AN OVERVIEW OF EMPIRICAL APPLICATIONS OF BUYER BEHAVIOR SYSTEM MODELS

John U. Farley, Columbia University
Donald R. Lehmann, Columbia University

Abstract

Results of five empirical system studies of consumer choice processes are compared for consistency in results. Applications include durables and nondurables, new and established brands and U.S. and foreign settings. Based on these comparisons, suggestions are made concerning research required for further development of the field.

Introduction

Choice models describing the processes used by consumers when buying brands of a product (Howard & Sheth, 1969; Engel, Kollat and Blackwell, 1973; and Nicosia, 1960) have been the subject of intensive research for over a decade. While much research effort has been focused on individual elements or subcomponents of the models (Farley, Howard and Ring, 1972), some empirical research has also been cast in the general framework of the overall model--initially with the goal of testing the feasibility of using model structures in their whole form (Farley & Ring, 1970), and later with the goal of providing insights and forecasts useful for managers designing and assessing marketing programs (Black & Farley, forthcoming). With most focus on versions of the Howard-Sheth choice model, these applications have involved durable as well as non-durable products, American as well as non-U.S. settings, and new as well as established products and brands. This paper attempts a synthesis of some common characteristics of a group of these studies in terms of:

1) model configuration and variable definition;
2) procedures for specification and parameter estimation;
3) problems to be solved before the field can yield its full potential.

The Studies

As Table 1 indicates, the five published studies selected for discussion here vary along several dimensions--size of the variable set (from 9 to 28 variables), setting (U.S., Argentina, and Kenya), type of products (durables and non-durables) and stage of development in terms of the product life cycle (products in test market and with established markets). Research design involved panels and waves of repeat interviewing of fresh samples. Several brands of a product were studied in some cases. Included are studies of:

1) A convenience food product reported in (Farley & Ring, 1970), as modified in (Farley & Ring, 1972).
2) A personal product reported in (Farley, Howard & Lehmann, 1974), and (Lehmann, O'Brien, Farley & Howard, 1974).
3) Paper products reported in (Katz, 1973), and (Farley & Katz, 1974).
4) Subcompact automobiles reported in (Farley, Howard & Lehmann, forthcoming).
5) A contraceptive product reported in (Black & Farley, forthcoming).

Model Configuration

The basic starting point for each study has been provided by a general flow-chart formulation of decision process models such as is shown in Figure 1, taken from (Farley & Ring, 1970).

Variable Configuration

The qualitative models generally incorporate two major variable groupings:

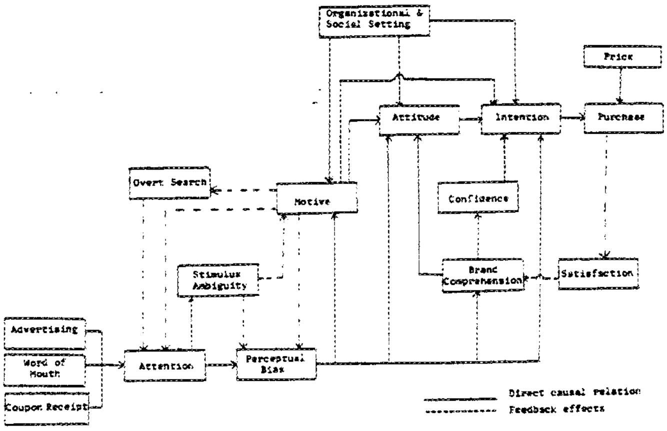
1) Jointly determined (endogenous) variables that are more or less common over studies. These can be

TABLE 1
Key Features of Data Collection For Research on Buyer Behavior Models

Product	Setting	Number of Variables in System			Stage of Product in Development Cycle	Data Collected Specifically for Systems Models?	Number of Waves of Interviews Used in Model	Number of Brands Modeled	Sample Size*	Type of Research Design	Date of Completion of Basic Modeling Work	Duration of Measurement Used in Model
		Endogenous	Exogenous	Total								
Convenience Food Product	U.S.	11	17	28	Test Market	No	1	1	691	Three waves of cross-sectional interviews combined and treated as one sample	1969	3 months
Paper Products	U.S.	5	16	21	Established Brands	No	2	2	222	Two waves of cross-sectional interviews	1974	6 months
Personal Product	Argentina	10	18	28	Test Market	Yes	2	1	200	Two-wave panel plus pre-introduction measurements	1971	9 months
Contraceptives	Kenya	3	6	9	Test Market	No	2	1	494	Two waves of cross-sectional samples plus pre-introduction measurements	1974	12 months
Automobiles	U.S.	4	8	12	National Introduction	No	5	8	1453	Five-wave panel, two pre- and three post-introduction	1975	18 months

*Indicates sample actually used in parameter estimation. In panel designs in which non-response cut sample size in later waves, this is the result in the earliest wave.

FIGURE 1
Conceptual Model of Buyer Behavior



identified in the flow-chart as variables which are affected by as well as affecting others in the system; 2) Pre-determined (exogenous) variables which are rather specific to the application at hand. These affect endogenous variables but are not in turn affected by other elements in the system.

The endogenous variables further divide into those variables which basically describe perception and learning processes, and those which describe decision making and cognition. In general, measuring and modeling the latter set has been more successful than the former. The exogenous variable set often includes three types of measures:
a) Lagged endogenous variables (especially behavior) which are statistically useful in removing spurious model elements at the point of parameter estimation;
b) Controllable variables--advertising, dealing, etc. These allow assessment of elements of the marketing program and prediction of the effects of program changes; and
c) Socio-demographic measures, usually gleaned from a rather large set of candidates which are specific to the market in question.

A more specific indication of the variable sets used in each of the five studies just described is shown in Table 2.

Model Specification and Parameter Estimation

Initial specification of models like those shown in Figure 1 at first involved a rather literal interpretation that causal links exist between points connected by arrows and flow in directions indicated by the arrowheads (Farley & Ring, 1970). This exercise yielded a set of general relationships generally equal in number to the number of endogenous system variables shown in Table 1. These general relationships were then cast into linear form and parameters were estimated with some sort of regression technique. More recently, the initial specification has been somewhat more flexible.

Empirically-Based Model Specification

In several cases, multivariate procedures other than regression have been used to provide additional situation-specific empirical bases for model specification. Multivariate procedures appear appropriate, of course, because of the complex variable structure, particularly among the jointly causal endogenous variables. For example, factor analysis, canonical correlation, cross-lag correlations, and AID techniques have been used as

TABLE 2
Variables Used in Buyer Behavior
Model Applications

	Application				
	Convenience Food Product	Personal Product	Paper Products	Compact Car	Contra-ceptives
Variables					
ENDOGENOUS VARIABLES					
Perception and Learning					
Attention or Awareness	X	X	X		X
Search for Information	X	X			
Stimulus Ambiguity	X	X			
Perceptual Bias	X	X			
Confidence	X	X		X	
Cognition					
Attitude	X	X	X	X	
Intention	X			X	
Motives and Choice Criteria	X				
Satisfaction	X	X			
Knowledge	X	X	X	X	X
Behavior					
Purchase or Trial	X		X		X
EXOGENOUS VARIABLES					
Past Behavior	X		X		X
Controllable Variables					
Advertising	X	X	X	X	
Coupons or Deals	X		X		
Distribution	X	X			
Price	X	X	X		X
Variety of Product Used		X	X		
Sociodemographic Measure					
Family and User Group Size	X		X		X
Region			X		
City or Area Size	X				
Sex of Respondents			X	X	
Age Variables	X	X	X	X	
Media Exposure	X	X	X		X
Education	X			X	X
Income	X	X		X	X
Employment Status	X				
Characteristics of Use Setting	X	X		X	
Social Characteristics of Users	X	X		X	X

specification aids (See Table 3). While the procedures have provided insights in some cases, the results also underline the critical importance of a priori model specification based on theory plus experience.

For example, factor analysis has been useful for assessing the dimensionality of the endogenous and exogenous variable sets separately (Katz, 1973). It is very useful to know that the endogenous variable set is generally of full dimensionality, because this fact assures that later regression work will be feasible. However, factor analysis has been less useful in specifying individual relationships.

Somewhat closer to the specification problem is canonical correlation, which can be used to form weighted compounds of related variables from two distinct groups--in this case the endogenous and exogenous variables (Farley & Ring, 1974; and Katz, 1973). This joint

TABLE 3
Multivariate Procedures Used in
Specification and Parameter Estimation
for Buyer Behavior Models Applications

	Application				
	Convenience Food Product	Personal Product	Paper Products	Compact Car	Contra-ceptives
Factor Analysis					
For Specification			X		
For Variable Reduction			X		
Cross-lag Correlation to Specify Causal Direction		X			
Canonical Correlation For Specification	X	X	X		
AID Analysis for Specification	X		X		
Regression for Parameter Estimation					
Ordinary Least Squares	X	X	X	X	X
Two-Stage Least Squares	X	X	X	X	

structure characteristic of canonical correlation has generally indicated the presence of expected variable groupings among either the endogenous or exogenous variable sets (attitude-awareness-intention-purchase, for example) but the procedure has not been useful in establishing the directions of causation. Cross-lag correlation has also been used to help establish direction of causation in panel studies in which measurements on the same individuals are available at different points in time (Lehmann, O'Brien, Farley & Howard, 1974). It now appears however, that the adjustment of the individual to experience or to new information is rather rapid relative to any practical inter-measurement period (e.g., one month), so cross-lag techniques have provided less specification guidance than had been hoped.

Various step-wise procedures have also been tested to help solve the specification problem (Katz, 1973; and Farley & Katz, 1976). These procedures, such as the Automatic Interaction Detector and step-wise regression, have the inherent disadvantage that they are single-equation techniques--that is, that they have a single-valued dependent variable. However, in practice AID has indicated the presence of more complex feedback structures than is usually specified in qualitative flow charts like Figure 1. This result indicates limited promise for recursive models except in the context of rigidly controlled experiments.

Finally, the problem of specifying parameter configuration (in contrast to variable configuration) has received only limited attention, although developmental work has indicated substantial potential. For example, analysis comparing results from a model with fixed parameters to a model in which parameters were allowed to vary over the range of values of explanatory variables showed substantial improvement in goodness of fit measures and allowed use of the segmented regressions to identify market segments defined within the buyer behavior system framework (Weinstein & Farley, 1974).

General Patterns of Parameter Estimates

As Table 3 indicates, a substantial amount of effort has gone into comparing results of equation-by-equation parameter estimation methods (e.g., ordinary least squares) with methods that take explicit account of inter-relationships among the system of endogenous variables (e.g., two stage least squares). The fact that OLS has been shown experimentally to stand up better than other methods to specification problems probably explains why results differ little qualitatively and in some cases are identical quantitatively for the different approaches (Farley, Lehmann & Howard, forthcoming). Rather consistent patterns emerge in terms of both goodness of fit and patterns of relationships among system parameters even though problems of parameter identification have arisen in connection with estimation (Lutz & Resek, 1972). There are also indications that results may vary systematically with the context in which data are gathered (Farley, Katz & Lehmann, 1976).

Table 4 shows goodness of fit measures for ordinary least squares equations common to several studies. Experience has shown consistently better fit for equations predicting a subset of the endogenous system variables--attention and awareness, intention or behavior, and knowledge (which appear in all four studies) as well as attitude and satisfaction. One result of this fact has been a tendency to cut back models in application to "working" versions which incorporate only these five variables as endogenous. These endogenous variables have two characteristics in common:
1) The measures (usually scales or actual records of one sort or another) correspond quite closely to the theoretical concept under study; and
2) There is a very substantial literature on the measurement of each variable.

While Table 4 also shows that the goodness-of-fit measures increase over time as experience with the models accrues, the coefficients of determination of even the well-fitting equations generally range from .2 to .5, and even the segmented regressions mentioned earlier are not beyond this range.

In terms of both marginal (pairwise) and partial measures, inter-relationships among pairs of this same subset of endogenous variables have been consistently positive and significant. Relationships involving other endogenous variables have been less predictable in terms of either sign or significance. Relationships between endogenous and exogenous variables (particularly socio-demographic measures) have generally been weak and have provided relatively little incremental explanatory power in the regression equations. In fact, the major utility of these variables has been in providing groupings for within-group averages (purchase or intention to purchase, attitudes, etc.) to be used for market segment identification (Farley, Howard & Lehmann, 1974).

Difficulties to be Dealt With

Our prediction is that these models will continue to be used and developed as both research tools and predictive devices. The broader the scope and ambition of the application, the greater the number of problems to over-

TABLE 4

Coefficients of Determination (R^2) for Ordinary Least Squares Equations Estimating Parameters of Three Systems Equations in Five Buyer Behavior Model Applications

	Contraceptives		Convenience Food Product		Paper Product*		Personal Product		Compact** Car
Endogenous Variable	Interviewing Wave 2	Wave 3	Unsegmented	Segmented	Brand A	Brand B	Interviewing Wave 2	Wave 3	
Attention or Awareness	.225	.236	.065	N.A.	.087	.081	.741	.948	N.A.
Knowledge	.265	.362	.028	N.A.	.165	.237	.665	.552	.086
Trial, Intention or Purchase	.519	.523	.153	.336	.338	.412	.328	.240	.242
Attitude	N.A.	N.A.	.231	.354	.315	.185	.264	.156	.043

* Averaged over 12 sample groups
**Averaged over three brands, five waves of interviewing and four sample groups

come. Certain endemic difficulties have arisen and have been openly discussed in applications:

Operationalization of the Constructs

One of the most difficult tasks is operationalizing the concepts designated by the model. Sometimes key constructs have been omitted by necessity--as in the compact car example where purchase was not used because only a handful of people purchased the product. In the case of the personal product in Argentina, intention was simply not included in the translated questionnaire. In other cases, omissions arose because sponsors were interested more in proven constructs than in complete models, or because data were used that were designed for purposes other than examining the models.

Real problems lie in measuring all of the constructs, of course. On one level is the issue of whether the constructs are uni- or multi-dimensional. The regression framework virtually dictates a single measure for each construct for parameter estimation. More basic problems occur with how to operationalize each individual construct into an item or set of items on a questionnaire. Even attitude and intention, which typically are measured on semantic differential scales, have been measured with different numbers of scale points and different anchors. Other constructs, such as brand comprehension, have been measured by the respondent's own perception of his comprehension, with a single yes-no objective question and with a sum of the number of correct answers to a series of multiple choice questions. Still other constructs, such as confidence, have been operationalized in numerous ways. As a result, comparability across studies is difficult. A set of relatively standardized measures is needed, and some work has been done in this direction (Katz, 1973).

Noise in Measurement

A major problem in using any set of data is assessing the level of noise in the measures. The early applications tended to result in coefficients of determination of cross-sectional regressions ranging from .05 to .5. While the .05 obviously can be improved, we also felt then that the .5 R^2's could be also increased to .8 or .9. Given the present state of measurement technology, it now appears that .6 is a much more likely upper bound for R^2.

The major reason that half the variance is essentially unexplainable is that the signal-to-noise ratio is relatively low. Put differently, the test-retest reliabilities of many of the measures (most attitude questions, for example), tend to produce correlations of about .7 and R^2's of .5. Since the test-retest reliability of the dependent variables is limited to R^2's of about .5, it is unreasonable to expect the regression equations to get higher R^2's.

Numerous reasons exist for the low reliability of the measures. One which we have struggled with is the problem caused by missing data which leads to a choice among substituting a mean value for the respondent or, as we have generally done, discarding him entirely. Another is the tendency of respondents to "halo" their answers by checking either favorable or unfavorable ratings for a given product regardless of the construct being measured. This implies that the statistical error structure of the endogenous variables may be very complex. Finally, there is a certain amount of random noise in the respondent's feeling toward the concepts, which means that on a 7-point scale, a response of 5 probably has at least 1 scale point random error associated with it.

Causal Priority and Timing of Measurements

The Howard-Sheth model implies a definite causal priority among the constructs. Unfortunately, this priority is very difficult to test in the context of the whole model, because it is almost impossible to design experiments which control all the endogenous elements. Consider, for example, the attitude to intention link. The "classic" way of testing the causal priority would be to measure both attitude and intention at 2 points in time and see whether the correlation of attitude at time 1 with intention at time 2 or attitude at time 2 with intention at time 1 is higher. This approach has three major difficulties. First, it is not clear how far apart the measurements should be taken to a) minimize measurement bias and b) match the time period in which a change in attitude becomes a change in intention. In fact, since for many of the links in the system the adjustment may take only seconds, such an approach is unfeasible practically. Second, even if the time periods were correctly chosen, the effects of other variables on the variables of interest must somehow be removed, and full experimental control on all of them is simply not feasible. Finally, it is likely that for some fraction of the people, the causal order is reversed at a given point in time because the feedback effects are dominant. Hence the real task is not to prove the causal order specified by the model, but rather to estimate the fraction of people for which it is true.

In contrast, survey data, which form the basis for most practical applications, are inherently non-experimental and strong assumptions must be made about the nature of the error structure and about the data-generating process in order to be able to estimate effects at all.

Individual Differences

The problem of inter-individual differences is both large and multi-faceted. One obvious problem is the measurements where a "4" may mean different things to different respondents. Normalization, the usual remedy, assumes that the number has no meaning, and in most cases that the variance of the responses has no meaning as well. We generally have used raw data and hence data normalization is one possible avenue of future research. Another problem of individual differences is the assumption that the marginal response to changes in the variables depends only on the level of the variables and not on the individual. While this is a practical necessity since the sytem cannot now be estimated statistically on an individual level, it is still bothersome, and perhaps some grouping of more or less homogeneous people would be useful prior to estimation. However, this process may be no more successful than the typical segmenting exercises based on demographic or personality variables have been in providing insight into the buyer behavior process.

Functional Form

The qualitative flow chart versions of the model do not specify functional forms of relationships among variables and much of the early work has recommended that non-linear functional forms be investigated. Unfortunately, given the noise level in the data and only modest non-linearity, it is difficult to determine whether the non-linear model will behave much differently from a linear approximation of the same function over a reasonable range of variable values. The use of orthogonal polynomials (Laroche, 1974) has in fact improved the fit of some of the relationships.

Non-Stationarity of the Model

Throughout the past five years, the model itself has been in a state of constant change (Farley, Howard & Ring, 1973). Changes have occurred in the operational definitions, both because of product situation and the availability of data. Similarly the model has changed to meet the situation, and it has also changed over time based on new findings in the literature, rethinking of the model, and accumulated empirical results of the studies like those described here. While this change is both appropriate and desirable, it has made comparability of results more difficult. The considerable criticism of individual studies as not being tests of "the" model also appear misaimed during a period of model modification.

Future Directions of Development

The future of examination and uses of models like the Howard-Sheth model should be different from the past in many ways. The model forms a useful organizing framework for data collection and analysis of situations where respondents process information about the product in question. In order for it to become more useful both as a simulation tool for decision makers and a tool for basic research, several things must happen:
1) Examination of alternative operational definitions must lead to agreement on a "best" set of definitions.
2) Explicit mathematical form or family of forms of the equations as well as segments must be specified.
3) Controllable decision variables must be tied more directly to the endogenous variables--e.g., advertising messages delivered to sales.

Progress toward all three of these goals is more likely through limited small scale studies of subsets of variables than through experimentation with the full model. Similarly, applications of different statistical and modeling methodologies will be useful in investigating these models. Progress will be facilitated if all concerned recognize that choice process models offer only one approach to the understanding of consumer behavior which is complimentary with many others.

References

Timothy R.L. Black and John U. Farley, "An Awareness-Knowledge-Trial Model of a Kenyan Contraceptive Test Market," Journal of Advertising Research, forthcoming.

James F. Engel, David T. Kollat and Roger D. Blackwell, Consumer Behavior (New York: Holt, Rinehart & Winston, 1973).

John U. Farley and John A. Howard, Control of "Noise" in Market Research Data (Lexington: Lexington Press, 1975).

John U. Farley, John A. Howard and Donald R. Lehmann, "A 'Working Version' Car Buyer Behavior Model," Management Science, forthcoming.

John U. Farley, John A. Howard and Donald R. Lehmann, "After Test Marketing, What?", Proceedings, Annual Meeting, Business and Economic Statistics Section, American Statistical Association, Detroit, December 27-30, 1970, 288-296.

John U. Farley, John A. Howard and Donald R. Lehmann, "Applying a Buyer Behavior Model to the Evaluation of a Test Market," Paper presented at Annual Meeting, Association for Consumer Research, Amherst, Massachusetts, August, 1970.

John U. Farley, John A. Howard and Donald R. Lehmann, "Evaluating Test Market Results: Buyer Behavior Analysis in Argentina," Journal of Business Administration, 5(2)(Spring, 1974), 69-88.

John U. Farley, John A. Howard and L. Winston Ring (eds.), Applied Buyer Behavior Theory (Boston: Allyn & Bacon, 1973).

John U. Farley and Jerrold P. Katz, "Toward a 'Model' for Survey Research on Repetitive Decision Making," working paper, Columbia University, 1976.

John U. Farley, Jerrold P. Katz and Donald R. Lehmann, "The Effect of Measurement Stimulus Set on Attitudes and Saliencies," working paper, Columbia University, 1975.

John U. Farley and L. Winston Ring, "An Empirical Test of the Howard-Sheth Model of Buyer Behavior," Journal of Marketing Research, 7(November, 1970), 427-438.

John U. Farley and L. Winston Ring, "'Empirical' Specification of a Buyer Behavior Model," Journal of Marketing Research, 11(February, 1974), 89-96.

John U. Farley and L. Winston Ring, "On L and R and HAPPISM," Journal of Marketing Research, 9(August, 1972), 349-353.

Robert Ferber, "On the Reliability of Purchase Influence Studies," Journal of Marketing, 19(January, 1955), 225-232.

John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior (New York: John Wiley & Sons, 1969).

Jerrold P. Katz, An Examination of Sample Survey Research in Marketing in Context of a Buyer Behavior Model, Ph.D. dissertation, Columbia University, 1973.

Michael LaRoche, "A New Approach to Non-Linear Consumer Behavior and Market Segmentation by the Use of Orthogonal Polynomials," unpublished doctoral dissertation, Columbia University, 1974.

Donald R. Lehmann, John U. Farley and John A. Howard, "Testing of Buyer Behavior Models," Proceedings, Second Annual Conference, Association for Consumer Research, College Park, Maryland, September 1-3, 1971, 232-242.

Donald R. Lehmann, Terrance V. O'Brien, John U. Farley and John A. Howard, "Some Empirical Contributions to Buyer Behavior Theory," Journal of Consumer Research, 1(December, 1974), 43-55.

Richard J. Lutz and Robert W. Resek, "More on Testing the Howard-Sheth Model of Buyer Behavior," Journal of Marketing Research, 9(August, 1972), 344-345.

David Weinstein and John U. Farley, "Market Segmentation and Parameter Inequalities in a Buyer Behavior Model," Journal of Business, 48(4)(October, 1975), 526-540.