Approaches to Modelling Consumer Preferences and Demand in Transportation

D. A. Hensher & P. B. McLeod

ABSTRACT

This paper is intended to review the approaches presently used in investigations of travel choice and demand, with particular emphasis on the responsiveness of methods to many of the current issues in transportation planning and policy. In view of the emphasis placed upon traffic restraint in overall policy, transportation studies are unable to predict how many people will be deterred from travelling under certain restrictions. A major aim is the highlighting of the behavioural approach as a more realistic approach to explaining and predicting traveller behaviour in contrast to the conventional urban travel demand model system.

INTRODUCTION

The movement of individuals in urban time and space, under the short run assumption of fixed residential and employment locations, contributes significantly to the total costs incurred by the individual, his family and the community in general. Scarce private and social resources are devoted to facilitating this urban movement. Any attempt to achieve a more desirable allocation of scarce resources between transport and non-transport uses and within transport, will be substantially

* We would like to thank the Bureau of Roads for providing the right working atmosphere during the period 1972-74 for investigating consumer preferences in urban trip-making.

aided by a greater understanding of the nature of the demand for travel (and the demand for various forms of passenger transport in particular).

In an attempt to 'explain' the travel demand process, and associated acts of choice,[1] multi-disciplinary research and planning efforts have, over the last twenty years, developed improved techniques (as research or operational tools) to explain various components of travel demand at various levels of regional segmentation.

The paper considers the properties of traditional and contemporary urban transport models in the area of travel demand models. The implications, problems and limitations of these models are discussed. Thus the paper essentially focuses on the structure of travel demand models. Several methods exist which are possible and sensible ways to organize a discussion of analytical structures, for example:

Simultaneous	versus	Sequential
Aggregate	versus	Disaggregate
Behavioural	versus	Non Behavioural
Deterministic	versus	Probabilistic
Direct	versus	Indirect
Choice Specific	versus	Choice Abstract

The main classifications used in this paper will be Aggregate versus Disaggregate and Simultaneous versus Sequential. Discussions along the lines of the other criteria

1. Exclusion of the land-use activity input results in a non-explicit modelling of travel as a derived demand commodity. Recently an attempt has been made to develop a procedure to explicitly consider land use decisions and build these into a theory of travel demand. See Brand (1970).

will be mentioned but will be subordinate to our main classification procedure. The models discussed are static and involve demand and choice only. Concepts of interdependence, system performance (supply) and land use are not discussed. In the discussion of each structure, mode choice and modal split are emphasised.

Approaches used at present in the operational context are essentially the traditional urban area wide approaches. Whilst these procedures have been subject to considerable modification to allow for improved functional forms and estimation techniques, the principles underlying the modelling procedures have changed very little. For example, travel behaviour or demand is still modelled as a series of sequential independent choices of trip generation and attraction, trip distribution, modal split and route assignment. Land use forecasting precedes travel forecasting as a separate step. This procedure will be referred to as the Urban Transportation Model System (UTMS). The discussion below of the various analytical structures, some of which are in their infancy, is designed to show that the entire demand model system needs restructuring to take account of the new set of requirements needed in traffic models. In particular, the model of travel demand must be behavioural and policy sensitive if it is to improve our ability to predict travel behaviour (short run and long run) under various investment programmes and to influence travel behaviour in desirable directions.

AGGREGATE DEMAND MODELS

The following brief summary of traditional consumer theory will aid the discussion of aggregate models. We begin with the utility function of the individual. The level of utility for the individual is a function of the quantity of goods consumed, q_x

$$U = U(q_1 \ldots q_n)$$

and this is subject to the budget constraint

$$\Sigma\, p_x\, q_x < Y$$

where p_x is the price of commodity x and Y is total money income. Since the q_x are uniquely determined by the p_x and Y, we can write the direct utility function $U(q_1, \ldots q_x, \ldots q_n)$ as the indirect utility function,

$$U = U(p_1, \ldots p_x, \ldots p_n, Y)$$

From this indirect utility function for the individual we can derive the individual's demand function which will be of the form

$$q_x = q(p_1, \ldots p_n, Y, T)$$

where q_x = quantity of commodity x, $p_1, \ldots p_n$ is a vector of commodity prices for x and its substitutes, Y = income and T = tastes (assumed constant). This is the demand function of an individual. Summation of these demand functions is possible in order to achieve market demand functions. Most approaches involve sectioning the market into different income (if you like, socio-economic) groups, tracing the effect of socio-economic variable changes on demand to them and aggregating back again. This gives an economic market demand function of the form, $Q = Q\,(P,S)$ where P is a vector of prices, S is a vector of socio-economic variables and $Y \varepsilon S$.

To adopt this general formulation to transport demand requires that we recognize transportation as a derived demand commodity. It is due to a demand for other goods, the consumption of which requires travel. In recognition of this, usually implicitly, trips are typically classified according to trip purpose (class of final good) and the demand for transport for each purpose is modelled separately.[1]

Furthermore when modelling aggregate travel demand the dependent variable is a volume of trips and a set of variables indicating spatial opportunities is added (A). Thus the demand function is of the form $To\overset{n}{d}m = go\overset{m}{d}m$ (L,S,A,) where $To\overset{n}{d}m$ is the volume of trips from origin o to destination d on mode m for purpose n, and L is a vector of level of service variables. For each P there exists a level of service vector $L = (p_x, t_x, cm_x, cn_x)$ reflecting the fact that there are several prices which the consumer pays to travel, these including money fare (p), time (t), discomfort (cm) and inconvenience (cn).

Agregate Sequential Models

As will be seen from above, the UTMS is a prime example of an Aggregate Sequential Model. Several aggregate sequential models have been commonly used in transport planning.[2] The most common sequence is the trip interchange modal split sequence where we sequentially model trip generation, trip distribution, modal split, and route assignment in that order. This corresponds to the choices; to travel, where to travel, what mode to travel on and what route to take.

Trip Generation (a) Production $T_o^n = g_1 (S_o)$

(b) Attraction $T_d^n = g_2 (A_d)$

1. Thusoverall we have a block recursive system whereby mobility choices (e.g. residential location) are modelled first. Given mobility choices, travel choices are modelled independently for each trip purpose.
2. For a good discussion of these, see Weiner (1969).

where T_o^n = volume of trips of purpose n leaving origin o

T_d^n = volume of trips of purpose n arriving at destination d

S_o = socio-economic variables

A_d = attraction variables or activity system variables for destination d.

Typical socio-economic variables include average annual income and the average number of vehicles owned, while typical attraction variables used are zonal populations, zonal employment, and intensity of retail and service employment.

Two structural trip production models have gained wide acceptance, these being the linear multiple regression model and the category analysis model, as methods of predicting the number of trips either generated by a zone or a household. Category analysis involves a simple non-variance specification procedure for categories of households by socio-economic characteristics and the application of associated expected or average trip rates for each category. Car ownership, household income and family structure are usually assumed to be the main determinants of the rate of tripmaking.[1]

<u>Trip Distribution</u> This second sequential step has the functional form:

$$T_{od}^n = g_3(T_o^n, T_d^n, L_{od})$$

where T_{od}^n= volume of trips from origin o to destination d for purpose n

T_d^n , T_o^n = results of trip generation stage

L_{od} = level of service variables between o and d.

A logical constraint on this function is that $\sum_o T_{od}^n = T_d^n$. In some applications an attempt is made to force this constraint because the structure of models used does not guarantee it. After adjustments to the original estimated T_d^n the model equations are applied again and iteration continues until an acceptable correspondence between $\sum_o T_{od}^n$ and T_d^n is obtained. The level of service variable is a composite variable over all

1. A reference to category analysis is Wooton and Pick (1973).

modes between o and d. In many cases however, only travel time by the highway is used as the level of service variable. A generalized cost which is a linear combination of travel time, distance and out of pocket costs is sometimes used.

The two most common functional forms for this sub-model are the gravity model and the intervening opportunities model. A typical version of the gravity model is as follows. (See Meyer and Strazheim 1972, chapt. 4 and Lamb 1970)

$$T_{od}^{n} = \frac{T_o^n \; A_d \; F_{od} \; K_{od}}{\sum_d A_d \; F_{od} \; K_{od}}$$

where T_{od}^{n} = trips produced in o and attracted to d for purpose n.

T_o^n = trips produced in o

A_d = attraction variables for d.

F_{od} = empirically derived travel time factors that are a function of the spatial separation between zones.

K_{od} = specific zone to zone adjustment factors to allow for incorporation of the effect on travel patterns of socio-economic or economic linkage not otherwise accounted for.

The intervening opportunities model is as follows. (Meyer and Strazheim 1972, chapt. 4.)

$$T_{od}^{n} = T_o^n \, e^{-L_n V_d^n} \quad (1 - e^{-L_n V_d^n})$$

where $V_d^n = \sum_K T_k^n$ = subtended volume or the volumes to all destinations already considered, i.e. reached before d.

L_n = empirical parameter representing the constant probability of a possible destination being accepted if it is considered.

k = all destination for which $t_{ok} < t_{od}$; t = time.

Basically the model allocates trips to destinations on an opportunity surface that has been rank ordered according to travel time from origin o.

Modal Split This third submodel is concerned with predicting the number of trips from origin to destination by a given mode. It takes the functional form,

$$T_{odm}^{n} = g_4 \, (T_{od}^{n}, L_{odm}, S_o, A_d)$$

where T_{odm}^{n} = trips by purpose n from origin o to destination d by mode m

T_{od}^{n} = trip produced in o and attracted to d for purpose n and comes from the trip distribution model

L_{odm} = level of service variables for mode m between o and d

S_o = socio-economic variables of o

A_d = attraction variables for d.

Typical socio-economic variables used include automobile ownership and income while level of service variables usually relate to time and cost.

Specifying the functional form has been limited to the following approaches. One of the earliest techniques used was that of the diversion curve whereby the percentage of trips via transit was related to the ratio of (or difference between) travel times on auto and transit.[1] Accessibility indices have also been used to rank the alternative modes where the index measures the ease with which any activity in the urban area can be reached from a particular zone on the specific transport system being considered. The percentage of trips by transit is then related to the ratio of the accessibility indices of the alternative modes (usually auto, transit). (See Weiner 1969). Regression techniques have enabled more complex models to be

1. A good example of the use of this technique is Houston (1969).

estimated of the form

$$P^n_{odm} = \frac{T^n_{odm}}{T^n_{od}} = g_5 \left(\frac{t_{oda}}{t_{odk}}, \frac{C_{oda}}{C_{odk}}\right)$$

or $$P^n_{odm} = \frac{T^n_{odm}}{T^n_{od}} = g_6 \left(t_{odk} - t_{oda}, C_{odk} - C_{oda}\right)$$

where P^n_{odm} = % of travel for purpose n between o and d on mode m (m = automobile, a; or transit, k)

t_{oda}, t_{odk} = travel time, auto and transit respectively

C_{oda}, C_{odk} = cost by automobile and transit respectively

Usually origin zones are stratified according to income level or automobile ownership and linear equations are developed for each subgroup. Several cost and time variables have been used and usually more than one is used. Time has been broken into in-vehicle time, waiting time, transfer time and access time while total cost has been decomposed into out-of-pocket costs, tolls, parking changes, fares, etc. Recently the following functional form has been used. (Shunk and Bouchard 1970.)

$$P^n_{odm} = 1/(1 + e^{h(L_{odm})})$$

where $$h(L_{odm}) = C + \sum_{l} a^l (t^l_{odk} - t^l_{oda}) + \sum_{l} b^l (C^l_{odk} - C^l_{oda})$$

Again times and costs are divided into several variables, while those relative level of service characteristics not measured by time and cost differences are accounted for by the constant (C) and the values of the parameters (a^l, b^l); $h(L_{odm})$ is interpreted as the difference in utility between transit and automobile travel.

In all the variations of the UTMS, route assignment is the last submodel, corresponding to the universal assumption that route choice is the last in the choice sequence. Very often choice of mode constrains choice of route, but when effective choice is available it will be a function of the level of service variables on the routes for each mode.

Aggregate Sequential Models : Problems with Sequential Modelling

Before considering the next model structure some comments are in order on the problems associated with aggregate sequential models.[1] As discussed above, the UTMS discusses the nature of urban passenger travel demand in terms of the decisions whether to travel or not, where to travel and how to travel. The traditional approach[2] to identifying travel demand models the the production of trips, the distribution of trips, mode and route choice as a series of implicit sequential independent choice processes, implying for example in the case discussed above that the decision to travel or not is prior to and bears no relationship to the factors influencing the selection of a method of transport for that trip. The absence of any circular flow or feedback is a major weakness of these earlier initial models despite any subsequent structural improvement to each of the separate demand submodels. Errors in a sub-model are of course compounded by any forward linkage. Trip production is described as a process relating the number of trips commencing or terminating in a particular location (usually a physically defined zone)to the land use and socio-economic characteristics of that location. Since the supply price of the activity for which travel is undertaken is not considered nor is the supply price of travel itself, trip production generates only potential demand. This independently derived area demand for trips is then allocated

1. This discussion has relevance to all sequential modelling and to the problem of aggregation in general.

2. Out of over 200 applications of the UTMS in the U.S.A., the best known studies contributing to methodological development are the Chicago and Detroit Area Transportation Studies.

between possible destinations, modes of transport and routes in a forward linkage manner. This magnitude of total trips is assumed completely unresponsive to the relative supply prices of destinations and modes, these prices only influencing the allocation of the fixed stock of trips. The final mechanism in the process involves the allocation to specific routes of a pre-determined number of trips between an origin and a destination by a particular mode. The actual volume of traffic between origin and destination is assumed perfectly inelastic with respect to mode and route characteristics (level of service variables).

This hierachical structure emphasising the forecasting of volumes of travel (to the neglect of a more fundamental concern with the explanation of the underlying factors that determine the response of the population when confronted with transportation choices) is largely a result of historical accident and computational convenience and lacks "common sense consistency". There is no fixed stock of trips to be allocated among modes, destinations and routes; rather total trip demand is elastic and will respond to changes in time, cost and other variables the traveller considers of influence. It is assumed that no relationship exists between the production and distribution of trips and the quality of the travel network. Areal activity is assumed to be independent of area accessibility and even if this level of activity is constant, the assumption that the propensity to travel is unaffected by the degree of accessibility is difficult to defend. It implies that propensity to travel is independent of present location, planned location and available methods of moving between the locations. The hypothesis of fixed trip ends is clearly difficult to defend. Contrary evidence is provided by the development of urban areas along transport routes, the decline of residential development with distance from the city centre and the response of industrial location to transport investment.[1]

Dissatisfaction with this sequential independent approach initially emanated from the non-engineering disciplines when the transport problem became recognised as a resource allocation problem. This provided a stimulus to the development of models which were premised on the assumption of simultaneous transport choices being a more logical modelling structure.

AGGREGATE SIMULTANEOUS MODELS

These models involve prediction of the volume of trips by origin, destination and mode with a single equation,

$$T^n_{odm} = g_7\ (L_{odM},\ S_o,\ S_d,\ A_o,\ A_d)$$

where T^n_{odm} = trips for purpose n by mode m from origin o to destination d

L_{odM} = level of service variables covering the set of alternative modes M

S_o, S_d = socio-economic variables describing o and d

A_o, A_d = attraction variables for o and d.

Whether the model is mode specific or mode abstract depends on whether the function g_7 is independent of the mode being evaluated. These models represent an attempt to model urban travel demand in a way which takes account of economic theories relating to consumer choice and preferences. Concepts of elasticity and cross elasticity are given prominent attention. A mode specific simultaneous model (Domenich and Kraft 1968) has been estimated in the linear log form and product exponential form. However,

1. Evidence from English data indicated an elasticity of trip production with respect to car travel time of .73 with a standard error of .19. (See Neuburger 1972, Appendix)

the mode abstract aggregate simultaneous model developed by Quandt and Baumol (1966) has been the most influential example of this model structure. The approach is based on the idea that the demand for travel by a mode between origin and destination is not dependent on the type of mode (e.g. car or train) and that besides the influence of general socio-economic and activity system variables the characteristics which describe the level of service each m ode offers will be key variables for explaining and predicting modal patronage. These characteristics will include time, cost, frequency and comfort. Although this approach is generally accepted today it was seen as novel when first used.[1] The model is as follows:

$$T_{odm} = a_o \, P_o^{a1} \, P_d^{\,a2} Y_o^{a3} \, Y_d^{a4} \, C_{odb}^{a5} \, t_{odb}^{a6} \, f_{odb}^{a7} \left(\frac{C_{odm}}{C_{odb}}\right)^{a8} \left(\frac{t_{odm}}{t_{odb}}\right)^{a_9} \left(\frac{f_{odm}}{f_{odb}}\right)^{a_{10}}$$

where T_{odm} = trips (all purposes) between o and d by mode m

P_o, P_d = populations at o and d

Y_o, Y_d = median incomes at o and d

C_{odb}, t_{odb}, f_{odb} = 'best' values of each level of service variable; cost, time, frequency (e.g. quickest, least expensive) between zones o and d regardless of which mode exhibits the best value for any of the variables. This vector constitutes the "Abstract Mode".

C_{odm}, t_{odm}, f_{odm} = level of service variables, time, cost, frequency for mode m between zones o and d. These are divided by the best values to indicate relative performance by mode m.

The model sometimes includes some institutional character indices which reflect the fact that cities (zones) with

1. The characteristics oriented theory of consumer behaviour was rigorously developed by Lancaster (1966 and 1971) although the basic idea had been discussed earlier. (See for example, Quandt 1956.)

a higher proportion of service, government and education industries give rise to more travel than cities (zones) relying predominately on manufacturing. Another form of aggregate simultaneous model is the modal share model of which the composite analytic demand approach is a good example. (McLynn and Woronka 1969.)[1] These models include two separable components, one to predict total trips from o to d and a second to predict the share of these trips on mode m. The McLynn form is,

$$T = g^{ln}\ (A_o,\ A_d,\ S_o,\ S_d,\ L_{odm})\ \frac{g^{lln}\ (L_{odm})}{\sum_m g^{lln}(L_{odm})}$$

where $g^{ln} = b_o\ P_o^{\ b_1}\ P_d^{\ b_2}\ Y_o^{\ b_3}\ Y_d^{\ b_4}\ L_{odm}^{\ \ b_5}$

$$g^{lln} = a_o\ C_{odm}^{\ a_1}\ t_{odm}^{\ a_2}\ f_{odm}^{\ \ a_3}$$

Although the attempts at simultaneous modelling were a significant step forward, all of the aggregate models discussed above have difficulties associated with them. These occur at two levels; namely, problems associated with aggregate models in general and problems relating to the specific formulations of models and sub-models. It is important to distinguish between structural weaknesses inherent in the models themselves and deficiencies imposed by data constraints, the latter being due either to the format of the data (for example, trip files in contrast to household files) or the contents of data such as the general absence of household and individual information, thus necessitating the use of zonal data.

AGGREGATE MODEL PROBLEMS : GENERAL

Several points need to be made regarding the use

1. The composite analytic demand model is an extension of the McLynn-Watkins (1965) cross elasticity demand model. A discussion of both models is presented in Hensher (1971).

of zonal data. The data bases used to develop and test aggregate models are generally collected at the level of the individual or household (these being behavioural analysis units) and then aggregated to the zonal or district level (areal units). The zones are presently defined as physical entities with no consideration of the nature of behavioural trip making entities they include. Frequent evidence suggests that the greater degree of variation in trip making behaviour occurs within these arbitrary zones rather than between zones. Zone based models replace individual information with zonal averages (mean travel time, mean zone income) and prediction then involves the application of these zonal averages uniformly to all the behavioural units in the zone. However, areal units are not homogeneous as this procedure implies. The dispersion of actual values about the mean can be great and it is these actual values that are relevant to analysing and predicting travel behaviour.[1] Intravariations are concealed in aggregation. A significant reduction in the problem of within zone variance can be achieved by concentrating on the demand for travel at the household and individual traveller level and setting up basic behavioural hypotheses on trip making relating to homogeneity criteria such as a socio-ecomic grouping.[2] Variables chosen as stratifiers should be those which lead to the most homogeneous grouping of individuals.

All aggregate models involve the problem of ecological correlation which results in the ecological fallacy. The problem arises out of the attempt to use aggregate data to describe the behaviour of individuals. (Robinson 1950). The following diagram can be used to demonstrate the ecological fallacy.[3] In each zone there is a positive correlation between

1. While two physical zones may have the same values for average socio-economic variables (e.g. income, age) and level of service variables (e.g. car time) they may have considerably different travel demand patterns due to differences in the dispersion of the actual values of the variables in the zones. Models based on zonal averages will wrongly predict the same travel demand for both zones.
2. This issue is discussed in greater detail in Hensher (1974).
3. This diagram is taken from DeDonnea (1971, pp. 33-35).

the proportion of people using the car and income, but a negative correlation between the mean proportion of car users in each zone and the mean income of each zone. A possible explanation of this might be the existence of better transport facilities and service in the higher income zones. An aggregate model using mean zonal income as an explanatory variable would create the impression that car use is an inverse function of the income level, whereas in terms of individual behaviour, increasing income will have a positive effect on car use, ceteris paribus.

AGGREGATE MODELS: SPECIFIC FORMULATION PROBLEMS

Apart from the sequential modelling issue, the sub-models of an aggregate recursive model like that discussed above (UTMS) have specific structural weaknesses as formulated.

As already discussed above, the trip production models used in aggregate sequential models do not include level of service variables as influences on trip generation and thus are mis-specified. Apart from the general issue of functional format (category analysis versus linear multiple regression) this mis-specification constrains the model to concentration on the 'income effect' of the conventional economic demand model.[1] Thus it reflects differences in the demand for trips which are independent of the price of making such trips. This effect can be measured by a number of relatively independent variables affecting the household or individual's overall socio-economic status such as income, age and stage in the family life cycle. Car ownership which is highly correlated with measures of socio-economic status is usually given separate inclusion in trip production models. Besides not considering the price of trip making, the model also ignores the derived demand nature of trip making by excluding variables describing the attraction of undertaking activities at various positions in urban space and time. The result is that the conventional trip production

1. See Green (1971) for a detailed discussion of economic consumer theory.

model is constrained, to generating predictions of potential demand only. However, it is important to be able to measure the realized demand for travel which will be related to measures of price (including accessibility). Associated with estimating realized or actual demand is the need to develop a supply relationship to enable the actual demand estimates to also be an equilibrium demand level.[1]

Regarding trip distribution, the specific formulation used most is the Gravity model of zonal interchanges. Several criticisms can be made of this model both theoretically and empirically. Firstly, it does not derive from a theoretical analysis of choice of destination behaviour[2] but is a model of spatial interaction evolved from the early work of the social physicists who argued that social phenomena could be explained by analogies to physical laws in this case Newton's law of gravity. The use of population and income city characteristics as attraction variables, and travel time or trip length as the transport system characteristic,result largely from data ease of collection and data fitting rather than a priori reasoning. This is associative rather than causal (behavioural). We would expect an effective model of trip distribution to contain several level of service variables describing the transport network operating between the two zones, some of these being attraction and some deterrence variables. As with the trip production model, the gravity model contains no supply side. There is no interaction with a supply relationship to produce equilibrium flows. Detailed criticism of the gravity model including documentation of its poor empirical performance has been presented by Heggie (1969). Briefly, Heggie argues that the gravity model overestimates the increase in travel that will be associated with an increase in population and overstates the amount of travel in densely populated

1. The concept of a supply function in modelling raises some important questions, but as stated previously, this paper concentrates on the demand analysis without delving into the nature of demand and supply interaction. However the need to eventually model the whole system must be kept in mind.

2. For a discussion of an attempt to give the gravity model a theoretical justification within the framework of utility theory, see Niedercorn and Bechdolt (1969; 1970; 1972) and Mathur (1970).

areas. Evidence from European Transport Studies suggest the inability of gravity models to forecast accurately with margins of error in the order of several hundred percent.

Aggregate modal split models whether estimated using the diversion curve, linear regression or the approach of Shunk and Bouchard (1970) have several deficiencies. These can be summarized as follows:

1. virtually no models explicitly considered the quality of service provided by alternative modes (e.g. comfort and convenience) but relied instead on descriptions of system performance (e.g. physical travel time and cost) to forecast demand;
2. data used was generally collected for other purposes than the analysis of mode choice, for example, origin-destination surveys. A bias toward highways is evident and the public transport data was often therefore statistically weak. A bias also exists in that transit data reflect travel patterns which have emerged in response to the relative deterioration of the quality of public transport over time. Conventional models were essentially 'locked in' to this situation and projections based on these data bases will clearly be inadequate for planning. A serious underestimate of patronage when forecasting for a new or vastly improved mode is implied;
3. models developed in one city were found to be difficult, if not impossible to apply to problems in another city. Since there is no reason to believe that the underlying mechanism of traveller behaviour varies significantly from city to city this shortcoming suggests that the structure of these conventional models did not adequately generalize the mode choice decision process.

Having considered problems associated with the various formulations of the sub-models of an aggregate sequential structure we must now turn to the specific formulations of aggregate simultaneous models. In this context we will discuss the abstract mode model of Quandt and Baumol mentioned above. The specification of the abstract mode model is based on the assumption that the cross elasticity between modes exists only in respect of variables that qualify as 'bests'. For any change in a level of service variable for any mode, that variable which was neither best before nor after the change has no effect on the travel demand for other modes in the system; it affects only the use of the given mode. Besides the general problems of zonal based data being unsuitable for the analysis of transport system characteristics (see previous discussion on zonal data) the abstract mode model has a problem concerning zonal data and elasticities. Because of the zonal aggregation, the estimated elasticities for various system characteristics (e.g. time, cost) are <u>average elasticities</u> of <u>average zone characteristics</u> whereas what we desire are average elasticities of individual consumer characteristics.

It will be noted that the abstract mode model is essentially an extended gravity model which explicitly allows for certain system characteristics. As such, it is subject to criticisms made earlier of the gravity formulation. The model does not allow for the fact that two areas having the same average socio-economic characteristics will have different modal splits due to differing dispersions of these characteristics. Several authors have made detailed criticisms of the abstract mode model. One important criticism is that the model ignores fundamental interaction between socio-economic variables and the transportation network. It simply assumes that attraction variables such as income and population are exogenous. (See Bergsman 1967.) The model assumes costs to be constant as demand (volume) increases i.e. a relatively elastic supply curve is assumed.

This assumption seems untrue of large urban areas, where firstly increasing demand is obtained via improved transport systems which are more expensive, and secondly, increased usage may strain capacity (especially at the peak) which would adversely affect frequency of departure. One of the posited merits of this abstract mode model is its ability to predict demand for a new unknown urban mode. Gronau (1969) has cast doubt on this fundamental advantage.

The problems associated with the use of zonal data and with the specific formulations used in aggregate models (sequential and simultaneous) have encouraged researchers to move in other directions. In particular, in recent years much research has been devoted to developing models at the level of the individual traveller in a choice context. These disaggregate models are discussed in the following sections.

Disaggregate behavioural modelling has been the subject of a good deal of research in recent years in the context of travel demand. However, in the main, the research and development has been confined to the sub-model of modal choice. Therefore, the next section gives a review and discussion of mode choice research as a means of illustrating the development of behavioural modelling.

DISAGGREGATE MODELS OF TRAVEL DEMAND

Two approaches may be made to disaggregate modelling. The first is to develop models where the dependent variable is a volume of trips for an individual or household. Such models are deterministic and have the same specification as aggregate models and thus the same specification errors. (Ben-Akiva, 1973). The second approach is to view travel demand at the disaggregate level in the choice context rather than in the traditional demand analysis framework.

In this approach behaviour is treated as probabilistic which is consistent with modern theories of discrimination and choice. (See Stopher and Lisco, 1970).

> "...choice behaviour is best described as a probabilistic not an algebraic phenomenon" (Luce, 1959, p. 14).

Travel choice can best be described therefore as a choice from a finite set of mutually exclusive and exhaustive alternatives and not as the selection of a quantity (volume) of a commodity (travel) in the deterministic economic demand theory framework.

This section will concentrate on these probabilistic disaggregate models of travel behaviour where probabilistic behaviour explains observations of different choices for the same set of observed independent variables. Several general points need to be made about these models. They may be constructed as sequential or simultaneous models and as with aggregate models, we use the assumption of a utility tree or separable utility function and negligible income effects which enables us to model independently a subset of the total array of choices facing the individual.[1]

Mobility and travel choices are assumed to be an independent branch of the consumer's utility function. Mobility choices (e.g. residential location) are assumed prior choices, leaving travel choices to be modelled separately on the assumption of fixed mobility choices. As with aggregate models, travel choices for different trip purposes are assumed to be independent. Disaggregate models operate at the level of the

1. For a discussion of separable utility functions see; Strotz (1957), Strotz (1959), and Green (1971, pp. 150-156).

individual and not the household, the latter being the conventional unit of economic analysis.[1]

Most research work on disaggregate probabilistic models has been at the sub-model level of mode choice. Therefore, discussion of specific formulations in this section will be concentrated on models of mode choice. To discuss these models we will consider the following choices to compose a trip probability:

1. trip frequency (f), (only relevant to disaggregate models)
2. destination zone (d)
3. mode of travel (m)
4. hour of day (h)
5. route (r)[2]

---

1. The selection of the individual as the basic decision unit in contrast to the household is based on the premise that it is the individual who maximises utility or some other function subject to household constraint. Household decision-making is more nebulous, and perhaps less appropriate. To illustrate this, it is important to have information on the number and availability of cars in the household, and the relevant importance of such household characteristics. But it is the individual as the traveller who finally selects a modal facility in the light of the availability of a car and other influences, maybe greater influences. It is his utility which we are trying to maximise in a travel choice context and we may or may not be maximising the utility of the household in so doing. The immediate advantage of the approach is that we should be able to consider the influence of household variables on the individual and the influence of variables directly related to the individual on the overall household choice process. It is a two-way problem. By adopting the household as the unit of analysis we have excluded an important aggregation problem, that of the summation of individuals in the household. There is little evidence to suggest that there is any more homogeneity between members of a household with respect to a particular issue than there is between individuals of entirely different households. The fact that we find socio-economic data on individuals as useful grouping criteria is testimony to this assertion.

2. Although this covers more stages than we discussed with aggregate models, this enlarged choice set enables a better overall picture of disaggregate models to be gained.

Disaggregate Sequential Models

We must make a decision as to what is the hierarchy of conditional decisions. We will look at a sequential structure following the order of travel choices listed above. Thus we begin with the marginal probability for frequency, this being

$$\text{Prob }(f:F) = \text{Prob }\{U_f \geq U_{f^1},\ f^1 \in Df\}$$

i.e. probability that frequency f is chosen out of the set of possible frequencies F is equivalent to the probability that the utility derived from f is equal to or greater than the utility associated with f^1 where f^1 can take any value in F. This is to say the individual maximizes utility in choice of frequency.

Having obtained a marginal probability for frequency we must choose a destination conditional upon (or given) the choice of frequency. This gives us:

$$\text{Prob }(d:Df) = \text{Prob }\{U_d|_f \geq U_{d^1}|_f,\ d^1 \in Df\}$$

where Prob (d:Df) is the conditional probability of selecting destination d from the set of alternative destinations Df which are consistent with the chosen frequency f.[1] $U_d|_f$ is the utility gained from d given f (i.e. it is a utility conditional on f). The interpretation of the equation is the same as that given for the marginal probability of frequency assumption. Using this terminology we can write the remainder of the sequential choices as follows:

$$\text{Prob }(m:M_{fd}) = \text{Prob }\{U_m|_{f,d} \geq U_m|_{f,d};\ m^1 \in M_{fd}\}$$
$$\text{Prob }(h:H_{fdm}) = \text{Prob }\{U_h|_{f,d,m} \geq U_{h^1}|_{f,d,m};\ h^1 \in M_{fdm}\}$$
$$\text{Prob }(r:R_{fdmh}) = \text{Prob }\{U_r|_{f,d,m,h} \geq U_{r^1}|_{f,d,m,h};\ r^1 \in R_{fdmh}\}$$

---

1. In more formal terms the set of alternative destinations can be partitioned according to frequency to give the vectors D_1-- D_f--D_n where each vector contains destinations consistent with the nominated frequency. Depending on the chosen frequency, one of these partitions becomes the destination choice set of alternatives. This partitioning concept is used considerably in the discussion.

At each stage the individual maximizes utility given the previous choices. Now we will have a set of independent variables X which are the variables (socio-economic, attraction and level of service) which determine the probabilities of various choices. This will be denoted by X_{fdmhr} and is a vector incorporating all the variables X for all relevant combinations of (f, d, m, h, r). Thus we get the result that:

$$\text{Prob } (f,d,m,h,r : FDMHR) = g(X_{fdmhr};\ fdmhr \in FDMHR)$$

which is associated with a utility function $U_{fdmhr} = U(X_{fdmhr})$.

Now we are using a sequential structure where each stage estimates conditional probability. The conditional probability of a particular choice, given other choices, will be a function only of a particular subset of the explanatory variables. Thus the sequential structure outlined above has the following independent utility function;

$$\begin{aligned} U_{fdmhr} &= U_f + U_{d|f} + U_{m|fd} + U_{h|fdm} + U_{r|fdmh} \\ &= U^f(X_f) + U^d(X_{fd}) + U^m(X_{fdm}) + U^h(X_{fdmh}) + U^r(X_{fdmhr}) \end{aligned}$$

where we have to develop expressions for

$$\text{Prob } (f) = g_f(\{X_f,\ f \in F\})$$

$$\text{Prob } (d|f) = g_d\ (\{X_{fd},\ d \in D_f\})$$

$$\text{Prob } (m|f,d) = g_m\ (\{X_{fdm}\ ,\ m \in M_{fd}\ \})$$

$$\text{Prob } (h|f,d,m) = g_h\ (\{X_{fdmh},\ h \in H_{fdm}\})$$

$$\text{Prob } (r|f,d,m,h) = g_r\ (\{X_{fdmhr},\ r \in R_{fdmhr}\})$$

At each of these stages composite variables will be needed. For example when estimating Prob (m|fd) we will be using the set of explanatory variables X_{fdm}. However, variables in this set will occur which relate to time of day and route characteristics. To estimate this functional relationship these variables must be composite variables defined over all routes and times. Thus we have;

$$X_{fdm} = C_1\ (X_{fdmh};\ h\ \in H_{fdm})$$

and $$X_{fdmh} = C_2\ (X_{fdmhr};\ r \in\ R_{fdmh})$$

where C_1, C_2 are composite functions. If we can specify a sequence of choices, and a rule for forming composite variables, given a separability assumption, a sequential travel demand model is possible.[1] As for actual work with a complete disaggregate sequential travel demand model, the Charles River Associates (1972) modelling of shopping trips is the best example. They developed a sequence of individual choice models based on the assumption that individuals first choose whether to travel, then where to travel, next what time to travel and last, what mode to use. The model estimates the conditional probabilities $P(d|f=1)$, $P(h|f=1,d)$, $P(m|f=1,d,h)$ assuming only one or zero shopping trips are made. It is important to note that the C.R.A. model allows for interaction between the various choices via the use of inclusive prices. This necessitates the calibration of the sub-models in the reverse order to the assumed order of individual's choice. This model is now considered in detail.

1. Mode Choice Sub-model. This reflects a binary choice between automobile and transit with the estimated probability being that of choosing automobile. Thus we have:

$$\frac{P(a|f = 1,d,h)}{1-P(a|f = 1,d,h)} = \exp\{\alpha + \sum_b \beta^s (L^s_{dha} - L^s_{dht}) + \sum_s c^s S_i{}^s\}$$

where $P(a|f = 1,d,h)$ = probability of choosing auto, a, given frequency f, destination d and time of day h.

L^s_{dha}, L^s_{dht} = level of service variables for automobile, a, and transit, t, for destination, d, at time of day, h.

S^s_i = socio-economic variables for household i.

---

1. As with aggregate models there are several choice sequences. The one used here is common but not unique. It should also be noted that composite variables also occurred in the aggregate models discussed above.

The socio-economic variables used include automobiles per worker, and occupation, while level of service variables were the descriptive system attributes of waiting time, in-vehicle time, operating, parking and fare costs.

2. Time of Day Choice Sub-Model. This is based on a binary choice of travelling both ways off-peak or at least one way during the peak. We have;

$$\frac{P(h|f=1,d)}{1-P(h|f=1,d)} = \exp\{\alpha+\beta\,(IP_{dz} - IP_{dp}) + \sum_s c^s\, S_i^s.\}$$

where $P(h\ f=1,d)$ = probability of making the trip at time h, given frequency and destination.

S_i^s = socio-economic variables for household i.

IP_{dz}, IP_{dp} = inclusive price of travelling to destination d at peak (p) and off-peak (z) times. These inclusive price variables allow interaction to be handled and are constructed as:

IP_{dz} = $\sum_s \beta^s L^s_{dhmz}$ where;

β^s = parameter from mode choice sub-model and

L_{dhmz} = level of service variable for mode m used during off-peak times to go to d at time of day h.

IP_{dp} is defined similarly to IP_{dz}. The socio-economic variables used are the sex of the head of house, and the number of pre-school children in the household.

3. Destination Choice Sub-Model. Only two destinations can be handled in the equation, written as:

$$\frac{P(d|f=1)}{1-P(d|f=1)} = \exp\{\alpha_1\,(IP_d - IP_{d^1}) + \alpha_2(A_d - A_{d^1}) + \alpha_3 S_i\}$$

where $P(d|f=1)$ = probability of choosing destination d given a shopping trip (f=1) is to be made.

A_d, A_{d^1} = attraction variables for destination d and alternative d^1

S_i = socio-economic variable for household i

IP_d, IP_{d^1} = inclusive prices of travelling to destinations d and d^1

$IP_d = \sum_s \beta^s L^s_{dhm}$ where β^s is the parameter estimated in the mode choice sub-model

IP_{d^1}, is defined in a similar way to IP_d. The CRA model used the fraction of total retail employment at each destination as the attraction variable, number of pre-school children in the household as the socio-economic variable, and only automobile level of service variables.

4. Trip Frequency Sub-Model. This analysed a binary choice of making zero or one trip for shopping each day, using the equation:

$$\frac{P(f=1)}{1-P(f=1)} = \exp(\alpha_1 IP_i + \alpha_2 IE_i + \alpha_3 Y_i)$$

where $P(f=1)$ = probability of making a shopping trip

Y_i = income of household i

IP_i = inclusive price for household i

IE_i = average shopping opportunity

Now $IP_i = \sum_d IP_d . P(d|f=1)$ where IP_d and $P(d|f=1)$ are obtained from the destination choice sub-model. $IE_i = \sum_d A_d . P(d\ f=1)$ where A_d and $P(d|f = 1)$ come from the destination choice model.

Although the CRA model is developed at the household level it illustrates the disaggregate sequential demand model and could be adopted to the individual traveller level.

Despite the improvement in modelling achieved by the disaggregate sequential structure there still remain problems concerning the assumption of a specific heirarchy of conditional decisions. When we are confronted with a complex decision to analyse which involves a large number of alternatives we often are able to simplify the task by replacing the complex decision with a set of sequenced decisions. Such sequential structure generation requires us to decompose the single complex decision into stages by partitioning the overall set of explanatory variables and alternatives. It has been shown that different partitions (i.e. sequential structures) produce different results. (Luce, 1959) Therefore, unless we have a priori reasoning to support a specific sequence, sequencing should be treated as a simplifying assumption, and we should test several possibilities. There being little a priori guidance as to the question of whether and how the individual simplifies a complex decision, and in the light of our previous criticisms of sequential modelling (see discussion of aggregate sequential models) it appears desirable to attempt simultaneous modelling of urban travel behaviour at the disaggregate level.

Disaggregate Simultaneous Models

Prior to the work of Ben-Akiva (1973) no serious attempt had been made to develop a disaggregate simultaneous travel demand model. CRA noted the desirability of a simultaneous structure, but elected to use a sequential structure as a means of reducing the estimation difficulties. Disaggregate simultaneous models have one equation to be estimated (like aggregate simultaneous models). A very large number of explanatory variables will be involved, whereas the sub-model equations of a sequential structure have as explanatory variables only subsets of the total set of explanatory variables included in the simultaneous equation.

A simultaneous structure involves estimation of the joint probability P(fdmhr : FDMHR), which is the probability of the individual choosing the combination frequency f, destination d, mode m, time of day h and route r, from all the possible combinations of frequency, destination,mode, time and route, which are possible (namely the set of alternatives FDMHR). It can be shown that for a simultaneous structure where each choice is interdependent on the other we must estimate the joint probability directly. It cannot be computed from conditional probabilities of the form: $P(m:M_{fdhr})$. (See Ben-Akiva, 1973, chapter 4.)

The simultaneous model discussed here can be written as:

$P(f|d,m,h,r)$
$P(d|f,m,h,r)$
$P(m|f,d,h,r)$
$P(h|f,d,m,r)$
$P(r|f,d,m,h)$

and this can be modified by introducing *a priori* reasoning about behaviour such as the conditional probabilities of destination and frequency not being conditional on the chosen route. Further modifications may be made. For example, if we were modelling a journey to work to the Central Business District, we may consider frequency of trip, destination and time of day to be constrained by prior mobility choices, leaving mode and route choice to be modelled simultaneously. Similarly, a trip to the theatre has a constrained time of day in most cases (e.g. evening).
Ben-Akiva has modelled the following structure;

$P(f)$
$P(d|f,m)$
$P(m|f,d)$
$P(r|f,d,m)$

This is typical of many models in that it is a mixture of sequential and simultaneous modelling with certain subsets of choices being modelled simultaneously (destination and mode choice in Ben-Akiva's model) but within an overall sequential framework.

Summary of Disaggregate Model Advantages

In general, disaggregate models are held to have the following advantages over aggregate models:

1. they avoid the problems associated with the use of aggregate or zonal data. In particular the avoidance of ecological fallacies of inference is a significant advantage;[1]
2. being constructed at the level of the individual these models may be aggregated to any required level. Furthermore, the study of individual behaviour and the use of plausible theories of consumer behaviour including probabilistic choice theory, will be an important guide to the appropriate method or criteria for aggregating data and models to develop more efficient aggregate models;[2]
3. because they study individual choice behaviour, the potential for transferability of the models between areas is higher and the same set of model structures may be used for multi-level planning;
4. these models offer a basis for inferring the values individuals place on various transport system characteristics such as the value of travel time savings. Changes in system characteristics affect travel choices. The relative magnitiudes of these impacts are an indication of the differential values individuals attach to the changes in characteristics;

---

1. See De Donnea (1971, pp. 31-35) as well as the discussion earlier in the paper.
2. For planning purposes aggregate models are needed but as yet we have no satisfactory method of aggregation that allows the individual to be the basic behavioural unit in spatial aggregation. This problem of aggregating a disaggregate model without loss of meaning is an important area for research. Aggregation need not use a spatial criteria : it may be based on characteristics of individuals (e.g. income, age) or some other non-spatial variable.

5. the individual offers a better opportunity to investigate causal relationships, being a natural unit of choice behaviour, as opposed to aggregation models which essentially rely on statistically derived correlations among zonal data.

BEHAVIOURAL VERSUS NON-BEHAVIOURAL MODELLING

We must distinguish between models which are based on the search for causal relationships and those based on statistical fitting where the objective is to develop relationships with high statistical correlations where the variables need not be causally related. The first type of model is usually called a behavioural or causal model, whereas the latter is an associative or non-behavioural model. It is also preferable to make a distinction between behavioural and causal models, where behavioural models are a subset of causal models which are characterized by models based on the analysis of attitudes and preferences and the direct measurement of human attitudes, preferences and reactions. The identification and measurement of psychological or subjective variables will be a feature of these modelling efforts. Thus a model may be causal but not behavioural and a causal model may be aggregate or disaggregate. While disaggregation need not imply behavioural modelling, behavioural modelling will imply disaggregation.

The behavioural approach to modelling choice behaviour is based on three basic premises. (Michaels, 1974.) Firstly, intrinsic needs of the individual motivate his spatial behaviour which involves physical movement to locations at which satisfaction of these needs can be achieved. Secondly, this spatial behaviour involves the individual in making choices, and his choice behaviour reflects his *subjective perception* of the various transport alternatives (or if you like spatial movement alternatives) available to him. His subjective perception need not be related to any objective criteria which can be applied to the set of alternatives. The final assumption is that although changes in the choice behaviour pattern will occur, the basic process by which individual choice decisions are arrived at will not change. Furthermore, within any defined population there are certain basic variables influencing this choice process which are universal. The implication is of course that if we can model the choice process and identify these fundamental variables, we will be in a position to predict the changing pattern of choice behaviour.

DISAGGREGATE BEHAVIOURAL MODELS OF MODE CHOICE

Probabilistic disaggregate models of mode choice were initially causal rather than behavioural in the sense in which behavioural is defined above.[1] Although several of these models have particular merit, our discussion will concentrate on the work of Lave (1969). Although his model is not particularly unique it has been selected because Lave gave specific attention to the need to make the *individual* the basic unit of analysis and recognized the desirability of incorporating perceptual variables such as comfort into the models.

These causal models did not utilize attitudinal data (i.e. data gathered via direct measurement of attitudes and preferences) and thus the behavioural content was limited. Lave did suggest that the desirable way to handle comfort and convenience was through the collection of attitudinal data.

---

1. See for example the following binary mode choice models developed in the period 1960-1972. Warner (1962); Quarmby (1967); Lisco (1967); Lave (1969); Stopher (1969); De Donnea (1971); Hensher (1972).

These models were based on the postulate that the individual's choice of mode would be a function of 1) personal characteristics of the individual such as income, tastes, age, sex and car ownership and; 2) characteristics of the alternative modes of transportation such as time, cost and comfort.[1]

Lave's model was binary, considering the choice between transit and automobile and had the following specification -

$$Y = \alpha_1 + \alpha_2\, kW\Delta T + \alpha_3\, \Delta C + \alpha_4 IDC_c + \alpha_5 A + \alpha_6 S$$

where:
- Y = probability of choosing transit
- W = hourly wage of commuter
- ΔT = time difference between auto and transit
- k = a factor the size of which indicates the individual's marginal preference for work versus leisure
- $kW\Delta T$ = marginal value (in money units) of saving commuting time
- ΔC = cost difference between auto and transit
- I = income
- D = distance
- C_c = comfort (binary variable)
- A = age
- S = sex

Car ownership, family size and family composition (the latter essentially being a measure of the pressure for car use; i.e. number of drivers in the family or number of cars per driver) were rejected as explanatory variables on theoretical grounds. (Lave, 1969, pp. 467-468.) Income enters the model via its influence on the commuter's perception of comfort. The model was calibrated using probit analysis and the dependent variable

1. This postulate also applies to the more truly behavioural models, based on measurement of attitudes, to be considered later in this section.

is the probability of the commuter using transit. It is suggested as a result that changes in the probability of transit usage in response to cost and time changes are quite small.[1,2]

Thus the Lave model and the others of the same type had the following general functional form;

$$PT = f(\Delta T, \Delta C, S_i)$$

where PT = probability of selecting transit from the binary set of alternatives; automobile/transit

ΔT = time differences (or ratios)

ΔC = cost differences (or ratios)

S_i = socio-economic variables for individual i

Attitudinal data was not included in the formulations and the only policy variables included were time and cost.[3] Comfort was tackled only by Lave and lack of data meant his method of handling comfort prevented it being used as a policy variable or as a prediction variable in the above model.

Following on these initial developments, attempts were made to develop far more comprehensive models of mode choice which were calibrated using attitudinal data and were truly behavioural according to our previous definition.[4] As a result of this

1. Lave (1969, p. 419) contains a summary table of empirical results. For example, relative time improvements of 5 minutes and 25 minutes respectively for bus would increase the probability of bus usage by 0.081 and 0.206 respectively.
2. A useful summary of this type of model (causal disaggregate) is given in Demetsky and Hoel (1972).
3. These may not be as flexible as is needed in terms of control variables because significant improvements in time usually require substantial investment programs associated with restructuring the transport network or/and introducing new modes.
4. Good examples of these models are to be found in; Allen and Isserman (1972); Demetsky and Hoel (1972); Golob (1970); Hartgen and Tanner (1970a, 1970b) and Sommers (1970).

work we can see developing a comprehensive theory of mode choice which incorporates attitudinal data and relevant socio-economic and demographic factors to produce a more accurate model of the *subjective mode choice process*. The general decision rule forming the basis of these models is that the individual traveller,

> ".....chooses the mode for which he *perceives* the least disutility or generalized cost"
> (Golob, 1970, p. 104).

An important idea here is that the traveller will assess available modes on the basis of his subjective perception of the characteristics of each mode (including time and cost), and therefore, at a general level, this approach is related to the theory developed by Lancaster (1966, 1971), which relies on the fundamental realization that goods are "n" dimensional in characteristics space. Utility is posited to be a function of the characteristics, while goods are the means of attaining certain desired combinations of characteristics. Thus the set of alternatives on which the consumer's preference ordering is defined must be conceived of as bundles of characteristics rather than, as conventional theory has it, bundles of goods. An implication of the consumption technology where goods produce characteristics is that the demand for a commodity is a derived demand. A relationship between goods as inputs and characteristics as outputs must be postulated and Lancaster utilizes a purely technical relationship analogous to the production function. This is the consumption technology. Of course this whole discussion of mode choice models implies a separable utility function or utility tree, whereby funds are allocated to broad expenditure classes (branches of the utility tree) one of which will be transportation and each allocation is spent optimally without reference to the other areas. Thus we can minimize disutility in transport independently of other areas (e.g. clothing, food).

Although Lancaster argues for objective criteria for deciding what the characteristics of a good are and for objective (physical) measurement of the quantity of a characteristic produced by a particular good, the behavioural mode choice models diverge from this position. They all utilize direct subjective measurement of the disutility (or utility) associated with the characteristics produced by the alternative modes available and in some cases the individual is allowed to specify the set of relevant characteristics. We now turn to the development of a disaggregate behavioural model of mode choice.

Trips on various modal alternatives produce certain combinations of characteristics and a mode will be selected because it provides that combination of characteristics which minimizes the disutility of making the trip.

Consider the choice between two modes, 1 and 2.[1] A mode 1 user can be represented by the following inequality;

$$DU_i^1 < DU_i^2 \qquad (1)$$

where DU_i^1 is the total disutility associated with mode 1 for user i and DU_i^2 is the total disutility associated with mode 2 for user i.

Using the assumption of additive utilities we can rewrite (1) as

$$\sum_{j=1}^{N} DU_{j,i}^1 < \sum_{j=1}^{N} DU_{j,i}^2 \qquad (2)$$

1. The linear model expounded here can only handle two modes at a time. It thus requires a multi mode alternative set to be analysed as a series of binary choices.

where $DU^{k}_{j,i}$ is the disutility associated with characteristic j of mode k(=1,2) by user i. The concept of using the individual's subjective perception of each alternative is important as it leads to the following argument common to all disaggregate behavioural models so far developed. The basic (independent) components of the perceived disutility of characteristic j, $(DU^{k}_{i,j})$ are;

(1) the degree of satisfaction (S) the particular tripmaker has regarding the ability of each mode to fulfil his individual requirement relating to the system characteristic j, and

(2) the degree of importance (I) the individual tripmaker attaches to the system characteristic j in his decision calculus. Thus we can express the disutility associated with a characteristic as;

$$DU^{k}_{i,j} = (I_{i,j})\ (S^{k}_{i,j}) \qquad (3)$$

where $I_{i,j}$ is the importance user i places on characteristic j and $S^{k}_{i,j}$ is a measure of the disutility the individual user i perceives to be associated with characteristic j on mode k. This formulation thus weighs the disutility of a characteristic by its importance in the decision calculus.[1] Note that the importance attached to a particular characteristic is assumed to be mode independent.

By substituting equation (3) into equation (2) we get;

$$\sum_{j=1}^{N} I_{i,j}\ S^{1}_{i,j} < \sum_{j=1}^{N} I_{i,j}\ S^{2}_{i,j} \qquad (4)$$

which can be rewritten as;

$$\sum_{j=1}^{N} I_{i,j}\ (S^{1}_{i,j} - S^{2}_{i,j}) < 0 \qquad (4a)$$

1. For a discussion of the relationship between utility theory and the formulation expressed in equation (3), see Hensher, McLeod and Stanley (1975).

Thus our criterion of mode choice can now be expressed as;

$$C_i = \sum_{j=1}^{N} I_{i,j} \; (S_{i,j}^{1} - S_{i,j}^{2}) \qquad (5)$$

Mode 1 is chosen if C_i is less than zero and mode 2 is chosen when C_i is greater than zero. The expressions on either side of equation (4) are disutility indices and derived from the individual's utility function, the specification of which is usually necessary to decide mode choice.[1] A problem arises in obtaining data which will enable us to compute these disutility indexes for each mode. We have described disutility in terms of the perception of modal characteristics. This is a purely subjective concept and depends on subjective values being attached to the variables of the disutility function.

A body of theory is available in the experimental psychology literature which has developed measurement techniques suitable for our problem; in particular the techniques of semantic differential and paired comparisons[2] enable us to question an individual directly about his attitudes and preferences toward those characteristics *we believe enter his disutility function*[3] and about the importance of each of these characteristics in his decision calculus.

The model outlined above assumes that individual travellers make an evaluation with respect to all of the

1. A mode may be superior in all characteristics in which case we would not need to specify the utility function to determine that it would be chosen.

2. Detailed discussion of these techniques and illustrations of their use in analysing the impact of various transport designs and in studying the demand responses to alternative transport proposals can be found in Golob (1972) and Vitt (1970).

3. This refers to the usual case where an a priori characteristics specification of the alternative modes is constructed by the researcher, and then presented to the individual being questioned.

attributes associated with the transport system. Hartgen and Tanner (1970a) have suggested that each individual about to make a decision will consider the specific attributes describing each mode and rank them according to a subjective importance hierarchy. The process will then be to group similar attributes together in a smaller number of relatively independent factors such as cost, comfort, convenience, safety and reliability. They argue that individual decisions are actually based on these factors rather than the individual attributes associated with them. The model of equation (5) then takes the following form;

$$C_i = \sum_q I_{q,i} \; (S^1_{q,i} - S^2_{q,i}) \qquad (6)$$

where $I_{q,i}$ is the composite importance of a factor q and will be a function of the attributes composing q as perceived by trip maker i. $S^k_{q,i}$ is the composite generalized cost (or disutility) associated with factor q for mode k as perceived by tripmaker i. $S^k_{q,i}$ will be a function of the attributes composing q. These formulations of the model mean that identical satisfaction differences have the same influence on mode choice at both high and low ends of the satisfaction scale. This could be countered by expressing the formula as;

$$C_i = \sum_q I_{q,i} \quad (1 - \frac{S^2_{q,i}}{S^1_{q,i}}) \qquad (7)$$

Semantic differentials can give us the satisfaction ratings for the various attributes of each mode for respondent (user) i. We must assume a linear relationship between satisfaction ratings, utility ratings and a disutility ratings.

At this stage we have a deterministic model of mode choice. However we will expect errors of omission (non exhaustive list of characteristics), errors of specification (the additive utilities assumption may not be correct), measurement errors and "irrational" (in terms of economic utility maximizing man) choice errors. These error possiblities and the argument developed previously regarding the placement of disaggregate models in the choice context suggest a probabilistic model. This can be done by relating the probability of choosing a particular mode (say transit) to the disutility indices differential. From equation (5) we would argue that $P_{1,i} \to 1.00$ as $C_i \to -\infty$, where $P_{1,i}$ is the probability that tripmaker i chooses mode 1. Thus our model will be of the form $P_{1,i} = f(\sum_{K=1}^{N} I_{i,j}(S_{i,j}^1 - S_{i,j}^2))$ where the exact nature of the functional relationship has to be decided upon. Some authors have used an S shaped relationship, while others have assumed the relationship between mode choice (expressed as a probability) and its explanatory variables was linear.[1] It has also been argued that users are unable to convey the true importance of various characteristics i.e. the $I_{i,j}$ responses are inaccurate. To overcome this we can eliminate the importances and estimate a linear probability model of the form

$$P_{1,i} = \alpha + \sum_{j=1}^{N} \beta_j \; (S_{i,j}^1 - S_{i,j}^2) + e \qquad (8)$$

This procedure will give us estimated weights β_j for the satisfaction differentials which best describe the pattern of mode choice observed in the group of individuals sampled, instead of relying on

1. Lisco (1967); Lave (1969); Stopher (1969); De Donnea(1971); Demetsky and Hoel (1972); and Hensher, McLeod and Stanely (1975) use an S-shaped relationship while Quarmby (1967) and Allen and Isserman (1972) fit a linear one. We will not discuss the relative merits of the two relationships except to note that the estimation techniques probit and logit analysis which fit S-shaped curves will keep the dependent variable within the legitimate probability range zero to one whereas linear regression will not. Also De Donnea has given apriori reasoning to support an S-shaped relationship where the probability of choosing an alternative mode moves asymptotically to one as the relative disadvantage of the current mode increases. The theoretical foundations of the S-shaped relationship are discussed in Hensher (1974a).

what the individual says is the impact of a satisfaction differential for a particular characteristic on his mode choice. It must be emphasized that the β_j's are statistical weights and are not analogous to or substitutes for the actual importance weights. The real issue, in the linear context, is whether equation (8) is a better representation of mode choice behaviour, than the specification,

$$P_{1,i} = \alpha + \sum_{j=1}^{N} \beta_j{}'_{i,j} \left(S^1_{i,j} - S^2_{i,j}\right) \qquad (9)$$

When we have the respondents answers to the "importance of characteristics" questions we are able to compute the disutility index directly from the survey data and then use a response surface to relate the percentage of trips by transit to the disutility index. (Hartgen and Tanner, 1970a).

The weights estimated in equation (9) above can be used for predicting the behaviour of new individuals facing various modes or the impact of an innovative mode. The estimated weights are applied to the survey data collected from the individuals concerned on the relevant explanatory variables. The mode choice models discussed above utilize only attitudinal variables as explanatory variables whereas the functional relationships we have discussed earlier in the chapter suggest that mode choice is influenced by socio-economic and demographic variables. A question arises as to how these are to be incorporated into the model. Considering linear models only we can establish two methods of inclusion of socio-economic variables. Firstly they may be entered in a simple additive fashion producing a model of the form;

$$P_{1,i} = \alpha + \sum_{j=1}^{N} \beta_j{}' \, I_{i,j} \left(S^1_{i,j} - S^2_{i,j}\right) + \sum_{a} \gamma_a M_{ai} + e \qquad (10)$$

where M_{ai} is the ath socio-economic variable for user i and γ_a

is the coefficient to be estimated with respect to this variable.

The second method is to estimate the model as it appears in equation (9) but to stratify according to the socio-economic variables and estimate the model for each stratification. If we adopt this approach the socio-economic environment influences the magnitudes of the estimated coefficients for the attitudinal explanatory variables. This second approach has been suggested as the most valid one.[1] One implication of equation (10) is that the value of time estimated from the coefficients of time and cost differences is the same for each socio-economic group and evidence suggests that this is not the case, in particular the value of time can be expected to vary with income. The second approach allows the value of time derived to vary between socio-economic groups.

CONCLUSION

Modelling is not required as an end in itself. It should be a constructive means of investigating those relationships on which sound policy decisions ought to be based. If not practically operational the model should make a positive contribution to our understanding of the phenomenon being studied. This review of modelling procedures has suggested that disaggregate behavioural models offer advantages on both these criteria. In particular the development of disaggregate behavioural modelling of mode choice has been a significant advance toward achieving these modelling goals for this particular sub-model. The review of analytical structures suggests that the relevant research task now is to extend this behavioural modelling approach to the whole set of choices comprising urban travel demand. This should be done in the context of a disaggregate simultaneous structure. The general neglect in Australia of research and development into behavioural approaches to transport planning must not continue.

1. See De Donnea (1971, pp.47-49), Stopher and Lavender (1972) and McLeod (1974).

REFERENCES

Allen, W. Bruce and Isserman, A. (1972). "Behavioural mode split", High Speed Ground Transportation Journal, Vol. 6, pp.179-199.

Ben-Akiva, M.E. (1973). "Structuring the Demand for Passenger Transport",unpublished Ph.D. Thesis, Department of Civil Engineering, M.I.T.

Bergman, J. (1967). "Comments on Quandt's and Baumol's demand model for abstract transport modes", Journal of Regional Science, Vol. 7, pp.83-85.

Brand, D. (1973). "Theory and method in land use and travel forecasting", Highway Research Record No. 422, pp.10-20.

Charles River Associates (1972). A Disaggregated Behavioural Model of Urban Travel Demand, (Boston: Charles River Associates).

De Donnea, F.X. (1971). The Determinants of Transport Mode Choice in Dutch Cities, (Rotterdam: Rotterdam University Press).

Demetsky, M.J. and Hoel, L.A. (1972). "Modal demand: A user perception model", Transportation Research, Vol. 6, pp.293-308.

Domenich, T.A. and Kraft, G. (1968). "Estimation of urban passenger travel behaviour: An economic demand model", Highway Research Record No. 238, pp.64-77.

Golob, T.F. (1970). "The survey of user choice of alternative transportation modes", High Speed Ground Transportation Journal, Vol. 4, pp.103-117.

Golob, T.F. et al (1972). "An analysis of consumer preferences for a public transportation system", Transportation Research, Vol. 6, pp.81-102.

Green, H.A. John (1971). Consumer Theory (London: Penguin Ltd.).

Gronau, R. (1969). "The demand for abstract transport modes: Some misgivings", Journal of Regional Science, Vol. 9, pp.153-157.

Hartgen, D.T. and Tanner, G.H. (1970a). Behavioural Model of Mode Choice (Research and Applied Systems Section, New York Department of Transportation).

Hartgen, D.T. and Tanner, G.H. (1970b). "Individual attitudes and family activities: A behavioural model of traveller mode choice", High Speed Ground Transportation Journal, Vol. 4, pp.439-467.

Heggie, I.G. (1969). "Are gravity and interactance models a valid technique for planning regional transport facilities", Operational Research Quarterly, Vol. 20, pp. 93-110.

Hensher, D.A. (1971). "Modal split and modal choice: A review of modelling procedures", unpublished paper, School of Economics, University of New South Wales.

Hensher, D.A. (1972). "The consumer's choice function: A study of traveller behaviour and values", unpublished Ph.D., University of New South Wales.

Hensher, D.A. (1974a). "A probabilistic disaggregate model of binary mode choice", in Urban Travel Choice and Demand Modelling, Special Report No. 12, ed. D.A. Hensher (Melbourne: Australian Road Research Board).

Hensher, D.A.,(1974b). "The problem of aggregation in disaggregate behavioural travel choice models with emphasis on date requirements", Issues in Behavioural Travel Demand Modelling and the Valuation of Travel Time, (Washington, D.C.: Transportation Research Board).

Hensher, D.A., McLeod, P.B. and Stanley, J.K. (1975). "Usefulness of attitudinal measures in investigating the choice of travel mode", unpublished paper.

Houston, Wynn F. (1969). "Shortcut modal split formula", Highway Research Record, No. 283, pp.48-56.

Lamb, G.M. (1970). "Introduction to transport planning - Part 4 trip distribution and modal split", Traffic Engineering and Control, April, pp.628-630.

Lancaster, K. (1966) "A new approach to consumer theory", Journal of Political Economy, Vol. 74, pp.132-157.

Lancaster, K. (1971). Consumer Demand: A New Approach, (New York: Columbia University Press).

Lave, C.A. (1969). "A behavioural approach to modal split forecasting", Transportation Research, Vol. 3, pp.463-480.

Lisco, T.E. (1967). "The value of commuters' travel time: A study in mass transportation", unpublished Ph.D., University of Chicago.

Luce, R.D., (1959). Individual Choice Behaviour (New York: Wiley)

Mathur, V.K. (1970. "An economic derivation of the gravity law of spatial interaction: A comment", Journal of Regional Science, Vol. 10, pp.403-405.

Meyer, J.R. and Strazheim, M.R. (1972). Techniques of Transport Planning. (The Brookings Institution).

Michaels, R.M. (1974). "Behavioural measurement: An approach to predicting transport demand", in Issues in Behavioural Travel Demand Modelling and the Valuation of Travel Time (Washington, D.C.: Transportation Research Board).

McLeod, P.B. (1974). "The role of subjective data in modelling mode choice", in Urban Travel Choice and Demand Modelling, Special Report No. 12, ed. D.A. Hensher (Melbourne: Australian Road Research Board).

McLynn, J.M. and Woronka, T. (1969). "Passenger demand and modal split models", N.E.C.T.P. Report 230.

McLynn, S.M. and Watkins, R.M. (1965). Multimode Assignment Model (Maryland: Talbird and McLynn Inc.).

Neuburger, H. (1972). "Transport demand models" (London: Department of the Environment) mimeo.

Niedercorn, J.H. and Bechdolt, B.V. (1969). "An economic derivation of the 'gravity law' of spatial interection", Journal of Regional Science, Vol. 9, pp.273-282.

______________________________(1970). "An economic derivation of the gravity law of spatial interaction: Reply", Journal of Regional Science, Vol. 10, pp.407-409.

______________________________(1972). "An economic derivation of the gravity law of spatial interaction: A further reply and a reformulation", Journal of Regional Science, Vol. 12, pp.127-136.

Quandt, R.E. (1956). "A probabilistic theory of consumer behaviour", Quarterly Journal of Economics, Vol. 70, pp.507-536.

Quandt, R.E. and Baumol, W. (1966). "The demand for abstract transportation modes: Theory and measurement", Journal of Regional Science, Vol. 6, pp.13-26.

Quarmby, D.A. (1967). "Choice of travel mode for the journey to work: Some findings", Journal of Transport Economics and Policy, Vol. 1, pp.273-315.

Robinson, W.S. (1950). "Ecological correlations and the behaviour of individuals", American Sociological Review, Vol. 15, pp.351-357.

Shunk, G.A. and Bouchard, R.J. (1970). "An application of marginal utility to travel mode choice", Highway Research Record No. 322, pp. 30-39.

Sommers, A.N. (1970). "Towards a theory of traveller mode choice", High Speed Ground Transportation Journal, Vol. 4, pp.1-8.

Stopher, P.R. and Lisco, T.E. (1970). "Modelling travel demand: A disaggregate behavioural approach: Issues and applications", Transportation Research Forum.

Stopher, P.R. and Lavender, J.O. (1970). "Disaggregate behavioural travel demand models: Empirical tests of three hypotheses", Transportation Research Forum.

Strotz, R.H. (1957). "The empirical implications of a utility tree", Econometrica, Vol. 25, pp.269-280.

Strotz, R.H. (1959). "The Utility tree - A correction and further appraisal", Econometrica, Vol. 27, pp. 482-488.

Vitt, J. et al (1970). "Determining the importance of user related attributes for a demand responsive transportation system", Highway Research Record No. 318, pp. 50-65.

Warner, S.L. (1962). Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice (Evanston: Northwestern University Press).

Weiner, E. (1969). "Modal split revisited", Traffic Quarterly, January, pp. 5-28.

Wooton, F.H.J. and Pick, G. (1973). "A model of trips generated by a household", Journal of Transport Economics and Policy, Vol. 1, No. 2.

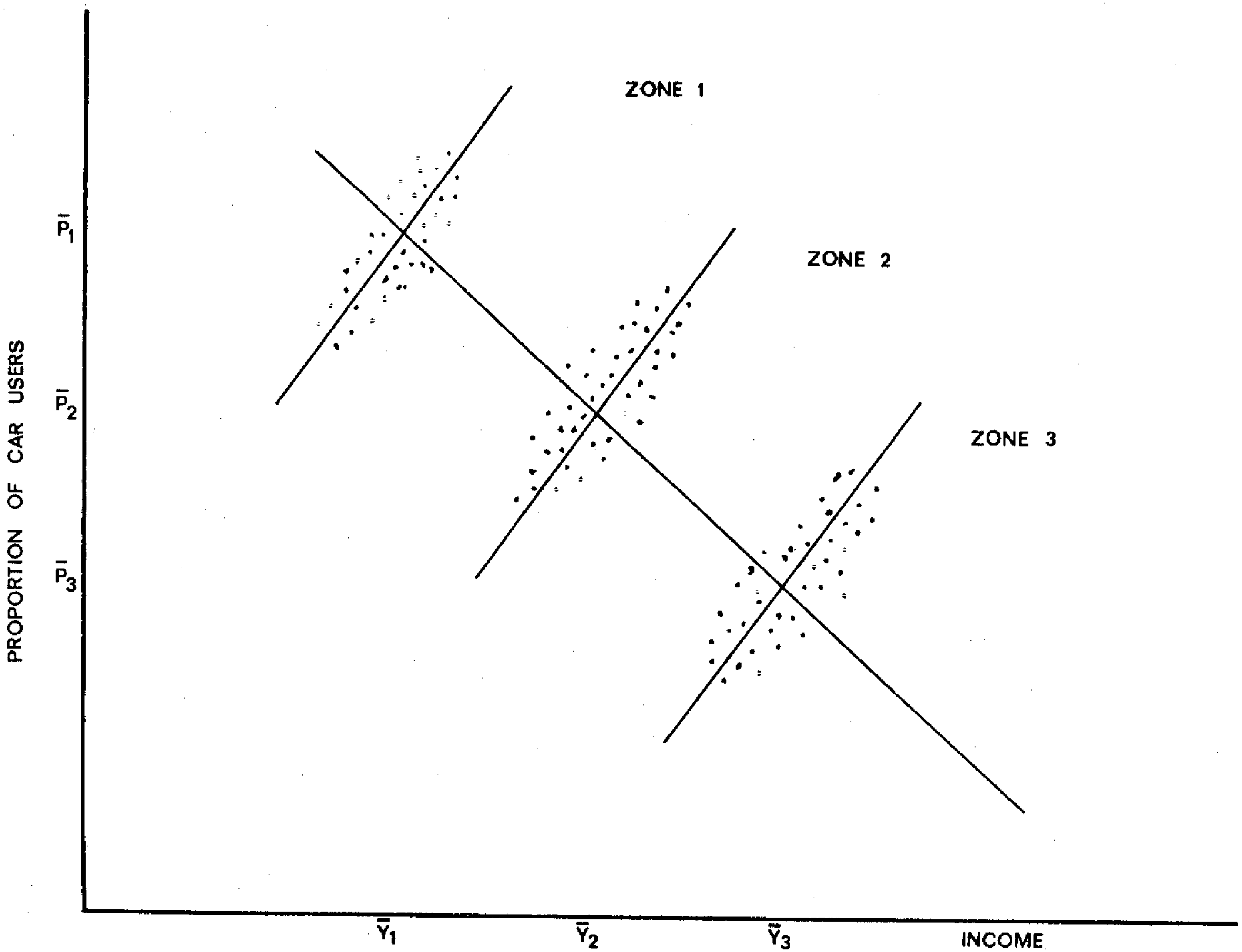

FIGURE 1 THE ECOLOGICAL FALLACY