A Threshold Model of Attribute Satisfaction Decisions

JAMES R. BETTMAN*

A threshold model of binary attribute satisfaction decisions is developed using Fishbein attitude model components. The model is tested using multivariate probit analysis and is supported by the data. Implications for attitude research and statistical methodology are discussed.

The purpose of this paper is to examine the information processing carried out by consumers in the course of making decisions about whether or not an attribute of a product is satisfactory to them. This type of decision is studied because it occurs frequently in decision net depictions of consumer choice processes. The literature on decision nets is reviewed in Bettman (1974a,b), so it will not be reexamined here. It is instructive, however, to examine the hypothetical decision nets shown in Figure 1. It can be seen from these examples that decision nets can involve several binary decisions about satisfaction with particular attributes. These binary choices are the subject of this paper.

Multiple-attribute attitude models also deal with information about product attributes. In particular, the Fishbein (1972) attitude model can be expressed as

\[ A_j = \sum_{i=1}^{n} a_i b_{ij} \]

where
- \( A_j \) = attitude (affect) toward some object \( j \)
- \( a_i \) = evaluation of attribute \( i \)
- \( b_{ij} \) = probability or likelihood that object \( j \) possesses attribute \( i \)
- \( n \) = number of attributes.

Studies using this model are reviewed in Wilkie and Pessemier (1973). Since these attitude models and decision net models both deal with information handling involving product attributes, it is reasonable to expect that the kinds of measures used in the attitude models, the \( a_i \) and \( b_{ij} \) measures, might be related to the processing involved in an attribute satisfaction decision. It should be noted that the measures used in both models cover up a great deal of detail of information processing. Certainly there are information manipulations on the part of the consumer underlying the binary satisfaction choices in decision nets or underlying some particular \( b_{ij} \) or \( a_i \) measure. That is, these measures summarize a great deal of processing of potential interest that is done by the consumer. Perhaps this is true because these models are often taken to be structural models, and not models of processing, particularly Fishbein models. However, to understand processing, it is important to examine these underlying manipulations in more detail. Thus, the viewpoint of this study is that since net models and Fishbein models both involve product attributes, perhaps we can learn more about the details of consumer information processing by using Fishbein type measures to model attribute satisfaction decisions, since this type of decision is very prominent in consumer decision nets.

In summary, nets use simple 0-1 branches. How do consumers determine if an attribute is satisfactory or not? Can one go deeper into the processing underlying a single branch? As argued above, one method for studying the 0-1 branches in decision nets is to model this decision in terms of the measures used in Fishbein models. In particular the \( a_i \) and \( b_{ij} \) variables (or some transformation of these variables) are used. Also, Bettman (1974a) has suggested that Fishbein variable measures might be useful for helping to determine branching in net models, since it is very difficult to obtain relatively objective measures for the 0-1 satisfaction decisions made (see Bettman, 1970 for an example of this difficulty). However, to do this one must know how to convert Fishbein measures into a 0-1 branching measure.

Thus, this paper develops a model of attribute satisfaction decisions using Fishbein-type measures in the next section. Then the methodology of the study designed to test the model is presented, and the analytical technique used, multivariate probit analysis, is outlined. Finally, implications for future research are presented.

* James R. Bettman is Associate Professor in the Graduate School of Management, UCLA.

©Copyright 1974 Journal of Consumer Research Policy Board
All Rights Reserved
A Threshold Model of Attribute Satisfaction Decisions

A MODEL OF ATTRIBUTE SATISFACTION DECISIONS

In this section a model is developed for determining when an attribute of a brand is seen as satisfactory by consumers, given measures describing that brand and others. Although the arguments are presented in terms of individuals, the analysis will later be carried out on an aggregate basis because of the relatively small number of observations available per individual.

The essential model to be used is a threshold model. An individual, when considering a brand using a decision net, must make decisions about whether or not he considers certain attributes to be satisfactory. The individual has some information, but he must decide whether or not that information is sufficient to convince him that an attribute is indeed satisfactory. Let us model this by a summary hypothetical construct called degree of conviction for each attribute. That is, the consumer has some degree of conviction that an attribute for a brand is satisfactory. If that degree of conviction is above a threshold value for that consumer, the attribute is judged to be satisfactory. If it is below the threshold, the attribute is seen as not satisfactory (a generalization of this is discussed below). The threshold value varies from individual to individual for a given attribute, and may vary from one attribute to another. Also, the degree of conviction that an attribute is satisfactory may be affected by several variables. Thus, to specify fully the threshold model, we must determine which variables should affect the degree of conviction that an attribute is satisfactory, and then develop a theory relating those variables specifically to degree of conviction—that is, theoretically expected signs for the coefficients.

Degree of conviction should be positively related to the degree of belief that the product possesses the attribute, as measured by \( b_{ij} \). This assumes that attributes are desirable and that one must believe an attribute is present before one can be satisfied with it. The assumption about attribute desirability may not hold in some cases, but should usually be true for most marketing applications.

If an attribute is undesirable, a simple scale transformation on \( b_{ij} \) can remedy the problem. An ideal point formulation such as that used by Bass, Pesse, and Lehmann (1972) causes problems unless the ideal point is at either extreme of the scale. Satisfaction with an attribute, however, is not synonymous with belief. If an individual rates an attribute very highly, if it is very "important" to him, then he may require a greater amount of "convincing" before determining that an attribute is satisfactory. This implies that the higher the value for \( a_i \), the evaluation of attribute \( i \), the lower the degree of conviction the attribute is satisfactory. Also, if there is a great deal of variation in beliefs (\( b_{ij} \)) for an attribute across the brands in a particular product category, then this variation may lead to increased feelings of risk for that attribute (Bettman, 1973). This risk should decrease the level of conviction for that attribute. Thus, the standard deviation of \( b_{ij} \) scores across brands for an attribute \( i \), \( \sigma_i \), should be negatively related to degree of conviction that attribute \( i \) is satisfactory.

If an individual is familiar with a brand, it is hypothesized that this will increase his conviction that the brand possesses all attributes satisfactorily—a "halo-effect." Familiarity is of course confounded with overall preference, and learning and reinforcement from experience for a brand, and one cannot infer causality here—it may be that one becomes familiar with brands that are perceived to have satisfactory attributes.

These hypotheses about degree of conviction can then be translated into statements about the outcome of the satisfaction decisions. The model, then, is expressed in the following hypotheses:

H1: The decision that attribute \( i \) is satisfactory for brand \( j \) is positively related to
(a) the degree of belief that brand \( j \) possesses attribute \( i \) (\( b_{ij} \)); and
(b) the degree of familiarity with brand \( j \) (\( f_j \)).
H2: The decision that attribute i is satisfactory for brand j is negatively related to 
(a) the evaluation of attribute i (a_i); and 
(b) the standard deviation of belief scores for attribute i across brands (σ_i).

Given these hypotheses, the next section discusses how the variables involved are measured, and in particular, the statistical technique appropriate for analyzing such a model, multivariate probit analysis.

METHODOLOGY

Data Measurement

The product category used for testing the models was toothpaste, chosen because several previous studies had used this product (Bass and Talarzyk, 1972; Bass and Wilkie, 1973; Cohen and Ahtola, 1971; and Sheth and Talarzyk, 1972). Five attributes were used in the study: whitening teeth, preventing cavities and decay, economical to use, freshening breath, and pleasant tasting. These had been found to be the relevant attributes in the earlier studies based on depth interviews, so they were used in this study. Seven brands were used: Ultra Brite, Pepsodent, Macleans, Crest, Closeup, Colgate, and Gleem. Subjects were 121 graduate and undergraduate students.

Each subject was asked to evaluate each attribute for each brand in terms of whether or not that brand was satisfactory for their needs for each attribute. These 0(not satisfactory)-1(satisfactory) scores provided the dependent measures of attribute satisfaction. Note that this measure does not arise strictly from observation of subjects to ascertain decision nets, although this type of decision about attribute satisfaction occurs frequently in decision nets. Hence, we can use such a measure as a tool for studying the processing underlying decision net representations even though we do not observe the nets themselves. In addition, each subject provided measures of the evaluation of attribute i (a_i) and beliefs (b_ij). These components were measured using Fishbein and Raven’s (1967) AB scales. These are semantic differential scales, with five adjective pairs in the A scale for measuring a_i (good-bad, clean-dirty, healthy-sick, wise-foolish, and beneficial-harmful) and five in the B scale for measuring b_ij (possible-impossible, true-false, existent-nonexistent, probable-improbable, and likely-unlikely). The concepts rated are, for example: “Having white teeth is:” (a_i); “The statement ‘Ultra Brite toothpaste whitens teeth’ is:” (b_ij). Five filler pairs were added for each scale, and the order randomized. Each pair is scored on a -3 to +3 bipolar scale, so the measures for a_i and b_ij, obtained by summing the scores for the five relevant pairs, range from -15 to +15. The standard deviation of belief scores (σ_i) was simply computed from the b_ij scores across the 7 brands. Finally, familiarity with each brand (f_j) was simply measured on a 0 (extremely unfamiliar) to 9 (extremely familiar) scale.

Given these measures, the data were analyzed for each attribute. The data were collapsed over brands, so that five analyses were run, one for each attribute, with 121(7) = 847 observations per analysis, since each subject rated 7 brands. This collapsing requires some explanation. The halo-effect hypothesis on familiarity allows us to collapse the data across brands when testing the model. That is, a different model will be estimated for each attribute, but not for each attribute-brand combination. There is no a priori reason to expect the mechanisms relating a_i, b_ij, and standard deviation of the b_ij to attribute satisfaction to differ across brands. The b_ij themselves should differ, but not the relationship of b_ij to satisfaction. One might expect that familiarity might relate to attribute satisfaction differently across brands, depending upon the b_ij value for each brand. The halo-effect hypothesis, however, indicates that this should not occur. To test strictly the halo-effect hypothesis the model should be run for all attribute-brand combinations, and the signs of the familiarity coefficients analyzed. However, because this would require 35 separate multivariate probit runs in the present case, it was felt that running five models as described above would be acceptable as a first step.

Note that this is a group-level analysis, rather than individual level. This is unfortunate because of the problems of heterogeneity inherent in subject populations. However, since the model involves four variables and data on only seven brands were collected from each subject, the model cannot be run at the individual level with much hope of meaningful results. Therefore, the results represent aggregated group behavior and not necessarily that of individual subjects. One would hope that the direction of effects, rather than their exact magnitudes, would not be overly distorted.

Multivariate Probit Analysis

Given the measures defined above, then a statistical model appropriate for estimating parameters showing how the 0-1 attribute satisfaction decision is affected by the variables under consideration is multivariate probit analysis (Kau and Hill, 1972; McKelvey and Zavoina, 1973). This model explicitly includes the concept of a threshold which may vary over attributes and subjects, as postulated in the model development section of this paper. Because this technique is not well known, a brief summary of its use will be given.

Let C_i be the degree of conviction that attribute i is satisfactory, the hypothetical construct referred to above. Then a linear version of the model outlined earlier is:

1 This model could be more complicated. One could test the effect of attribute determination (Alpert, 1972), for example.
\[ C_i = \gamma_{10} + \gamma_{11} b_{i1} + \gamma_{12} f_{i1} + \gamma_{13} a_{i1} + \gamma_{14} \sigma_{i1} + e_{i1} \]  

(2)

where the \( \gamma \)'s are coefficients to be empirically estimated and \( e_{i1} \) is the error term. The signs of these coefficients are hypothesized in H1 and H2 above: \( \gamma_{11}(+) \), \( \gamma_{12}(+) \), \( \gamma_{13}(-) \), and \( \gamma_{14}(-) \). Now we must examine how multivariate probit analysis estimates these coefficients (See Kau and Hill, 1972 for a parallel development).

Multivariate probit analysis assumes that there is a threshold \( C^*_i \) for degree of conviction that attribute \( i \) is satisfactory. A subject will decide that an attribute is satisfactory only if his degree of conviction \( C_i \) is above the threshold \( C^*_i \) (subscripts for subjects are left off for convenience). Suppose \( Y_i \) represents the attribute satisfaction decision for attribute \( i \) (i.e., \( Y_i = 1 \) if attribute \( i \) is considered satisfactory and \( Y_i = 0 \) if it is not). Then theoretically, if \( C_i \) and \( C^*_i \) were known,

\[
Y_i = 1 \quad \text{if} \quad C_i \geq C^*_i \\
Y_i = 0 \quad \text{if} \quad C_i < C^*_i.
\]

(3)

But \( C^*_i \), the threshold, is assumed to be normally distributed across the sample of subjects. It is not known for an individual subject. This distribution can be standardized to be a random variable \( N(0,1) \). Then the probability that an observed value for \( Y_i \) is 1 (attribute \( i \) is considered satisfactory) for an arbitrary individual is the probability that the threshold is less than \( C_i \). If \( F(x) \) represents the standard cumulative normal distribution evaluated at the point \( x \), then the above arguments imply

\[
\begin{align*}
\text{Prob} (Y_i = 1|C_i) &= F(C_i) \\
\text{Prob} (Y_i = 0|C_i) &= 1 - F(C_i).
\end{align*}
\]

(4)

Although we cannot observe the threshold, we can observe the 1-0 satisfactory-not satisfactory decisions, \( Y_i \). Given the above probabilities, we can develop a likelihood function for the observed pattern of satisfaction decisions in the sample, therefore. Since \( C_i \) is a function of the \( \gamma \)'s, then the likelihood function depends on the \( \gamma \)'s. Thus, we can estimate the \( \gamma \)'s by finding those values of the \( \gamma \)'s which maximize the likelihood function. Since the partial derivatives of this function with respect to the \( \gamma \)'s are nonlinear, iterative techniques are used to solve the system obtained by setting the partial derivatives equal to zero (see McKelvey and Zavoina, 1971). Note that this technique allows us to estimate the \( \gamma \)'s without observing actual values of \( C_i \). After these maximum likelihood estimates of the \( \gamma \)'s have been developed, one can test the hypothesis that \( \gamma_{11} = \gamma_{12} = \gamma_{13} = \gamma_{14} = 0 \) with a likelihood ratio test. This tests whether the model is appropriate in total.

\[ \text{by using } \chi^2 \text{ as a variable. However, the simple linear form} \]

\[ \text{was used because it seemed to fit the data best, and more importantly, it examined the effect of each of the Fishbein model measures individually.} \]

This test of fit of the overall analysis is performed on the statistic \( (-2) \log \text{likelihood function} \) which is distributed approximately as a \( \chi^2 \) with 4 degrees of freedom for this example. In addition, one can develop t-tests for the significance of individual coefficients and likelihood-ratio tests for subsets of coefficients.\(^2\)

Since multiple regression with a 0-1 dependent variable could also be used to analyze this data, let us examine why multivariate probit analysis is preferable. From a philosophical point of view, the major attraction of the probit model is the explicit treatment of the threshold concept, as the threshold concept is inherent in the model developed. However, there are also compelling statistical reasons for preferring the multivariate probit approach. First, it is well known that applying regression where the dependent variable is dichotomous leads to heteroscedastic disturbances (Johnston, 1963, p. 227; Kau and Hill, 1972, p. 267). Second, the expected value of the error term is not zero. Given a linear model, for some values of the independent variables the predicted value will be greater than one (less than zero), and hence positive (negative) deviations from the actual values of the dependent variable (one or zero) are impossible (Johnston, 1963, p. 229; McKelvey and Zavoina, 1973). Finally, if the dependent variable is discrete, the error term will not be distributed normally (McKelvey and Zavoina, 1973). Other advantages of the multivariate probit model are given by Kau and Hill (1972). Despite the departures from assumptions noted above, multiple regression is quite robust, so the results obtained for any particular analysis if regression were used may not differ substantially from those of the probit model. Therefore, the data were also analyzed using least squares multiple regression with the 0-1 dependent variable.

Thus, ten analyses were run, one multivariate probit analysis and one multiple regression analysis for each of the five attributes. Again, it must be emphasized that the form of the analysis implies we are modeling group behavior rather than individual behavior.

RESULTS AND CONCLUSIONS

The results of the multivariate probit analysis for each of the five attributes are shown in Table 1, where \( \gamma \) coefficients and t-values are presented. A one-tailed t-test was used to test significance of the coefficients, since direction was hypothesized. The multiple regression results are not presented, since they were very

\(^2\) McKelvey and Zavoina (1973) derive the same results from a different viewpoint. They assume that the data actually satisfy a regression model on an underlying scale, that is, that the model of equation 2 is in fact satisfied. However, the categorical data is assumed to arise from the fact that we cannot measure \( C_i \) adequately. Thus, multivariate probit analysis is an attempt to estimate the true relationship of the independent variables to \( C_i \) as it would be measured on the hypothetical underlying interval scale, given the imperfect 0-1 measures available.
similar to the results in Table 1. Although actual
coefficient magnitudes differed, all signs were the same,
statistical significance was almost identical, and the
relative sizes of the coefficients across the four independ-
ent variables were very similar for the two analyses.
Thus, although the regression model assumptions are
violated, the results need not be different. However,
a cautionary note is in order. The assumptions are in-
deed violated, and substantive conclusions drawn from
the two kinds of analysis need not coincide. Indeed,
McKelvey and Zavoina (1973) present an example of
voting behavior of Congressmen for the 1965 Medicare
bill where the substantive conclusions do differ.

The results of the multivariate probit analysis to a
large extent bear out the hypotheses. Beliefs \((b_{ij})\),
familiarity \((f_j)\), and standard deviation in beliefs \((\sigma_i)\)
all have the appropriate sign and are statistically signif-
icient at at least \(p < .10\) for each analysis. For the
evaluative term \((a_i)\), four of five signs are as hypothe-
sized, and the one that is not is not significantly differ-
ent from zero. However, the significance levels as a
whole are not as impressive for the \(a_i\) variable. Overall,
the equations are each significant at beyond \(p < .01\).

These results show that there is a relationship between
the measures used in Fishbein models and satisfaction
decisions of the kind found in decision nets. It is a first
step at trying to add more detail to information processing
that has been assumed away in both decision net and attitude models. Thus, attribute satisfactory-not satisfactory choices seem to depend on evaluation, beliefs, and familiarity, but also on variation in beliefs over brands. Thus, there are brand comparison processes that are hidden by the normal decision net model, which assume only one brand at a time is passed through the net. This study points out the need for much
more work on the processes underlying our model
measures, as done here for satisfaction decisions like
those found at decision net branches. However, the
predictor variables used here still cover up a great deal of
processing by consumers, and to understand con-
sumer choices fully, more detailed models of these
variables are needed. For example, the measure \(b_{ij}\)
summarizes a great deal of processing of communica-
tions or other information that has determined an
individual's beliefs. Understanding how \(b_{ij}\) is influenced
would thus seem to be of great importance in under-
standing how advertising affects choice. One could
potentially model \(b_{ij}\) using a protocol-decision net ap-
proach or perhaps even using linear learning models.
Also, the effects of learning and reinforcement from ex-
perience are not treated at all directly in any of the
measures. The point of this paper is to show that this
kind of strategy, of building more detailed models of
components in other models, is both fruitful and im-
portant.

d A second point for discussion deals with the struc-
ture of the models presented. It was assumed that
satisfaction decisions are binary—an attribute is or is not
satisfactory. It might make more sense to postulate a
threefold classification—satisfactory, undecided, and
not satisfactory. The undecided category might be re-
lated to information seeking behavior. This model im-
plies two thresholds dividing the value for level of
confidence into three regions. In addition, the regions
are ordered, from rejection to indecision to acceptance.
This type of model relates directly to the Sherif attitude
model, with its latitudes of acceptance, rejection, and
indifference (Sherif, Sherif, and Nebergall, 1965). It
also could be applied to new product situations, where
accept, reject, undecided decisions are possible. If an

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
</tr>
<tr>
<td>Whitensh</td>
<td>.125</td>
</tr>
<tr>
<td></td>
<td>-.05</td>
</tr>
<tr>
<td>Cavity Prevention</td>
<td>-.150</td>
</tr>
<tr>
<td></td>
<td>-.73</td>
</tr>
<tr>
<td>Economy</td>
<td>.125</td>
</tr>
<tr>
<td></td>
<td>1.19</td>
</tr>
<tr>
<td>Breath Freshness</td>
<td>.156</td>
</tr>
<tr>
<td></td>
<td>1.14</td>
</tr>
<tr>
<td>Taste</td>
<td>.072</td>
</tr>
<tr>
<td></td>
<td>.65</td>
</tr>
</tbody>
</table>

\(^a\) Distributed as \(\chi^2\) with 4 degrees of freedom.

\(^b\) The first entry is the estimated coefficient. The entry directly below the coefficient is the t-value.

\(^c\) Significant at \(p < .10\) (one-tailed test).

\(^d\) Significant at \(p < .05\) (one-tailed test).

\(^e\) Significant at \(p < .01\) (one-tailed test).
dual threshold model could be estimated, then factors leading to placement in the regions could be examined, as done in this study for a single threshold model. McKelvey and Zavoina (1971) present a procedure for estimating this type of model. As more categories are added, however, one might expect the multiple regression model to perform quite well. However, as noted above, the two sets of analyses need not give similar results.

REFERENCES


