Structure of Disaggregate Behavioral Choice Models

Stein Hansen, Møre Og Romsdal Distriktshøgskole,
Molde, Norway

This paper reviews the foundations of some of the choice models most frequently used in transportation planning and outlines the strengths and weaknesses of these approaches in the analysis of travel behavior. The first part deals with algebraic utility theory. The foundations of the textbook approach are briefly reviewed and an evaluation is made of the characteristics and economics of time allocation models. The different algebraic utility structures implied by the algebraic demand models most frequently found in practice are discussed. The second part of the paper attempts to link economic utility theory to that approach developed in mathematical psychology, and the distinction is made between fixed and random preference models. The practical models in this field are derived from a probabilistic choice approach, and the development of the well-known logit formula is briefly outlined. Certain similarities to the separability properties discussed in the first part of the paper are indicated. The paper closes with suggestions of the direction of further development of simultaneous models or new theoretical support for particular choice sequences or both.

Modern analyses of travel behavior have primarily been concerned with choice rather than demand as a point of departure. However, travel demand is frequently used to label travel choice models. As a consequence, the relation—or lack of such—among utility, choice, and demand should be understood by analysts who determine what traveler preferences are and evaluate transportation policy and investment schemes.

Two analytical approaches are available for the description of individual choice behavior: algebraic and probabilistic.

Although probabilistic elements play an important role in any transportation planning model, a distinction is made in this paper based on the underlying behavior assumptions as expressed in the consistency axioms of choices. Thus, econometric models based on traditional microeconomics and extensions thereof are considered algebraic, whereas choice models based on thresholds in choice, random utility indicators, and "almost optimizing behavior" are considered probabilistic.

ALGEBRAIC THEORIES OF CONSUMER CHOICE

Foundations

The microeconomic theory of choice deals with a decision rule by which consumer purchases are made under given market conditions. It links desires and action and provides the means for transforming utility restrictions into demand properties. Since demands are observable but utility is not, any check on theory requires translating assumptions on the latter into properties of the former. Then if individual demands do not have these properties, the theory does not give an adequate explanation of individual behavior.

The decision unit in traditional microeconomic theory acts in a pure exchange economy with n commodities. The unit is described by his consumption set $X = (X_1, \ldots, X_n)$, which is a closed, convex, and bounded subset of commodity space S; his preferences U, which is a complete, continuous, twice differentiable, and strictly convex preordering of X; and his initial endowments $\overline{X} = (\overline{X}_1, \ldots, \overline{X}_n)$, which is a vector in S.

The behavior of the decision unit is derived from these assumed characteristics: He regards all commodity prices, $p = (p_1, \ldots, p_n)$, fixed regardless of his own actions, and he chooses the greatest element for U in his budget set, $X \epsilon S | p \cdot X = p \cdot \overline{X}$. Consequently, the indifference map of choices in this model is compatible with the conclusion that unique and continuous demands exist and express the equilibrium point for the consumer in the sense that maximum utility is attained (15, 19, 23).

As a consequence of the various constraints just introduced, the demand functions must satisfy the following properties:

1. Reallocations of the budget due to income and price changes respectively must continue to exhaust total income (the adding-up property);
2. Multiplying all prices and income with the same factor should leave demands unaltered (the homogeneity property);
3. Demand for a specific commodity cannot increase as its price increases and all other prices remain unchanged, and income changes (raises) just enough to compensate for the price increase (the negativity property); and
4. The compensated cross-demand effects are symmetric,

$$\delta X_i / P_{j_{U=\overline{U}}} = \delta X_j / P_{i_{U=\overline{U}}}, \text{ for all } i \neq j \tag{1}$$

This ensures integrability or choice consistency and rules out the possibility that demand functions (or choice functions) are such that a sequence of price and income changes will lead the consumer through a series of positions, each of which is preferred to the previous one, but which in the end lead back to the starting point (the symmetry property).

The pure microeconomic choice theory presented here is not sufficient to specify an operational model. More specific behavioral assumptions are needed for that purpose. Before an assessment is made of the demand models in applied consumer choice economies, a couple of other recent approaches to the deterministic microeconomic analyses of consumer choice are reviewed.

Characteristics and Consumer Demand Theory

It has been argued that the pure microeconomic theory of choice does not offer a satisfactory account of why some goods are consumed more than others or why some goods are not purchased at all. A further difficulty arises with the introduction of new goods. This creates particular difficulties in constructing cost of living index numbers and in accepting further consumption of outdated commodities by a group of homogeneous individuals.

Lancaster (21) suggested that these difficulties can be lessened by regarding the elements of the set of alternatives by which the consumer orders his preferences U as bundles of characteristics c associated with goods X rather than as bundles of goods—consequently, U(c). Thus, for example, the various means of travel from a given home base to a given work base constitute a closely related group of goods because they, and they alone, supply the characteristics with respect to arrival time at work and commuting comfort.

Formally, let g be a fixed number representing the total number of characteristics attainable from all goods in the economy. Let c_j represent the objectively measurable quantity of the j th characteristic and $c = (c_1, \ldots, c_g)$. With each commodity bundle X is associated a specific vector of characteristics such that

$$c = h(X) \tag{2}$$

The consumer decides on purchases by maximizing U(c) subject to Eq. 2 and the usual budget constraint $p \cdot X = R$. Defining

$$U(X) = U[h(X)] \quad (3)$$

and assuming the existence of the basic utility model properties, we can derive demand functions having properties similar to those discussed above. Certain problems may arise, however.

A unique bundle of characteristics does not necessarily imply a unique bundle of goods. No problems arise so long as the number of distinct goods does not exceed the number of characteristics, but modern complex economies are probably characterized more by goods than by characteristics, and this is the world we set out to model. In this case, the quantity of none of the goods would be uniquely determined. One consequence of the goods-characteristics model is then that the goods-demand curves may be perfectly elastic at a given price. The commodity demands would then be demand correspondences, which in terms of the theory of the previous section would follow from a relaxation of the strictly convex assumption to one of weakly convex indifference curves in commodity space.

Little is known at present of the practical importance of the Lancaster approach. But it surely has some interesting theoretical properties that make it possible to illuminate economic problems that are insoluble by traditional means.

Microeconomic Theories of the Allocation of Income and Time

Some recent developments in microeconomic theories of consumer choice have focused on the time allocation problem and have recognized that leisure covers time used for consumption, commuting, and sleeping, which are necessary activities in order to perform further work (2, 8, 11, 12, 16). The increasing interest in this field is probably due to the idea that in wealthy countries people behave as if time is a scarce resource.

In attempting to construct a "general theory of the economics of time allocation," Bruzelius (8) proposes to integrate the traditional consumer choice theory, discussed in the previous section, with a similarly pure theory for time allocation. This is motivated from the shortcomings of both theories. The pure theory of time allocation rests on a utility function defined for time activities only. The quantities are measured in time units. Utility is maximized subject to a time resource constraint only.

In general, the utility generating activities are connected with both time and goods; i.e., the consumer will usually not indulge in something that is a pure good or a pure time activity. A general theory should require that the consumer allocation problem be described in terms of the 2 dimensions, the simpler extreme problems being special cases.

The "general model" suggested by Bruzelius (8, pp. 9-15) is as follows: The utility function

$$U(X_1, \ldots, X_n, T_1, \ldots, T_n) \quad (4)$$

where

X_i = quantity of good i and
T_i = time used along with the use of X_i,

is maximized subject to the following constraints:

$$\sum_{i=1}^{n} P_i X_i - r_w T_w - V \leq 0 \quad (5)$$

where

P_i = price of good i,
r_w = wage rate,
T_w = work time, and
V = exogenous income.

Equation 5 expresses the economic budget constraint

$$\sum_{i=1}^{n} T_i + T_w - \overline{T} \leq 0 \qquad (6)$$

where $\overline{T}$ = total time. Equation 6 expresses the time resource constraint. In case of an inequality, the constraint is closed by a slack variable T_{n+1}.

$$g_i(X_i,\ T_i) \leq 0, \text{ for } i = 1, \ldots, n \qquad (7)$$

Equation 7 expresses physical relations between the time and the good variables that enter into the activity-producing process.

$$X_i \geq 0,\ T_i \geq 0,\ T_w \geq 0 \qquad (8)$$

Equation 8 expresses the nonnegativity constraints on the endogenous variables in the model.

To compare this model with the traditional consumer demand theory, we consider the first order conditions for maximum utility (8, p. 13). The interpretation of these conditions can be carried out in a variety of ways depending on the explicit character of the physical relation (Eq. 7). The following explicit version of Eq. 7 is chosen for illustration:

$$g_i = a_i X_i - T_i \leq 0 \qquad (9)$$

where it is assumed that $X_i = 0 \Leftrightarrow T_i = 0$. This can be interpreted to say that to each amount of the good X_i there is a minimum of time that must be allocated to it, but this minimum may be exceeded. Or, to look at it the other way, associated with each level of T_i there is a maximum amount of X_i, but the consumer may choose a lower level. According to this model,

1. The marginal utility of X_i should equal the marginal utility of monetary outlays plus the marginal utility of saving time in producing the particular activity multiplied by the number of units of time a_i required as a minimum per unit of X_i; and
2. The marginal utility of time in activity i should equal the marginal utility of time as a resource plus the marginal utility from saving time in commodity i multiplied by 1 (because of the choice of Eq. 9).

This approach has additional features that should be appreciated in applied economics. If the utility function (Eq. 4) is written in terms of the utility generating activities $Z_1, \ldots, Z_n$,

$$U = U(Z_1, \ldots, Z_n) \qquad (10)$$

then Eq. 7 may be viewed as a household production function. Although the X_i and T_i have been treated as scalars above, X_i actually is a set of market goods, $X_{i1}, X_{i2}, \ldots, X_{in}$, used in producing Z_i, and similarly for T_i. This theoretical approach yields not only information on which market goods are close substitutes and which are not [in a way similar to that described by Lancaster (21)] but also justification for the use of weakly separable utility functions. This property of the model implies that the marginal rate of substitution between any 2 factors (markets goods and time) producing Z_i

is independent of the quantity of any good not used in this particular process or, equivalently, the ratio of the marginal utilities of the 2 factors depends only on the factors used in that particular production process (27). Consequently, this approach gives theoretical justification for reducing the number of cross effects to be quantified in a planning context.

Price of Time

The concept of the price of time has initiated a lot of research by people involved in transportation planning. In accepting the modeling techniques reviewed here, one must make a clear distinction between the price of time as a resource and the value or price of saving time.

The first of these stems from the fact that the consumer may regard time as a scarce resource (e.g., the constraint in Eq. 6) and expresses the willingness to pay to have an additional unit of time were this possible.

The value of time saving concerns the willingness of the consumer to pay to have time reduced in one activity in order to allocate it to some other activity. In principle there is no reason why this price should not vary from activity to activity or from consumer to consumer and be either higher or lower than the price of time as a resource.

Utility and Demand in Deterministic Models

We can now assess the fruitfulness of the algebraic modeling approach. Starting from the demand functions described earlier, we conclude that there are n income responses and n^2 price responses that are of immediate interest to the analyst. That is, data for estimation purposes must be sufficient to yield $n(n + 1)$ pieces of information if the demand equations are to be estimated without further a priori information. The properties of these demand functions come in handy in this context because the data needs are considerably reduced as a consequence of the a priori restrictions imposed by these properties.

The homogeneity property gives n restrictions, the adding-up property gives $n + 1$ restrictions, the symmetry property gives $\frac{1}{2}n(n - 1)$ restrictions, and the negativity property gives n inequalities. If we ignore the inequalities, the unrestricted $n(n + 1)$ responses are thus reduced to $(n - 1) \cdot (\frac{1}{2}n + 1)$, and that obviously is a considerable improvement with respect to basic data needs. Still, however, there are likely to be too many simply because n is usually large and data are seldom plentiful.

As a consequence, our discussion of some explicit demand functions will be related both to the consistency aspect and to the question of practical application. I intend not to provide a complete list of demand models applied in transport economics but to compare basic differences in the behavioral structure of a few frequently used models in comparative statistics. By far the simplest demand function to be used is

$$X_i = b_i(R/P_i), \text{ for } i = 1, \ldots, n \qquad (11)$$

where R = income or total expenditure, and X_i and P_i have been defined already. Obviously, the coefficient $b_i \geq 0$ and $\sum_{i=1}^{n} b_i = 1$. This set of functions implies a utility function of the form

$$U = \prod_{i=1}^{n} X_i^{\beta_i} \qquad (12)$$

where β_i = structural coefficient. This model implies the following demand properties:

1. All budget elasticities of demand are unity implying straight Engel curves through the origin (Engel curves express demand solely as a function of the consumer's income);
2. Expenditure on each commodity is a constant, and when R is given all the own-price-elasticities are equal to -1;
3. It follows from the specification (Eq. 11) that all cross elasticities are 0; and
4. The Slutzky equations

$$\frac{\partial X_i}{\partial P_j} = \left[\frac{\partial X_i}{\partial P_j}\right] U = \text{constant} - X_j \cdot \frac{\partial X_i}{\partial R} \tag{13}$$

make it clear that because of properties 1 and 3

$$\left[\frac{\partial X_i}{\partial P_j}\right] U = \text{constant} > 0$$

which means that all pairs of commodities are net substitutes.

The model (Eq. 11) is clearly inconsistent with the empirically well-established Engel's law, which states that the proportions of the budget devoted to certain groups of commodities vary considerably as the budget changes (7, p. 1173). This is a strong argument against the application of the model.

Another simple (from the econometric point of view) class of demand functions are those that are linear in $P_1, \ldots, P_n$, R (or can be transformed into a linear form). The most obvious is

$$X_i = \sum_{j=1}^{n} \frac{P_j}{P_i} a_{ij} + b_i \frac{R}{P_i} \tag{14}$$

where a_{ij} = structural coefficients, for i, j = 1, ..., n. The theory of consumer demand developed in the above implies that there exist numbers $s_1, \ldots, s_n$ such that Eq. 14 can be written as

$$X_i = s_i + \frac{b_i \left[R - \sum_{j=1}^{n} P_j \cdot s_j\right]}{P_i} \tag{15}$$

This model, developed by Stone, is known as the linear expenditure system, and has been one of the most important in empirical demand studies (15, pp. 315-318).

Equation 15 says that expenditure on commodity i can be divided into 2 parts: the purchase of a fixed quantity s_i(survival minimum) and a constant fraction b_i of what is left after all the bare survival quantities of all commodities have been bought. The demand functions imply a utility function of the form

$$U = \prod_{i=1}^{n} (X_i - s_i)^{b_i} \tag{16}$$

This model implies that, if $R > \sum_{i=1}^{n} P_i s_i$, then all commodities are normal (positive income elasticities), all pairs of commodities are net substitutes (see definition below Eq. 13), and the demand for each commodity is inelastic with respect to its own price (15, pp. 315-318).

The model is capable of behaving more in accordance with Engel's law than the simple model (Eq. 11). Although the Engel curves still are straight lines, they do not pass through the origin but rather through the point $s_1, \ldots, s_n$. Thus, it is perfectly possible for the budget share for food, for example, to decrease as the budget increases.

Another demand model often seen in applied transport economics is the log-linear.

$$\log X_i = \log C_i + \sum_{j=1}^{n} a_{ij} P_j + b_i \log R \tag{17}$$

where C_i = constant, for $i, j = 1, \ldots, n$. This is a constant elasticity model where the homogeneity restriction requires $\sum_{i=1}^{n} a_{ij} = -b_i$. The difficulty with Eq. 17 and several other applied demand models is that it is either extremely difficult or impossible to find a traditional static algebraic utility model from which the complete set of demand functions chosen can be derived. [Several papers contain an exercise in deriving explicit demand functions from utility functions (4, 9).]

Only the trivial case where $b_i = -a_{ii}$ and $a_{ij} = 0$ for $i \neq j$, which turns Eq. 17 into Eq. 11, is capable of making this derivation easily come through. Equation 17 can, however, almost be derived from an indirect additivity type of utility model (7, p. 1204).

Given the direct utility function in terms of X_i and the logically derivable demand functions $X_i(R, P_1, \ldots, P_n)$, an indirect utility function

$$U = U[X(R, P)] \tag{18}$$

is implied that relates the maximum utility attainable to the exogenously determined level of prices and income. Since any such function can be interpreted as the dual of the direct utility function, minimizing it subject to given P and R will lead to the demand equations.

Only in a special case will directly and indirectly additive utilities occur in the same model. Assuming additive indirect utilities

$$U = \sum_{i=1}^{n} U_i \left(\frac{P_i}{R}\right) \tag{19}$$

implies strong behavioral constraints. Brown and Deaton (7, p. 1201) showed that for all indirectly additive models the uncompensated cross-price elasticities are identical for all goods affected and depend only on the good whose price has changed:

$$(\partial X_i/\partial P_j)(P_j/X_i) = (\partial X_k/\partial P_j)(P_j/X_k), \text{ for all } i, k \neq j \tag{20}$$

It is, however, worth noting that Brown and Deaton (7, p. 1203) conclude that in all relevant respects the linear expenditure system, which implies linear demands, is superior to the indirect additive utility model.

The last travel demand model to be commented on is closely related to a well-known variant of the gravity formula, which can be derived from entropy maximization (30).

$$X_{ik} = A_i B_k e^{-C_{ik}/\beta_0} \tag{21}$$

where

X_{ik} = total travel (for all households) from i to k,
A_i = number of households at i,
B_k = structural coefficient, and
C_{ik} = cost of a round trip from i to k.

Assuming an integrated logarithm utility model,

$$U = U_o + \sum_k [(\beta_k + \beta_o) X_k - \beta_o X_k \log X_k] \quad (22)$$

where

U_o = constant,
$\beta_o, \ldots, \beta_k$ = structural coefficients, and
X_k = number of household trips to k.

Beckmann and Golob (4) have shown that Eq. 22 leads to

$$X_k = e^{\beta_k/\beta_o - C_{ik}/\beta_o} = \beta_k \cdot e^{C_{ik}/\beta_o} \quad (23)$$

If all households at i have identical utility functions, the aggregate gravity formula (Eq. 21) follows. This particular model does, however, violate the nonsaturation axiom since the marginal utilities $\partial U/\partial X_k$ approach $-\infty$ for large values of X_k.

Equation 22 is an additive utility model that leads to very simple demand functions (Eq. 23). No cross effects are assumed to exist; consequently, the model is incapable of dealing with some of the most urgent policy problems in transportation planning today.

Our review of algebraic demand models has indicated that either practical models are directly based on the theory or they are designed so that one or more of the theoretical properties can be subjected to empirical testing. Despite the common basis in algebraic utility theory, the models in use may appear surprisingly dissimilar and reflect quite different assumptions regarding the reactions of the decision unit to price and income changes resulting from policy decisions.

As a consequence, one should be somewhat careful when postulating econometric travel demand models. More efficient models may result once the aim of the study is clearly defined and the behavioral assumptions on which the explicit model is to be based have been chosen. The first question the analyst will face in choosing his set of assumptions is, Will the choice of assumptions influence the major conclusions to be drawn from the analysis? Only a couple of such problems frequently faced by travel demand analysts are discussed here. Should a theory of travel choice behavior be mode specific or mode abstract? Are separability assumptions acceptable? Are sequential choice assumptions acceptable?

The first question has been discussed in previous works (3, 6, 9, 26, 28). A mode-specific model treats each travel mode as a specific commodity with its own demand schedule. This is principally in line with the traditional theory of consumer choice. A mode-abstract model, on the other hand, regards travel by different modes between 2 points in space as distinct observations appearing in the same econometric equation. This modeling approach is philosophically in line with Lancaster's characteristics approach (21).

Assuming that the same independent variables are all relevant and the only relevant variables in both models, we can construct a mode-abstract model to be a special case of a mode-specific model including several modes, provided the regression coefficients of each explanatory variable can be assumed to be mode independent. This hypothesis can be tested by means of Chow's equality test (10), which can be applied to the mode-specific model.

The assumption of separability is essential when the travel market is analyzed alone or when travel and housing are treated as one commodity subgroup to be distinguishable as a group of commodities and services. We have earlier indicated that acceptance of the household production functions in utility models justifies the use of weakly separable utility functions in demand studies. If we can also assume that the household production functions are homogenous to the first degree, the number of parameters of the family of demand functions for the market goods is drastically reduced, and simpler and more manageable demand (choice) models become available (27).

Whether to simplify further by introducing even stronger separability assumptions is a question of the trade-off between realism and computational ease, although shortage of data may force further simplification. One should, however, always ask what consequences further separability assumptions will have on analytical conclusions. The following types of separability are frequently found in the literature (15).

1. Pearce separability implies that the marginal rate of substitution (MRS) between any 2 goods within a given group (travel and housing or perhaps only travel) is independent of the quantity of any good but those 2.
2. Homogenous separability (want independence) implies homothetic indifference surfaces for a given group with respect to origin. In other words, the demand elasticities of each good within the group with respect to expenditure on the particular group (travel) is unity. This particular type of separability requires that one must never group luxuries, near luxuries, and necessities.
3. Strong separability implies that the MRS between any 2 goods in any distinct groups (travel and food) is independent of any good in any third group (clothing).
4. Additive separability implies the existence of continuous functions $v_1, \ldots, v_n$ such that, for all feasible X, $U(X) = \sum_{i=1}^{n} v_i(X_i)$.

It is frequently assumed that "what consumers in fact do is to set aside or commit sums of money for broad general purposes, and decide at the appropriate time on the detailed disposition of these sums" (15, p. 153).

Separability assumptions give a further possible justification for such a budgeting procedure in which the decision to commit a sum of money to a particular purpose is taken, not on the basis of detailed knowledge or prediction of the prices of individual goods on which it is to be spent, but rather on a notion of the general level of those prices.

Green (15, pp. 154-156) shows that only homogenous separability will meet the requirement that a 2-stage budgeting procedure of the "within-group type" be consistent in the sense that it leads to the same optimal vector of quantities as if one had found directly the quantities by means of the general 1-stage budgeting procedure in traditional choice analyses.

From the outline earlier in this paper, the reader may be bothered by one of the implications of homogenous separability—that of unity demand elasticities with respect to group budgets. However, the budget constraints may be adjusted so that a linear expenditure model appears. As indicated earlier, this model implies demand elasticities with respect to expenditure that may perfectly well be consistent with Engel's law.

Having chosen among the various degrees of separability to justify simpler models, a traffic analyst may be faced with the next question, Are the behavioral theories reviewed compatible with a specific order in which the various travel choices follow each other?

Although to introduce separability assumptions on theoretical grounds seems worthwhile, the question of sequential assumptions is hardly compatible with the static models reviewed above. One may perhaps argue that the sequence chosen in most urban transportation planning models is a consequence of the time horizon relevant to each choice. That is, choice of home residence is a long-run decision to the household, whereas choice of route to travel along is a short-run decision. Consequently, different sets of variables should explain these choices, and treating them separately may be both practical and theoretically justifiable. However, the question of which choice is made first does not seem to be compatible with the static algebraic choice models above.

Even though static theory seems to be incompatible with particular travel choice sequences, a sequential estimating procedure may be strongly recommended if it can be shown that the parameter values to be estimated will not be influenced by the choice of choice sequence. Such independence is perhaps present when the decision-maker is facing very simple decisions, and perhaps homogenous separability can be assumed. In such a case, the practical model can be significantly simplified, and research to clarify the role of the particular sequence chosen ought to be given high priority before decisions are made with regard to further model developments.

If parameter estimates can be shown to be sensitive to the choice of sequence in traffic models, this should lead planners to seriously reconsider the use of present urban transportation models in selecting transportation policies and perhaps to concentrate on developing dynamic utility maximizing models based on a utility tree approach.

PROBABILISTIC ANALYSES OF CHOICE

Foundations

Efforts to test the validity of algebraic choice theory have not provided it with an overwhelming amount of support. One possible explanation is that observable conclusions of the theory have not been correctly interpreted in light of the data base used in testing. The consumer may certainly misjudge his actual preferences or permit them to be altered by random shocks. By recognizing such possibilities, analysts may give new implications of utility maximization to provide more appropriate foundations for empirical tests.

The assumption should be that consumer behavior has a probabilistic consistency and not a deterministic consistency. Several recent authors have approached the analyses of choice behavior by describing it as a probabilistic rather than a deterministic phenomenon. Two basically different theories may form the basis for a probabilistic choice theory.

One deals with a consumer whose preferences obviously exist and can be assumed to be fixed, but he himself is not completely aware of what they are. Nevertheless he must still make decisions even when facing such uncertainty. On such occasions the consumer cannot always be expected to pick the utility maximizing bundle from his budget set. The consumer makes errors in determining his optimal commodity bundle. The probabilistic models developed on this basis are referred to as fixed preference models.

Alternatively, suppose that the consumer's preferences themselves are subject to random shocks. Thus, a sudden traffic accident may increase his desire relative to other commodities for better safety devices, or a sudden inconvenient delay may change his commuting pattern. Randomness is present, but for a different reason than in the fixed preference models, and models based on these premises are referred to as random preference models. Katzner (19, pp. 161-167) has briefly formalized the distinction between these 2 basic approaches.

In the fixed preference models, each choice does not necessarily represent a utility maximizing point in commodity space. Consequently, the functions relating the chosen commodity bundles to prices, income and the random term that shows deviations from optimum choices, cannot always be interpreted as demand functions. A fixed preference model may be required to yield as a result of repetitive choices an "average" commodity bundle compatible with the utility maximizing bundle. Observing only one choice in commodity space that violates basic demand properties consequently does not suffice to refute demand theory.

Assuming a random preference model implies random demands since each choice is such that maximum utility is attained. Empirically, even such a model may lead to the perhaps incorrect rejection of demand theory.

Using data for a limited time period may yield irrationality as a conclusion, although the reason for the observed changes in behavior is due to the random elements influencing the consumer's preferences. In reality, each choice comes from a different utility maximizing relation.

To model this assumed optimizing behavior in practice is an extremely difficult task, and practical choice models have therefore chosen a much simpler point of departure. Rational behavior within the framework of a pure theoretical random preference model may contradict the basic axioms in the more pragmatic probabilistic choice models to be presented in the next section. Hildenbrand (18, pp. 414-420) has, for example, shown that rational behavior on the average within a rather general pure theoretical random preference model may be inconsistent with the basic choice axiom of Luce to be discussed later. Stochastic transitivity, according to Marschak's meaning (27, p. 318), is another assumption violated by the rational individual in Hildenbrand's theory.

Probabilistic Choice Models in Practice

In the application of demand or choice theory to practical problems, the usual procedure is to define a limited time period of analyses for which cross-sectional choice data are collected from a random sample of individuals. The preferences of the individuals in this population can be described partly in nonrandom terms that reflect representative tastes and partly in random terms that reflect individual idiosyncracies in taste, whatever the reasons for these are.

One fairly general model of this kind has been presented by McFadden (25, pp. 9-11). An individual in the population faces J alternatives, each described by a vector of attributes X_j. The individual has a utility function that can be written in the form

$$U = V(X) + E \tag{24}$$

where V is nonrandom reflecting representative population tastes and E is random reflecting the individual idiosyncracies in tastes for each attribute vector X. The probability that an individual drawn at random from the population will choose alternative i among the J alternatives, then, equals

$$\begin{aligned} P_i &= P_r\,[V(X_i) + E_i \geq V(X_j) + E_j, \text{ for all } j \neq i] \\ &= P_r\,[E_j - E_i < V(X_i) - V(X_j), \text{ for all } j \neq i] \end{aligned} \tag{25}$$

Charles River Associates (9) showed that explicit models based on the assumption that each individual maximizes his utility and further based on Eq. 25 are derivable from a probabilistic choice theory first developed by Luce (22) and Marschak (24).

The basic starting point in this theory of individual choice behavior is a choice axiom (25, p. 7). The most important implication of this choice axiom is the independence-of-irrelevant-alternatives condition. Originally developed by Arrow (1) in an algebraic context, this condition is that a comparison of 2 alternatives according to some algebraic criterion like preference should be unaffected by the addition of new alternatives or the subtraction of old ones (recall the various separability definitions given earlier).

The probabilistic version of this axiom should require that the ratio of the probability of choosing one alternative to that of choosing the other not depend on the total set of alternatives available, and this is exactly what is implied by Luce's choice axiom. Only the ratio of the 2 probabilities and not the probabilities themselves is invariant to changes of the irrelevant alternatives (note the similarity to the concepts of separability discussed above).

Another property of Luce's choice model regards transitivity. The choice axiom is a probabilistic version of the transitivity axiom in deterministic choice theory. The Luce model can also be shown to imply the existence of a ratio scale (22, p. 23) that is unique except for its unit and independent of any assumptions about the structure of the set of alternatives. Let T be a finite set such that, for every $S \subset T$, P_s is defined. Let the elements in T be the numbers 1, 2, ..., i, j, ..., J. Then,

$$P_s(i) = \frac{U(i)}{\sum_{j \in S} U(j)} \tag{26}$$

Our interest is now concentrated on the explicit probability function for the random variables E_i and E_j and the random variable $E = E_i - E_j$.

Assuming that $V(X_i)$ and $V(X_j)$ are linear in their unknown parameters, we can show that a wide variety of functional forms for the probability function are consistent with random utility theories of binary individual choice. The set includes the frequently used logit, probit, and truncated linear models. The logit model results if (a) $E = E_i - E_j$ has the logistic cumulative distribution and (b) E_i and E_j are statistically independent of the identical reciprocal exponential distribution, which is a distribution frequently used in the study of extreme values (20, pp. 332 and 344),

$$P_r(E_i \leq E^*) = e^{-e^{-E^*}} \tag{27}$$

The logit in the binary case is defined as

$$\log \frac{P_{ij}(i)}{1 - P_{ij}(i)} \tag{28}$$

and the following probability function is derived:

$$P_{ij}(i) = \frac{1}{1 + e^{-E}} \tag{29}$$

where $P_{ij}(i)$ means the probability of choosing i from a set $\{i, j\} \subset T$. Substituting from Eq. 25, we can write

$$P_{ij}(i) = \frac{1}{1 + e^{-[V(X_j) - V(X_i)]}} = \frac{1}{1 + \dfrac{e^{V(X_i)}}{e^{V(X_j)}}} = \frac{e^{V(X_j)}}{e^{V(X_j)} + e^{V(X_i)}} \tag{30}$$

Choosing e^V as the explicit form for the positive valued function U in Eq. 26 reveals that the logit model is consistent with Luce's ratio scale, which Marschak's has called the strict-utility function (24, p. 322).

Before evaluating this particular choice model, let us examine a more general approach, multiple choices. Assuming $V(X_j)$ and $V(X_i)$ are linear in their unknown parameters α and assuming the distribution properties of the random terms E_i are the same as in the binary case result in the multinominal logit formula (9, pp. 5.15-5.28):

$$P_J(i) = \frac{e^{\alpha' X_i}}{\sum_{j=1}^{J} e^{\alpha' X_j}} \tag{31}$$

This model may be called an explicit form of the strict-utility function in the multiple choice sense (24, p. 324).

It has been established (9, p. 5.19) that the assumption that the random utility function has a reciprocal exponential distribution is equivalent to the independence-of-irrelevant-alternatives axiom. This means that the odds $P_J(i)/P_J(k)$ of choosing alternative i over alternative k are independent of the presence or absence of third alternatives. This is easily seen if we look at the model in the following way:

$$\frac{P_J(i)}{P_J(k)} = \frac{\dfrac{e^{V_i}}{\sum_{j=1}^{J} e^{V_j}}}{\dfrac{e^{V_k}}{\sum_{j=1}^{J} e^{V_j}}} = e^{V_i - V_k} \tag{32}$$

from which the multinominal logit follows:

$$\log \frac{P_J(i)}{P_J(k)} = V_i - V_k = \alpha'(X_i - X_k) \tag{33}$$

The probabilistic choice model (Eqs. 24 and 25) can also be derived from a different set of assumptions regarding individual choice behavior. Point of departure is the economics-of-time model in the first part of the paper. To use this model in analyzing aspects of travel choice, let X_1 be the number of visits to a particular spot, X_2 the number of car trips, and X_3 the number of transit trips. Consequently, $X_1 = X_2 + X_3$. Individuals are assumed to maximize utility and will always choose the mode with which the largest Lagrangian value (derived from first order condition for maximum utility) is associated.

Introducing a set of rather strict separability assumptions can show that the individual chooses the alternative with the lower generalized cost. If individuals are drawn at random from a population, a random element should be added to the generalized cost formulas. By assuming the same statistical properties as for the strict utility choice model, Bruzelius has shown that the same explicit econometric choice models are derived (logit, probit, and so on).

The extensive analyses by Charles River Associates referred to above reject the multiple-choice generalizations of random utility models, where other probability distributions of the random utility elements are assumed, as analytically intractable or otherwise impossible to work with. (Charles Lave, University of California, Irvine, has in private communication expressed the same ranking based on computational efficiency.) For this reason, the multinominal extensions of the frequently used probit model and the truncated linear model are not discussed in this paper. An extensive discussion of these models is given in another report (9).

The conclusion to be drawn, before further discussions of the strengths and weaknesses of the logit model, is that the binary logit model is the only binary probability model for which the multinominal extension is practical at present.

The reliance on the choice axiom, which makes the model rest on the independence-of-irrelevant alternatives condition, is the principal strength as well as the principal weakness of the logit model. A similar conclusion follows from a critical examination of the separable time-allocation model developed by Bruzelius.

There is nothing in the separability property implied by Luce's axiom above that limits the discussion to subsets regarding only the various aspect of travel choice. The independence property holds for any subset, and the analogy to separability in deterministic utility models should be noted. The weakness of relying on this independence assumption in the logit model is not necessarily worse than relying on a similar separability assumption in algebraic choice models.

Luce evaluates his choice axiom in concluding his analyses of individual choice behavior (25, pp. 131-134). It could well happen that the basic choice axiom will hold when a situation is analyzed one way but not when it is viewed another way. The problem in practice is to know when a subject decomposes a decision into 2 or more stages; this is again the problem of knowing how a subject conceives the alternatives. The validity of this probabilistic choice theory seems to depend on the definition of alternatives, and alternatives should be defined in such a way that a subdivision of the decision into

2 or more stages is of no importance with respect to the final result. The only purpose of such a staging procedure should be to simplify the practical work with the model. These comments are very similar to those found in the section on the separability properties of deterministic choice models. That independence may be an implausible, strong assumption is excellently illustrated in the Charles River Associates report (9, pp. 5.25-5.26):

> Suppose an individual faces the alternatives of one auto mode and one bus mode, and chooses the auto mode with probability 2/3. Now suppose a second bus mode is introduced which follows a different route, but has essentially the same attributes as the first bus mode. Intuitively, the individual will still choose the auto mode with probability 2/3, and will choose either of the bus modes with one-half the probability 1/3 of choosing some bus mode, or 1/6. However, the independence of irrelevant alternatives condition requires that the relative odds of choosing the auto mode over either of the bus modes be 2 to 1, implying the probability of choosing the auto mode drops to 1/2 and the probability of choosing each bus mode is 1/4. The reason this result is counter-intuitive is that we expect the individual to lump the two bus modes together, not treat them as "independent" alternatives.

This example suggests that application of the strict-utility model should be limited to multiple-choice situations where the alternatives can plausibly be assumed by the decision-maker to be distinct and independent. Care must then be taken in specifying the available alternatives and decision-making structure when this multiple-choice model is used.

A simplifying consequence of the strict-utility model is that new modes, routes, or destinations may be introduced without recalibration of the model once the parameters have been estimated. The new choice aspects are introduced simply by the addition of new terms to the denominator of the particular strict-utility function in question.

The consequence of the choice axiom is that the odds with which the previous alternatives are selected are independent of the introduction of new alternatives. The probabilities of choosing the previous alternatives will, of course, decrease when new alternatives appear, but the old odds remain unchanged as a consequence of the independence of irrelevant alternatives (9, pp. 5.20-5.23).

This model property is a valuable simplifying aspect provided the problem on which the model is applied is considered simple enough to be modeled by means of this probabilistic choice approach.

The Charles River Associates model, referred to above, uses the independence property in a way that leads to an indirect travel demand model. This means that the logit model is applied to each choice, and a particular choice sequence is implied. Tests of parameter sensibility to alternative choice of sequences are not plentiful, and the procedure is difficult to evaluate against a simultaneous approach where the logit model is applied to the joint probability of the various travel aspects. However, Ben-Akiva (5) certainly confirms the suspicion that parameter estimates seem to be sensible to choice of sequence. It seems natural, therefore, to approach the sequencing problem from a decision-tree point of departure and, thus, have the preferred sequence as a result of utility maximization. This would, in principle, do away with the purely technical problems of different sequences leading to different parameter estimates.

The final point to make in this evaluation is perhaps rather academic but nonetheless of interest to those working with the econometrics of travel-choice analyses. Two different theoretical developments have been shown to yield the same econometric choice model. Unless the behavioral assumptions are explicitly stated from the very start, the logit model will be underidentified in the sense that further a priori information is needed to tell what we are really "explaining" by means of the econometric model.

CONCLUSION

Algebraic theories have been a basic point of departure in formulating choice and demand models. The discussions regarding stochastic or absolute consistency in choosing have to some extent been confused by too little precision in formulating probabilistic models. Recent works (14, 18, 19, 25) have clarified these aspects of choice analyses, and the present state of the art indicates that only probabilistic models are worthwhile in practice. As long as such models are developed from a rather traditional micro-economic platform, there is not so much theoretical evidence in favor of the Luce-Marschak approach as one may think by studying the present travel choice literature. The importance of the Luce-Marschak models has primarily been to challenge economists to take another look at the world, and as a consequence the theoretical basis of probabilistic choice models now in use has been clarified with respect to strengths and weaknesses and certain desirable simplifications have been theoretically justified. This regards first of all certain weak assumptions of choice independence or separability. Variables should be carefully defined so that the separability properties necessary for model operation do not violate the realism of the model. The sequences chosen in present travel demand and choice models do not seem well founded in theory, and it is expected that this field will be looked into more closely in the future. The outcome of this work may be a sounder theoretical basis as justification for particular sequences or a switch to simultaneous models.

ACKNOWLEDGMENT

The author is deeply indebted to Åke Andersson, University of Gøteborg; Nils Bruzelius, University of Stockholm; Frank Koppelman, Massachusetts Institute of Technology; Antti Talvitie, University of Oklahoma; John Pipkin, Northwestern University; and Ole H. Hansen, Møre og Romsdal distriktshøgskole, for their many invaluable suggestions and corrections. The opportunity to discuss an early draft of the paper with the staff members of the Transportation and Urban Analyses Department at the General Motors Research Laboratories led to some significant improvements and is highly appreciated. Needless to say, all remaining mistakes are the sole responsibility of the author. Research that led to this paper was made possible by funds provided by Bank of Norway's Economic Research Fund and The Memorial Fund of Professor Wilhelm Keilhau.

REFERENCES

1. Arrow, K. J. Social Choice and Individual Values. Wiley, New York, 1951.
2. Becker, G. A Theory of the Allocation of Time. Economic Journal, 1965.
3. Beckmann, M., and Golob, T. A. A Utility Model for Travel Forecasting. Transportation Science, 1971.
4. Beckmann, M., and Golob, T. A. On the Metaphysical Foundations of Traffic Theory: Entropy Revisited. Paper presented to the 5th International Symposium on Theory of Traffic Flow and Transportation, Berkeley, Calif., 1971.
5. Ben-Akiva, M. E. Structure of Passenger Travel Demand Models. M.I.T., PhD thesis, 1973.
6. Brand, D. Travel Demand Forecasting: Some Foundations and a Review. HRB Spec. Rept. 143, 1973, pp. 239-281.
7. Brown, A., and Deaton, A. Surveys in Applied Economics: Models of Consumer Behaviour. Economic Journal, 1972.
8. Bruzelius, N. Economic Theories of the Allocation of Time: A Survey. University of Birmingham, M. Soc. Sci. dissertation, 1972.
9. A Disaggregated Behavioral Model of Urban Travel Demand. Charles River Associates, Inc., Cambridge, Mass., Rept. 156-2, 1972.
10. Chow, G. C. Tests of Equality Between Sets of Coefficients in Two Linear Regressions. Econometrica, 1960.
11. de Donnea, F. X. The Determinants of Transport Mode Choice in Dutch Cities. Rotterdam Univ. Press, 1971.
12. DeSerpa, A. C. A Theory of the Economics of Time. Economic Journal, 1971.

13. Georgescu Roegen, N. The Pure Theory of Consumer Behaviour. Quarterly Journal of Economics, 1936.
14. Georgescu Roegen, N. Threshold in Choices and the Theory of Demand. Econometrica, 1958.
15. Green, H. A. J. Consumer Theory. Penguin Modern Economics, 1971.
16. Gronau, R. The Value of Time in Passenger Transportation: The Demand for Air Travel. NBER, New York, Occasional Paper 109, 1970.
17. Hensher, D. Review of Studies Leading to Existing Values of Travel Time. Commonwealth Bureau of Roads, Australia, 1973.
18. Hildenbrand, W. Random Preferences and Equilibrium Analyses. Journal of Economic Theory, 1971.
19. Katzner, D. W. Static Demand Theory. Macmillan Series in Economics, London, 1970.
20. Kendall, M. G., and Stuart, A. The Advanced Theory of Statistics. C. Griffin and Co., London, Vol. 1, 1969.
21. Lancaster, K. Consumer Demand: A New Approach. Columbia University Press, New York and London, 1971.
22. Luce, D. Individual Choice Behavior. Wiley, New York, 1959.
23. Malinvaud, E. Lectures in Microeconomic Theory. North Holland Publ. Co., Amsterdam and London, 1972.
24. Marschak, J. Binary Choice Constraints and Random Utility Indicators. In Mathematical Methods in the Social Sciences (Arrow, K. J., Karlin, S., and Suppes, P., eds.), Stanford University Press, 1959.
25. McFadden, D. The Theory and Measurement of Qualitative Consumer Behavior: The Case of Urban Travel Demand. University of California, Berkeley, Working Paper 1, Sept. 1972.
26. McGillivray, R. G. Demand and Choice Models in Modal Split. Journal of Transport Economics and Policy, 1970.
27. Muth, R. Household Production and Consumer Demand Functions. Econometrica, Vol. 34, No. 3., 1966.
28. Ruijgrok, C. J. Unpublished note. Netherlands Institute of Transport, Rijswijk, Feb. 1973.
29. Uzawa, H. Preference and Rational Choice in the Theory of Consumption. In Mathematical Methods in the Social Sciences (Arrow, K. J., Karlin, S., and Suppes, P., eds.), Stanford University Press, 1960.
30. Wilson, A. G. Entropy in Urban and Regional Modelling. Pion Ltd., London, 1970.