ and a consumer's (aggregate) preference ranking is $b_1 > b_3 > b_2$ ($>\, \equiv$ preferred to), we write:

$$(1) \qquad P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

as postmultiplying (b_1, b_2, b_3) by P yields $b_1 > b_3 > b_2$: $(b_1, b_2, b_3)\ P = (b_1, b_3, b_2)$, Thus the rows of P correspond to the rank and the columns correspond to the choice alternatives. The permutation matrix P merely reorders the elements in the row vector from (b_1, b_2, b_3) to (b_1, b_3, b_2). $\mathcal{P}$ denotes the set of all permutation matrices of order m$(|\mathcal{P}| = m!)$ i.e., the set of all conceivable rankings of the m choice objects. For future reference we also need an obvious extension of permutation matrices: doubly-stochastic matrices. These matrices have all entries between 0 and 1 and row and column sums all equal to one (see, for example, equation 2 below). $\mathcal{S}$ denotes the set of all doubly-stochastic matrices.

Model Input

The following inputs are used to determine the individual's aggregate preference ranking:

1) A set of n attributewise preference rankings say $P_1, P_2, \ldots, P_k \ldots, P_n$ ($P_k \in \mathscr{P}$). If the consumer ranking is weak, ties can be accommodated by entering $\frac{1}{t_v}$ for the entries of the permutation matrix corresponding to the v-th tie class (t_v is the number of brands in that class). For instance, if the ranking is $(b_1, b_3) > b_2$ (b_1 and b_3 are tied), we write:

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix} \tag{2}$$

Note that the row and column sums are still all 1 as in the strict ranking case, eq. (1).

2) A set of n normalized weights, one for each attribute, denoted $w_1, w_2, \ldots, w_n$. Weights can be determined externally by direct questioning or generated by a variety of known procedures [17]. Any of these procedures may be used to provide weights which become input to the model, where $\sum_{k=1}^{n} w_k = 1$ and $0 \leq w_k \leq 1$.

Model Specification

Any aggregation process must deal with the problem of scale heterogeneity. The possibility of a different psychological metric for each attribute is a common problem among multiattribute models. Here, only ordinal evaluations of the brands on each attribute are required. And, in processing these various attributewise evaluations, it is assumed that the consumer follows a linear aggregation process.[1]

$$S = \sum_{k=1}^{n} w_k P_k \tag{3}$$

While alternative processing models have been proposed, the linear compensatory hypothesis has played a central role in the multiattribute modelling literature. Empirical support for the hypothesis has been reviewed by Slovic and Lichtenstein [20].

In general, S is a doubly-stochastic matrix in $\mathscr{S}$. For instance, if

$B = (b_1,b_2,b_3)$, $A = (a_1,a_2,a_3)$ and $w_1 = w_2 = 1/4$, $w_3 = 1/2$,

$$\underset{(b_1,b_2,b_3)}{P_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \underset{(b_2,b_1,b_3)}{P_2} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \underset{(b_2,b_3,b_1)}{P_3} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

(4) $$S = 1/4P_1 + 1/4P_2 + 1/2P_3 = \begin{bmatrix} 1/4 & 1/4 & 1/2 \\ 3/4 & 1/4 & 0 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

The aggregate preference ranking, say P*, which "best" approximates the doubly-stochastic ranking S provides a predicted preference ordering among the choice alternatives for each consumer. The individual's predicted aggregate preference ordering, P*, can be found by finding the permutation matrix P which solves the following minimal distance problem:

(5) $$\min_{P \in \mathscr{P}} d(P,S)$$

where d is some distance in $\mathscr{S}$. In particular, it can be shown [11] that assuming d to be the Euclidean or city-block metric leads to the same solution, and that this is also the solution to the following linear assignment problem:[2]

(6) $$\max_{P \in \mathscr{P}} \sum_{i,j} s_{ij} P_{ij}$$

A Geometric Interpretation

The minimal distance problem in equation (5) seeks the brand preference ranking which most closely matches the aggregated attributewise rankings in S. It also leads to a convenient geometrical representation of the problem. In the case where m=3 and n=3, the problem has a simple two-dimensional representation (Figure 1). If we choose d to be the city-block metric, we write equation (5) as:

(7) $$\min_{P \in \mathscr{P}} z = \sum_{(i,j)} |P_{ij} - s_{ij}|$$

In our previous example (eq. 4), the solution is:

$$P^* = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{i.e. } b_2 > b_3 > b_1$$

and $z(P^*) = 3/4 + 1/2 + 1/2 = 7/4$.

Geometrically, this solution can be found by projecting S orthogonally onto the closest face of the polytope $\mathscr{A}$, and repeat this procedure until we reach a face of order 0, that is, a vertex ($b_2 > b_3 > b_1$, here); or until no improvement in z can be found, if there are several optima. In the latter case, we could have P**, say, if b_1 and b_3 were tied for second place (see Figure 1). If all three brands were tied, we would get point C in the center of this polytope. Again, referring to Figure 1 above, we see that the wedge-shaped portions centered at C delineate the regions of $\mathscr{A}$ which lead to the corresponding vertex being chosen.

The Preference-Choice Link

At this point, a simple deterministic view of consumer choice might lead us to predict that a consumer would choose his most preferred brand, if no constraints were placed in his way. But the evidence presented above suggests that we would frequently be incorrect. In this case, a strict determinist might then argue that the choice reflects a shift in the consumer's preferences. An alternative interpretation which leads to a prediction of the likelihood of choice is presented here.

The Stochastic View

The alternative view argues that the consumer's attitude structure governs not actual choice on any single purchase occasion but the relative frequency of choice over repeated purchase occasions. If we allow for repeated purchasing, then it seems unreasonably restrictive to force the consumer always to choose the brand he states as most preferred. One may agree with this statement but argue that the circumstances under which each individual choice is made would determine the actual choice on each occasion. However, a prediction of every single choice may not be necessary for managerial decision making; rather a

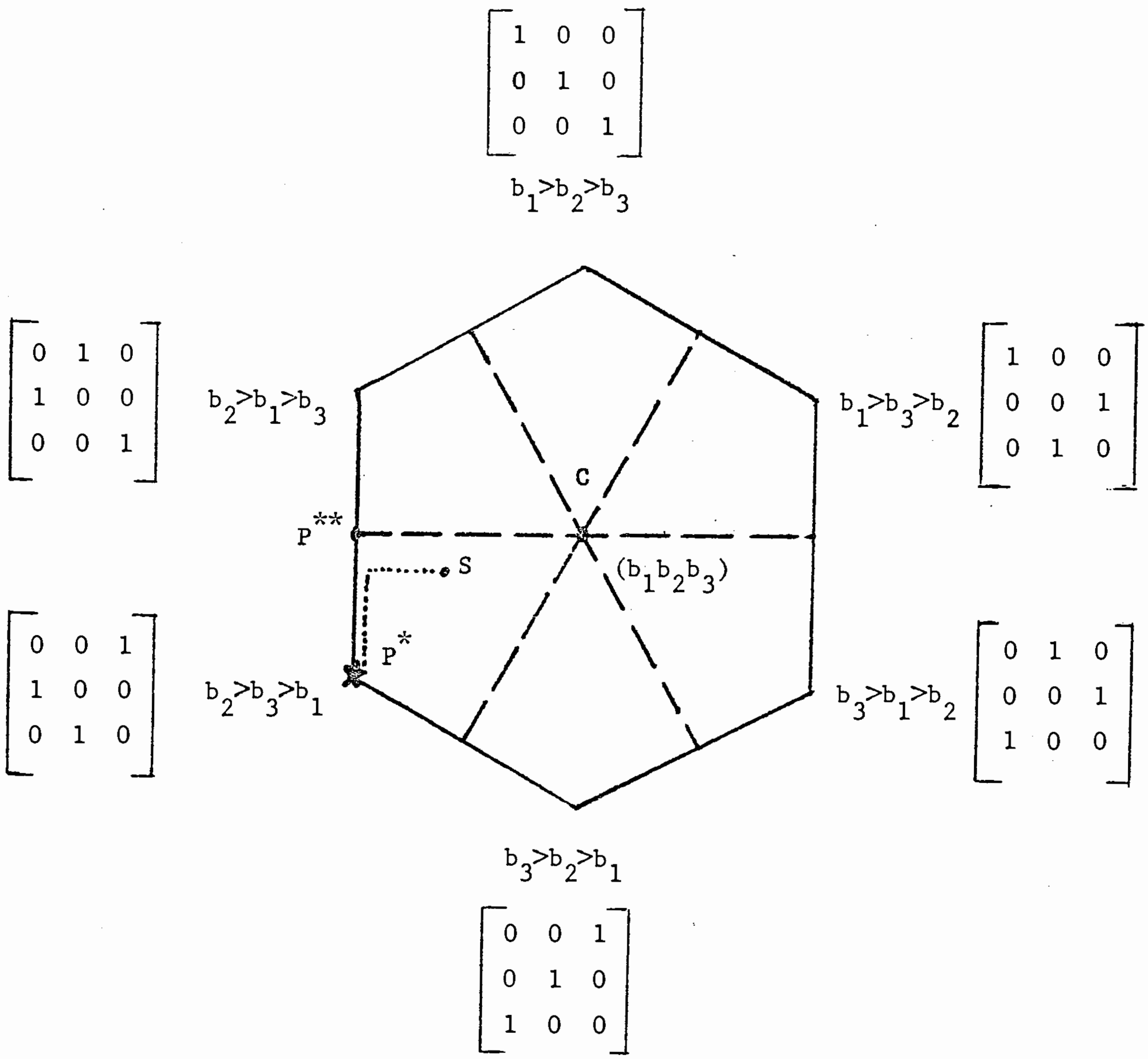

Figure 1. Geometrical Representation for $\mathscr{A}_{3\times 3}$.

prediction of the likelihood of choice may be sufficient. That is, the model should provide information about the probability with which each of the brands will be chosen on any given purchase occasion. Our avowed determinist then may argue that such information is only to be found in an examination of the circumstances surrounding each choice. The alternative view offered here is that it is contained in the attitudinal structure of the consumer. That is, the individual's evaluation of the alternatives can be used to predict the relative frequency of choice over repeated purchase occasions.

The Stochastic Model

In the model the consumer's likely switching pattern is reflected in the set of attributewise rankings of the alternatives he reveals. This information is captured in the aggregate stochastic ordering matrix S, e.g., eq. (4). The dilemma of the consumer faced with the failure of any one brand to dominate on all salient attributes is reflected in the model. In the above example (eq. (4)), row three reflects the fact that b_3 does not dominate on all attributes and, in fact, has an equal probability of being ranked second or third, given the importance assigned to the third attribute. In particular, the first column of S reflects the relative probability of any one of the brands being the most preferred brand. That no alternative has a probability of one indicates that no alternative in the choice set dominates on all the relevant attributes. Thus, column one of S provides a procedure for estimating the relative frequency with which each of the brands will be chosen, when switching behavior can be solely attributed to the multiplicity of attributes.

However, this solution still falls short of fully integrating all the factors involved in the decision process: in particular, it ignores the presence of uncertainty. Realistically, at least two sources of uncertainty are faced by the consumer: (1) internal uncertainty about, which attributes are relevant to his choices, what weight should they be given, and how each choice alternative performs on each attribute scale, and (2) external uncertainty due to environmental influences such as availability, dealing, etc. Internal uncertainty can be dealt with directly in the model. Formally, this means that both the attribute weights w_k and the attribute preference ordering P_k are random variables whose distribution reflects the amount of uncertainty faced by the consumer. Then

$$S = \sum_{k=1}^{n} w_k P_k \tag{8}$$

is also a random variable in $\mathscr{A}$, whose density function (see Figure 2) depends on the distribution of n, w_k, and P_k.

Assuming that a consumer would always choose his most preferred brand, external uncertainty acts to divert the consumer's choice to a less preferred alternative. If we are unable to explicitly include all of these random shock factors into our models we must fall back to a prediction of the likelihood of choice. And, it is the authors' contention that the consumer's attribute-preference structure provides the information on which to make this prediction.[3]

Specifically, as noted above, we can partition $\mathscr{A}$ into m mutually exclusive and jointly exhaustive regions, one for each ranking with brand i leading (i = 1,2,...,m) as the solution vertex, for some d. For instance, if d is the city-block metric, we would have the situation represented in Figure 2. Integration of the density function of S over each choice region yields probabilities of choice of each brand b_i. Practically, these probabilities define a multinomial process for choice over B, the brand set, and the probability of choice of brand i, say Π_i, is a function of the size of the "b_i-chosen" region and the shape and location of the density pattern. These probabilities represent estimates of a consumer's likelihood of choice on any given purchase occasion.

Model Validation

To be useful in predicting consumer behavior, the model must be amenable to empirical testing. An application of the model to the prediction of preference requires only (i) an individual's set of preference rankings on each

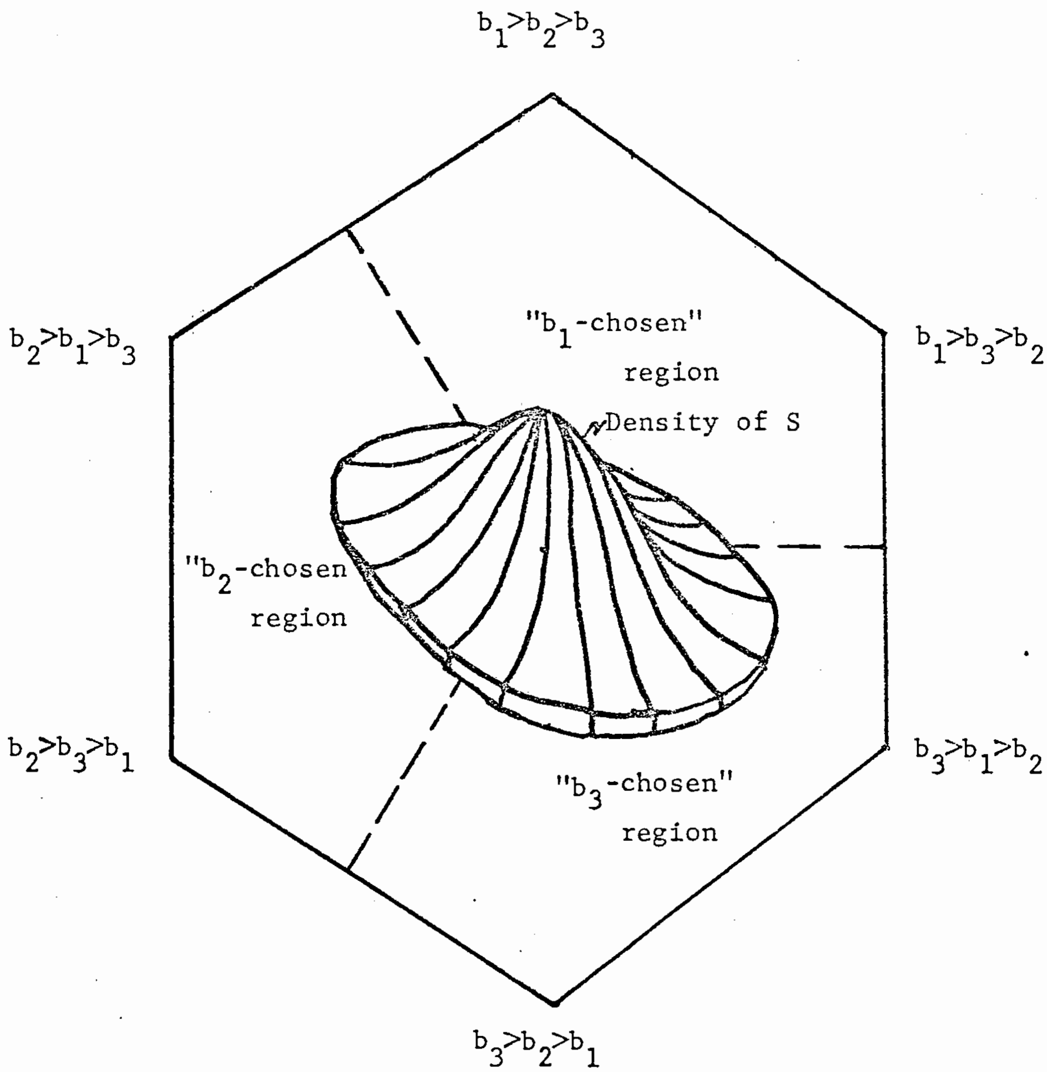

Figure 2. Geometrical Representation of Uncertainty Over The Choice Regions

attribute and (ii) an individual's weighting of the attributes. With this data it is possible to find that ranking of the brands which comes closest to, in an already specified sense, the individual's aggregate preference ranking. Thus, it is possible to test the model by comparing the predicted against the stated rank order preference of the brand set across individuals. However, as has been argued, this ignores the stochastic nature of consumer choice behavior. What is needed are good predictions of the likelihood of choice. If switching behavior is to be solely attributed to the multiplicity of attributes, then the model provides a prediction of the frequency of choice of each brand from the first column of S. However, if we are to integrate this effect with the effect of uncertainty in the choice situation, then in theory what is required to estimate the choice probabilities is the joint distribution of P_k and w_k, i.e., which yields a derived density function for S.

Eliciting the density function directly from an individual would be an arduous cognitive task. However, the individual does provide some information about his uncertainty in the input data. By allowing the consumer to express ties in the attributewise rankings, P_k, it is possible to approximate the density pattern of S by a discrete distribution. Each tie indicates a consumer's perceptual uncertainty about the appropriate positioning of the tied brands on a given attribute. In the absence of additional information, each strict ranking consistent with the individual's stated weak ranking would be assumed equally probable.[4] For instance, in the case of complete uncertainty, i.e., all brands tied on each attribute, it can easily be shown [11] that assuming each strict P_k ordering as equally likely for all k (and w_k equal), yields a uniform distribution for S over $\mathcal{S}$ and

$$(9) \qquad E[S] = \begin{bmatrix} 1/n & 1/n & \dots & 1/n \\ 1/n & 1/n & \dots & 1/n \\ \vdots & \vdots & \ddots & \vdots \\ 1/n & 1/n & & 1/n \end{bmatrix}$$

Thus, consumer revelation of uncertainty appears in the attributewise rankings, P_k's. In order to evaluate the distribution of S, we then need to consider each strict ordering compatible with a given weak ordering on each attribute and compute S for each such strict ordering. To compute the predicted outcome from the model, we would solve the resulting problem Min d(S,P) for each different S obtained. By looking at the brand chosen first in the model, we can assign frequencies of occurrence to S in each b_i-chosen region resulting from the solution of the associated linear assignment problem. These frequencies are used to compute the Π_i, probability of brand i being chosen on any given purchase occasion. In the absence of any alternative heuristic way to determine the distribution of S, this solution assumes little about the consumer's preferences beyond the information provided in the attributewise rankings. The validity of this assumption is empirically tested in the next section by comparing the model's choice predictions with observed consumer choice

Model Predictions

The analysis reported here is based on the experimental data generated by Bass, Pessemier, and Lehmann [4]. The details of the experiment and the information collected are reported in their article. Subjects were required to make choices from a set of eight soft drinks four days a week for a period of three weeks. In addition to making consumption choices among the brand set, participants completed questionnaires supplying information about their preferences for the brands, beliefs about the brands on seven attributes, and the perceived importance of each of the attributes. Complete preference, attitudinal, and choice information were obtained for two hundred and sixty-four participants.

Preference Predictions

The most frequently reported criterion for multiple-attribute model evaluation is preference prediction. Table 1 is a confusion matrix showing the conditional probabilities of stated preference rank given the predicted rank from the model. The model predictions are based on the subjects' attributewise rankings of the alternatives on each of seven attributes: carbonation, calories, sweetness, thirst-quenching, popularity, packaging, and after-taste.[5] Each attribute was weighted equally.[6] These results are similar to those reported in previous studies [4,5]. The predictions are best for first preference and last preference. The probability that a brand will be preferred first or second given that the model predicts it is most preferred, is about .6.

It is interesting to note that little information was lost in using only ordinal ranking information for each attribute. This can be seen by comparing the model predictions (Table 1) to preference order predictions for each subject using a Fishbein type attitude score developed for each brand (Table 2). Comparison of the results reported in Tables 1 and 2 suggests some important conclusions. The procedure for collecting belief measures using Likert scales provides only ordinal information about the brands. An assumption of cardinality must be made to use the Fishbein model to combine the attributes to obtain an attitude score. Judging from the results reported in these tables, it appears that this assumption is unnecessary. This finding was corraborated by further comparisons utilizing reduced attribute lists and alternative attribute weighting schemes, i.e., equal weights vs. self explicited weights. A similar pattern prevailed throughout these trials. Thus an assumption of equal distance between alternative rankings on each attribute scale provides as much predictive power as the stronger assumption required by the Fishbein type linear compensatory models.

Additional information about an individual's attitude structure is provided by the value of the objective function, z (see eq. (7)), obtained from the model. z is a measure of the concordance of the consumer's rankings of the choice alternatives across the attributes. z can take on values from m for complete lack of agreement to 0, for complete agreement of the rankings across all salient attributes, where m is the number of choice alternatives. Thus, z indicates the closeness of the predicted preference ordering P^* to the subject's ideal point represented by S. In this sense, z could be considered a loyalty index. A lack of concordance would show up as a larger z-value and

Table 1

Conditional Probability of Stated Preference Rank vs. Predicted Preference Rank Using the Minimal Distance Model

		Subject Stated Rank								
		1	2	3	4	5	6	7	8	
Predicted Rank	1	.520	.189	.110	.075	.047	.028	.020	.012	1.000
	2	.189	.343	.169	.102	.087	.067	.024	.020	1.000
	3	.091	.220	.272	.193	.114	.071	.024	.016	1.000
	4	.126	.110	.185	.256	.146	.091	.059	.028	1.000
	5	.028	.051	.138	.150	.283	.165	.114	.071	1.000
	6	.020	.028	.067	.094	.213	.287	.185	.106	1.000
	7	.008	.020	.008	.071	.055	.150	.362	.327	1.000
	8	.016	.039	.043	.059	.063	.146	.217	.417	1.000

Table 2

Conditional Probability of Stated Preference Rank vs. Predicted Preference Rank Using Fishbein Model*

		Subject Stated Rank								
		1	2	3	4	5	6	7	8	
Predicted Rank	1	.538	.212	.121	.068	.038	.008	.008	.008	1.000
	2	.227	.383	.163	.095	.061	.034	.027	.011	1.000
	3	.114	.220	.299	.193	.102	.042	.019	.011	1.000
	4	.038	.083	.163	.299	.159	.155	.072	.030	1.000
	5	.049	.042	.095	.125	.242	.205	.136	.106	1.000
	6	.015	.038	.076	.136	.197	.239	.170	.129	1.000
	7	.008	.019	.064	.038	.114	.193	.356	.208	1.000
	8	.008	.008	.023	.053	.087	.117	.212	.492	1.000

*The Fishbein Model as used here refers to a linear compensatory model of the form $A_j = \sum_{i=1}^{7} w_i B_{ij}$, where w_i represents the weight attached to attribute i, B_{ij} is the individual's belief about choice alternative j on attribute i, and A_j is the Fishbein attitude score for alternative j.

suggest a greater likelihood that there would be brand switching than would be anticipated for small z-values resulting from similar attributewise rankings of the alternatives. The distribution of z-values among the subjects is shown in Table 3. The average z-score was 2.6. Thus, z provides a measure of the expected level of switching across brands in the product category. In his theory of stochastic preference, Bass [2] recognized the differential level of switching for various product categories by introducing a brand loyalty factor. In this context, z serves as a similar measure.

Choice Predictions

While the multiattribute models provide reasonably good predictions of the stated preference order for brands a much harder test is the ability to predict actual choice. As noted above, it may be unrealistic to expect an attitude model to be able to predict choice accurately on individual purchase occasions. Rather, what is needed is a model which can provide accurate predictions of the probabilities of choice. These probabilities reflect the relative frequency with which a consumer will purchase any one of the brands in the market.

Many different estimation procedures could be used for predicting choice probabilities. Five separate models which can be implemented at the individual level and do not require parameter estimation are evaluated here.[7] The first

Table 3
Distribution of z Scores

< 1.5	.0%
1.5 - 2.0	8.3
2.0 - 2.5	38.2
2.5 - 3.0	35.4
3.0 - 3.5	13.8
3.5 - 4.0	4.3
> 4.0	0.0
	100.0%

naïve model assumes only knowledge of the number of brands available and assigns equal probabilities to the choice alternatives. The second model assigns probabilities equal to each brand's market share. This amounts to assuming that, in the absence of any information about individual's preferences one would predict that a consumer is likely to choose each brand in proportion to its popularity, i.e., market share. The third model uses information obtained directly from each subject about his relative preference for the brands. After stating his most preferred brand, the subject was asked to state how much extra he would pay to buy a six pack of his most preferred brand instead of his next most preferred brand. This procedure generated a dollar metric value which, when compared with the price of the preferred brand, provides a measure of the relative strength of preference for each of the brands. The fourth model used the Fishbein-type model's attitudinal scores scaled to sum to one as an estimate of the choice probabilities. Finally, the Blin-Dodson model was used to estimate the brand choice probabilities. The estimation of the density function of S was carried out through the procedure described in the Model Validation section. Each strict ordering of the brands compatible with the subjects' attributewise orderings was assigned equal probability. Iteratively sampling these orderings and solving the resulting assignment problem provided an estimate of the density function of S which in turn yields probability estimates. One hundred iterations were made to estimate each subject's choice probabilities.

Each model's prediction was compared to the subjects' actual frequency of choice revealed over the twelve purchase occasions. Four separate criteria were used to evaluate the accuracy of the models' predictions. The average value for each model's predictions across all subjects is reported in Table 4. The dollar metric and the Blin-Dodson model consistently fare best on each criterion. The first two measures, rank order correlation and product moment correlation, merely provide a measure of the degree of correlation between the actual and the predicted frequency of choice without providing any comparison of the magnitude of the difference. The latter two measures take into account the magnitude of the difference.

The real measure of performance for each model lies not in its average score but in its ability to consistently provide better predictions of individual choice probabilities. Each model's predicted choice probabilities were compared to each consumer's observed frequency of choice. The number of times each model fared best on each criteria is shown in Table 5. The Blin-Dodson model outperformed the others on all but one criterion. In terms of rank order correlation, the dollar metric measure fared best; whereas on all other criteria, product moment correlation, average absolute deviation,

Table 4

Average Value for Specified Measure Across 264 Subjects for Each Model

	Measure			
Model	Rank Order Correlation*	Prod.Moment Correlation**	Avg.Absolute Deviation#	Information Measure##
Equal Probabilities	.597	.000	.142	24.227
Market Share	.517	.456	.121	18.355
Dollar Metric	.674	.565	.112	8.120
Fishbein	.590	.505	.127	19.666
Blin-Dodson	.596	.533	.115	4.005

*Spearman's rank order correlation between predicted ranking and ranking based on frequency of choice over the twelve purchase occasions.

**Pearson's product moment correlation between predicted probability and actual frequency of choice.

#Average absolute deviation between the predicted brand choice probability and the actual frequency of choice.

##An information statistic $E = -2N \sum_{i=1}^{m} f_i \ln (\hat{p}_i/f_i)$ where $\hat{p}_i$ is the predicted choice probability, f_i is the actual frequency of choice, N is the number of observations, and m is the number of choice alternatives.

and the information measure the Blin-Dodson model more accurately predicted the subjects' actual frequency of choice.

Table 5

Number of Individuals For Whom Model Was Best, Given Specified Measure

Model	Measure: Rank Order Correlation	Measure: Prod. Moment Correlation	Measure: Avg. Absolute Deviation	Measure: Information Measure
Equal Probabilities	30	4	10	1
Market Share	26	46	40	22
Dollar Metric	100	86	89	89
Fishbein	43	29	23	9
Blin-Dodson	65	99	102	143

Conclusion

The relevance of attributes and the importance of measuring and evaluating them for predicting choice is clear. What has remained obscure is a way of linking consumer attitudes to choice. It has been argued here that the ultimate criteria for models of choice should be actual choice rather than stated preference. Stochastic models of buyer behavior have been proposed to understand and analyze actual purchasing data. The aim of this paper has been to provide a basis on which to relate attitudinal measures to choice behavior, thus providing a behavioral basis for the stochastic models of buyer behavior.

Whatever the cause for brand switching, the multiplicity of attributes, situational influences, or uncertainty in the choice process, a model of choice should provide information about a consumer's strength of preference for each of the market brands. One measure of the consumer's preferences is his choice probabilities. The model presented here provides good estimates of these probabilities when compared with four alternative models. The model links consumer attitudes and choice and provides a basis for reconciling some of the differences in two important but divergent research traditions. Also, by gaining an understanding of the relationship between consumers' attitudes toward a brand and the behavior of that brand in the marketplace, marketing management can do a better job of coordinating its marketing program.

Footnotes

1. This aggregation requires an assumption about the distance between alternative rankings on each attribute. The model assumes that all rankings are equidistant from one another. This assumption is much weaker than the assumption made in expectancy-value models, that consumers can judge the distances between alternatives and that these distances can be measured by a Likert scale.

2. The nature of this solution is more fully discussed by Bernardo and Blin [9].

3. This is not to say that such influences as advertising, pricing, etc. should not be included directly in multiattribute models of brand choice where possible.

4. This assumption is equivalent to the Laplace criterion in games against nature. Without additional information the assumption of equal probability represents maximal uncertainty. Given a commonly accepted measure of information [20] this assumption is equivalent to maximizing entropy.

5. These measurements were taken at the start of the third week of the experiment.

6. There is evidence in the literature that differential weighting doesn't improve and may even hurt model predictions [6]. Beckwith and Lehmann [7] concluded that using weights estimated for each individual, even perfectly measured, provided only modest improvements in predicting preference. They noted that the variance of individual perceptions of the choice alternatives tended to be larger for the attributes judged to be more important. Equal weights were assigned to each attribute to avoid inflating the relative importance of the attributes and to avoid confusing the influence of differential weighting with other issues.

7. A set of models which require parameter estimation were used by Pessemier, Burger, Teach, and Tigert [18] to predict the relative frequency of brand purchase. The mean absolute deviations they report in many cases exceed those reported in Table 4.

References

1. Axelrod, Joel N. (1968), "Attitude Measures That Predict Purchase," Journal of Advertising Research, Vol. 8, pp. 3-18.

2. Bass, Frank M. (1974), "The Theory of Stochastic Preference and Brand Switching," Journal of Marketing Research, Vol. 11 (February), pp. 1-20.

3. Bass, Frank M. (1975), "Analytical Approaches In The Study of Consumer Purchase Behavior and Brand Choice," Paper No. 520, Institute For Research in Economics and Management, Purdue University.

4. Bass, Frank M., Edgar A. Pessemier, and Donald R. Lehmann (1972), "An Experimental Study of Relationships Between Attitudes, Brand Preference, and Choice," Behavioral Science, Vol. 17, No. 6 (November), pp. 532-541.

5. Bass, Frank M. and W.W. Talarzyk (1972), "An Attitude Model For The Study of Brand Preference," Journal of Marketing Research, Vol. 9 (February), pp. 93-96.

6. Bass, Frank M. and William L. Wilkie (1973), "A Comparative Analysis of Attitudinal Predictions of Brand Preference," Journal of Marketing Research, Vol. 10, pp. 262-269.

7. Beckwith, Neil E. and Donald R. Lehmann (1973), "The Importance of Differential Weights in Multiple Attribute Models of Consumer Attitude," Journal of Marketing Research, Vol. 10, pp. 141-145.

8. Belk, Russell W. (1975), "Situational Variables and Consumer Behavior," Journal of Consumer Research, Vol. 2, No. 3 (December), pp. 157-164.

9. Bernardo, John and J.M. Blin, "A Mathematical Model of Consumer Choice Among Multiattributed Brands," Working paper, Graduate School of Management, Northwestern University, 1975, revised 1976.

10. Blin, Jean-Marie and Joe A. Dodson (1977), "A Multiple Criterion Decision Model For Repeated Choice Situations," in S. Zionts, ed., Multiple-Criteria Problem Solving: Theory, Methodology, Practice, New York: Springer-Verlag.

11. Blin, Jean-Marie (1976), "A Linear Assignment Formulation of the Multi-Attribute Decision Problem," Revue Francaise d'Automatique d'Informatique et de Recherche Operationnelle, Vol. 6, No. 2 (June).

12. Brown, George (1952-53), "Brand Loyalty - Fact or Fiction?" Advertising Age, Vol. 23 (June 9, June 30, October 6, December 1, 1952), Vol. 24 (January 25, 1953).

13. Fishbein, Martin (1967), "A Behavior Theory Approach to the Relations Between Beliefs About an Object and the Attitude Toward the Object," in Martin Fishbein, ed., Readings in Attitude Theory and Measurement, New York: Wiley, pp. 389-99.

14. Green, Paul E. and Yoram Wind (1973), Multiattribute Decisions in Marketing, Hinsdale, Illinois: The Dryden Press.

15. Lancaster, K.J. (1966), "A New Approach to Consumer Theory," Journal of Political Economy, Vol. 74, pp. 132-157.

16. Massy, William F., David B. Montgomery, and Donald F. Morrison (1976), Stochastic Models of Buyer Behavior, Cambridge, Massachusetts: M.I.T. Press.

17. Pekelman, Dov and Subrata K. Sen (1974), "Mathematical Programming Models for the Determination of Attribute Weights," Management Science, Vol. 20 (April), pp. 1217-1229.

18. Pessemier, Edgar A., Philip Burger, Richard Teach, and Douglas Tigert (1971), "Using Laboratory Preference Scales to Predict Consumer Brand Purchases," Management Science, Vol. 17 (February), pp. 371-385.

19. Rosenberg, Milton J. (1956), "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, Vol. 53 (November), pp. 368-72.

20. Shannon, Claude E. and Warren Weaver, The Mathematical Theory of Communication, The University of Illinois Press, Urbana, 1964.

21. Slovic, P. and S. Lichtenstein (1971), "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment," Organization Behavior and Human Performance, Vol. 6, No. 6 (November), pp. 649-744.

22. Wilkie, William L. and Edgar A. Pessemier (1975), "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, Vol. 10 (November), pp. 428-441.