Market Response Measurement

using the

Multinomial, Multiattribute Logit

Choice Model

by

Peter Manning Guadagni

B.S., University of California, Riverside
(1977)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

MASTER OF SCIENCE IN MANAGEMENT

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1980

© Peter Manning Guadagni

Signature of Author_____
Alfred P. Sloan School of Management
May 15, 1980

Certified by_____
John D.C. Little
Thesis Supervisor

Accepted by_____
Michael S. Scott Morton
Chairman, Departmental Graduate Committee

ABSTRACT

MARKET RESPONSE MEASUREMENT
USING THE MULTINOMIAL, MULTIATTRIBUTE LOGIT
CHOICE MODEL

by
Peter Manning Guadagni

Submitted to the Alfred P. Sloan School of Management on

May 15, 1980, in partial fulfillment of the requirements

for the degree of Master of Science.

Universal Product Code (UPC) scanning is the key to a new marketing data source. Unfortunately, most existing marketing models do not have a structure which can utilize the richness of UPC data. A new model is proposed which uses individual purchase choices from members of a UPC scanner panel to measure consumer response to marketing variables.

The development and theory of the multinomial logit (MNL) model as well as goodness of fit measures are discussed. An adaptation of the MNL is proposed where multiple purchase choices of a sample of population members are used to get a disaggregately calibrated measure of aggregate response. The model is then tested using UPC panel and sales data from the coffee category. The model performed well enough in this initial implementation to indicate that it may be a useful analysis tool. Specifically, the model was able to measure the impact of price, promotion, and individual consumer loyalties as well as the more subtle effect of a promotional purchase's impact on subsequent buying.

ACKNOWLEDGEMENTS

CHAPTER 1

INTRODUCTION

Management scientists have been building models to analyze marketing data for years. These models have given managers substantial insight into individual and aggregate consumer behaviour. Unfortunately, the ultimate usefulness of these models has always been inhibited by the quality of the data used to calibrate them.

Traditionally, marketers have relied upon warehouse withdrawals, store audits, and consumer diary panels to supply data on competitive sales and consumer purchase dynamics. Warehouse withdrawal and store audit data can answer market status questions (see Little(1979)). That is, they give management information on total and regional market shares as well as indicate the extent of seasonalities in sales. However, the extent of aggregation in the compilation of this type of data has smoothed out the peaks and valleys of sales which comprise the effect of a promotion or price change at the individual store level. Use of this data in market response measurement is therefore limited to the analysis of only the most widespread and pervasive marketing activities, and even their usefulness here has been debated (Shoemaker and

Pringle(1980)).

Purchase diaries do not suffer from the effects of aggregation but have other problems. The most frequently cited one concerns the ability of the panel to represent the general population accurately. The rigors of diary keeping discourage the bulk of the population from participating, thereby increasing the likelihood of a bias. Moreover, extended diary keeping tends to make the panelists extremely aware of prices and therefore possibly more sensitive than the rest of the population to small price differentials. Perhaps a more serious problem with the usual national diary panel is that it gives no indication of the environment in which a purchase is made. In other words, you learn the price a consumer paid but not the prices of other products on the shelf. This is important since the significance of the price of a purchased product can only be evaluated if the price of competing products is also known.

Universal Product Code (UPC) scanner technology may be the key to an improved data source. Product movement in a scanner equipped store may be inexpensively monitored on a weekly, daily, or even hourly basis. Thus it is possible to observe the effects of a temporary price cut or end aisle display at the individual store level. In addition,

it is the analyst, not the data supplier, who has the choice of how and where to aggregate, thereby making the data much more flexible in its use.

What is perhaps more important is the fact that many scanners have the ability to collect panel data. A participant in a scanner panel need only indentify himself/herself as such to the check-out clerk rather than recording the purchase information by hand. This advantage undoubtedly reduces the bias of the data as it is a less, obtrusive method and participation is so much easier. Futhermore, scanner panel data is augmented by store movement data giving the researcher access to the store environment which may be used in the analysis of individual purchase behaviour.

Scanner data does have disadvatages which the user should be aware of. A panelist's purchases are only recorded for purchases made in the scanner store which he/she is a panelist in, and only if he/she remembers to identify himself/herself as a panelist. As only loyal store shoppers have been recruited as panelists this problem has been minimized. However, there does exist a problem for categories such as health and beauty aids where a substantial number of purchases may be made in drug and discount stores. It is our belief that these problems are

relatively minor so long as the researcher is aware of them.

UPC scanners can thus supply data which will allow management to understand how consumers respond to marketing actions in a way never before possible. Unfortunately, most existing marketing models do not have a structure which can accommodate the richness of scanner data. The use of these models means that many of the benefits of scanner data will be lost. The challenge to marketing scientists then is to develop models which can take advantage of the fertility of UPC data.

In this paper we shall develop and test such a model. The proposed model combines individual purchase observations with the competitive environment in which the purchases were made to produce a disaggregately calibrated measure of the aggregate response to the marketing mix. Both product and individual consumer attributes may be used in the model to explain purchase decisions.

CHAPTER 2

THE MULTINOMIAL LOGIT MODEL

A multinomial logit (MNL) model has been suggested by
several authors (McFadden (1973), Silk and Urban(1978), and
Gencsh and Recker (1979)) as a reasonable model of
individual choice behaviour. The derivation of the model
assumes that individuals make choices to maximize their
utility. That is, given a set of alternatives S, an
individual will choose the alternative, k S, which holds
for him the greatest utility.

The utility of an alternative for and individual is
assumed to consist of: (1) a deterministic component which
can be measured as a function of the alternative's
attributes and (2) an unobserved random component.

Utility can thus be written:

$$u(k) = v(k) + e(k) \qquad (1)$$

where:

$u(k)$ = the utility of alternative k to the individual

$v(k)$ = the deterministic component, and

$e(k)$ = the random component

Further, the deterministic component, $v(k)$, will be

taken to be the following linear function of observed attributes of k:

$$v(k) = \sum_{i \varepsilon I} b(i)x(i,k) \qquad (2)$$

where:

I = set of observed attributes

b(i) = weight given to attribute i

x(i,k) = value of attribute i, for alternative k.

The probability that an individual will prefer and thus choose alternative k, from an alternative set S, is simply the probability that the individual will derive more utility from k than from any other alternative in S. This may be written mathetically as the following:

$$P(k:S) = Prob\{u(k) > u(j) \text{ for all } j \varepsilon S, j \neq k\} \qquad (3)$$

where:

P(k:S) = Probability of choosing k, from alternative set S

Breaking (3) into the random and fixed components of utility we get:

$$P(k:S) = Prob\{(v(k)+e(k)) > (v(j)+e(j)) \text{ for } j \varepsilon S, j \neq k\} \qquad (4)$$

or

$$P(k:S) = Prob\{(v(k)-v(j)) > (e(k)-e(j)) \text{ for } j \varepsilon S, j \neq k\} \qquad (5)$$

Finally, it has been shown by McFadden(1973) that if

the random errors, e(j), are independently, identically
distributed with the Weibull distribution then (5) takes
the following form:

$$P(k:S) = \exp(v(k)) / \sum_{j \varepsilon S} \exp(v(j)) \qquad (6)$$

This is what is referred to as the multinomial logit
model.   A more detailed derivation of this type of strict
utility model and its properties   may   be   found   in
McFadden(1973) or Theil(1971). The choice theory to which
it is related may be found in Luce(1959).

The most important output of the model is not a
measure   of utility but an estimate of the relative
importance of various factors in determining utility.   The
factors need not contribute to utility themselves but may
be surrogates for other factors which are difficult to
quantify.   The   importance of each factor, represented by
the b(i) attribute weightings, may be estimated using a
maximum likelihood procedure. The particular program used
in this research was developed at M.I.T by C.F. Manski and
Moshe Ben-Akiva and is documented in Ben-Akiva(1973).
McFadden(1973) has shown that the maximum likelihood
technique yields estimators which are consistent,
asymptotically efficient, and are unique under very general

conditions.

A property of the MNL model is the independence of irrelevant alternatives (IIA). The IIA property states that the relative probability of choosing one alternative over another should be unaffected by the attributes of or the presence or absence of a third alternative. Mathetically this means:

P(a:Sl)/P(b:Sl) = P(a:S2)/P(b:S2)   (7)

..nere:

Sl = {a,b}

S2 = {a,b,c}

Any significant violation of the assumptions of the MNL model will cause the IIA property to fail to hold. Generally, the violations may be traced to the MNL assumption that the the random utility component is independent across alternatives and independent of the observed attributes (McFadden, Train, and Tye(1977)). Intuitively, the error term requirements in the MNL are analogous to the requirements for residuals in least squares regression. If either the error terms in the MNL or the residuals in least squares regression do not behave properly the estimated coefficients will be biased.

There are two common sources of this problem. The

first is improper specification of the model. Since error terms are defined as the difference between the true utility and observed utility different specifications of observed utility will result in different error terms. If an attribute omitted from the model is not an independent random variable the error terms will probably not be independent either. Therefore, it is possible for the IIA property to hold or fail to hold on the same set of alternatives depending upon the model specification. See McFadden, Train, and Tye(1977) for a more in depth treatment of this type of problem.

Failure of the IIA property can also be caused by a sufficiently heterogenous alternative set. This condition implies choices may be made in a hierarchical manner where decision makers first choose between widely differing groups of alternatives and then pick an alternative within the chosen group. See appendix A for a more complete discussion of this type of IIA problem.

The IIA property is necessary to interpret the attribute weightings as cross elasticities. The value of a MNL model is therefore dependent upon proper specification of the model and the selection of a homogeneous alternative set.

The data requirements for model calibration are a set of choice observations and alternative attributes for each attribute analyzed. The alternative attributes need not apply to all alternatives but it is necessary that each alternative apply at least some of the time to a chosen alternative. The observed choices may be made by one individual or by different individuals within a population. In the case of the latter it is necessary that the population be homogeneous with respect to its members' attribute weightings.

The MNL model is an extremely flexible model of choice behaviour. It is founded on a tested theory of individual choice behaviour and has been shown to be valid in a variety of settings. The primary use of the model thus far has been to forecast demand for new transportation services (see McFadden(1973) and Ben-Akiva(1977)).

However, it has been used in a marketing context by Silk and Urban(1978) to predict the market share of new packaged goods, and by Gensch and Recker(1979) to evaluate the relative importance of supermarket attributes.

CHAPTER 3

PROPOSED MODEL

At its most basic level a product's market share represents the fact that some proportion of consumers have chosen that product over its competitors some proportion of the time. If we can understand how and why individuals choose one product over another we should gain insight into the reasons for a products relative success. By considering each product as an alternative in a set composed of it and its competitors the MNL model will give us a structure suitable for the analysis of consumer purchase choices.

Silk and Urban(20) recognized this and have used a logit model to predict the market share of new packaged goods prior to test market using consumer preference data. Jones and Zufryden(1978) and (1979) have proposed a logit based model for analysis of consumer purchase data. Customer characteristics, marketing mix variables, store characteristics and the competitive environment are all used in the model to explain purchase behaviour. While their model is a step in the right direction, the effects of aggregation and a shortage of data restrict its usefulness.

In order to use the MNL model effectively to analyze the effects of various marketing variables on consumer behaviour extensive data are needed. The model requires information on the decision maker, the chosen products, all of the competing products, and on the current environment of the market. With the emergence of UPC scanner data the data requirements for effective use of the MNL can often be met. Thus perhaps for the first time it is possible to use a dissagregate MNL model to evaluate several elements of the marketing mix in an adequate way. However, some ground work must be laid first.

The first step in the modeling process is to decide whether the parameters will be estimated for each individual or across the population. The advantage of calibrating the model on an individual by individual basis is freedom from the assumption of population homogeneity. Valid implementation of the MNL requires the population be homogeneous with respect to the relative importance its members give to the various attributes. When the model is run on one individual's choices this requirement is satisfied trivially. The disadvantage of individual by individual analysis is the difficulty in evaluating the importance of each individual's attribute weightings to aggregate response. Such analysis might result in a theoretically honest model which is so difficult to

interpret as to be managerially useless.

A model which utilizes the separate decisions of individual population members does necessitate the assumption of homogeneity. It also carries with it the danger of cross-sectional effects confounding the results. The advantage is that the model will give a disaggregately calibrated measure of the aggregate response to marketing mix variables. Futhermore, problems of heterogeneity and cross-sectional effects may be handled by dividing the population into homogeneous segments and running the model on each segment. This may increase the difficulty of interpreting the results but will help insure that those results are meaningful. Ultimately, the data requirements of the model usually rule out calibration by individual. A general rule of thumb for logit modeling states that a minimum of 100 observations are needed to yield unbiased results (Mcfadden(1973)). Seldom have enough purchase observations been collected for one individual to meet these suggested minimums.

Once the decision is made on whether to calibrate the model by individual or by population the model builder must next define an alternative set. While this may seem like a minor task it may prove to be the most difficult. Whether or not the model violates its underlying assumptions and

thus the ultimate utility of the model depend on a satisfactory selection of an alternative set.

The most obvious choice is the various products within a category. However, the IIA property will fail if all alternatives do not compete with each other on the same level (see Appendix A). There are some categories where this may be the case. Take the ready to eat cereal category as an example. If an individual eats pre-sweetened cereal half the time and bran cereal the other half then his predicted probability of choosing a particular pre-sweetened cereal over a bran cereal will vary depending on whether one, two, three, or more pre-sweetened cereals are available.

Bran and pre-sweetened cereals would seem to be two separate subcategories. A hierarchical logit model has been proposed to handle this sort of problem. In the hierarchical logit a probability would be assigned to the choice between bran and pre-sweetened and then purchase probabilities assigned to each member of the bran and pre-sweetened categories.

Given a category in which all products directly compete with each other, the various brands may seem to be an appropriate alternative set. In this case a problem can

arise if there are multiple sizes. The concern here is how to define an alternative's attributes. For example, if price is an attribute, which price should be used the price of the size with the cheapest unit cost or the price of the most popular size? In categories where there is high size loyalty either specification is likely to cause problems.

In general, the problems of defining the alternative set will vary with the product class being studied. We would propose a simple but pragmatic rule which may prove useful in a variety of product categories. Define the alternative set in accordance with the supermarket trade's pricing policies. If all sizes of each brand have the same unit price and are always promoted together, define the alternatives at the brand level. If, on the other hand, different sizes of a brand have different unit prices and are usually not promoted together then define the alternative to be a brand and size combination. This sort of policy is an indication that a brand's various sizes are being treated as separate products and thus are likely to be perceived as such by the consumer. Some customers may exhibit definite loyalties to particular sizes or brands but this does not necessarily indicate that the resulting alternative set is so heterogeneous as to result in failure of the IIA property. Proper modeling can usually capture the effects of differing loyalties well enough to insure

that the mo⁻el's assumptions are not violated.

Some researchers may wish to limit the alternatives to
each customers evoked set. By evoked set we mean those
products which the individual confines his/her purchases
to. This is a possible modification and one which does
seem to have considerable appeal. However, it is our
experience that proper specification of the model will
yield choice probabilities that are small. for products
outside each individual's evoked set. Actually, a purchase
probability close to 0 is more plausible than the apparent
impossibility of purchase implied by a 0 probability.
Restriction of alternatives to each individual's evoked set
also complicates the modeling of a product's migration into
or out of an evoked set.

The analyst may specify attributes in a MNL model much
the same way he would in a regression model. It is
important to remember that attributes in a MNL must apply
directly to the alternatives. This requirement is imposed
by the model's basic structure of purchase probability
being related to the observed utility of an alternative
which in turn is a function of the alternative's
attributes. Characteristics of the decision maker must be
modeled so as to apply to the alternatives. For example, a
loyalty variable may change with the decision maker, but

the changing loyalty is expressed as an attribute of an alternative for the decision maker.

The use of dummy variables as attributes for all but one of the alternatives is has an important advantage. In other words for each alternative but one we define an attribute which has a value of 1 at every observation for that alternative and 0 for all others. Such a dummy attribute captures any uniqueness un an alternative not captured by other explanatory attributes, at least insofar as describing average choice behaviour over all observations. A maximum likelihood estimate of the the model with only the dummy variables as attributes insures that the predicted probabilities will be equal to each alternative's market share for the entire sample. This turns out to give the model some desirable properties which will be discussed later. It should be noted that using a dummy variable for all of the alternatives causes a singularity which makes the estimation procedure impossible, hence the dummy for one of the alternatives should be excluded from the attribute set. The alternative whose dummy is excluded may be thought of as having an attribute weighting of zero for the excluded dummy. The managerial interpretation of the coefficients of these variables is the product's franchise. That is, the population's underlying preference for a product

independent of price, promotion, advertising, individual consumer loyalties, and other explanatory attributes included in the model.

The proposed model then takes the following form:

$$P(k:S) = \exp(v(k)) / \sum_{j \in S} \exp(v(j)) \tag{8}$$

where:

$P(k:S)$ = the probability that a consumer will choose product, k, out of a set of products S, given the observed product atributes, $x(i,k)$

$$v(j) = \sum_{i \in I} b(i)x(i,j) + b(0,j) \tag{9}$$

$b(i)$ = attribute weight for attribute i

$x(i,j)$ = value of attribute i for product j

$b(0,j)$ = weight of dummy attribute for product j, $b(0,j)$ is defined to be 0 for one and only one alternative

Where the $b(i)$'s and $b(0,j)$'s are estimated using the the decisions of members of a homogeneous population. It should be noted that an observation subscript is implicit in all the notation. That is, the attribute values and therefore the $v(j)$'s and choice probabilities may change with each different observation.

In choosing attributes to add to the model it is important to remember that the model is based on the theory

of   utility   maximization.   All attributes   should   then
contribute to utility in some way, or should be surrogates
of   unquantifiable   attributes   which   are   surmised   to
contribute to utility.

The proposed model is thus a versatile tool allowing
the consideration of diverse attributes. The variables may
be   independent   of   the   decision   maker,   as price and
promotion, or they may take on different values for each
decision maker,   such   as   loyalty.   A carefully specified
model will   yield   accurate measures   of   the   relative
importance   of various marketing   mix   variables   in
determining consumer purchase choices. These measures will
give an indication of how the market responds to changes in
price, promotion, or advertising.

CHAPTER 4

GOODNESS OF FIT MEASURES

Essential to the construction of any model is an index
of how well that model performs. A goodness of fit measure
will tell the model builder how successful he has been in
explaining the observations. It may be interpreted as an
indication of the reliablity of the estimated parameters
and may be used to compare the quality of different models.

The log likelihood value, a standard output of most
logit programs, could be used as such a measure. The log
likelihood value is the log of the estimated probability of
the given observations occuring   given   the   estimated
coefficients.   The closer that value is to zero the better
we have done in explaining the data.   However, the log
likelihood value has no lower bound and tends to decrease
with the number of observations. This makes its use
d'fficult in determining whether a model has a good or poor
fit.   What is needed is a bounded measure of fit where a
particular value will indicate a particular quality of fit
regardless of the number of observations. A familiar
measure of fit which is bounded is the R squared in linear
regression.   The R squared uses the relative magnitude of
the residuals to formulate a measure which has a range from

zero to   one.   Unfortunately,   logit   models   predict
probabilities,   and   so   makes the calculation of residuals
and hence   a R squared   inappropriate.   A model   which
predicts probabilities   requires a measure which indicates
how reasonable those predictions are as   estimates   of   the
actual but unknown probabilities.

A measure of fit which has   been   suggested   is   a
'batting   average' where we find the proportion of time the
alternative with   the   highest   choice   probability   was
actually   chosen,   this   measure is also known as the first
preference recovery. This index does provide an   intuitive
aid useful in judging a model's performance and can be
useful in explaing   the   quality   of   a   model   to   a
non-quantitative   manager.   Its structure, however, belies
the probabalistic nature of the MNL model and   thus   should
not be used exclusively. Indeed, one would wonder what to
think of a model which gave a maximum predicted probability
of .5 but yielded a perfect first preference recovery.

One way recognize   the   probabilistic   nature   of   the
model in a goodness   of fit measure is to sum the choice
probabilities for   each   alternative   across   observations.
The result will be a prediction of the number of times each
alternative   will   be chosen, which can be converted into a
market share prediction by dividing each prediction by   the

number of observations. By comparing the actual with the predicted market shares we should have a resonable measure of the quality of the model. Unfortunately, if dummy variables are used for all but on of the alternatives the model is constrained to predict the actual market shares perfectly accurately. Even if this were not the case the measure would not be able to distinguish between reasnonable fits. As an example consider the cases in table 4-1. The first case would be the result of a model where only dummy variables were used as alternative attributes. The second case might be the result of a model with a more complete set of attributes. Note that both cases yield perfect predictions of the actual market shares. Case 2, however, gives us a much better indication of when each alternative will be chosen. The second model gives us more information.

In order to capture this difference Hauser(1978) developed a set of statistics based on information theory. These statistics use our prior knowledge and the knowledge gained from the tested model to give a bounded measure of fit and an indication of the significance of the tested model.

In order to use these goodness of fit measures we must first define the following:

TABLE 4-1

POSSIBLE PROBABILITIES OF CHOICE

CASE 1:

PREDICTED PROBABLITIES OF CHOICE

ALTERNATIVE

OBSERVATION	A	B
1	0.20	0.80
2	0.20	0.80
3	0.20	0.80
4	0.20	0.80
5	0.20	0.80
6	0.20	0.80
7	0.20	0.80
8	0.20	0.80
9	0.20	0.80
10	0.20	0.80
TOTAL:	2.00	8.00
PREDICTED MARKET SHARE:	20%	80%

CASE 2:

PREDICTED PROBABILITIES OF CHOICE

ALTERNATIVE

OBSERVATION	A	B
1	0.025	0.975
2	0.025	0.975
3	0.900	0.100
4	0.025	0.975
5	0.025	0.975
6	0.025	0.975
7	0.025	0.975
8	0.900	0.100
9	0.025	0.975
10	0.025	0.975
TOTAL:	2.00	8.00
PREDICTED MARKET SHARE:	20%	80%

NOTE: ALTERNATIVE A WAS CHOSEN ON THE 3RD AND 9TH OBSERVATION,
B WAS CHOSEN ON ALL OTHER OCCASIONS. THE RESPECTIVE ACTUAL MARKET
SHARES IS 20% AND 80%

PRIOR ENTROPY (H)

$$H = -1/N \sum_{i=1}^{N} \sum_{k=1}^{M(i)} p(i,k) \ln(p(i,k)) \tag{11}$$

EXPECTED INFORMATION (EI)

$$EI = 1/N \sum_{i=1}^{N} \sum_{k=1}^{M(i)} P(i,k) \ln(P(i,k)/p(i,k)) \tag{12}$$

OBSERVED INFORMATION (OI)

$$OI = 1/N \sum_{i=1}^{N} \sum_{k=1}^{M(i)} d(i,k) \ln(P(i,k)/p(i,k)) \tag{13}$$

where:

H = the uncertainty relative to the null model

EI = expected information from the tested model relative
to the null model

N = number of observations

M(i) = Number of alternatives for the ith observation

p(i,k) = probability of choosing alternative k.
on observation i, under the null model

P(i,k) = probability of choosing the alternative, k
on observation i, under the tested model

d(i,k) = 1 if alternative k, on observation i, is chosen
0 otherwise

Hauser uses these three statistics to develop measures
of usefulness, accuracy, and significance. The usefulness

measures   are constructed by comparing the expected and the observed information with the prior entropy:

$$G = EI/H \tag{14}$$

$$USQRD = OI/H \tag{15}$$

G is said to be the expected proportion of uncertainty explained by the tested model, while USQRD (pronounced U squared)   is   the   observed   proportion   of   uncertainty explained by the tested model. The statistics are measures of usefulness in that they show how useful the tested model is in explaining   the   residual   uncertainty   of   the   null model.   Both   of   these   statistics are analogous to the R squared of regression in that they have a range of 0 to 1. They   will   take   on   a   value   of zero if the tested model yields predicted probabilities identical to   those   of   the null model. If the tested model makes perfect predictions (probability of 1   that   the   chosen   alternative   will   be chosen) both measures will be 1.

The accuracy measure is defined to be   the   ratio   of observed information to expected information.

$$ACCURACY = OI/EI \text{ or } USQRD/G \tag{16}$$

The   interpretation   of   accuracy   is   how   well   the probabilistic model   is   able   to   explain   real   world observations.

Finally, the significance measure uses the fact that OI is the log likelihood ratio of the tested model to the null model divided by the number of observations. Therefore by defining, S, as follows:

$$S = 2*N*OI \qquad (17)$$

we have a statistic which Wilkes(22) has shown to have a chi squared distribution if the null model is a restriction of the tested model. Degrees of freedom for the statistic are equal to the difference in degrees of freedom between the tested and null models. As stated earlier the calibration of the model with only dummy variables will result in predicted probabilities equal to the market shares. Therefore, the proposed model will always meet Wilkes' requirement when the market shares are used as a null model. Similarly, since a logit model run with no attributes will result in predicted probabilities equal for all alternatives, Wilkes' result will hold for any logit model when the equally likely null model is used. A large chi squared statistic indicates that the tested model is significantly better than the null model. Figure 4-2 shows the value of each of these measures for the model in case 2 of table 4-1. The null models used are the market share model (case 1) and the equally likely model.

Part of the appeal of these measures is the fact that any null model may be used. This makes it possible to

measure the impact of each added attribute by using the model without that attribute as the null model. Using the measures in this way will tell the model builder whether the addition of a variable significantly improves the model as well as indicating how much of the remaining uncertainty was explained by the variable. It should be pointed out that the magnitude of the measures will depend on the quality of the null model. The more naive the null model is, the larger the magnitude of prior entropy and thus the easier it is to explain a substantial portion of the uncertainty. An example of this is seen in figure 4-2, where the USQRD is larger when the equally likely model is used. A possible problem in using these measures is the difficulty in knowing the maximum value we can expect G or USQRD to be. A theoretical maximum of 1 exists but the highest achievable value is dependent on the degree of randomness in the behaviour we observe. This in fact is also true of the R squared, but it is an attribute so often forgotten we felt it warranted mention here. It is difficult to know at this time how random consumer purchase behaviour is. However, repeated application of the model in the future should help establish guidelines on the magnitude of values which should be expected from a fully specified model.

A problem with these measures for our uses is that

FIGURE 4-2

EXAMPLES OF HAUSER'S GOODNESS OF FIT MEASURES

CASE A: MARKET SHARE USED AS NULL MODEL
         i.e. $p(i,k) = ms(k)$, where $ms(k)$ is market
         share of alternative k

OI = 0.12

EI = 0.12

USQRD = 0.87

G = 0.87

S = 2.46

H = 0.06

CASE B: EQUALLY LIKELY MODEL USED AS NULL MODEL
        i.e. $p(i,k) = 1/$(number of alternatives)

OI = 0.65

EI = 0.65

USQRD = .94

G = .94

S = 13.00

H = 0.69

they require choice probabilities for each alternative and observation. Calculation of these probablities can be time consuming on even a powerful computer and thus expensive. Fortunately, Hauser(1978) shows that under conditions which hold on our proposed model all of these statistics may be calculated from the log likelihood values of the null and tested models.   As the log likelihood values are usually part of the standard output of logit programs this represents a great advantage.

Hauser has shown that:

$$USQRD = 1 - L(t)/L(n) \qquad (18)$$

where:

L(t) = log likelihood of the tested model

L(n) = log likelihood of the null model

Which implies:

$$OI = H(1 - L(t)/L(n)) \qquad (19)$$

choice probabilities are constant across observations for the market share and equally likely models their H and L(n) values may be calculated easily. Finding the OI, USQRD, and S statistics is thus fairly easy when one of those two models is used as a null hypothesis. It should be noted that the formulation of USQRD in equation (17) shows that it is equivalent to the rho squared

recommended for use as a fit index by Domencich and McFadden(). To further simplify things we can use the following results from Hauser(1978).

$$OI(t2:t1) = OI(t2:n) - OI(t1:n) \tag{20}$$

$$H(t1) = H(n) - OI(t1:n) \tag{21}$$

$$USQRD(t2:t1) = OI(t2:t1)/H(ti) \tag{22}$$

where:

OI(a:b) = observed information of model a relative to model b

H(a) = entropy under model a

USQRD(a:b) = observed propertion of uncertainty explained by model a relative to model b

n = simple null model

t1,t2 = complex tested models

Therfore, we can find the OI, USQRD, and S statistics for any complex model relative to any other complex model if we first find those statistics for those models relative to a simple model such as the market share model.

We now need a simple formula for the calculation of EI and ACCURACY. We do this by using Hauser's result that:

$$OI(t2:n) - OI(t1:n) = EI(t2:n) - EI(t1:n) \tag{23}$$

if the model is constrained to predict accurate market shares. Since this is the case under the proposed model EI and OI must always differ by an additive constant for any

tested and null models. Therefore, if we can show OI = EI for any tested and null models then OI = EI for all tested and null models. This is done in appendix b using the market share model as the tested model and the equally likely model as the null. Under the proposed model we are guaranteed perfect accuracy in the sense of Hauser's measure, making the EI and G statistics redundant. We thus have a method of calculating all of Hauser's goodness of rit statistics without having to calculate the predicted probabilities of choice for the tested model.

Although, these measures give a good indication of how well the model fits the data they may not help in diagnosing problems.   The model builder may know he has a bad fit but not why the fit is bad. Contingency tables may be useful as a diagnostic tool. By breaking the population into segments, based on demographics or some other criteria, we can see how well the model is doing in these segments. If the model systematically over or under predicts in some groups the analyst may want to add variables which help explain the difference between segments. Another solution to the problem would be to break the population up into those segments and then run the model on each of these subpopulations separately.

CHAPTER 5

DATA DESCRIPTION

The data used in this research consists of regular ground coffee sales and panel data from four Kansas City scanner equipped supermarkets. It was collected by Selling Areas-Marketing Incorporated (SAMI) and put on-line by Management Decision Systems, Inc (MDS) using EXPRESS a high level information analysis language developed by MDS. The data collected covers the time period from June 6, 1978 to March 12, 1980.

The sales data contains weekly item movement for all products in the ground coffee category as well as the shelf price for each item each week.

The panel consists of 2,000 members whose purchase transactions were colleted and separated into categories. Each member identifies himself/herself as a panelist to the check out clerk before his/her purchases are scanned. The clerk then keys in a code specific to that panelist so that each product he/she buys will be recorded with its price and attributed to him/her. There is a problem in that purchases made by the panelist outside the store in which he/she is a scanner panelist in will not be recorded. This difficulty

has been minimized by the recruitment of only loyal store shoppers as panelists. There is also a problem in panelists forgetting to identify themselves. It is difficult to know the extent of this problem. Missed transactions should only have an effect on the analysis in so far as they may cause our loyalty attribute to be misspecified. Out of the 2,000 panel members a total of 1,000 members were both category users and actively participating in the panel from December 12, 1978 to February 13, 1980, the researched time period. Eighty-five panelists were randomly selected from the population of active panelists to serve as a test group.

The major brand regular ground caffinated coffee purchases of the test group were used to develop and test the model. This represented 1,470 purchase transactions.

In addition to the UPC data, information on local newspaper advertising by the four stores has been collected and was used to identify individual store feature activity.

CHAPTER 6

MODEL SPECIFICATION AND ANALYSIS

The first step in implementing the model is the definition of an alternative set. In the coffee category this process is relatively straight forward. We began by restricting the alternatives to ground caffeinated products. This choice was based on research done on the structure of the coffee market by Urban, Johnson, and Brudnick(1979). Their findings were that consumers first choose between instant and ground and then between caffeinated and decaffeinated products. Next it was observed that switching between brands and between sizes in this data is frequent enough to indicate that consumers do not make their choices in a hierarchical manner. If consumers first chose a brand and then a size or first choose a size and then a brand we would not expect to find the observed switching behaviour. Therefore, the alternative set was constructed so that the primary choices would be brand and size combinations. The small size of a brand is considered a different alternative from the large size. Although some customers may display a strong loyalty to a particular size or brand it was felt that these characteristics could be modeled into the estimation procedure in a way which would not violate the IIA

property.

This selection of an alternative set is consistent with the way in which grocers treat the coffee category. Sizes within a brand of coffee are rarely promoted together and similarly price per ounce levels for the various sizes within brand are not necessarily at parity. Therefore, our alternative set selection avoids ambiguities in the specification of attribute values. This alternative set definition results in a group of possible choices consisting of ten products. As the data is proprietary the names of the specific brands involved have been disguised. The number of times each alternative was purchased by the test group and its disguised name is listed in table 6-1. The first nine alternatives in the list were assigned a dummy variable attribute.

To make interpretation of the results as easy as possible the decision was made to do the estimations with all of the purchases of all 85 members of the test group. This decision was supported by our belief that the population is homogeneous with respect to their attribute weightings in a fully specified model.

The most obvious product attribute which might affect

TABLE 6-1

ALTERNATIVE NAMES AND MARKET SHARES

ALTERNATIVE	NUMBER OF PURCHASES	MARKET SHARE
SMALL BRAND A*	260	17.7%
LARGE BRAND A*	105	7.1%
SMALL BRAND B*	23	1.6%
SMALL BRAND C*	516	35.1%
LARGE BRAND C*	149	10.1%
SMALL BRAND D*	52	3.5%
LARGE BRAND D*	5	0.3%
SMALL BRAND E*	255	17.3%
LARGE BRAND E*	97	6.6%
SMALL BRAND F	8	0.5%
TOTAL:	1,470	100.0%

* DENOTES BRAND IS REPRESENTED WITH A DUMMY VARIABLE ATTRIBUTE

NOTE: BRAND B AND F ARE ONLY CARRIED IN SMALL SIZES

the purchase decision is price. Other things being equal consumers are thought to prefer the item with the lowest price. The price variable used is the product's depromoted price per ounce at the time of purchase. Depromotion is necessary to avoid the confusion of price effects with those of display and advertising during a promotion. The depromoting is done by adding the depth of the price cut d ring a promotion to the observed shelf price. If there was no promotion in effect at the time of purchase then the observed shelf price was used. See appendix C for a listing and definition of this and all other variables used in the analysis. The results of the model with the price variable are given in figure 6-1.

The negative coefficient of the price attribute confirms our intuition that a reduced price increases the probability of purchase. However, while the new model is significantly better than the market share model the low USQRD value indicates it gives us little new information. This is probably due to the fact that major coffee brands are usually priced at parity with the only significant deviations occurring at times of promotion. This means there will be little variation between alternatives in our depromoted price variable, making it difficult to explain the observed choice behaviour.

FIGURE 6-1

NUMBER OF ATTRIBUTES: 10
NUMBER OF ALTERNATIVES: 10
NUMBER OF OBSERVATIONS: 1,470

LOG LIKELIHOOD: -2,649.19

ATTRIBUTE	COEF EST	STD ERR	T STAT	NORM	RANGE
PRICE	-9.53	2.16	-4.73	0.18	.12-.22
SMALL.A*	3.64	0.36	10.39	–	–
LARGE.A*	2.71	0.37	7.36	–	–
SMALL.B*	1.42	0.42	3.40	–	–
SMALL.C*	4.44	0.36	12.26	–	–
LARGE.C*	3.19	0.37	8.67	–	–
SMALL.D*	2.35	0.40	5.94	–	–
LARGE.D*	-0.03	0.59	-0.05	–	–
SMALL.E*	3.70	0.36	10.17	–	–
LARGE.E*	2.67	0.37	7.20	–	–

MEASURES OF FIT WITH MARKET SHARES USED AS NULL MODEL:

OI = 0.01
USQRD = 0.00
S = 18.97

* DENOTES DUMMY VARIABLE ATTRIBUTES

NOTE: BY NORM WE MEAN THE MODAL VALUE FOR THAT
ATTRIBUTE'S VALUES ACROSS ALL OBSERVATIONS AND
ALTERNATIVES. RANGE IS THE RANGE OF VALUES THE
THE ATTRIBUTE MAY TAKE ON.

As deals are important in stimulating purchases for the dealt brand, the addition of a promotion attribute should increase the model's ability to explain behaviour. Unfortunately, there is no direct measure of promotion in the data. We do have information on sales, price, and advertising, though. Because it is reasonable to expect a promotion to be accompanied by a price change, heavy item movement, or advertising we may infer a deal is in effect by the presence one of these. Since any of these occurences may not be a reliable indication alone, we will identify a product as being on promotion only if two of these three things are present. The promotion attribute is implemented as a dummy variable which indicates whether the product was identified to be on promotion at the time of the purchase. The results of the model with the promotion variable added are in figure 6-2.

The coefficient of the added attribute shows that the presence of promotions greatly increases the probability of purchase. Furthermore, the addition of the promotion variable yields a model which is significantly better and one which explains a substantial amount of the uncertainty of the price only model. Both of these facts indicate the importance of deals in consumer purchase choices.

FIGURE 6-2

NUMBER OF ATTRIBUTES: 11
NUMBER OF ALTERNATIVES: 10
NUMBER OF OBSERVATIONS: 1,470

LOG LIKELIHOOD: -2,216.03

ATTRIBUTE	COEF EST	STD ERR	T STAT	NORM	RANGE
PRICE	-6.27	2.23	-2.81	0.18	.12-.22
PROMOTION	1.92	0.07	28.43	-	-
SMALL.A*	2.81	0.36	7.73	-	-
LARGE.A*	1.96	0.37	5.28	-	-
SMALL.B*	1.30	0.42	3.10	-	-
SMALL.C*	3.86	0.36	10.62	-	-
LARGE.C*	2.62	0.37	7.09	-	-
SMALL.D*	1.98	0.40	4.98	-	-
LARGE.D*	-0.27	0.58	-0.46	-	-
SMALL.E*	2.94	0.37	8.02	-	-
LARGE.E*	1.97	0.37	5.28	-	-

MEASURES OF FIT WITH MARKET SHARES USED AS NULL MODEL:

OI = 0.30
USQRD = 0.17
S = 885.30

MEASURES OF FIT WITH PREVIOUS SPECIFICATION USED AS NULL MODEL:

OI = 0.29
USQRD = 0.11
S = 866.33

* DENOTES DUMMY VARIABLE ATTRIBUTES

By using a dummy variable to indicate promotion we are implicitly assuming that promotion has only one effect, say due to special display or advertising. There has been research (see Little et al. (1977)) which indicates that a promotion with no price change associated with it will never the less result in increased sales. This does not imply that the effect of promotion is independent of the depth of the price cut. Quite likely the response to a promotion has two components. A fixed component due to the effects of special display and advertising, and a variable component dependent on the depth of the price cut. The dummy variable should have captured much of the constant component, but we need to add an attribute which will capture the variable effect. For this purpose we define a price cut variable which represents the the difference between the observed price per ounce before the promotion and during it. It takes on a value of zero if there is no promotion in effect. The results of the model with the price cut variable added are given in figure 6-3.

Again, the result of a positive coefficient for the new attribute agrees with our expectations that the larger the price cut the greater the probablity of purchase. As far as the magnitude of the effect is concerned we can see

FIGURE 6-3

NUMBER OF ATTRIBUTES: 12
NUMBER OF ALTERNATIVES: 10
NUMBER OF OBSERVATIONS: 1,470

LOG LIKELIHOOD: -2,194.44

ATTRIBUTE	COEF EST	STD ERR	T STAT	NORM	RANGE
PRICE	-14.77	2.60	-5.68	0.18	.12-.22
PROMOTION	1.39	0.10	13.27	-	-
PRICE CUT	18.72	2.90	6.46	0.03	.00-.06
SMALL.A*	3.01	0.37	8.25	-	-
LARGE.A*	2.13	0.37	5.71	-	-
SMALL.B*	1.62	0.42	3.83	-	-
SMALL.C*	4.07	0.37	11.13	-	-
LARGE.C*	2.87	0.37	7.70	-	-
SMALL.D*	2.41	0.40	5.98	-	-
LARGE.D*	0.14	0.58	0.24	-	-
SMALL.E*	3.05	0.37	8.29	-	-
LARGE.E*	2.15	0.37	5.76	-	-

MEASURES OF FIT WITH MARKET SHARES USED AS NULL MODEL:

OI = 0.32
USQRD = 0.17
S = 928.48

MEASURES OF FIT WITH PREVIOUS SPECIFICATION USED AS NULL MODEL:

OI = 0.01
USQRD = 0.00
S = 43.17

* DENOTES DUMMY VARIABLE ATTRIBUTES

that it is relatively minor as a price   cut   equal   to   the
total price will not even match the contribution to utility
of a promotion with no price cut.

What is especially interesting about   the   results   is
the difference in the magnitudes of the price and price cut
attribute weightings.   While   this   difference may not be
statistically significant, it does appear that reducing the
price while the product is on promotion will   result   in   a
greater   response   than   an   equal   price   reduction   not
accompanied by   a   promotion.   An   explanation   for   this
difference in elasticities is that consumers are more aware
and   thus   sensitive   to   prices   of   products   which   are
receiving some   special   treatment   like   advertising   or
display.

Thus far we   have   been   treating   the   population   as
though   each   member had the same underlying preference for
the various   products.   Certainly   this   is   not   a   valid
assumption.   Each   consumer   will   display   differing
preferences for the different products.   We   capture   this
effect   with the use of a loyalty variable which is assumed
to be a reasonable surrogate for preference.   To   construct
this   variable we first calculate each alternative's market
share for each customer in the 20 week period prior to   the

calibration period. These market shares are used as each customer's loyalty value for each alternative at the time of his/her first purchase of the period the model was calibrated on. This loyalty measure is then exponentially smoothed with each subsequent purchase in the following manner:

$$L(i,j,k) = .75*L(i,j-1,k) + .25*d(i,j-1,k)$$

$L(i,1,k)$ = customer i's market share of product k during the 20 weeks prior to the model calibration period

$L(i,j,k)$ = customer i's loyalty value for product k on observation j

$d(i,j,k)$ = 1, if customer i chose product k on observation j
0, otherwise

Loyalty is thus an attribute of an alternative but one which is dependent on the particular decision maker and his/her past behaviour.

Exponential smoothing with a constant of .25 is used to construct the loyalty measure because it is felt that loyalties can change faster than a market share might imply but not as fast as a larger smoothing would imply. The results of the model with the loyalty attribute included are given in figure 6-4.

As might be expected the addition of the loyalty variable gives a much better fit. Not only is the model significantly better with the loyalty measure than without

FIGURE 6-4

NUMBER OF ATTRIBUTES: 13
NUMBER OF ALTERNATIVES: 10
NUMBER OF OBSERVATIONS: 1,470

LOG LIKELIHOOD: -1,441.28

ATTRIBUTE	COEF EST	STD ERR	T STAT	NORM	RANGE
PRICE	-16.26	3.20	-5.09	0.18	.12-.22
PROMOTION	1.84	0.12	14.76	-	-
PRICE CUT	24.17	3.43	7.05	0.03	.00-.06
LOYALTY	4.37	0.14	32.01	-	.00-1.00
SMALL.A*	1.71	0.37	4.55	-	-
LARGE.A*	1.42	0.38	3.75	-	-
SMALL.B*	1.47	0.43	3.39	-	-
SMALL.C*	2.27	0.38	5.98	-	-
LARGE.C*	1.61	0.39	4.15	-	-
SMALL.D*	1.38	0.44	3.16	-	-
LARGE.D*	0.23	0.59	0.39	-	-
SMALL.E*	1.66	0.38	4.37	-	-
LARGE.E*	1.29	0.38	3.88	-	-

MEASURES OF FIT WITH MARKET SHARES USED AS NULL MODEL:

OI = 0.83
USGRD = 0.46
S = 2,434.79

MEASURES OF FIT WITH PREVIOUS SPECIFICATION USED AS NULL MODEL:

OI = 0.51
USGRD = 0.25
S = 1,506.32

* DENOTES DUMMY VARIABLE ATTRIBUTES

it,   the   current   specification   explains   nearly half the
uncertainty of the market share model.

A similar concept to loyalty is that of the evoked
s⌐t.   The evoked set is composed of those products which
the consumer confines his/her purchasing activity to.   A
consumers evoked set   must therefore contain the products
he/she is most familiar with.   A less familiar product, one
outside the evoked set,   will   likely   be   perceived   as   a
greater   risk   and accordingly have different elasticities.
To investigate this possibility we may define different
attribute variables for products in the evoked set and
products not in it.   Evoked set promotion will be defined
to be the same as the current promotion variable for evoked
products but will always be zero for unevoked products.
Similarly, promotion for unevoked products will be the same
as the current promotion variable if the product is not in
the customers evoked set and zero otherwise.   The price and
price cut attributes are similarly defined for evoked and
unevoked products.   A product is defined to be in the
evoked set in this study if its loyalty value is greater
than .02.   This means a product can not be in the evoked
set if it has not been purchased within the past 13
purchase occasions and probabably will not be in it if it
has not been purchased in the last 10 occassions.   The

results of the model where these new variables have replaced the original price, promotion, and, price cut attributes may be found figure 6-5. Goodness of fit measures based on the last specified model are omitted here as the prior model was not a restriction of our current formulation.

Separating the marketing variables in to attributes which represent products which are either in or out of the evoked set does give us more information. If the true values of the elasticities of evoked and unevoked products were equal we would not expect this to happen.

At first glance the results of the model seem counterintuitive. The larger magnitudes of the price and promotion coefficients for the unevoked products imply that consumers are more sensitive to the price and promotion activities of these products than of evoked products. But with a little thought we can see this is actually what we should expect. Since consumers have little experience with unevoked products they are more dependent on price or the presence of a promotion in making a purchase than they would be with an evoked product where they have knowledge of the products quality (captured with the loyalty variable) to guide them. When an evoked product is dealt, though, consumers special note of the reduced price and

FIGURE 6-5

NUMBER OF ATTRIBUTES: 16
NUMBER OF ALTERNATIVES: 10
NUMBER OF OBSERVATIONS: 1,470

LOG LIKELIHOOD: -1,388.70

ATTRIBUTE	COEF EST	STD ERR	T STAT	NORM	RANGE
E PRICE	-14.85	3.24	-4.59	0.18	.12-.22, 0
N E PRICE	-24.36	3.37	-7.24	0.18	.12-.22, 0
E PROMOTION	1.62	0.14	11.30	—	—
N E PROMOTION	2.29	0.21	10.69	—	—
E PRICE CUT	27.23	4.30	6.34	0.03	.00-.06, 0
N PRICE CUT	22.19	4.74	4.68	0.03	.00-.06, 0
LOYALTY	3.05	0.18	17.04	—	.00-1.00
SMALL.A*	1.28	0.39	3.33	—	—
LARGE.A*	1.00	0.39	2.55	—	—
SMALL.B*	1.49	0.44	3.35	—	—
SMALL.C*	1.69	0.39	4.34	—	—
LARGE.C*	1.12	0.40	2.79	—	—
SMALL.D*	1.22	0.45	2.69	—	—
LARGE.D*	0.56	0.59	0.94	—	—
SMALL.E*	1.15	0.39	2.96	—	—
LARGE.E*	0.85	0.39	2.16	—	—

MEASURES OF FIT WITH MARKET SHARES USED AS NULL MODEL:

OI = 0.86
USQRD = 0.48
S = 2,539.97

NOTE: PREVIOUS SPECIFICATION IS NOT A RESTRICTION OF TESTED
      MODEL, HENCE FIT MEASURES ARE NOT CALCULATED USING IT.

NOTE: THE 0 FOLLOWING THE RANGE OF SOME OF THE
ATTRIBUTES INDICATES THAT THE ATTRIBUTE MAY ALSO
TAKE ON A VALUE OF 0.

hence are very responsive to it. This results in a larger
price cut response to products which are evoked than to
ones which are not. The implication of these results is
that brands with smaller franchises (i.e.   brands which
remain unevoked by the majority of the population) may be
able to increase their sales by being especially
competitive in their pricing and promotional policies.

Deals may increase a product's sales in the short   run
but this does not imply the activity will increase the
products franchise. To know whether it will or not we must
estimate the long term effect of a promotion. We can do
this by using a lagged promotion attribute which will show
how a purchase for a product on promotion affects the
probability that the product will be purchased the next
purchase occasion. As there may be different effects
depending on whether the product was in the customers
evoked set at the time of the purchase or not we   implement
the   attribute with two dummy variables. One   dummy
indicates that the product was purchased on promotion on
the last purchase occasion while it was in the customer's
evoked set; the other variable indicates that the product
was purchased on promotion but was not a member of the
evoked set at the time. The results of the model with
these dummies included are given in figure 6-6.

FIGURE 6-6

NUMBER OF ATTRIBUTES: 18
NUMBER OF ALTERNATIVES: 10
NUMBER OF OBSERVATIONS: 1,470

LOG LIKELIHOOD: -1,370.22

ATTRIBUTE	COEF EST	STD ERR	T STAT	NORM	RANGE
E PRICE	-14.99	3.28	-4.57	0.18	.12-.22, 0
N E PRICE	-23.73	3.41	-6.96	0.18	.12-.22, 0
E PROMOTION	1.66	0.14	11.58	-	-
N E PROMOTION	2.24	0.22	10.41	-	-
E PRICE CUT	27.36	4.24	6.45	0.03	.00-.06, 0
N PRICE CUT	22.99	4.78	4.81	0.03	.00-.06, 0
E LAG PROM	-0.74	0.13	-5.75	-	-
N E LAG PROM	-0.37	0.28	-1.32	-	-
LOYALTY	3.51	0.20	17.44	-	.00-1.00
SMALL.A*	1.35	0.39	3.49	-	-
LARGE.A*	1.04	0.39	2.65	-	-
SMALL.B*	1.46	0.45	3.26	-	-
SMALL.C*	1.73	0.39	4.43	-	-
LARGE.C*	1.12	0.40	2.81	-	-
SMALL.D*	1.14	0.46	2.81	-	-
LARGE.D*	0.57	0.60	0.95	-	-
SMALL.E*	1.21	0.39	3.10	-	-
LARGE.E*	0.86	0.39	2.20	-	-

MEASURES OF FIT WITH MARKET SHARES USED AS NULL MODEL:

OI = 0.88
USQRD = 0.48
S = 2,576.92

MEASURES OF FIT WITH PREVIOUS   SPECIFICATION USED AS NULL MODEL:

OI = 0.01
USQRD = 0.01
S = 36.95

* DENOTES DUMMY VARIABLE ATTRIBUTES

The negative coefficients of the two lagged promotion variables seemingly indicate that a promotional purchase decreases the liklihood of a subsequent purchase. This effect was noted by Dodson, Tybout, and Sternthal(1978) in research on the impact of dealing and deal retraction on brand switching. This may not be what our results actually imply, though. In evaluating the impact of a promotional purchase we must remember our modeling of loyalty results in a measure which increases for any product purchase. As the model implies that loyalty increases the probability of a subsequent purchase the determination of a promotional purchase's impact must therefore take into account both the positive and negative effects of that purchase. A way to estimate the net effect of a promotional purchase is to compare the probability of purchase of the product prior to a promotional purchase and after the purchase. This is done in figure 6-7 for a product which is evoked at the time of purchase and one which is not. This analysis indicates a promotional purchase has almost no effect on the purchase probability of an evoked brand but does significantly increase the chances of a purchase of an unevoked product. This could be interpreted to mean there is no long term effect of promotions for popular products. On the other hand, promotions for less popular or new

FIGURE 6-7

CARRY OVER EFFECTS OF PROMOTIONAL PURCHASES

CASE 2: PROBABILITY OF PURCHASING SMALL BRAND A WHEN IT IS EVOKED
BEFORE AND AFTER A PROMOTIONAL PURCHASE. NO PROMOTION IS
IN EFFECT FOR ANY ALTERNATIVE. PRICE IS $.18/OZ FOR ALL
ALTERNATIVES.

BEFORE		LOYALTY VALUES	AFTER	
SMALL A	.35		SMALL A	.51
LARGE A	.20		LARGE A	.15
SMALL B	.00		SMALL B	.00
SMALL C	.20		SMALL C	.15
LARGE C	.25		LARGE C	.19
SMALL D	.00		SMALL D	.00
LARGE D	.00		LARGE D	.00
SMALL E	.00		SMALL E	.00
LARGE E	.00		SMALL E	.00
SMALL F	.00		SMALL F	.00

PROBABILITY OF PURCHASE

BEFORE* = .322   AFTER = .321

CASE 2: PROBABILITY OF PURCHASING SMALL BRAND A WHEN IT IS NOT EVOKED
BEFORE AND AFTER A PROMOTIONAL PURCHASE IS MADE, GIVEN NO PROMO-
TION IS IN EFFECT FOR ANY ALTERNATIVE. PRICE FOR ALL ALTERNA-
TIVES IS $.18/OZ.

BEFORE		LOYALTY VALUES	AFTER	
SMALL A	.00		SMALL A	.25
LARGE A	.20		LARGE A	.15
SMALL B	.35		SMALL B	.26
SMALL C	.20		SMALL C	.15
LARGE C	.25		LARGE C	.19
SMALL D	.00		SMALL D	.00
LARGE D	.00		LARGE D	.00
SMALL E	.00		SMALL E	.00
LARGE E	.00		LARGE E	.00
SMALL F	.00		SMALL F	.00

PROBABILITY OF PURCHASE

BEFORE* = .019   AFTER = .161

products can potentially have a positive impact on   future

sales.

Another possible effect of promotions is that they

encourage consumers to to buy on promotion. If this is

indeed the case promotions may have a positive impact on

short term sales but a negative impact on long term

profits. To investigate this possibility we can make a

distiction between the effect of a promotional purchase on

the subsequent purchase occassion when there is and when

there is not another deal available. This has been

implemented by replacing the current lagged promotion

attributes with four new dummy variables which indicate

whether the product was previously bought on promotion,

whether it was a member of the evoked set at the time of

the purchase, and whether it is currently on promotion.

The results of the model with this modification are given

in figure 6-8.

Once again it is difficult to evaluate the meaning of

the coeficients because of the multitude of interactions.

Hence, we again compute the probability of purchase prior

to the promotional purchase and after it. The results of

these computations are in figures 6-9a and 6-9b. In the

first case where there is no subsequent promotion available

FIGURE 6-8

FINAL MODEL SPECIFICATION

NUMBER OF ATTRIBUTES: 20
NUMBER OF ALTERNATIVES: 10
NUMBER OF OBSERVATIONS: 1,470

LOG LIKELIHOOD: -1,368.44

ATTRIBUTE	COEF EST	STD ERR	T STAT	NORM	RANGE
E PRICE	-15.44	3.28	-4.71	0.18	.12-.22, 0
N E PRICE	-23.77	3.41	-6.98	0.18	.12-.22, 0
E PROMOTION	1.87	0.15	12.27	–	–
N E PROMOTION	2.30	0.22	10.62	–	–
E PRICE CUT	28.10	4.24	6.63	0.03	.00-.06, 0
N PRICE CUT	22.74	4.77	4.76	0.03	.00-.06, 0
E LAG PROM,PROM	-1.32	0.21	-6.73	–	–
E LAG PROM,NPROM	-0.43	0.16	-2.77	–	–
N E LAG PROM,PROM	-1.14	0.35	-3.21	–	–
N E LAG PROM,NPROM	0.47	0.34	1.36	–	–
LOYALTY	3.58	0.20	17.54	–	.00-1.00
SMALL.A*	1.31	0.39	3.39	–	–
LARGE.A*	1.01	0.39	2.58	–	–
SMALL.B*	1.46	0.45	3.23	–	–
SMALL.C*	1.69	0.39	4.32	–	–
LARGE.C*	1.08	0.40	2.69	–	–
SMALL.D*	1.13	0.46	2.48	–	–
LARGE.D*	0.58	0.60	0.97	–	–
SMALL.E*	1.16	0.39	2.98	–	–
LARGE.E*	0.82	0.39	2.08	–	–

MEASURES OF FIT WITH MARKET SHARES USED AS NULL MODEL:

OI = 0.88
USQRD = 0.49
S = 2,596.48

* DENOTES DUMMY VARIABLE ATTRIBUTES

NOTE: PREVIOUS SPECIFICATION IS NOT A RESTRICTION OF TESTED
MODEL, HENCE FIT MEASURES ARE NOT CALCULATED USING IT.

for an evoked product the probability of purchase increases after a promotional purchase has been made. However, if there is a subsequent promotion available the probability of purchase relative to the probabilty prior to making a promotional purchase goes down. A consumer is less likely to buy a product on promotion if he/she had previously bought that product on promotion than if he/she had not! A possible explanation for this is that when a customer buys an evoked product on promotion he/she expresses his/her gratitude with increased loyalty to the product. However, if the customer observes that the product is on promotion during two consecutive purchase occassions he/she may interpret the reduced prices as a sign of reduced quality. The reaction to this perceived decrease in product quality is a reduction in the probability of purchase. Another explanation is that the first time a customer sees a special display for a product it has certain shock effect and therfore is especially effective in stimulating a purchase. However, if the consumer sees a display the next purchase occassion it won't seem quite as special and thus not quite as effective in stimulating purchases. The implication of the first explanation is that promotions do have a positive long term effect for popular products so long as they are not so frequent as to cause consumers to question the products quality. The second explanation

FIGURE 6-9a

EFFECT OF A PROMOTIONAL PURCHASE FOR AN EVOKED PRODUCT
(SMALL BRAND A IS EVOKED)

CASE 1: PROBABILITY OF PURCHASING SMALL BRAND A BEFORE AND AFTER A PROM-
OTIONAL PURCHASE IS MADE, GIVEN NO PROMOTION IS IN EFFECT. PRICE
FOR ALL ALTERNATIVES IS $.18/OZ. NO ALTERNATIVE IS ON PROMOTION
OR HAS A PRICE CUT.

BEFORE		LOYALTY VALUES	AFTER	
SMALL A	.35		SMALL A	.51
LARGE A	.20		LARGE A	.15
SMALL B	.00		SMALL B	.00
SMALL C	.20		SMALL C	.15
LARGE C	.25		LARGE C	.19
SMALL D	.00		SMALL D	.00
LARGE D	.00		LARGE D	.00
SMALL E	.00		SMALL E	.00
LARGE E	.00		LARGE E	.00
SMALL F	.00		SMALL F	.00

PROBABILITY OF PURCHASE

BEFORE* = .32   AFTER = .39

*THIS WOULD HAVE BEEN PROBABILITY
OF PURCHASE AT TIME OF PROMOTIONAL
PURCHASE HAD NO PROMOTION BEEN
OFFERED

CASE 2: PROBABILITY OF PURCHASING SMALL BRAND A BEFORE AND AFTER A PROM-
OTIONAL PURCHASE IS MADE, GIVEN A PROMOTION IS IN EFFECT. PRICE
IS $.18/OZ FOR ALL ALTERNATIVES. NO PROMOTION IS IN EFFECT FOR
ANY ALTERNATIVE EXCEPT SMALL BRAND A. PRICE CUT IS $.03/OZ.
BEFORE AND AFTER LOYALTIES ARE THE SAME AS IN CASE 1.

PROBABILITY OF PURCHASE

BEFORE** = .877   AFTER = .800

**THIS IS WHAT PREDICTED PROBABILITY
OF PURCHASE WAS AT TIME OF PROMO-
TIONAL PURCHASE.

implies that excessively frequent promotions will result in a reduced response to them but do not necessarily have a detrimental effect on brand image. In either case the implication is that a strategy of heavy promotion may not be wise for popular products.

The situation is somewhat different for products not in the evoked set at the time of a promotional purchase. In this case substantial gains are made in the probability of purchase both when the product is not currently on promotion and when it is. These gains are primarily due to the fact that the product enters the evoked set after the promotional purchase is made and hence is attributed greatly increased loyalty in percentage terms. What is especially interesting and in sharp contrast to the evoked product's case is the increased likelihood of purchase when the product is currently on promotion as compared to the case when it is not. A plausible explanation for this result is that while a promotional purchase causes a product to enter the consumer's evoked set he/she still views its purchase as being somewhat risky. A further promotion thus helps the product overcome its perceived riskiness. Promotions for less popular products should have positive long term effects regardless of their frequency. These products may in fact benefit from a

FIGURE 6-9b

EFFECT OF A PROMOTIONAL PURCHASE FOR AN UNEVOKED PRODUCT
(SMALL BRAND A IS NOT EVOKED)

CASE 1: PROBABILITY OF PURCHASING SMALL BRAND A BEFORE AND AFTER A PROM-
OTIONAL PURCHASE IS MADE, GIVEN NO PROMOTION IS IN EFFECT. PRICE
FOR ALL ALTERNATIVES IS $.18/OZ. NO ALTERNATIVE IS ON PROMOTION
OR HAS A PRICE CUT.

BEFORE		LOYALTY VALUES	AFTER	
SMALL A	.00		SMALL A	.25
LARGE A	.20		LARGE A	.15
SMALL B	.35		SMALL B	.26
SMALL C	.20		SMALL C	.15
LARGE C	.25		LARGE C	.19
SMALL D	.00		SMALL D	.00
LARGE D	.00		LARGE D	.00
SMALL E	.00		SMALL E	.00
LARGE E	.00		LARGE E	.00
SMALL F	.00		SMALL F	.00

PROBABILITY OF PURCHASE

BEFORE* = .019   AFTER = .303

*THIS WOULD OF BEEN PROBABILITY
OF PURCHASE AT TIME OF PROMOTIONAL
PURCHASE HAD NO PROMOTION BEEN
OFFERED.

CASE 2: PROBABILITY OF PURCHASING SMALL BRAND A BEFORE AND AFTER A PROM-
OTIONAL PURCHASE IS MADE, GIVEN A PROMOTION IS IN EFFECT. PRICE
IS $.18/OZ FOR ALL ALTERNATIVES. NO PROMOTION IS IN EFFECT FOR
ANY ALTERNATIVE EXCEPT SMALL BRAND A. PRICE CUT IS $.03/OZ.
BEFORE AND AFTER LOYALTIES ARE THE SAME AS IN CASE 1.

PROBABILITY OF PURCHASE

BEFORE** = .282   AFTER = .567

**THIS IS WHAT PREDICTED PROBABILITY
OF PURCHASE WAS AT TIME OF PROMO-
TIONAL PURCHASE.

strategy of frequent promotion.

As the results in figure 6-8 represent the final specification of the model it would be appropriate to review the implied effects of the basic marketing variables of price, promotion, and price cut. The basic relationships between these variables are the same as when they were discussed earlier. The constant component of promotion is still much more important than the variable component. A price reduction results in an increased probability of purchase whether or not it is accompanied by a promotion. It should be noted, though, that the price and price cut elasticities implied by the coefficients are the same for unevoked but different for evoked products. The reason for this differential in evoked products was attributed to the increased price awareness of consumers toward specially displayed or advertised products. In the case of unevoked items the special promotional treatment may cause the customer to temporarily consider buying the product resulting in an increased purchase probability. But the consumer is still wary of unfamiliar brands and evaluates their price as critically as if it were not on promotion.

To fully understand the meaning of the attribute

weightings   for   the   various marketing variables we should
examine the elasticities they imply. Unfortunately,   these
elasticities   are   dependent on the value of the attributes
for all alternatives and on   the   particular   product
considered.   Therefore, it is impossible to give one number
which   indicates   the   impact of an incremental change in a
product attribute.   A   way   we   can   get   an   idea   of   an
attribute's   impact is to hold the values of all but one of
the attributes constant and   then   examine   the   change   in
predicted   probabilities   as   the   remaining   attribute   is
varied.   The implied response curves from this analysis are
given in figure 6-10a when the   representative   alternative
is evoked and in figure 6-10b when it is not.

As predicted earlier the effect of a promotion is much
greater than the effect of either price or price cut.   The
results of this also indicate that response is dependent on
the   current   status   of the market.   That is, the purchase
probability before an incremental change   in   an   attribute
will   affect the response to that change.   Response will in
general be highest when the prior   probability   is   in   the
middle   range   (.25   to   .75),   implying the S shaped curve
frequently   found   in   aggregately   measured   response
functions.   Evaluation of attribute coefficients should
take this   into   account.   For   example,   although   the
magnitude of the price cut coefficient for evoked products
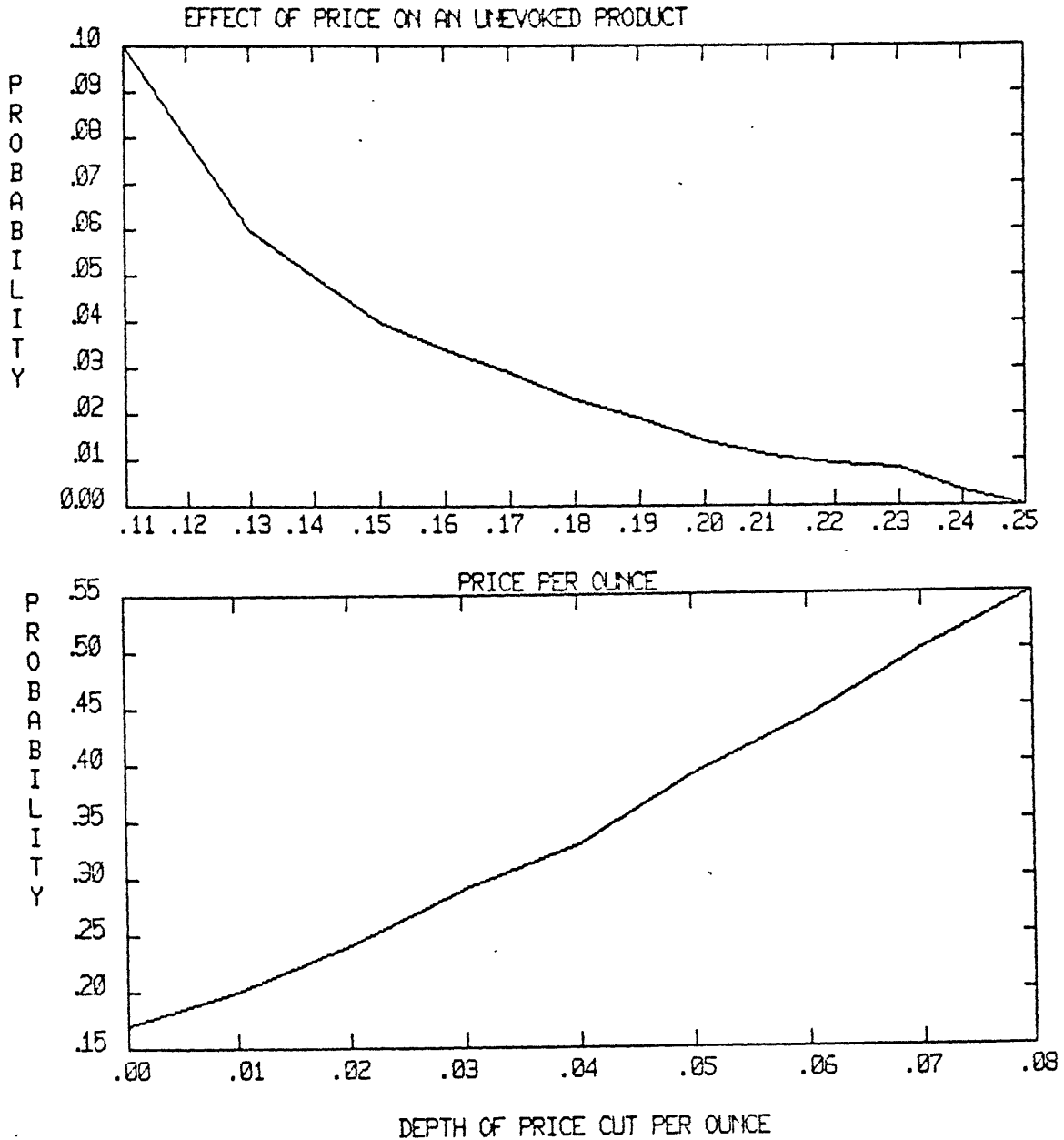
FIGURE 6-10a
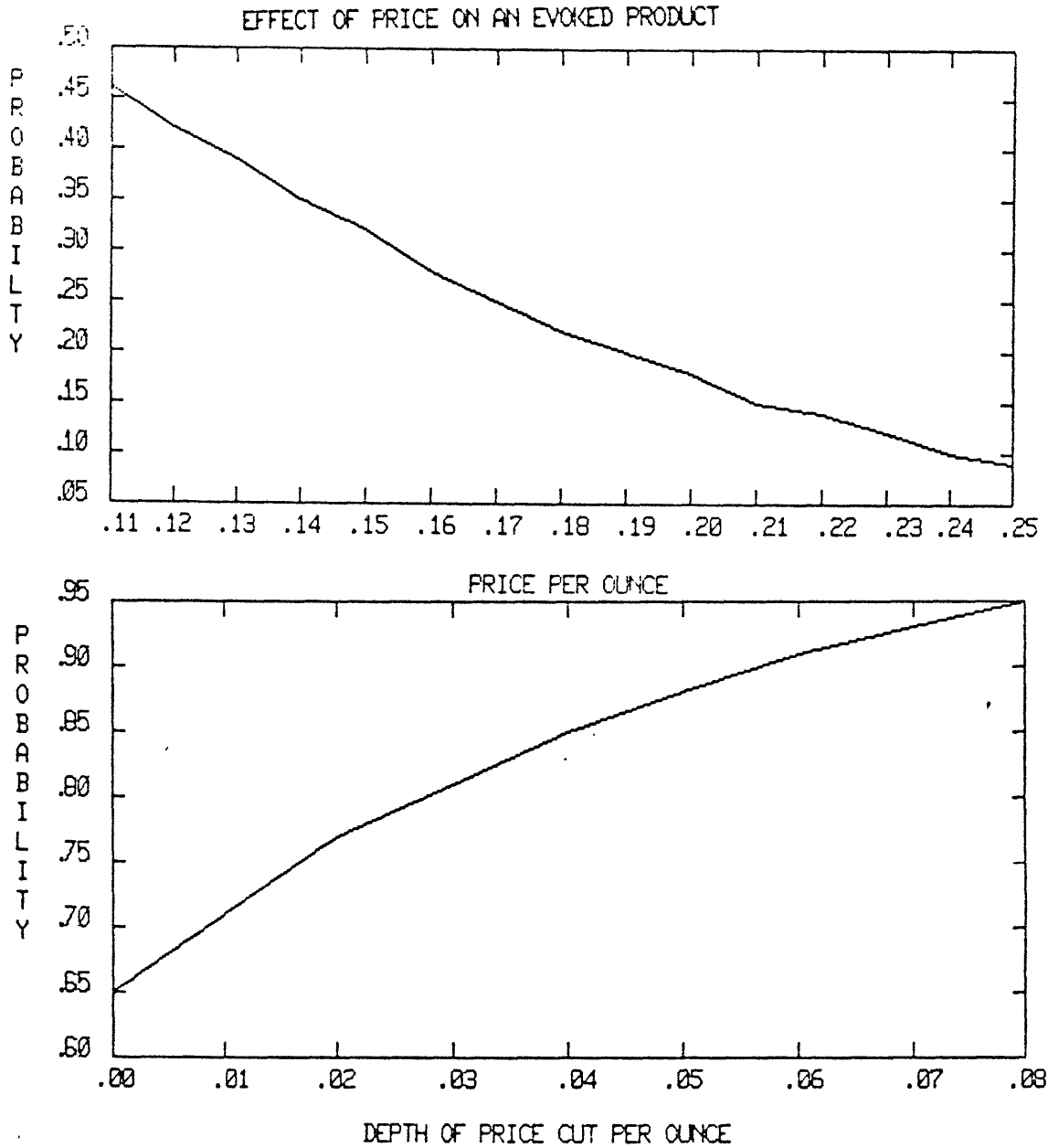
EFFECT OF PRICE ON AN EVOKED PRODUCT

NOTE: WHEN PRICE WAS VARIED ALL OTHER ALTERNATIVES HAD A PRICE OF $.18/OZ. PRICE CUT, PROMOTION, AND LAGGED PROMOTION VARIABLES WERE 0 FOR ALL ALTERNATIVES. WHEN PRICE CUT WAS VARIED ALL ALTERNATIVES HAD A PRICE OF $.18/OZ AND LAGGED PROMOTION VALUES OF 0. A PROMOTION IS IN EFFECT FOR SMALL BRAND A BUT NOT FOR ANY OTHER ALTERNATIVE. IMPACT OF PROMOTION CAN THUS BE FOUND BY COMPARING PROBABILITY OF PURCHASE FOR A PRICE OF $.18/OZ WITH A PRICE CUT OF 0. LOYALTY VALUES ARE THE SAME AS IN FIGURE 6-9a PRIOR TO A PROMOTIONAL PURCHASE.

is nearly double that of the price attribute the presence
of a promotion boosts the probability so high that the
added effect of a given price reduction may not be any more
than the effect of an equivalent price reduction
unaccompanied by a promotion. This is indicated in figure
6-10a where the absolute increase in probability is the
same for a price drop from 18 cents per ounce to 15 cents
as it is for a price cut increased from 3 cents off per
ounce to 6 cents off.

The model does not give us constant elasticities which
are easily interpretable. It is doubtful that this would
even be a desirable feature, as we should expect
elasticities to change with different circumstances.   It
will, however, give us a means to evaluate alternative
pricing and promotional strategies. Furthermore, we take
into account the competitive environment and a product's
position with in that environment when we use the model to
evaluate alternative strategies. Management can therefore
use it simulate response to alternate marketing mixes under
various competitive assumptions, allowing them to fully
understand all implications of a possible strategy. The
model is also well suited to measurement. This is because
the model will automatically control for the dominant
effects of price and promotion so that more subtle effects

FIGURE 6-10b

EFFECT OF PRICE ON AN UNEVOKED PRODUCT

PRICE PER OUNCE

DEPTH OF PRICE CUT PER OUNCE

NOTE: WHEN PRICE WAS VARIED ALL OTHER ALTERNATIVES HAD A PRICE OF $.18/OZ. PRICE CUT, PROMOTION, AND LAGGED PROMOTION VARIABLES WERE 0 FORR ALL ALTERNATIVES. WHEN PRICE CUT WAS VARIED ALL ALTERNATIVES HAD A PRICE OF $.18/OZ AND LAGGED PROMOTION VALUES OF 0. A PROMOTION IS IN EFFECT FOR SMALL BRAND A BUT NOT FOR ANY OTHER ALTERNATIVE. IMPACT OF PROMOTION CAN THUS BE FOUND BY COMPARING PROBABILITY OF PURCHASE FOR A PRICE OF $.18/OZ WITH A PRICE CUT OF 0. LOYALTY VALUES ARE THE SAME AS IN FIGURE 6-9b PRIOR TO A PROMOTIONAL PURCHASE.

such   as the long term impact of promotion can be analyzed.

CHAPTER 7

CONCLUDING REMARKS

We began this paper recognizing the   need   for   models
which can accomodate the richness of UPC data.   In response
to this need we adapted the multinomial logit model to the
analysis of scanner panel data.   The   result   was   a   model
which   combines   purchase   information with the competitive
environment in which a purchase was made   to   estimate   the
relative   importance of different product attributes in the
purchase decision.

Implementation of this model yielded estimates of   the
relative   importance of price, promotion, and preference(as
measured by loyalty) in consumer choices for ground coffee.
The long term effect of promotion   was   also   measured   and
found   to be positive in most cases.   In all, we were able
to explain   about   half   of   the   uncertainty   surrounding
.consumer coffee purchases.   It is difficult to say if there
is room for improvement in the implementation as we have no
notion   at   this time of how random consumer choices are in
this category.   Results of the implementation in the coffee
category were   intriguing.   The   analysis   indicated   that
consumers   were   relatively   insensitive   to   minor   price
fluctuations of products in their evoked set.   Loyalty   was

found   to   be an important factor in the purchase decision,
though.   Promotion was indicated to be highly effective   in
stimulating   purchases,   both   for   products   in   and   for
products out of the evoked set.   Most   of   a   promotion's
effect   came from the constant 'specialness' component of a
deal rather than the price   dependent   variable   component.
Promotions   were   found to have a positive long term effect
in most cases, however, the   long   term   effectiveness   was
especially pronounced in products which were not previously
in   the   customer's   evoked set.   This implies that dealing
should be an important component of the   marketing   mix   of
new and less popular products.

While this model has been tested in only one   category
the   results   from   that analysis are encouraging enough to
indicate that it may be a valuable tool in the   measurement
of market response.   The model's ability to control for the
effect   of   major   marketing   variables makes it especially
attractive for the measurement of subtle product attributes
such as advertising quality.   This use, while   not   tested,
is   appealing since the measure of copy effectiveness would
be directly   related   to   the   ad's   ability   to   stimulate
purchases.   In summation, we believe the proposed model to
be an attractive tool which when coupled with   UPC   scanner
data   can   begin   to   answer   long   standing   questions   on
consumer behaviour.

APPENDIX A

INDEPENDENCE OF IRRELEVANT ALTERNATIVES

Significant heterogeniety in the alternative set can cause failure of the IIA property. As an example of how this can occur consider an alternative set composed of some beverages say coffee, tea, and coke. Now consider an individual who prefers each alternative equally so that there is a one third probability that he/she will choose any one of them. Suppose that a fourth alternative, pepsi, is added and that the individual considers pepsi equivalent to coke but not a substitute for coffee or tea. The choice probabilities are thus 1/3, 1/3, 1/6, and 1/6 for coffee, tea, coke, and pepsi. Since the probability of choosing coke over the probability of choosing coffee has changed it is clear that the IIA property has failed.

The problem is that coke and coffee are not from the same level of competition. That is, consumers theoretically do not make a choice between coke and coffee. They first choose between coffee and soft drinks and then if soft drinks are chosen coke is a possible alternative. Therefore, if the IIA property is to hold all alternatives must be directly competing with each other.

APPENDIX B

PROPOSITION: OI = EI for market model relative to equally
                likely model, when market shares are
                constrained to be accurately predicted.

PROOF:

$$OI = 1/N \sum_{j=1}^{N} \sum_{i=1}^{NA} d(i,j) \ln(ms(i)/(1/NA))$$

$$= 1/N \sum_{i=1}^{NA} NP(i) \ln(ms(i)/(1/NA))$$

$$= 1/N \sum_{i=1}^{NA} NP(i)(\ln(NP(i)/N) + \ln(NA))$$

$$= \sum_{i=1}^{NA} (NP(i)/N)(\ln(NP(i)/N) + \ln(NA))$$

$$= \sum_{i=1}^{NA} ms(i)(\ln(ms(i) + \ln(NA))$$

$$= 1/N \sum_{j=1}^{N} \sum_{i=1}^{NA} ms(i) \ln(ms(i)/(1/NA))$$

$$= EI$$

WHERE:

OI = observed information
EI = expected information
N = number of observations
NA = number of alternatives
NP(i) = number of times product i was chosen
ms(i) = market share of product i
d(i,j) = 1, if alternative i was chosen on observation j
       = 0, otherwise

APPENDIX C

VARIABLE SPECIFICATIONS

PRICE	= observed shelf price per ounce at time of purchase + PRICE CUT
PROMOTION	= 1, if 2 of the following are present at time of purchase: advertising, price change, unusual movement = 0, otherwise
PRICE CUT	= difference between observed shelf price per ounce before and during the promotion, if promotion is in effect = 0, otherwise
LOYALTY	= $l(i,j,k)$, where current purchase is ith purchase for customer j $l(i,j,k)$ is defined on page 46
evoked set	alternative is in evoked set if loyalty value is greater than .02 at the time of purchase
E PRICE	= PRICE, if alternative is in evoked set. = 0, otherwise
N E PRICE	= PRICE, if alternative is not in evoked set = 0, otherwise
E PROMOTION	=. PROMOTION, if alternative is in evoked set = 0, otherwise
N E PROMOTION	= PROMOTION, if alternative is not in evoked set = 0, otherwise
E PRICE CUT	= PRICE CUT, if alternative is in evoked set = 0, otherwise
N E PRICE CUT	= PRICE CUT, if alternative is not in evoked set = 0, otherwise

LAG PROM

= 1, if this alternative was purchased
last purchase occasion and was on
promotion at that time
= 0, otherwise

E LAG PROM

= LAG PROM, if alternative was in
evoked set at time of purchase
= 0, otherwise

N E LAG PROM

= LAG PROM, if alternative was not
in evoked set at time of purchase
= 0, otherwise

E LAG PROM,PROM

= E LAG PROM, if alternative is
currently on promotion
= 0, otherwise

E LAG PROM,PROM

= E LAG PROM, if alternative is not
currently on promotion
= 0, otherwise

N E LAG PROM,PROM

= N E LAG PROM, if alternative is
currently on promotion
= 0, otherwise

N E LAG PROM,NPROM = N E LAG PROM, if alternative is not
currently on promotion
= 0, otherwise

BIBLIOGRAPHY

1. Ben-Akiva, M., "Program for Maximum Likelihood Estimation of the Multinomial Logit Model," unpublished paper, Transportation Systems Division, Department of Civil Engineering, Massachusetts Institute of Technology, January 1973.

2. _____ and Lerman, S., "Disaggregate Travel and Mobility Choice Models and Measures of Accessiblity," Third International Conference on Behavioural Travel Modeling, Australia, April 1977.

3. Dodson, J.A., Tybout, A.M., and Sternathal, B., "Impact of Deals and Deal Retraction on Brand Switching," Journal of Marketing Research, February 1978, p. 72-81.

4. Domencich, T., and McFadden, D., "A Dissaggregated Behavioural Model of Urban Travel Demand," Report No. CRA-156-2, Charles River Associates, Inc., Cambridge, Massachusetts, 1972.

5. Ehrenberg, A.S.C., Repeat-Buying: Theory and Applications, N. Holland Publishing Co., 1972.

6. Gensch, D.H., and Recker, W.W., "The Multinomial, Multiattribute Logit Choice Model," Journal of Marketing Research, February 1979 p. 124-132

7. Green, P.E., Carmone, F.J., and Wachpress, D.P., "On the Analysis of Qualitative Data in Marketing Research," Journal of Marketing Research, February 1977 p. 52-59.

8. Hauser, J.R., "Testing the Accuracy, Usefulness, and Significance of Probabilistic Choice Models: An Information Theoretic Approach," Operations Research, Vol. 26, No. 3, May/June 1978.

9. Jones, J.M., and Zufreyden, F.S., "An Approach for the Study of Retail Strategy Impact and Market Segmentation for Consumer Non-Durable Products," UCLA Working Paper Series, Center for Marketing Studies, Paper #68, November 1978.

10. _____, "Implementing a Decision Relevant Multivariate Model of Consumer Purchase Behaviour," UCLA, March 1979.

11. Little, J.D.C., "Decision Support Systems for Marketing Managers" <u>Journal of Marketing</u>, Summer 1979, p. 9-26.

12. _____, et. al., Unpublished field research.

13. Luce, R.D., Individual Choice Behaviour, New York: John Wiley and Sons, 1959.

14. McDonald, C., "What is the Short-Term Effect of Advertising?", British Market Bureau Limited, London.

15. McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behaviour," in P. Zarembka, Ed., <u>Frontiers in Econometrics</u>, New York: Academic Press, 1973

16. _____, Train, K., and Tye, W.B., "An Application of Diagnostic Tests for the Independence of Irrelevant Alternatives Property of the Multinomial Logit Model," <u>Transportation Research Record</u>, Vol. 637(1977) p. 39-45.

17. Shoemaker, R., and Pringle, L.G., "Possible Biases in Parameter Estimation with Store Audit Data," <u>Journal of Marketing Research</u>, February 1980 p. 91-96.

18. Silk, A.J., and Urban, G.L., "Pretest Market Evaluation of New Packaged Goods: A Model, and Measurement Methodology," <u>Journal of Marketing Research</u>, May 1978, p. 171-191.

19. Theil, H., <u>Principals of Econometrics</u>, Wiley: New York, 1971.

20. Urban, G.L., Johnson, P.L., and Brudnick R.H., "Market Entry Strategy: A Hierarchical Modeling and Consumer Measurement Approach," W.P. No. 1103-80, The Marketing Center, M.I.T., December 1979.

21. Wilkes, S.S., <u>Mathematical Statistics</u>, New York: Wiley 1962.