DENNIS H. GENSCH and WILFRED W. RECKER*

The authors argue that for the cross-sectional multiattribute approach to choice modeling, the multinomial logit is theoretically and empirically superior to the more commonly used regression approach. Other choice methodologies also are discussed briefly in relation to logit. The difference between individual level (where regression is appropriate) and cross-sectional analysis is recognized. Most marketing managers, because of their research goals, will be using a cross-sectional approach. The derivation of the logit from an underlying behavioral model of choice is illustrated. It is this underlying behavioral model of choice that provides logit with several conceptual advantages for modeling a multiattribute choice structure.

The Multinomial, Multiattribute Logit Choice Model

The concept of evaluating a decision, product, or service as a function of its attributes is a rather universally accepted approach which has been implemented in such fields as economics (Fishburn, 1967, 1968; McFadden, 1973; McGuire and Weiss, 1976; Theil, 1970), engineering (Gustafson et al., 1971; Huber 1968; Turban and Metersky, 1971), finance (Slovic et al., 1972), medicine (Huber et al., 1969), and social psychology (Dawes, 1971; Fishbein, 1972, MacCrimmon, 1973; Rosenberg, 1956). Readers need only refer to an extensive review article (Wilkie and Pessemier, 1973) and past issues of this journal to see the prominence of the multiattribute approach in current marketing research. Marketing practitioners are now attempting to implement the multiattribute approach for their specific problems.

The goals of the research determine whether an individual or cross-sectional multiattribute analysis is appropriate. Major uses of the individual approach involve trying to understand the information processing done by particular individuals making a decision (Bettman et al., 1975) or predicting the actual choice of particular individuals (Keeley and Doherty, 1972; Slovic et al., 1972). Because the analysis is by individual rather than by alternative, regression has been successfully and correctly used for situations in which the number of alternatives exceeds the number of attributes.

Marketing practitioners usually are more interested in statistics indicating group tendencies or preferences than in sets of unique statistics for each individual. The desire for aggregate statistics implies cross-sectional rather than individual analyses. When they are attempting to predict the choice distribution of a population or when their interest centers on the diagnostic information about an attribute's relative influence on preference for the *total population,* cross-sectional analysis is appropriate. However, when analyzing the crucial diagnostic information in a cross-sectional multiattribute study in which the dependent variable is a rank or intervally scaled measure of the preference for an alternative, the practitioner observes that the statistical technique most used in marketing research, often with reservations, is regression (Bass and Wilkie, 1973; Sheth and Talarzyk, 1972). One thus might conclude that regression, despite its limitations, is perhaps the best or only technique available.

The authors attempt to show that, for the purpose of deriving diagnostic information from multiattribute hypotheses, a reasonable model is one in which individuals compare pairs of choice alternatives on the basis of their *perceived differences* in satisfaction on

*Dennis H. Gensch is Professor, School of Business Administration, University of Wisconsin at Milwaukee. Wilfred W. Recker is Visiting Senior Analyst, Transportation and Urban Analysis Department, Research Laboratories, General Motors Corporation.

Journal of Marketing Research
Vol. XVI (February 1979), 124–32

those attributes possessed by all alternatives adjusted by absolute levels of satisfaction on attributes specific to one alternative. For example, consider the comparison of two grocery stores. Both have the same set of attributes with the one exception that store 1 does not have a meat counter and store 2 does. For the attribute "quality of meat" an absolute value is added to the preference score of store 2. The rest of the preference score for store 2 and the entire preference score for store 1 consist of perceived differences in satisfaction on the common attributes. Such a conceptualization considers multiattribute model specfications in which evaluations between alternatives are treated as differences between attribute scores (Bass and Wilkie, 1973). The method also allows each individual in the sample to have a *different* set of available or feasible alternatives.

The purpose of this article is to illustrate the usefulness of the logit model in a real-world situation in which a set of consumers are asked to rate the grocery stores they frequent most. As in many real-world studies, the researcher is faced with the prospect of individuals having different sets of stores (alternatives) from which they make a choice; also, each of the stores can have a different set of relevant attributes. Because the logit model is derived from an underlying behavioral model of choice, the logit is suited to choice situations that present problems for other covariance techniques such as regression, discriminant analysis, and factor analysis which are often used in choice modeling. In addition, the authors argue that, because of the underlying choice model from which the logit is derived, the diagnostic information such as the elasticities of the attributes provides more meaningful measures of reality than the diagnostic information from other covariance choice models.

In the following sections, the logit model's derivation from an underlying behavioral model of choice is presented briefly, the basic approaches to choice modeling are categorized and logit is discussed in relation to other covariance approaches, the actual application of the logit to real-world data and a comparison with a regression model on the same data are provided, and finally an interpretation of the diagnostic information provided by the logit model is discussed. In conclusion, a summary is given of the basic features of the logit model that make it particularly well suited for a covariance analysis of the choice process with cross-sectional data.

Multinomial Logit Method

One can hypothesize that an individual decision maker's overall preference or ranking of a choice alternative is a function of the utility which the alternative holds for the individual.

One also can hypothesize that an individual's utility for an alternative is separable into two components: (1) a deterministic component measured in terms of expressed attitudes toward that alternative, and (2) an unobserved random component. The source of the deterministic component is consistent with rational choice theories in psychology and economics (Manski, 1973) and requires no elaboration here. The source of the random component, however, is different from the "repeated measurements" source of error in psychological studies (Bock and Jones, 1968). Manski (1975) lists "omitted structure," the need to make measurement in terms of proxy or "instrumental variable," and "cross-sectional preference variation" as logical bases for such a concept (which he refers to as "stochastic utility").

In essence the assumptions underlying the error term in logit analysis reflect the complexity and richness of the choice process by recognizing that a model of the choice process seldom will be fully specified in terms that can be measured accurately and which identify all of the current and historical attributes that really influence the choice process. ***In reality, most models of choice are underspecified and this fact should be taken into account in the analysis.***

Utility is thus written

$$U_i^k = V_i^k + \varepsilon_i^k \tag{1}$$

where:

U_i^k = the utility of alternative k to individual i,
V_i^k = the deterministic component, and
ε_i^k = the random component, which is assumed to be independent and identically distributed across all individuals i.

Consistent with the majority of other multiattribute models (Turban and Metersky, 1971), a form for the deterministic component which is linear and additive in terms of the attribute scores can be assumed:

$$V_i^k = \sum_{j \in S^k} a_j^k x_{ij}^k + \sum_{j \in S} b_j x_{ij}^k \tag{2}$$

where:

x_{ij}^k = score given by individual i to the k^{th} alternative on the j^{th} attribute,
a_j^k = utility weight reflecting the importance of the j^{th} attribute defined uniquely for the k^{th} alternative,
b_j = utility weight reflecting the importance of the j^{th} generic attribute defined consistently for all alternatives,
S^k = set of attributes relevant to alternative k, which are not common to all other alternatives,
S = set of attributes common to the description of all available alternatives.

It is postulated that an individual will *prefer* the choice alternative perceived to have the greatest utility. The probability that individual i will prefer alternative k from a set of available alternatives A_i, denoted by $P_i(k\colon A_i)$, consequently can be written in utility terms as

(3) $P_i(k:A_i) = \text{Prob}\{U_i^k > U_k^l \quad \text{for all } l \in A_i, l \neq k\}.$

In light of the division of utility in expression 1 into random and deterministic components, the preference probability can be rewritten

$$(4) \qquad P_i(k:A_i) = \text{Prob}\{(V_i^k + \varepsilon_i^k) > (V_i^l + \varepsilon_i^l) \quad \text{for all } l \in A_i, l \neq k\}$$

or

$$(5) \qquad P_i(k:A_i) = \text{Prob}\{(V_i^k - V_i^l) > (\varepsilon_i^l - \varepsilon_i^k) \quad \text{for all } l \in A_i, l \neq k\}.$$

This specification of choice probability in terms of utility differences for each individual automatically takes care of the scale differences problem reflected by differences in scale means across individuals.

If it is assumed that an individual, in making decisions so as to maximize the overall utility of his or her choice, is concerned with maximum values of the unobserved variables contained in the random component of utility, as well as with such values in the deterministic component, then Gumbel (1954) has shown that for a class of distributions (including the normal distribution) the random terms ε_i^l are independently identically distributed with the Weibull (Gnedenko extreme value) distribution, i.e.,

$$(6) \qquad \text{Prob}\{\varepsilon_i^l \geq w\} = \exp(-\exp(-w)).$$

With the assumption represented by equation 6, it is easily shown that the probabilistic choice model, equation 5, takes the form:

$$(7) \qquad P_i(k:A_i) = \exp(V_i^k) \Big/ \sum_{l \in A_i} \exp(V_i^l).$$

A detailed derivation of this type of strict utility model, called the multinomial logit model, is provided by Thiel (1969) and discussions of the rationale for and properties of such models are provided by McFadden (1968, 1972, 1973) and Thiel (1971).

A specific structure of perceptual comparison among choice alternatives is implied by the multinomial logit model. To better visualize this structure, rewrite expression 7 as

$$(8) \qquad P_i(k:A_i) = 1 \Big/ \left(1 + \sum_{\substack{l \in A_i \\ l \neq k}} \exp(V_i^k - V_i^l)\right),$$

where V_i^k and V_i^l are defined in expression 2.

The a_i^k and b_i attribute rating coefficients in the multinomial logit choice model can be estimated by using maximum likelihood (McFadden, 1968) or maximum score (Manski, 1975) techniques. McFadden (1968) has demonstrated that the maximum likelihood estimators are consistent, asymptotically efficient, and are unique under very general conditions. Furthermore, the asymptotic normality of these estimators allows the use of t-statistics for asymptotic tests of coefficient significance. Currently several user-oriented maximum likelihood multinomial logit computer programs are available, such as the one used in the present research (Cambridge Systematics, Inc., 1974).

Logit Related to Other Techniques

There are several basic approaches to choice modeling. Lexigraphic (Russ, 1971) and hierarchical (Recker et al.) approaches attempt to consider the sequence in which attributes enter the decision process. Multidimensional scaling attempts to map the entire preference space (Lehmann, 1971). Choice usually is determined in terms of a metric distance from an "ideal point." Recently linear programming has been suggested as means of relating alternatives within the preference space (Pekelman and Sen, 1974; Shocker and Srinivasan, 1974). A third basic approach is a covariance approach that attempts to relate attribute scores to preference measures of the alternatives. The most commonly used of these techniques are regression, multiple discriminant analysis, and logit and probit analysis. A comprehensive discussion of the relative merits and the situational appropriateness of the basic approaches would be worthwhile, but is beyond the scope of this article. Rather than compare the covariance approach with alternative forms of choice modeling, the authors briefly differentiate the logit approach from the other commonly used covariance techniques.

The basic structure of the multinomial logit model can be used for a number of purposes; for example, a logit formulation has been shown to be one method of providing "logically consistent" explanatory variables for an econometric model of market share where the explanatory variables were product ratios (McGuire and Weiss, 1976). However, the logit approach is particularly well suited for choice models. Within this general framework, the authors argue that logit is the most appropriate technique for cross-sectional multiattribute modeling using a covariance approach. Logit and probit are virtually identical for a dichotomous dependent variable; logit can be expanded to more than two values on the dependent variable, whereas probit is strictly limited to a two-value dependent variable. Recently approximation techniques for expanding probit to multiple choice situations have been proposed (Lerman and Manski, 1976). The computer codes are complex and not generally available; furthermore, it is not clear whether the more general form of the error term in the multinomial probit will actually lead to substantially different or better predictions than those of the logit.

Multiple discriminant analysis (MDA) is a fundamentally different type of model in that it is a *classification* rather than a choice model. The other techniques basically assume one population making choices based on their evaluation of the independent variables. MDA assumes *several distinct populations,* each having a different pattern of scores on the independent variables. Further distinctions between the fundamentally very different multinomial logit and MDA approach are well articulated for the two-group case by Westin and Watson (1975, p. 283). Extensions of their arguments to the general multigroup case are straightforward.

Gensch et al. (1975) give a detailed discussion of four reasons why regression is conceptually incompatible with cross-sectional multiattribute data. The following empirical section shows that because the multinomial logit model is conceptually a statistical technique derived from an underlying behavioral model of choice, it has predictive and diagnostic superiority over a regression model for the data set used.

EMPIRICAL RESULTS: LOGIT VS. REGRESSION

The empirical results presented for model comparison are related to choice of store for grocery shopping. The same data have been analyzed by using different formulations of regression and logit models. Gensch et al. (1975) present more than one formulation of both the regression and logit model. Here, one formulation of each model is considered sufficient to convey the differences in empirical results. Though comparison of model performance is emphasized, the model results also offer insights into the store selection process underlying most brand choice and product purchase models. Researchers working on problems of estimating brand or product purchasing decisions have stated that considerably more work must be done on understanding store preferences *before* meaningful work on brand or product purchasing decisions can proceed (Carman, 1970; Rao, 1969).

The data analyzed were collected from a mail survey sent to a random sample of 1500 households in Buffalo, New York. The return rate was 22.5%. For each household, descriptions of grocery stores visited most often and attitudes toward these stores were elicited. The attitudinal data included satisfaction ratings on a set of prespecified attributes and evaluations of up to four frequently visited stores with respect to each of the attributes. The attitudinal data were in the form of 7-point semantic differentials. The respondent also was asked to indicate the frequency of shopping at each store mentioned and the time required for the trips to each store (as well as the return travel time, if different). Attribute ratings were elicited by asking the respondents to mark a 7-point scale with descriptors ranging from "excellent" to "extremely poor" for each store on each attribute. The attributes are listed in Table 1.

The grocery stores were classified according to brand name for nine major chain supermarkets. In addition two categories, "small independent market" and "neighborhood store," were used to classify other "nonchain" stores.

Table 1
LIST OF ATTRIBUTES

1.	Easy to find a parking spot
2.	Easy to get home from store
3.	Easy to get to store from work
4.	Easy to get to store from home
5.	It is near other shops I use
6.	They have convenient hours
7.	They have reasonable prices
8.	They have a good variety of items
9.	The meat quality is good
10.	The produce quality is good
11.	Easy to find things in store
12.	Store has large selection of goods
13.	Store accepts credit cards
14.	It is easy to cash checks there
15.	Easy to return or exchange goods there
16.	Store has items others don't

Regression Model

Several regression models were estimated to serve as bases for comparison with the logit estimations. These estimations were obtained by using the "screen" regression approach of Furnival and Wilson (1971) which finds the optimal subsets of explanatory variables for increasing numbers of explanatory variables *without* the constraint that the optimal subset at any level contain the optimal subset at the next lowest level as in stepwise regressions. It also lists the R^2 values for the "next best" subsets which usually reveal the somewhat arbitrary nature of the optimal subset. The decision of how many variables to include in the model was made on the basis of adjusted R^2 values. For cases in which the adjusted R^2 were virtually identical, the t-statistic was used as a criterion and the set in which the t-values were most significant was chosen.

The dependent variable in the regression models (and also in the logit models) was related to the actual frequencies of trips to the stores. The reason for selection of this dependent variable is twofold. First, it is well known that the relationships found in relating attributes to consumer intentions are often very different from the relationships found in relating the same attributes to *actual* decisions. Second, the frequencies contain more interval information about the store visits than does a rank order. Other formulations of the models using alternative dependent variables, such as dollars spent per store, could be analyzed in a manner analogous to that used for the trip frequency dependent variable.

In the present case a regression model using a dichotomous dependent variable has a substantially better R^2 value than alternative regression formulations tested (Gensch et al., 1975). This regression formulation is also similar in structure and thus most directly comparable to the logit model. In the questionnaire, each respondent gave frequencies for up to four stores; thus, for each respondent and for each pair of stores with different trip frequencies reported (up to $(4 \cdot 3)/2 = 6$ pairs per respondent), a 1-0 dependent variable was created where the 1 was assigned to the store with the higher frequency and 0 to the other. The independent variables consisted of differences between the satisfaction ratings for the two stores on the attributes of Table 1 and between the perceived times required to get to the stores. The model thus has the form of

$$(9)\quad y_{ik} \text{ is } \begin{Bmatrix} 1 \text{ if } k > l \\ 0 \text{ if } k < l \end{Bmatrix} = \beta_1(x_{k1} - x_{l1}) + \beta_2(x_{k2} - x_{l2}) + \ldots + \beta_n(x_{kn} - x_{ln}) + \alpha$$

where k and l are the shopping frequencies associated with alternatives k and l. Table 2 shows the results for the model defined in equation 9 for a subset of five variables and, though low, the R^2 values obtained were significantly greater than those obtained in the previous two regression models.

Unfortunately, the beta coefficients (diagnostic information) really should not be interpreted because the dichotomous dependent variable forces the error term to violate two of the assumptions underlying the linear model. As Goldberger (1964, p. 249) points out, it is inherent in the model using a dichotomous dependent variable that the classical assumption of homoscedasticity is violated. More important, Thiel (1971, p. 628–30) illustrates that in order for the error term to have an expected value of zero it would have to take on specific values with probabilities greater than one or less than zero. Thus, though one may use the foregoing model for predictive purposes (R^2), an interpretation of the beta coefficients is not advisable because they were generated in a manner that violates assumptions underlying the linear model.

Table 2
REGRESSION MODEL

$R^2 = .282$
Adjusted $R^2 = .271$

Variable	Coefficient (B)	Coefficient (beta)	t
Perceived time to store	−.109	−.325	−5.99
Easy to get to store from home	.046	.146	2.62
They have convenient hours	.047	.126	2.68
Easy to find things in store	.042	.115	2.44
Easy to return or exchange goods there	.048	.136	2.32
Constant	.460		

Multinomial Logit Models

The utility weights (i.e., a_j) in the deterministic component of utility specified by equation 2 were estimated by using maximum likelihood techniques for the probabilistic choice model in the form of equation 8. Because in this application all stores were evaluated on the same set of attributes, the sets of alternative-specific attributes, S^k, for all k, are null sets. The dependent variable was the probability that an individual would shop *most* frequently at store chain k. The set of relevant alternatives for each individual consisted of only the stores where the individual actually shopped. Because this logit formulation uses differences in attribute perceptions, observations where the individual shopped at only one store, only at stores of the same chain, or at stores belonging to different chains but with the same frequency were not used in the model estimation. These restrictions and the withholding of some observations because of missing data reduced the sample size to just over 100 individuals.

Presented in Table 3 are the b_j coefficients for all attributes j for which the coefficients are significantly different from zero at the 95% confidence level; all attributes with coefficients not significantly different

Table 3
LOGIT MODEL 1
(SAMPLE SIZE = 96)

Variable	Coefficient estimate	t
Perceived time to store	−1.486	−4.23
Easy to return or exchange goods there	0.842	2.57
Easy to get to store from home	1.336	2.92
Easy to find things in store	0.546	2.29
The meat quality is good	0.342	1.93

Ratio of choices predicted correctly = 0.87.
Ratio of individuals predicted correctly = 0.81.

Store	Percent of choices store was used most frequently	Prediction ratio
1	6.2	0.89
2	21.2	0.77
3	2.7	1.00
4	9.6	0.93
5	2.1	1.00
6	9.6	0.86
7	4.1	0.83
8	21.9	0.94
9	2.1	0.67
10	13.0	0.84
11	7.5	0.91

from zero were excluded in the estimations. The coefficient values are listed in order of their t-statistics.

Although the dichotomous regression model did not yield interpretable diagnostic data, its goodness-of-fit was significantly better than that of other regression formulations tested. The authors therefore attempt to relate the goodness-of-fit achieved by the logit model to that of the dichotomous regression model.

Because the R^2 coefficient of determination goodness-of-fit measure is inappropriate for nonlinear models, reliance must be placed on other goodness-of-fit indices to validate model results. The result of a chi square test (χ^2 = 84.4 for 5 d.f.) offers firm rejection of the joint null hypothesis that all coefficients of the logit models presented herein are equal to zero. In addition, two measures based on the percentage of actual choices predicted "correctly" by the model are provided.

The first measure, the ratio of choices predicted correctly, is the proportion of times the individuals' predicted probability of the chosen alternative was greater than the predicted probability of a nonchosen alternative. As is shown in Table 3, the model correctly predicted the choice between the observed chosen alternative and a nonchosen alternative in 87% of such choice pairings. This same measure of performance, when computed for the dichotomous regression model, was only 71%. Thus the logit provided a 22.5% (16/71) improvement in prediction. This measure also was disaggregated by alternative chosen and the results are shown in Table 3. In this case the disaggregate prediction ratios indicate a uniformly high predictive power across all alternatives. In practice, the finding of a specific alternative with a relatively high frequency and relatively low predictive ratio often will suggest to the researcher a unique attribute of the alternative currently unspecified by the logit model.

The second measure, the ratio of individuals predicted correctly, is the proportion of individuals for which the predicted probability of the chosen alternative was greater than that of every other relevant alternative. The results in Table 3 indicate that the model correctly predicted the total choice set of 81% of the sample. In fact, even when restricted to only those variables included in the regression model, the logit model maintains its high prediction ratios, predicting 87% of the choices and 80% of the individuals correctly. These results again indicate 22.5% and 27% improvements in prediction by the logit. This performance is all the more impressive in light of the insignificance of the coefficient associated with the rating of one of the predictor variables, parking facilities ($t = 0.61$).

Management policy questions relating to what impact a change in an independent variable can be expected to have on the probability of a particular store being the most frequently visited receive considerable guidance from an examination of the elasticities and cross-elasticities associated with the predictor variables for each alternative. An elasticity is a dimensionless quantity defined by economists as the percentage change in the dependent variable which would result from a 1% change in an independent variable:

$$E_{ij}^{kl} = \frac{\partial P_i(k:A_i)}{\partial X_{ij}^l} \frac{|X_{ij}^l|}{P_i(k:A_i)} \tag{10}$$

where:

E_{ij}^{kl} = elasticity of the probability that alternative k is chosen from set A_i by individual i with respect to variable j of alternative l.

By calculating the deviations of the choice probabilities in equation 7 and substituting into equation 10, one can write the individual elasticities of the probability of choice as

$$E_{ij}^{kl} = |X_{ij}^l| \sum_{q\epsilon A} P_i(q:A) \frac{\partial V_i^k}{\partial X_{ij}^l} - \frac{\partial V_i^q}{\partial X_{ij}^l} \tag{11}$$

The elasticities for $l = k$ are referred to as direct elasticities, or simply elasticities, because they represent the changes in probability of choosing an alternative with respect to perceptions toward the attributes of the alternative itself; the elasticities for $l \neq k$ are referred to as cross-elasticities, and can be interpreted as the sum of the changes in the probabilities of choosing any of the other stores with respect to changes in perceptions. The distribution of these changes in probabilities among the remaining stores is not determined.

Aggregate elasticities can be developed from the individual elasticities to measure the overall sensitivities of choice probabilities to uniform percentage changes in explanatory variables for all individuals. It is easily shown that these aggregate elasticities are given by

$$E_j^{kl} = \frac{\sum_{j=1}^{M} [P_i(k:A_i)\, E_{ij}^{kl}]}{\sum_{i=1}^{M} P_i(k:A_i)} \tag{12}$$

where:

E_j^{kl} = aggregate elasticity of the probability of choosing alternative k with respect to uniform changes in the perceived evaluation of alternative l on attribute j for all $i = 1, ..., M$ individuals.

As can be seen in the foregoing formulas, the logit model implies that the magnitude of the effect from the particular predictor variable depends on the level of the dependent choice variable and not solely on the estimated coefficients. This is a more realistic representation of the choice process than the regression approach which ignores the level of the variables (Gensch et al., 1975).

Table 4 provides the actual elasticities associated with store 2 (the major supermarket chain in the survey area) and store 6 (a chain of "discount" supermarket/department stores typically located in the suburban fringes of a city).

A first observation is that the results are intuitively satisfying for several reasons which include: knowledge of the relative locations, store 2 has a typical grocery store inventory layout whereas store 6 has a mixture of food and nonfood goods in the same floorspace, and returning goods at store 2 is relatively quick at an office adjacent the checkout counters whereas returns at store 6 are accomplished at a centralized "customer service" counter and require the completion of forms detailing the return.

The elasticities indicate that store 6 will increase its probability of becoming the most frequented store more than store 2 will, for the same percentage increase in any of the five attributes listed in Table 4. This pattern again appears to be a realistic representation of the choice environment. Table 3 indicates that currently store 2 is the first choice of 21.2% of the sample compared with 9.6% for store 6. Obviously, it should be easier to get a 10% improvement on a current level of 9.6% than on a current level of 21.2%. As previously noted, the logit elasticities, in contrast to the regression elasticities, take into account the current level of the dependent variable and thus in the authors' opinion provide the model user with more realistic information. From *a priori* knowledge of the choice situation, the logit model appears to be providing reasonable results.

A second observation is related to the influence of a 10% increase or decrease in the ease of getting to the store from home. A 10% increase for store 2 is associated with a 9% increase in the probability store 2 will be most frequently visited; similarly, a 10% increase for store 6 is associated with a 19% increase in the probability. However, the cross-elasticities indicate a 10% decrease is associated with a 60–70% reduction in the probability of being visited most frequently. In a wide variety of regression and logit runs the ease of getting to the store was a significant independent variable. Indeed, this finding is generally provided as empirical justification for gravity models for assessing new store location potential. The foregoing elasticity results suggest that the attribute "ease of access" and, to a lesser degree, the attribute "perceived time to store," are very important in determining whether the store is in the feasible set an individual will consider visiting, *but* these two attributes in relation to other attributes are less important in determining store choice within the feasible set.

This finding seems to indicate that consumers tend to treat travel time as a satisficing criterion. If travel time is within a certain tolerance the store is considered in the decision set; the decision as to which store within the decision set to frequent seems then to be reached on the basis of other variables. Thus gravity models would be most efficient when there are minimum competitive store locations within the same threshold travel time overlay. Use of a linear function of time or distance to closest store stocking each brand as a means of attempting to account for differences in availability prior to an analysis of brand purchase data is probably an incorrect procedure.

To derive specific managerial implications in terms of actions that a manager of store 6 could take to make his store more competitive with store 2, one would need an estimate of the costs and savings involved in changing the samples' perception of a given attribute by one rating unit from its current level. Shocker and Srinivasan (1974) provide a good discussion of techniques for managerially operationalizing a consumer perception space. For example, assume that the manager of store 6 estimated that by cutting back his meat counter by 50% he could install a more convenient counter for exchanging goods. He estimates over the long run he will operate the return counter and reduced meat counter at a lower cost than the current meat counter. He estimates that this change will lower consumers' perception of meat quality by one unit but increase their perception of

Table 4
ELASTICITIES AND CROSS-ELASTICITIES FOR TWO STORES CALCULATED FROM LOGIT MODEL 1

	Store 2		*Store 6*	
Variable	*Elasticity*	*Cross-elasticity*	*Elasticity*	*Cross-elasticity*
Perceived time to store	−0.327	2.34	−0.404	1.31
Easy to get to store from home	0.922	−7.24	1.93	−6.10
The meat quality is good	0.229	−1.59	0.377	−1.23
Easy to find things in store	0.366	−2.84	0.789	−2.38
Easy to return or exchange goods there	0.532	−4.18	1.30	−3.75

return convenience by 1.5 units. The elasticities suggest that this move would be associated with an increase in store 6's probability by about 9%.

CONCLUSIONS

Multinomial logit analysis has been proposed as a more appropriate structure for the multiattribute analysis of the choice behavior of *populations* of individuals. On a theoretical basis, the multinomial logit formulation is purported to be superior to linear approaches in that: (1) it can be developed from a behavioral utility theory framework, (2) the bounds on choice are implicitly incorporated in the model form, (3) differences among individuals in terms of alternatives available to them are considered, (4) it allows individuals to consider unique attributes per alternative, (5) the elasticities of demand are consistent with accepted tests of diminishing marginal utility, threshold, and saturation, and (6) the error term recognizes that the choice model is probably underspecified.

In terms of applications, multinomial logit analysis considers the multichoice case as well as the binary choice case, and its estimation eliminates problems of scale differences and unique attribute meanings among choice alternatives. Empirically it has been shown to result in better fit than the regression approach. In addition, the diagnostic information obtained is judged to be in a form that is useful to the decision maker by being disaggregate by attribute and choice alternative.

REFERENCES

Bass, Frank M. and William L. Wilkie. "A Competitive Analysis of Attitudinal Predictions and Brand Preference," *Journal of Marketing Research,* 10 (August 1973), 262–9.

Bettman, James R., Noel Capon, and Richard J. Lutz. "Cognitive Algebra in Multi-Attribute Attitude Models," *Journal of Marketing Research,* 12 (May 1975), 151–64.

Bock, R. and L. Jones. *Measurement and Prediction of Judgment and Choice.* San Francisco: Holden Day, 1968.

Cambridge Systematics, Inc. "Multinomial Logit Estimation Package: Program Documentation," Version 2, Mod. 1, Cambridge, Massachusetts, 1974.

Carman, James M. "Correlates of Brand Loyalty: Some Positive Results," *Journal of Marketing Research,* 7 (February 1970), 67–76.

Dawes, R. M. "A Case Study of Graduate Admissions: Applications of Three Principles of Human Decision Making," *American Psychologist,* 26 (1971).

Fishbein, Martin. "The Search for Attitudinal-Behavioral Consistency," *in* Joel B. Cohen, ed., *Behavioral Science Foundations of Consumer Behavior.* New York: Free Press, 1972, 245–52.

Fishburn, Peter. "Methods of Estimating Additive Utilities," *Management Science,* 13 (March 1967).

———. "Utility Theory," *Management Science,* 14, (January 1968).

Furnival, George M. and Robert W. Wilson. "Regression by Leaps and Bounds," presented at Joint Meeting of the Institute of Mathematical Statistics and Biometrics Society at the Pennsylvania State University, University Park, Pennsylvania, April 21–23, 1971.

Gensch, Dennis H., Thomas F. Golob and Wilfred W. Recker. "A Multinomial, Multi-Attribute Logit Choice Model," working paper 197, School of Business, University of Wisconsin–Milwaukee, August 1975.

Goldberger, Arthur S. *Econometric Theory.* New York: John Wiley & Sons, Inc., 1964.

Gumbel, E. J. "Statistical Theory of Extreme Values and Some Practical Applications," *National Bureau of Standards Applied Mathematics,* Series 20, No. 33, 1954.

Gustafson, D. H., G. K. Pai, and G. C. Kramer. "A Weighted Aggregate Approach to R & D Project Selection," *AIEE Transactions,* 3 (March 1971).

Huber, George. "Multiplicative Utility Models in Cost Effectiveness Analyses," *Journal of Industrial Engineering,* 19 (March 1968).

——— V. Sahney, and D. Ford. "A Study of Subjective Evaluation Models," *Behavioral Science,* 14 (November 1969).

Keeley, S. M. and M. E. Doherty. "Bayesian and Regression Modeling of Graduate-Admission Policy," *Organizational Behavior and Human Performance,* 8 (October 1972).

Lehmann, Donald R. "Television Show Preference: Application of a Choice Model," *Journal of Marketing Research,* 8 (February 1971), 47–55.

Lerman, Steven R. and Charles F. Manski. "An Estimator for the Generalized Multinomial Probit Choice Model," paper presented at 56th Annual Meeting, Transportation Research Board, Washington, D.C., August 1976.

Luce, R. Duncan and P. Suppes. "Preference, Utility, and Subjective Probability," *in* R. D. Luce, ed., *Handbook of Mathematical Psychology,* Vol. 3. New York: John Wiley & Sons, Inc., 1965, 249–410.

MacCrimmon, Kenneth R. "An Overview of Multiple Objective Decision Making," *in* J. L. Cochrane and M. Zeleny, eds., *Multiple Criteria Decision Making.* Columbia: University of South Carolina Press, 1973.

Manski, Charles F. "A Stochastic Utility Model of Choice," unpublished paper, School of Urban and Public Affairs, Carnegie-Mellon University, 1973.

———. "Maximum Score Estimation of the Stochastic Utility Model of Choice," *Journal of Econometrics,* 3 (August 1975), 205–28.

McFadden, Daniel. "Conditional Logit Analysis of Qualitative Choice Behavior," *in* P. Zarenlka, ed., *Frontiers in Economics.* New York: Academic Press, 1973.

———. "Quantal Choice Analysis: A Survey," presented at NSF-NBER Conference on Decision Rules and Uncertainty, March 22–23, 1972, University of California, Berkeley. Unpublished paper, Department of Economics, University of California, Berkeley.

———. "Revealed Preferences of a Government Bureaucracy," unpublished paper, Department of Economics, University of California, Berkeley, 1968.

McGuire, Timothy W. and Doyle L. Weiss. "Logically Consistent Market Share Models," *Journal of Marketing Research,* 13 (August 1976), 296.

Pekelman, Dov and Subrata K. Sen. "Mathematical Programming Models for the Determination of Attribute

Weights," *Management Science*, 20 (April 1974), 1217–29.

Rao, Tanniru R. "Consumer's Purchase Decision Process: Stochastic Models," *Journal of Marketing Research*, 6 (August 1969), 321–9.

Recker, W. W., D. H. Gensch, and T. F. Golob. "SEQUEL: A Sequential Model of Choice Behavior," submitted to *Management Science*.

Rosenberg, Milton J. "Cognitive Structure and Attitudinal Affect," *Journal of Abnormal and Social Psychology*, 53 (November 1956), 368–72.

Russ, F. A. "Evaluation Process Models and the Prediction of Preference," *in* D. M. Gardner, ed., *Proceedings of the Second Annual Conference*, Association for Consumer Research, 1971, 256–61.

Sheth, Jagdish N. and Wayne W. Talarzyk. "Perceived Instrumentality and Value Importance as Determinants of Attitudes," *Journal of Marketing Research*, 9 (February 1972), 6–9.

Shocker, Allan D. and V. Srinivasan. "A Consumer-Based Methodology for the Identification of New Product Ideas," *Management Science*, 20 (February 1974), 921–37.

Slovic, P., D. Fleisnner, and W. S. Bauman. "Analyzing the Use of Information in Investment Decision Making: A Methodological Proposal," *Journal of Business*, 45 (April 1972).

Thiel, Henri. "A Multinomial Extension of the Linear Logit Model," *International Economic Review*, 10 (October 1969), 251–9.

———. "On the Estimation of Relationships Involving Qualitative Variables," *American Journal of Sociology*, 76 (1970), 103–54.

———. *Principles of Econometrics*. New York: John Wiley & Sons, Inc., 1971.

Turban, E. and M. L. Metersky. "Utility Theory Applied to Multi-Variable System Effectiveness Evaluation," *Management Science*, 17 (August 1971).

Westin, Richard B. and Peter L. Watson. "Reported and Revealed Preferences as Determinants of Mode Choice Behavior," *Journal of Marketing Research*, 12 (August 1975), 282–9.

Wilkie, William L. and Edgar A. Pessemier. "Issues in Marketing's Use of Multi-Attribute Attitude Models," *Journal of Marketing Research*, 10 (November 1973), 428–41.