Working Paper No. 263

# BEHAVIOURAL TRAVEL THEORIES
# MODEL SPECIFICATION
# AND THE RESPONSE ERROR PROBLEM

H.C.W.L.Williams*
and
J.D.Ortuzar**

* School of Geography,
University of Leeds,
LEEDS LS2 9JT,
England.

** Institute for Transport Studies,
University of Leeds,
LEEDS LS2 9JT,
England.

October 1979.

C O N T E N T S

ABSTRACT

WILLIAMS, H.C.W.L. and ORTUZAR, J.D. (1979) Behavioural travel theories, model specification and the response error problem. Leeds: Univ. Leeds.


Conventional or first generation transport models have for some time been heavily criticised for their lack of behavioural content and inefficient use of data; more recently second generation or disaggregate travel demand models based on a theory of choice between discrete alternatives have also been viewed critically. Firstly, it has been argued that implemented structures - and particularly the Multinomial Logit model - have not been sufficiently general to accommodate the 'interaction' between alternatives; and secondly, and perhaps more importantly, that the underpinning theory, involving a perfectly discriminating rational man ('homo economicus'), endowed with complete information is an unacceptable starting point for the analysis of behaviour. In this paper the potential errors in forecasting travel response arising from theoretical misrepresentation are investigated; more generally, the problems of inference and hypothesis testing in conjunction with cross-sectional models are noted.

A framework is developed to examine the consequences of the divergence between the behaviour of individuals in a system, the observed, and that description of their behaviour (which is embedded in a forecasting model) imputed by an observer, the modeller. The extent of this divergence in the context of response to particular policy stimuli is examined using Monte Carlo simulation for the following examples:

i) Alternative assumptions relating to the structure of models reflecting substitution between similar alternatives.
ii) Alternative decision-making processes.
iii) Limited information and satisficing behaviour.
iv) Existence of habit and thresholds in choice modelling.

The method has allowed particular conclusions to be made about the importance of theoretical misrepresentation in the four examples. More generally, it highlights the problems of forecasting response with cross-sectional models and draws attention to the problem of validation which is all too often associated solely with the goodness of statistical fits of analytic functions to data patterns.

## 1. INTRODUCTION

In the absence of specific information on travel and locational response to transport system changes it has become almost universal practice to infer the propensity of individuals to modify their behaviour from trip patterns revealed at a single cross section. Explanations and theories of traveller behaviour become associated with the interpretation of dispersion - the variability exhibited when individuals with similar observable characteristics are associated with different travel related options.

It is a fundamental assumption of the cross sectional approach that a measure of the response to (incremental) change may be assessed from demand functions simply by determining their derivatives with respect to the policy variables in question. Thus, if we write the demand model as a functional relation

$$y_s = f( \dots \underline{Z}_r \dots ; \underline{\theta}) \qquad r, s = 1, 2, \dots \tag{1.1}$$

between the probability $y_s$ of occupying a particular state s, the vectors of attributes $\underline{Z}_r$, $r = 1, \dots, s, \dots$, and the vector of parameters $\underline{\theta}$, the response $\delta y_s$ to a policy stimulus, identified by the combination of incremental changes in the components $(\delta Z_r^\mu)$, may be written

$$\delta y_s = \sum_r \sum_\mu \frac{\partial f_s}{\partial Z_r^\mu} \cdot \delta Z_r^\mu \ . \qquad s = 1, 2, \dots \tag{1.2}$$

$y_s$ and $\delta y_s$ might, for example, refer to the probability and its change of selecting a particular travel related option. We shall write this in the matrix notation.

$$\underline{\delta y} = \underline{g}(\underline{Z}, \underline{\theta}) \cdot \underline{\delta Z} \tag{1.3}$$

in which $\underline{g}$ contains the partial derivatives of the demand function.

The assumption that a realistic stimulus-response relation may thus be derived from the cross sectional model (1.1) has been elevated to such a level of faith that the notion of validity of a travel response forecasting model is often interpreted exclusively in terms of the statistical goodness-of-fit achieved between the base predictions and the observed travel behaviour. We shall argue below that the traditional use of the cross-sectional approach, which has transcended the significant differences between two generations of travel demand models, has important implications for response forecasting and more particularly for the development of behavioural hypotheses which are associated with a demand function.

The popularity of the so-called second generation demand models conceived in terms of probabilistic consumer choice concepts has, in part, been founded on the successful marriage of an explicit theory of behaviour with a micro-representation permitting an efficient statistical analysis of travel dispersion at the level of the individual traveller.

While the merits of the 'conventional' micro-approach have been widely recognized, there is a broad and growing body of literature inspired by what are considered to be the deficiencies of the commonly adopted model representations, and in particular the theoretical basis underpinning them (see, for example, Burnett and Thrift, 1979; Stopher and Meyburg, 1976; Heggie, 1978; Heggie and Jones 1978; Banister, 1978). While these criticisms may be discussed under the generic heading of behavioural misrepresentation, there are two particularly important classes of comment which are worthy of note, relating to : the definition and characterization of travel related options which are deemed relevant to the choice and response contexts; and secondly, the nature and characterization of the decision process attributed to an individual. In the latter context it is argued that the notion of a perfectly discriminating 'rational man' endowed with complete information is an unrealistic starting point for the analysis of travel behaviour.

In this paper we seek to assess the <u>consequences</u> of behavioural misrepresentation. The incentive to study this issue comes not only from the broad-based assault on the assumptions underpinning both first and second generation models, but also from the results of a number of studies which have been able to compare model predictions with actual behaviour, thus providing an appropriate consideration of model validity.

Specific examples of the very considerable discrepancies between forecasts and behavioural response, although scarce and largely confined to modal or route share contexts, are not confined to any particular class of model. First generation, or aggregate models, tend to produce more frequent examples, no doubt because of their widespread application over two decades. One fairly typical example is that of the Park-and-Ride service for Maidstone, in which it was found that

"... in each case the model 'predicts' a demand about double the patronage actually observed" (Kent County Council, 1976)

More relevant perhaps in this context are the results of the inspired, but particularly unsuccessful, Zones and Collar experiment in Nottingham (Vincent and Layfield, 1977) carefully monitored by government bodies, where the effect of time-penalty policies fell dramatically short from expectations.

Second generation models do not escape this plight; a disturbing example is documented by Train (1978) who records considerable differences between observed demand and 'predicted' results produced by a sophisticated model applied before the opening of the Bay Area Rapid Transit (BART) system in the United States.

I appears that no better results are obtained using models based on <u>stated</u> rather than <u>revealed</u> preferences. Papaoulias and Heggie (1976), for example, found that after taking into account compensating errors, the prediction errors in their park-and-ride model were 52% for week day forecasts and 33% for Saturday forecasts.

In this paper we consider the potential errors which can arise in response forecasts from the mis-specification process. Four specific examples concerning misrepresentation of the behavioural hypotheses underpinning choice models are taken to illustrate the approach. As we are not in the possession of actual response behaviour we have simulated data pertaining to choice contexts on the basis of particular specifications of the behavioural process. A model(s) has then been applied to this data and its (their) predictions compared with the simulated response. It is a central concern to determine whether two or more competing models of behaviour can generate equally acceptable 'statistical fits' at the cross section, and differ significantly in their response predictions. The issue of mis-specification, and the role played by the <u>observer</u> in this process, are discussed in detail in Section 2 together with the simulation test method through which 'response errors' are observed.

In Sections 3 through to 6, four examples of misrepresentation/ mis-specification are considered, and deal in turn with:

(i) The selection of a demand model which is inappropriately structured in relation to its accommodation of the perceived similarity between travel related alternatives. We shall, in particular, be able to discuss some of the theoretical frailties of the Multinomial Logit model and assess their importance.

(ii) alternative models of the decision process whereby individuals are considered to identify a preferred alternative.

(iii) the relaxation of the assumptions of perfect information and discrimination in the choice process.

(iv) the existence of habit in choice contexts and its influence on response analysis.

While it is a matter for empirical consideration to determine the extent to which such aspects as habit, information limitations and satisficing behaviour assert themselves in a real world context, it is our intention to investigate how they impinge on the modelling process itself and thereby reflect on the notion of validity of a demand response forecasting model.

## 2. DECISION CONTEXTS, FRAMES OF REFERENCE AND THE SIMULATION OF BEHAVIOUR.

> "The theory of rational choice behaviour asserts that a decision maker can rank possible alternatives in order of preference, and will always choose from available alternatives the option which he considers most desirable, given his tastes and the relative constraints placed on his decision making, such as his level of income or time availability. Suitably modified to take account of the psychological phenomena of learning and perception errors, this theory . . . forms the foundation of modern economic analysis" (Domencich and McFadden, 1975).

As McFadden (1975) has further remarked

> ". . . classical economic analysis makes the assumption of homo economicus virtually tautological, if an object is chosen then it must maximise 'utility' <u>as the chooser perceives it</u>" (our emphasis).

In this sense any aspects of traveller behaviour which are deemed 'irrational' or 'idiosyncratic' are but apparent and simply the consequences of inaccurate or inappropriate descriptions of behaviour in the frame of reference selected by the <u>observer</u>.

One of the objects of this paper is to investigate the extent to which an <u>observer</u> of travel related behaviour can impose particular and inappropriate perceptions of that behaviour on the modelling process, and the consequences of this for response forecasting. In order to examine this question we introduce two <u>perspectives</u> A and A*, each of which involves a description of behaviour within a selected frame of reference.

The former, $\Lambda$, entails a description of behaviour adopted by an observer (the modeller) who is endowed only with the information in traditional cross sectional surveys. $\Lambda^*$, on the other hand, is a perspective of relative priviledge and will involve a more 'realistic' description of the individual decision process (it does not concern us yet whether this information is relevant and how it is obtained). In the examples to follow this chapter, $\Lambda^*$ will, in fact, involve simple refinements to the assumptions normally employed for the generation of choice models.

We shall now assume that $\Lambda$ and $\Lambda^*$ are characterized by the following aspects of a choice context, for each individual i

$$\boxed{\Lambda^* : \{\underline{d}^*; \underline{A}^*(s^*,\underline{Z}^*) \; ; \psi^*;Q^*\}_i} \qquad \boxed{\Lambda : \{\underline{d}; \underline{A}(s,\underline{Z}) \; ; \psi \; ; Q\}_i}$$

in which we denote

| | | |
|---|---|---|
| $\underline{d}$, $\underline{d}^*$ | : | descriptors (attributes) of individual decision makers i |
| $\underline{A}$, $\underline{A}^*$ | : | the sets of alternatives out of which a selection is considered by individual i. Each alternative will be characterized by state descriptors s and s* respectively (which might relate, for example, to a formal specification of an individual travel tour or journey), and each state will be characterized by the sets of attributes $\underline{Z}^*$ and $\underline{Z}$. |
| $\psi$, $\psi^*$ | : | the set of constraints to which the individual is subject. These may include : travel time, cost, family interdependence constraints, in addition to search time constraints. |
| Q, Q* | : | the set of objectives motivating the choice process. |

These formal descriptors are introduced simply to characterize the two perspectives. In the formation of mathematical expressions for predicting response these aspects are incorporated into explicit models $D_i^*$ and $D_i$ of the decision process. In the simulation tests on mis-specification which we shall describe below, the focus of attention will be on the differences between the two perspectives.

Now, if the available information in a perspective does not allow each individual to be deterministically assigned to the appropriate alternative (with no residual dispersion or misallocation in the base year) it is necessary to invoke probabilistic choice concepts in which an individual is considered to be sampled from a population, each member of which is associated with the same measured characteristics. The probabilistic choice model involves the aggregation over individual decisions and will be resolved in the form of an analytic function, an algorithm, or a simulation process.

Cast in these terms, there is considerable scope for misrepresentation of the behavioural theory underpinning a choice model, firstly in the description of the choice context; secondly, in the specification of the decision model, and thirdly, in the aggregation process over individuals. Let us examine some of the assumptions embedded in the process by which conventional choice models, and the Multinomial Logit model in particular, are generated within random utility theory. An explicit statement of these assumptions will be useful later for comparative purposes.

(i) Choice making populations are identified with individual segments of the transport markets. Constraints are handled either explicitly or implicitly through the use of proxy variables.

(ii) All individuals $I_i$ in a given market segment S (we assume that certain obvious constraints are catered for, eg. the possession of a driving licence prior to being a car driver), have the same deterministic choice set A, containing members $A_1, \ldots, A_\rho, \ldots, A_N$.

(iii) The objectives of each individual are resolved in the formation of utility functions $U_\rho(\underline{Z}, \underline{\theta})$ which are used to record preferences.

(iv) The decision model is simply one of utility maximisation:

$D_i$ : Individual i will select $A_\rho$ if

$$U^i_\rho > U^i_{\rho'} \; ; \; \forall A_{\rho'} \in A \qquad (2.1)$$

(v) Dispersion can arise either because attributes of the individual and/or of the alternatives are unobserved, and/or there is taste variation over measured factors in the attribute set $\underline{Z}$. Formally, this is accommodated by decomposing utility functions $U_\rho$ into a 'representative' component $\bar{U}_\rho$ and a component $\varepsilon_\rho$ which is a deviation from the group average.

Thus

$$U^i_\rho = \bar{U}_\rho + \varepsilon^i_\rho \qquad \forall\, I_i \in S,\ \forall\, A_\rho \in A \tag{2.2}$$

(vi) The aggregation process is performed by assuming the 'residuals' $\underline{\varepsilon}$ to be distributed randomly over the population S. Specification of the joint probability distribution function $f(\varepsilon_1, \ldots, \varepsilon_\rho, \ldots, \varepsilon_N) \equiv f(\underline{\varepsilon})$ allows the choice probabilities $P_\rho$ to be determined by integration over that portion of utility space $R_\rho$ for which condition (2.1) holds:

$$P_\rho = \int_{R_\rho} d\underline{\varepsilon}\, f(\underline{\varepsilon}) \tag{2.3}$$

(vii) If the random components $(\varepsilon_1, \ldots, \varepsilon_\rho, \ldots, \varepsilon_N)$ are identically and independently distributed according to Weibull functions $W(0,\sigma)$ with standard deviation $\sigma$, the aggregation process (2.3) may be resolved analytically to yield the Multinomial Logit model (Domencich and McFadden, 1975; Cochrane, 1975).

$$P_\rho = \frac{e^{\theta \bar{U}_\rho}}{\sum_\rho e^{\theta \bar{U}_\rho}} \tag{2.4}$$

in which

$$\theta = \frac{\pi}{\sigma\sqrt{6}} \tag{2.5}$$

(viii) The representative utilities $\bar{U}_\rho$ are expressed as linear functions of the attributes $\underline{Z}$, and the parameter set $\underline{\phi}$ (to be determined)

$$\bar{U}_\rho = \sum_\mu \phi_\mu Z^\mu_\rho \tag{2.6}$$

(ix) The response $\underline{\delta P}$ arising from a change $\delta\underline{Z}$ is traced directly through Equations (2.6) and (2.4), assuming constant parameters $\theta$ and $\underline{\phi}$.

Now, the adoption of the Multinomial Logit (MNL) model does not <u>necessarily</u> imply a commitment to the theory outlined above, as it is well known that the model can be derived from other theoretical standpoints (see, for example, Thrift and Williams, 1979). There may also be other explanations of dispersion which are statistically consistent with the model. We shall seek such explanations, which are based on assumptions underpinning models of behaviour which can be approximated in a statistical sense by the MNL function.

In order to investigate the implications of different behavioural perspectives a set of simulation experiments have been designed in which a data set consisting of population choices are generated according to $\Lambda^*$. Choice probabilities are determined by simulating the outcome of the decision models $D_i^*$ for all members in the choice making population S. A model, which will be considered to be underpinned by a perspective $\Lambda$, is then estimated using this data, and used to predict response under conditions of change. Modelled forecasts are then compared with those separately generated using simulation according to $D^*$, and the deviation or error between the response forecasts is recorded. The experimental scheme is depicted in Figure 2.1.

The model adopted will usually be of Multinomial Logit form, and we shall provisionally identify with the perspective $\Lambda$ the stages outlined in steps (i) to (ix) above. The perspective $\Lambda^*$ will involve the relaxation of one or more of the assumptions employed in these steps. The simulation tests described in Sections 3,5 and 6 will, in fact, involve relaxations of those assumptions associated with steps (vii), (ii) and (ix), respectively, and in each case we shall juxtapose the different aspects of the perspectives $\Lambda$ and $\Lambda^*$. In Section 4 we shall comment on the modification of steps (iii) and (iv).

In each case we seek alternative explanations and models which are statistically consistent with given patterns of choice behaviour, and which result in different response forecasts. This difference will be referred to as the response error, and is a measure, as we shall later discuss, of the consequences of failing to discriminate between alternative behavioural hypotheses at the cross section.

In the simulation experiments to be described below sets of data which consist of the discrete probability distribution $[P_1^*, \ldots, P_\rho^*, \ldots P_N^*]$, identified by a vector $\underline{P}^*$, are generated by Monte Carlo methods. Each member of the population $S^*$ of size M is sampled and assigned to a particular alternative $A_1, \ldots A_N$ according to the outcome of the decision rule D* associated with the perspective A*. The proportion $M_\rho/M$ which 'select' $A_\rho$ is then identified with $P^*_\rho$ when M is large. What was considered 'large' was the subject of a separate investigation. (The resolution of choice models in this way, using Monte Carlo simulation has also been considered by Albright et. al. (1977), Ortuzar (1978), Robertson (1978), Robertson and Kennedy (1979), Manski and Lerman (1978)).

The vector $\underline{P}^*$ is now considered to be the travel related choice data available to the modeller. Having accounted for any variation in behaviour arising from <u>observed</u> differences in individual choice contexts - that is having identified a choice making population S - the modeller might now seek to attribute the remaining dispersion in $\underline{P}^*$ to the dispersion in perceived utility. If then a logit model

$$P_\rho = \frac{e^{\theta \bar{U}_\rho}}{\sum_\rho e^{\theta \bar{U}_\rho}}$$

is adopted to account for the variation in the data $\underline{P}^*$, and found to provide a good statistical fit, the estimated parameter θ will then be consistent with an interpretation of $\underline{P}^*$ having been generated by Weibull utility distributions with means $\{\bar{U}_1, \ldots \bar{U}, \ldots \bar{U}_N\}$ and common standard deviation σ given by Equation (2.7).

$$\sigma = \frac{\pi}{\sqrt{6}\,\theta} \qquad (2.7)$$

As $\underline{P}^*$ will, in fact, have been generated according to an alternative set of hypotheses, it is clear that a good fit to base patterns is in this case consistent, in a statistical sense, with at least two interpretations of the observed dispersion or variation in behaviour. It now remains to see whether the response predicted by the model used by the observer (the Multinomial Logit function in the case above), provides an accurate approximation to the response of the population S achieved by simulation from D* according to a revised set of external conditions - interpreted as a policy stimulus.

We are now in a position to consider an example of structural misspecification, which involves the use of a model insufficiently refined to accommodate the perceived similarity between alternative choices.

## 3. MODEL STRUCTURES AND THE SIMILARITY OF TRAVEL-RELATED SUBSTITUTES

The construction of travel choice models for contexts in which the alternatives are endowed with degrees of 'similarity' has resulted in a series of ambiguities and inconsistencies which have recently been resolved within the framework provided by random utility theory (Williams, 1977; Langdon, 1976; Daly and Zachary, 1978; McFadden, 1979; Haussman and Wise, 1978). It is now well known that if the Multinomial Logit model is indiscriminantly applied to choice contexts, involving multiple modes, multiple routes, mode-route or location-mode combinations, the cross elasticities obtained, and therefore the response properties of the model, will often be blatantly unacceptable. In fact the Multinomial Logit model has often been applied in conjunction with utility functions which are inconsistent with its formation (Williams, 1977; 1979).

Within random utility theory in which models are generated from equations (2.1)-(2.3), the structure of that model is directly and uniquely determined by the utility functions, and a number of parameters which characterise that structure will be embodied in the variance-covariance matrix $\underline{\Sigma}$. The elements of $\underline{\Sigma}$ are defined by

$$\Sigma_{\rho\rho'} = E(\varepsilon_\rho \varepsilon_{\rho'}) \tag{3.1}$$

in which $E(\cdot)$ denotes the expectation value. The Multinomial Logit model, generated from identical and independent Weibull distributions, is characterised by the matrix with elements

$$\Sigma_{\rho\rho'} = \sigma^2 \delta_{\rho\rho'} \tag{3.2}$$

in which

$$\begin{aligned} \delta_{\rho\rho'} &= 1 \text{ if } \rho = \rho' \\ &= 0 \text{ otherwise} \end{aligned} \tag{3.3}$$

A number of model structures have now been proposed which accommodate varying degrees of similarity or correlation, ranging from: the Nested or Hierarchical Logit (HL) model; the Cross-Correlated (CCL) Logit model (Williams, 1977) and the General Extreme Value (GEV) model proposed by McFadden (1979); to the general but computationally unwieldy Multinomial Probit model (Daganzo et al., 1977). While the Multinomial Logit (MNL) model which entails no aspects of correlation or similarity may be theoretically restrictive it is important to know the extent of the potential misspecification. We shall examine this issue for a particular choice context which involves a combination of choice dimensions X and Y - say location and mode, or mode and route. If $X_\mu$, $\mu=1,2, \ldots$ represents the set of alternatives available in the X dimension and $Y_\nu$, $\nu=1,2, \ldots$ that in the Y dimension the total set of available alternatives $\{\ldots A_\rho \ldots\}$ is composed of the combinations $\{\ldots X\mu Y\nu \ldots\}$.

The utility function governing choice between these alternatives will be taken to be of the traditional form (Williams, 1977).

$$U(\mu,\nu) = U_\mu + U_\nu + U_{\mu\nu} \quad \forall \mu\varepsilon X, \quad \nu\varepsilon Y \tag{3.4}$$

in which $U_\mu$ and $U_\nu$ are components which are specific to the choice dimensions X and Y respectively, while $U_{\mu\nu}$ is an interaction term. Thus in a location (X) and mode (Y) choice context, $U_\mu$ would refer to components which vary over locations but not modes, $U_\nu$ vary over modes but not locations, while $U_{\mu\nu}$ are components, perhaps transportation costs, which vary over locations and modes.

If now $U(\mu,\nu)$ is written in terms of representative utilities and residuals

$$U(\mu,\nu) = \bar{U}(\mu,\nu)+\varepsilon(\mu,\nu) \tag{3.5}$$

in which

$$\bar{U}(\mu,\nu) = \bar{U}_\mu+\bar{U}_\nu+\bar{U}_{\mu\nu} \tag{3.6}$$

and

$$\varepsilon(\mu,\nu) = \varepsilon_\mu+\varepsilon_\nu+\varepsilon_{\mu\nu} \tag{3.7}$$

the variance-covariance matrix can be expressed in terms of the following expectation value

$$\Sigma_{\mu\nu,\mu'\nu'} = E(\varepsilon_\mu+\varepsilon_\nu+\varepsilon_{\mu\nu},\varepsilon_{\mu'}+\varepsilon_{\nu'}+\varepsilon_{\mu'}+\varepsilon_{\mu'\nu'}) \tag{3.8}$$

We shall now assume that $\varepsilon_\mu,\varepsilon_\nu$ and $\varepsilon_{\mu\nu}$ are separately identically and independentaly distributed, with

$$E(\varepsilon_\mu\varepsilon_{\mu'}) = \sigma_X^2\delta_{\mu\mu'} \tag{3.9}$$

$$E(\varepsilon_\nu\varepsilon_{\nu'}) = \sigma_Y^2\delta_{\nu\nu'} \tag{3.10}$$

$$E(\varepsilon_{\mu\nu},\ \varepsilon_{\mu'\nu'}) = \sigma_{XY}^2\ \delta_{\mu\mu'}\delta_{\nu\nu'} \tag{3.11}$$

with all cross terms (such as $E(\varepsilon_\mu\varepsilon_\nu)$) vanishing. The matrix elements of $\underline{\underline{\Sigma}}$ now become

$$\Sigma_{\mu\nu,\ \mu'\nu'} = \sigma_X^2\ \delta_{\mu\mu'} + \sigma_Y^2\ \delta_{\nu\nu'} + \sigma_{XY}^2\ \delta_{\mu\mu'}\ \delta_{\nu\nu'} \tag{3.12}$$

Models and their underpinning utility functions can be referred to in terms of the utility coordinates $(\bar{U}_X,\ \bar{U}_Y,\ \bar{U}_{XY};\ \sigma_X,\ \sigma_Y,\ \sigma_{XY})$. A set of special cases may be identified by the structure of the similarity between alternative choices, which are reflected in the variance-covariance matrix $\underline{\underline{\Sigma}}$. These are shown in Figure 3.1. In the first case, both $\sigma_X$ and $\sigma_Y$ are zero and a diagonal $\underline{\underline{\Sigma}}$ matrix results. This case will correspond to the Multinomial Logit (MNL) model if the utility functions are drawn from independent Weibull distributions. It is clear that the use of the utility function (3.4) in a MNL model of the form

$$P_{\mu\nu} = \frac{e^{\Delta\bar{U}(\mu,\ \nu)}}{\sum\limits_{\mu\nu} e^{\Delta\bar{U}(\mu,\ \nu)}} \tag{3.13}$$

will be inconsistent because the appropriate $\underline{\underline{\Sigma}}$ matrix, corresponding to that utility function is not of the diagonal form involved in the generation of the model (3.13).

In the second special case, the component $\sigma_Y$ vanishes and the structure of the variance-covariance matrix may be represented in the 'tree' or 'hierarchical' form shown in Figure 3.1(b). The two parameters $\sigma_X$ and $\sigma_{XY}$ allow different degrees of cross-substitution between intra- and inter-branch alternatives. That is, between $X_\mu Y_\nu$ and $X_\mu Y_{\nu'}$ in the former case, and between $X_\mu Y_\nu$ and $X_{\mu'} Y_{\nu'}$ in the latter. The Hierarchical Logit (HL) model (Williams, 1977; Daly and Zachary, 1978) which is a special case of the General Extreme Value model proposed by McFadden (1979) is an appropriate structure to accommodate this form of correlation between alternatives. It has the following analytic structure

$$P_{\mu\nu} = \frac{e^{\beta(\bar{U}_\mu + \bar{U}_{\mu *})}}{\sum_\mu e^{\beta(\bar{U}_\mu + \bar{U}_{\mu *})}} \cdot \frac{e^{\Delta(\bar{U}_\nu + \bar{U}_{\mu\nu})}}{\sum_\nu e^{\Delta(\bar{U}_\nu + \bar{U}_{\mu\nu})}} \tag{3.14}$$

with

$$\bar{U}_{\mu *} = \frac{1}{\Delta} \ln \sum_\nu e^{\Delta(\bar{U}_\nu + \bar{U}_{\mu\nu})} \tag{3.15}$$

and

$$\beta \leq \Delta \tag{3.16}$$

The conditions (3.15) and (3.16) must be satisfied if the model is to be consistent with utility maximisation. The inequality (3.16) is of particular importance and its violation may imply cross-elasticities of the wrong sign. This violation has, in fact, been observed in conventional transportation models (Williams and Senior, 1977). The relationship between the dispersion parameters $\beta$ and $\Delta$ and the utility distribution parameters $\sigma_X$ and $\sigma_{XY}$ are given as follows (Williams, 1977)

$$\Delta = \frac{\pi}{\sqrt{6}\, \sigma_{XY}} \tag{3.17}$$

$$\beta = \Delta\left(1 + \frac{6}{\pi^2} \sigma_X^2 \Delta^2\right)^{-\frac{1}{2}} \tag{3.18}$$

When $\sigma_X = 0$ the model collapses to the Multinomial Logit form which is characterized by the single parameter $\Delta$.

The alternative special case, with $\sigma_X = 0$ and $\sigma_Y \neq 0$ corresponds to the $\underline{\Sigma}$ matrix and pictorial representation shown in Figure 3.1(c). The functional form is simply derived from Equation (3.14), by interchanging the X and Y dimensions and relabelling.

In summary, the first three cases in Figure 3.1 involve utility maximisations in which the variance-covariance matrices $\underline{\Sigma}$ are special cases of the 'cross-correlated structure' with a $\underline{\Sigma}$ matrix and pictorial representation summarized in Figure 3.1(d). The Cross-Correlated Logit (CCL) model has been proposed as an approximation to the exact three parameter utility maximising model derived from the utility function (3.4). Its detailed form is given by Williams (1977), and further considered by Ortuzar (1979).

We are now in a position to present a set of simulation tests on structural mis-specification which are designed to examine the following questions:

(i) How good an approximation is the three parameter CCL model to the exact model generated from Equation (3.4) through utility maximisation?

(ii) What potential errors are made by invoking the single parameter MNL and two parameter HL models, which accommodate restricted degrees of simularity between alternatives, to an appropriate three parameter specification?

The simulation tests involve the perspectives : $A^*$ and $A$ specified by

| | |
|---|---|
| $A^*: D^*(\bar{U}^*_X, \bar{U}^*_Y, \bar{U}^*_{XY}; \sigma^*_X, \sigma^*_Y, \sigma^*_{XY})$ | $A : D(\bar{U}_X, \bar{U}_Y, \bar{U}_{XY}; \sigma_X, \sigma_Y, \sigma_{XY})$ |

Data are generated by direct simulation from utility functions of the form (3.4). The four models (MNL, two alternative HL, and CCL) are then adopted for assessing goodness of fit and response error. In the MNL model and alternative hierarchical forms, the respective parameters $\Lambda$ and $(\beta, \Lambda)$ were estimated by maximum likelihood. (The adoption of one of these models may be interpreted as accepting either $\sigma_X$ or $\sigma_Y$ or both to be zero). For the Cross-Correlated Logit model, which we re-emphasise is an approximation to the exact choice model generated from (3.4) which embodies HL and MNL forms as special cases, the parameters were theoretically determined (Ortuzar, 1979).

In the first set of tests the corresponding pairs of representative utility values $\bar{U}^*_X$, and $\bar{U}_X$ etc., were taken to be identical and the simulation tests involved variation of the coordinates $(\sigma^*_X, \sigma^*_Y, \sigma^*_{XY})$. A standardization or 'normalization' condition is used to bound the joint variation of these quantities, and is of the form

$$\sigma^{*2}_X + \sigma^{*2}_Y + \sigma^{*2}_{XY} = \text{constant} \tag{3.19}$$

A particular coordinate $(\sigma^*_X, \sigma^*_Y, \sigma^*_{XY})$ corresponds to a simulation test. To illustrate the possible combinations of these three components we can appeal to a property of an equilateral triangle for which the sum of perpendicular distances to the three sides from an interior point is a constant, equal to the height of the triangle. Any test point may thus be identified with a point in or on the boundary of the triangle as shown in Figure 3.2(a). At interior points a three parameter model (such as the CCL function) is necessary to capture the full range of cross substitution implied by the utility function (3.4). On the boundaries CB and CA the alternate Hierarchical Logit models for which $\sigma^*_Y = 0$ and $\sigma^*_X = 0$ respectively are appropriate (see Figures 3.1(b) and 3.1(c)). It is only at the vertex $C(\sigma^*_X = \sigma^*_Y = 0)$ that the Multinomial Logit is an appropriate specification.

Four particular coordinate test points were taken, as shown in Figure 3.2(b), in addition to test points randomly sampled from within the triangle. In all tests two alternatives were taken in each of the X and Y dimensions, allowing a four alternative choice model to be generated. The general performance of the four models for these test points is shown in Figure 3.2(c).

A particular set of base and response results for the four models used to fit data generated from test point 2 in Figure 3.2(b), is shown in Figure 3.2(d). We have restricted ourselves to a comparative quality assessment of the fit, in that it is the relative performance of the models in which interest lies.

Under conditions of change, points in the second and fourth quadrants of the right hand side of Figure 3.2(d) are deemed pathological because the change in behaviour predicted by the model is opposite to that simulated. This type of behaviour has been associated with the violation of the condition (3.16) in Hierarchical Logit model specifications. If, for example $\sigma_X^* >> \sigma_Y^*$ then the specification Y/X corresponding to $\sigma_Y > \sigma_X$ will usually involve pathological behaviour. The condition appears also to depend on the underlying representative utility values.

The results of the simulations are consistent with the following conclusions:

(i) the Cross-Correlated Logit (CCL) model is a good theoretical approximation to the three parameter utility function (3.4). The superiority to other logit forms is especially apparent when the three quantities $\sigma_X$, $\sigma_Y$ and $\sigma_{XY}$ are rather different from zero and from each other. Its estimation is complex (Williams, 1977) and for this reason does not commend itself.

(ii) the Multinomial Logit (MNL) model performs reasonably well at interior points of the triangle (it is considerably more robust than the authors had anticipated) and this is particularly true when $\sigma_X^* \simeq \sigma_Y^*$. Its maximum error occurs near the sides of the triangle (except in the immediate region of point C) where the 'independence from irrelevant alternatives' property is a considerable impediment.

(iii) Out of the three alternative structures (MNL and two HL forms), the model which provides the best base fit provides a good estimate of the response to change.

In a second series of tests the restriction ($\overline{U}_X^* = \overline{U}_X$; etc) was lifted and the effect of omitting a particular component - by putting $\overline{U}_X = 0$ for example - was determined. All results were inferior to their counterparts in the first test series, as would be expected because the number of 'degrees of freedom' of the model specifications in A had been reduced. The performance of the HL models was particularly suspect and pathological behaviour became more prevalent.

In the two series of tests we have examined the capability of extended members of the logit family to accommodate the structure of similarity between alternatives embodied in the utility function (3.4). We could, of course, have appealed to the Multinomial Probit function as a model appropriate to (3.4), and indeed to more general utility forms. Horowitz (1978) has, in fact, examined the potential mis-specification problems of the <u>Multinomial Logit</u> model compared with a 3-alternative probit model. The general probit model is, however, rather cumbersome in its estimation and currently restricted to 10-15 options.

Although the simulation tests have been confined to a 2 x 2 example (4 options) it is our opinion that the conclusions above would not be qualitatively modified when the number of alternatives are increased, because the structure of the variance-covariance matrix itself has been the focus of the mis-specification tests. (The dependence of the results on the number of options could, of course, be tested). Our results above, and particularly those in (iii), would seem to lend some theoretical/ numerical support for the suggestion of Ben-Akiva (1977b) that the results of alternative Hierarchical Logit and Multinomial Logit models could be compared and the appropriate model selected according to the estimated value of the 'similarity' parameter, which in our notation is $\underset{\Delta}{\underline{\beta}}$.

## 4. THE MULTICRITERION PROBLEM, SATISFICING AND THE GENERATION OF MODELS

In this section we shall comment on two related issues: firstly, on alternative models of the decision process by which choices are considered to be made; and secondly, on the rationalisation of demand model structures.

It should be remembered that one of the prime motivations behind the construction of alternative models of the <u>decision process</u> D* - based on, for example, elimination by aspects (Tversky, 1972) - was precisely the limitations of simple scaleable choice models typified by the Multinomial Logit function. It was recognised in the early 1960's that such simple models would often portray restrictive and unrealistic properties of cross-substitution. The defects of these models have sometimes been attributed to the optimisation framework underpinning them, rather than the particular assumptions used in conjunction with this framework.

The development of more general model structures, outlined in Section 3, which can accommodate varying degrees of similarity between alternatives, and thereby exhibit more realistic cross substitution properties, has <u>in a certain sense</u> removed some of the original justifications for the construction of alternative decision models. This does not, of course, imply that the currently adopted model of the decision process and the assumptions through which the more general random utility models have been achieved are necessarily appropriate;

also, it is always desirable to examine competing frameworks.

Alternative models D* through which individuals or organisations reach decisions have been discussed at large in the public administration, management science, decision theory, psychology and economics literature. The concept of satisficing and the theory of 'bounded rationality' proposed by Simon (1955) is perhaps the most widely discussed alternative to an optimisation framework. As Eilon (1972) has remarked

> " ... Optimisation is the science of the ultimate: satisficing is the art of the feasible. The optimiser sets off in a single minded fashion to determine the best solution to a given problem in given circumstances ... The satisficer on the other hand, acquiesces with the proposition that it is seldom possible to define the ultimate in unambiguous terms, and that it is sufficient to do well enough .."

When the notion of satisficing is applied to travel related decisions and particularly those involving location (Heggie, 1978 ; Thrift and Williams, 1979), the decision model is closely bound up with the acquisition of information in a search process. Such models are treated in terms of sequential decision making. As Young and Richardson (1978) have pointed out, a search structured on attributes of choices should be distinguished from one based on alternatives. In the former, attributes are selected in turn and alternatives are 'processed' and maintained or rejected in the search process, depending on the value of these attributes; while in the latter, alternatives are selected in turn and their 'bundle of attributes' examined. At any stage of the process, alternatives which do not satisfy norms or other constraints are eliminated. In a complex decision process of selecting a house say, both strategies of appraisal exist simultaneously.

In order to draw out some of the formal distinctions between the so-called compensatory models (such as those employed in Section 3) and non-compensatory models (as exemplified by the Elimination-by-Aspects model of Tversky, 1972) we shall consider the decision process in terms of the formal solution of a multicriterion problem (see also Eilon, 1972). An individual contemplating a decision is considered to have a set of goals or objectives and a set of constraints. How could he resolve this problem? (and how does he?)

We shall formally state the multicriterion problem as follows:

$$\begin{aligned}
&\text{Max } \{f_1(Z_1^1) \; \ldots\ldots \; f_1(Z_N^1)\} \\
&\{\text{options}\} \\
&\text{Max } \{f_\mu(Z_1^\mu) \; \ldots\ldots \; f_\mu(Z_N^\mu)\} \\
&\{\text{options}\} \\
&\text{Max } \{f_M(Z_1^M) \; \ldots\ldots \; f_M(Z_N^M)\} \\
&\{\text{options}\}
\end{aligned} \tag{4.1}$$

subject to the vector of constraints

$$\underline{g}\,(\underline{Z}) \leq \underline{b} \tag{4.2}$$

in which $f_\mu(Z_\rho^\mu)$ is the value of a criterion function associated with the attribute $Z_\rho^\mu$ accompanying alternative $A_\rho$. For example, we might be interested in finding an option in the N membered set $\{A_1, \ldots, A_N\}$ which simultaneously minimises travel time, minimises cost, maximises comfort, safety, etc. These attributes might, in addition, be requried to satisfy 'absolute' constraints such as (4.2).

If a single alternative is found which simultaneously optimises the M functions and whose attributes are feasible in terms of (4.2) then an unambiguous optimal solution is obtained. There will, in general, however be conflicts between objectives - that is, options will be superior in some respects and inferior in others - and this of course gives the multicriterion problem its flavour.

Eilon (1972) has discussed several approaches to the resolution of (4.1) and (4.2) two of which are particularly relevant to the problem of constructing a decision model. The first is the objective trade-off strategy, which forms the basis of Cost Benefit Analysis, in which a <u>single</u> objective function

$$F = F(f_1, f_2, \ldots f_M)$$

is formed and the appropriate option extracted. If the $f_\mu$ functions are simply the attributes $Z_\rho^\mu$, or linear transformations of them, F may be written

$$F = \{\sum_\mu \alpha_\mu \, Z_1^\mu; \; \ldots\ldots \; \sum_\mu \alpha_\mu \, Z_\rho^\mu, \; \ldots\ldots \; \sum_\mu \alpha_\mu \, Z_N^\mu\} \tag{4.3}$$

and the conventional type of linear trade-off problem must be addressed. The trade-off parameters $\underline{\alpha}$ are determined from either the stated or revealed preferences of the decision maker.

This process of solution which allows some attributes to <u>compensate</u> others corresponds to the traditional decision model employed in Section 3. An alternative approach to the solution of (4.1) and (4.2) is through the conversion of some or all of the objectives to constraints ('norms' or 'thresholds'). That is we might require

$$Z_1^{\mu}, \;\ldots., \; Z_p^{\mu}, \;\ldots.., \; Z_N^{\mu} \geq \overline{Z}^{\mu} \qquad (4.4)$$

in which $\overline{Z}_{\mu}^{\mu}$ is a minimum or maximum satisfactory value for $Z^{\mu}$. In satisficing approaches the distinction between goals and constraints is a moot point.

Various forms of satisficing models may be generated by converting all (or in certain cases some) of the constraints into norms and by establishing a structured search for the desired alternative in conjunction with an elimination strategy. The search structured around the sequential consideration of alternatives, and attributes are but two examples of this heuristic proces. The process itself refers to an individual, and the difficult problem is that of aggregation over different individuals. Explicit models of the search process are accompanied by the rank order by which each individual processes attributes or alternatives. (Gensch and Svestka, 1978). In Tversky's (1972) model, which is a probabilistic elimination by aspects model the ordering by which alternatives are addressed is implicitly determined through probabilities proportional to the utility contribution of the various attributes. It is easily demonstrated, and indeed well known (Tverksy, 1972); Makowsky et al, 1977) that the final option chosen can be very sensitive to the selected ordering.

In order to assess the extent to which alternative decision models can influence response forecasts we can juxtapose the perspectives.

| A*:D: sequential elimination by aspects | A:D: linear compensation utility model with optimisation |
|---|---|

The authors are proposing to implement the HIARC model introduced by Gensch and Svestka (1978) which has actually been compared with a logit formulation. HIARC is a deterministic model which generates aggregate estimates of norms from individual responses, where the responses consist of the sequence through which attributes are considered, the perceived value of attributes and the preferred option. In the present simulation experiment the orderings and thresholds (or norms) are <u>input</u> in A* to generate data. It is too early to say how the response error will depend on the details of D*.

Before considering the influence of search process and satisficing behaviour further, we shall comment briefly on alternative rationalisations of model structures. In the early 1970's it was argued that the structure of the Hierarchical Logit models, such as that in equation (3.14) arose through a sequential decision process, while the Multinomial Logit function was the outcome of a simultaneous process. Now it is known that both have a rationalisation in terms of <u>optimal</u> decision making involving utility functions of particular structures. However, as Brand (1973) has suggested, nested or hierarchical models appear to have some rationalisation through a genuine sequential decision model at the micro-level. The exact details of this justification are as far as the authors are aware still to be resolved. It is conjectured that just as the four structures shown in Figure 3.1 can be rationalised on the basis of utility maximisation, they will also <u>all</u> be rationalised on the basis of 'Elimination-by-Aspects' according to the rank orderings of attributes of all alternatives and the special cases these may take.

## 5. INFORMATION, DISPERSION AND CHOICE SET GENERATION

In conventional probabilistic choice models based on utility maximisation each individual confronted by a choice is considered to have the same deterministic choice set available. There is a growing literature which confirms the widely held view that in location (of jobs and residence) and travel choice contexts individuals act under a restricted knowledge of alternatives and their attributes (see, for example, the papers by Kirby, 1979; Thrift and Williams, 1979; Richardson, 1978 and the references cited therein). Indeed the geographical concept of a mental map is a recognition of the spatial heterogeneity of information (Gould and White, 1974, Young and Richards, 1978).

Models which explicitly recognise these aspects of the choice process have tended to emphasise the search process, in conjunction with aspiration levels and satisficing behaviour, and a well developed mathematical theory of this process is available as detailed by Weibull (1978) and others.

In this section we shall examine information limitations from a perspective relevant to the cross section in which the search process and imperfect knowledge are accommodated through the random generation of choice sets (which are simply collections of alternatives). The search process will be described in terms of its outcome, which is characterised by a distribution over the population S of the number of alternatives searched before a choice is made. We wish to investigate the consequences of assuming complete information - identical and deterministic choice sets for all individuals (the perspective A), when the actual choice process involves a distribution of choice sets (the perspective A*).

In order to generate models involving heterogeneous choice sets we shall relate the probability $P_\rho$ of selecting an alternative $A_\rho$, to the probability of selecting $A_\rho$ from all possible choice sets containing it. We can write, after Manski (1977)

$$P_\rho = \sum_c P_\rho(c)P(c\varepsilon C) \tag{5.1}$$

in which are defined the following:

C: the set of all available choice sets

$P(c\varepsilon C)$: the probability of selecting choice set c from the set C

$P_\rho(c)$: the probability of selecting alternative $A_\rho$ from the choice set c.

It is useful to consider a further decomposition in which the set of all choice sets C is partitioned according to the number of members in each subset c. Equation (5.1) may now be written

$$P_\rho = \sum_{n=1}^{N} \sum_{c\varepsilon c(n)} P_\rho(c).P(c\varepsilon c(n)).P(c(n)\varepsilon C) \tag{5.2}$$

Here c(n) defines the set of all choice sets with n members and $P(c\varepsilon c(n))$ is the probability of drawing a <u>particular</u> choice set with n numbers from c(n). $\sum_{c\varepsilon c(n)}$ denotes summation over all choice sets of size n.

Let us consider this decomposition for the N choices $A_1, \ldots, A_\rho, \ldots, A_N$. Associated with N alternatives there are a maximum of $2^N-1$ available choice sets which can be arranged in terms of the subsets c(1), ... c(n), ... c(N) as follows:

$$
\begin{array}{ll}
c(1) & \{A_1\}, \ldots, \{A_\rho\}, \ldots, \{A_N\}; \ldots \\
\vdots & \\
c(n) & \{A_1 A_2 \ldots A_n\}, \ldots, \{A_{N-n+1}, \ldots, A_{N-1}, A_N\}; \ldots \\
\vdots & \\
c(N) & \{A_1 A_2 \ldots A_N\}
\end{array}
$$

There are clearly ${}^{N}C_{n}$ choice sets in the set $c(n)$. This decomposition of C may be organised on a tree structure as shown in Figure 5.1

The conventional assumption in random utility models is, as we stressed above, that all members of S select an alternative from the single set $c(N)$. That is

$$
\begin{aligned}
P(c(n) \varepsilon C) &= 1 \quad \text{if } n = N \\
&= 0 \quad \text{otherwise}
\end{aligned}
\tag{5.3}
$$

This will be a special case of a choice set distribution considered below.

In order to generate a choice model the probabilities $P(c \varepsilon C)$, which have been decomposed into $P(c \varepsilon c(n))$ and $P(c(n) \varepsilon C)$, must be specified in terms of the size of the sets and the attributes of its members. A decision model must also underpin the process of selecting $A_\rho$ from the sets c containing it.

Two perspectives will be formed based on these two facets of the choice process. $A^*$ will involve a knowledge of dispersion due to the heterogeneity of information interpreted analytically through the probability functions of choice set generation. In A it will be assumed that all members of S scrutinise one available choice set $\{A_1 \ldots A_N\}$ containing all alternatives. Both perspectives will involve a decision model based on utility maximisation - individuals select what they consider to be the best alternative from the sample which each considers.

We can now summarise the simulation experiments through the following description of the perspectives:

| | | | |
|---|---|---|---|
| $A^*$: | Random generation of choice sets | $D^*$: | Utility maximisation from $W(o,\sigma^*)$ |

| | | | |
|---|---|---|---|
| A: | All members have the full choice set available | D: | Utility maximisation from $W(o,\sigma)$ |

In two series of simulation tests, the data $\underline{P^*}$ will be generated from parametric probability distributions - the probability of a choice set of size n being selected being a truncated Binomial function with parameter q. That is, $P(c(n) \varepsilon C)$, which we shall write $P(n|N,q)$, will be taken as

$$P(n|N,q) = {}^{N}C_{n} q^{n}(1-q)^{N-n}(1+\xi_q) \quad n=1, \ldots, N \qquad (5.4)$$

$\xi_q$ is a factor which normalises the distribution due to truncation at n=1 - the probability of selecting a set with zero options is not considered relevant when a choice will actually be made. $\xi_q$ is thus given by

$$\xi q = \frac{(1-q)^N}{\sum_{n=1}^{N} {}^{N}C_{n} q^{n}(1-q)^{N-n}} \qquad (5.5)$$

The mean choice set size selected, which is given by $\sum_{n=1}^{N} nP(n|N,q)$, is approximately equal to Nq (except for q=0 when it tends to unity). The forms of the discrete probability function $P(n|N,q)$ for different q values are shown in Figure 5.1. (Continuous curves are used to distinguish the various functions.) The two special cases corresponding to q=1 and q=0 are worthy of note. In the former

$$\begin{aligned} P(n|N,1) &= 1 \text{ if } n=N \\ &= 0 \text{ otherwise} \end{aligned} \qquad (5.6)$$

which corresponds to the homogenous and complete choice set assumption of conventional theory. For q=0, on the other hand

$$\begin{aligned} P(n|N,0) &= 1 \text{ if } n=1 \\ &= 0 \text{ otherwise} \end{aligned} \qquad (5.7)$$

all individuals select a single alternative from the set of alternatives $\{A_1, \ldots, A_\rho, \ldots, A_N\}$.

Having specified the probability of selecting a particular set c(n) it is necessary to examine the probability of selecting a given set of alternatives $\{\ldots\}_n$, containing n members. In the first series of simulation tests (A), it is assumed that the ${}^{N}C_{n}$ member sets in c(n) are equally probably - that is the probability of selecting a set of alternatives is a function of its size alone. In the second series of simulations (B), the probability of selecting any member $\{\ldots\}_n$ from c(n) will be a function of the mean utilities of its constituents.

In each set of tests five alternatives (N=5) were considered, and the Weibull curves adopted in A* had a standard deviation σ* of 5 units (the difference between the largest and smallest mean utilities was 1 standard deviation). The data $\underline{P}^*(q)$ were generated by sampling choice set sizes randomly from $P(n|N,q)$. In tests (A) choice sets cεc(n) were drawn from a uniform distribution. Utility values were then sampled for the n members from Weibull (and other) distributions. The maximum utility option was recorded, and the process repeated for the M members of S.

For all values of $0<q<1$ it was found that the logit function

$$P_\rho(q) = \frac{e^{\theta(q)\overline{U}_\rho}}{\sum_\rho e^{\theta(q)\overline{U}_\rho}}$$

provided an excellent fit to the base data $P^*(q)$. The variation of the estimated dispersion parameter $\theta(q)$ with q is shown in Figure 5.2(a). For q=1 the complete set $\{A_1, \ldots, A_\rho, \ldots, A_N\}$ is selected by all members of S and the estimated parameter (q=1) may be compared with its exact value $\theta = \frac{\pi}{\sqrt{6}\sigma^*} = 0.257$. A reduction of q, which has the effect of introducing smaller choice sets, results in an increase in the standard deviation $\sigma(q)$ imputed by the observer in accordance with

$$\sigma(q) = \frac{\pi}{\sqrt{6}\theta(q)} \tag{5.9}$$

The agreement between the model predictions $\underline{P}$ and the simulated results $\underline{P}^*$ in the base system are shown for different q values in Figure 5.3(a).

We can now compare the simulated and predicted response, $\underline{\Delta P}^*$ and $\underline{\Delta P}$ respectively, when the utility values $|\overline{U}_1, \ldots, \overline{U}_\rho, \ldots, \overline{U}_N|$ are modified. In fact a single value $\overline{U}_3$ was altered to allow the direct and cross elasticities to be measured. In Figure 5.3(b) we plot the predicted share modifications from the base system against their simulated counterparts for the five alternatives $A_\rho$: $\rho=1, \ldots, 5$. Although the agreement is slightly erratic, the important point to note is the lack of a systematic over- or under-prediction of response.

In the above set of tests (A), the probability of drawing any set of alternatives $\{\ldots\}_n$ with n members was given by

$$P(c\varepsilon c(n)) = \frac{1}{{}^N C_n} \tag{5.10}$$

It might be more realistic to assume that the probability of selecting a particular set $\{\ldots\}_n$ is related to the characteristics of the alternatives in the set - perhaps the utility values, or the spatial configuration. In the second series of tests (B), $P(c\varepsilon c(n))$ involves a discounting factor which reduces the probability of an alternative $A_\rho$

being selected for membership of any set in accordance with the difference between its mean utility $\bar{U}_\rho$ and that of the maximum, max $\{\bar{U}_1, \ldots, \bar{U}_N\} \equiv \bar{U}^+$. That is, the relative odds of membership of $A_1, \ldots, A_\rho, \ldots, A_N$ in a set of any given size are taken as

$$e^{-\kappa(\bar{U}^+ - \bar{U}_1)}, \ldots, e^{-\kappa(\bar{U}^+ - \bar{U}_\rho)}, \ldots, e^{-\kappa(\bar{U}^+ - \bar{U}_N)}$$

respectively. When $\kappa=0$, all members are equally likely to be chosen, while as $\kappa$ increases there is an increasing likelihood that utility options with high mean values will be sampled. (In the actual construction of the tests the options were considered to be spatially arranged and $\kappa$ identified with a spatial discounting factor.)

There are now two parameters characterising information heterogeneity which are identified with the two levels of the tree structure in Figure 5.1. The binomial parameter q governs the choice set size distribution, while $\kappa$ governs the probability of selecting a given set $\{\ldots\}_n$ from $c(n)$.

We can now enquire, as before, if the logit function $\underline{P}(q,\kappa)$ with an estimated dispersion parameter $\theta(q,\kappa)$ - now a function of q and $\kappa$ - provides a good statistical fit to $\underline{P}^*(q,\kappa)$, in the base and response contexts. For the base system the fit is impressive, deteriorating only slightly for higher values of $\kappa$. The variation of $\theta(q,\kappa)$ with q and $\kappa$ is shown in Figure 5.2(b) and the base vectors $\underline{P}^*(q,\kappa)$ and $\underline{P}(q,\kappa)$ are compared in Figures 5.3(c) and (e). Under conditions of change, however, there is a systematic overprediction of modelled response for values of $\kappa$ greater than zero, as indicated in Figures 5.3(d) and (f). We can now summarise the results and implications of these tests as follows.

As all combinations of selected parameters $\{q,\kappa\}$ generate data sets $\underline{P}^*$ which are well fitted by the logit function we may say that the dispersion reflected in the estimated parameter $\theta(q,\kappa)$ is consistent with utility maximisation under both full and partial information assumptions. That is, $\underline{P}^*$ is consistent with the following interpretations:

(i) Complete information and preference dispersion from utility distributions with standard deviation given by

$$\sigma(q,\kappa) = \frac{\pi}{\sqrt{6}\,\theta(q,\kappa)} \tag{5.11}$$

(ii) Partial information characterised by the distribution parameters q and $\kappa$; in conjunction with random utility functions of standard deviation $\sigma^*$. For given $\sigma^*$, the $(q,\kappa)$ combinations which generate the same dispersion as $\sigma(q,\kappa)$ defined by equation (5.11) may be read off Figure 5.2(b).

Further, if information is uniformly heterogeneous (i.e. $\kappa=0$) no systematic response error will be involved with logit forecasts. On the other hand if the data is the outcome of a process characterised by preferential tendencies for particular choice sets to be selected, logit forecasts will involve a systematic over prediction of the response to policy stimuli. As the observer in $\Lambda$ has no knowledge whatsoever what process did generate $\underline{P}^*$ it is not possible to discriminate between the interpretations (i) and (ii), and there must remain a certain theoretical indeterminacy of the response forecasts.

Clearly the above simulations may be generalised in a number of ways, and it would be interesting to ascertain how the response error was dependent on further detailed aspects of choice set generation and decision models. This is however beyond the scope of this paper. Before leaving the theme of choice set generation however, we wish to relate the above considerations to two further models, namely: the DOGIT model introduced by Gaudry and Dagenais (1979) and a location model discussed by Kirby (1979).

The DOGIT model is a generalisation of the logit model to accommodate varying degrees of interaction between alternatives. Its formation, originally achieved by means of transformation theory, has been given a behavioural interpretation and derivation by Ben-Akiva (1977a) who considered individuals to either be captive to a particular alternative or to have the full choice set available to them. The set C is thus considered to contain the N sets associated with c(1) and the single set c(N) as shown in Figure 5.1(b). The spectrum $P(c(n)\varepsilon C)$, shown in Figure 5.1(d) consists of the two 'spikes' at n=1 and n=N.

If the probability of selecting the individual choice sets is taken to be

$$P(c\varepsilon C) = \left\{\frac{\mu_1}{1+\sum\limits_{\rho=1}^{N}\mu_\rho}, \ldots, \frac{\mu_\rho}{1+\sum\limits_{n=1}^{N}\mu_\rho}, \ldots, \frac{\mu_N}{1+\sum\limits_{\rho=1}^{N}\mu_\rho}; \frac{1}{1+\sum\limits_{\rho=1}^{N}\mu_\rho}\right\} \tag{5.12}$$

for which the odds of being captive and non-captive are

$$P(c(n=1)\varepsilon C) = \frac{\Sigma\mu\rho}{1+\Sigma\mu_\rho}\ ;\ P(c(n=N)\varepsilon C) = \frac{1}{1+\Sigma\mu_\rho} \tag{5.13}$$

the DOGIT model is readily derived from equation (5.1) and is given by

$$P_\rho = \frac{e^{\theta\overline{U}_\rho} + \mu_\rho \sum_\rho e^{\theta\overline{U}_\rho}}{(1+\mu_\rho)\sum_\rho e^{\theta\overline{U}_\rho}} \tag{5.14}$$

The parameters $\underline{\mu}$ may be taken to be functions of the attributes associated with the alternatives.

For the particular simulations performed in this section 'lumpy' distribution spectra for $P(c(n)\varepsilon C)$ - which are appropriate to the DOGIT model have not been adopted. We should not therefore be particularly surprised that the modifications to the logit function due to captivity, which inspired the formation of the DOGIT, have failed to assert themselves in numerical tests. These modifications will inevitably become more important as the number of alternatives become small (as, for example, in modal choice contexts).

We turn finally and briefly to Kirby's model (Kirby, 1979) which involves search within locationally defined choice sets. In its simplest form the model incorporates a range function $\phi(r)$ which denotes the probability of an individual selecting a house up to range r from his or her place of work. (The housing market and choice making populations are actually considered stratified and r is considered to be expressed in the terms of generalised cost). Individuals are now considered to confine their search within their selected range and to choose a zone of residence within it with a probability proportional to the number of houses of the relevant type within the zone.

It is clear that the model is a special case of the choice set generating processes defined in equation (5.1) above. For suitably defined choice sets the functions $P(c\varepsilon C)$, or $P(c\varepsilon c(n))$ and $P(c(n)\varepsilon C)$, may be related in a straightforward way to the range function. The Kirby model appears in fact to be formally equivalent to the intervening opportunities model, although the behavioural descriptions underpinning them are distinct.

## 6. HABIT, THRESHOLDS, HYSTERESIS AND TRAVEL RESPONSE

Many authors have remarked on the relevance and role of habit, learning and "triggers" in the decision process accompanying location (i.e. migration) and travel choice behaviour (see, for example, Banister

1978; Heggie, 1978; Hensher, 1975; Goodwin, 1977, 1979 and the references cited therein). In spite of the widespread recognition of the influence of habit, few empirical or theoretical results concerning this phenomenon have been forthcoming, although the papers by Blase (1979) in the former class, and those of Goodwin (1977, 1979) and Wilson (1976) are noteworthy.

The work of Goodwin is of particular interest and our intention is to offer some elaboration of the ideas presented by that author within the framework of perspectives developed in this paper. We shall adopt a random utility model of binary choice incorporating habit developed by Goodwin.

The existence of habit, or what might be considered as inertia accompanying the decision process of the individual is possibly the most insidious of behavioural aspects which represent divergencies from the traditional assumptions of a 'rational being' for its existence appears <u>directly</u> in the <u>response</u> context. In order to examine the effects and implications of habit it is appropriate to return once more to the assumption underpinning the conventional cross-sectional approach.

In Figure 6.1(a) the familiar S-shaped curve relevant to binary choice is reproduced. We can think of this in terms of a continuum of populations $S(\bar{U}_2-\bar{U}_1)$ characterised by an <u>imputed</u> utility difference distributed along the curve. For a given difference $\bar{U}_2-\bar{U}_1$ there exists a single population S with a <u>fixed</u> proportion of members associated with the alternatives, and identified by a point on the curve. Under conditions of change $(\bar{U}_2-\bar{U}_1 \rightarrow \bar{U}_2'-\bar{U}_1')$ the population $S(\bar{U}_2-\bar{U}_1)$ will simply acquire the characteristics of $S'(\bar{U}_2'-\bar{U}_1')$ observed in the base system. The <u>response</u> is determined from the cross-sectional <u>dispersion</u>.

An implication of this assumption is that the response to a particular policy or change will be exactly reversed if the stimulus is removed. The stimulus-response relation is symmetric with respect to the sign and size of the stimulus.

These features are consistent with the 'rationality' assumption attributed to an individual, prepared to <u>continually</u> monitor his present and alternative options. They must be modified in the presence of habit, although we should re-emphasise that the necessity for modification is not a refutation of 'homo-economicus' but simply a re-interpretation by the observer of the actions and behaviour of the individual. To understand the required modifications, we consider a single population S subjected to a continuous modification in characteristics of the alternatives $A_1$ and $A_2$ which are reflected in $\overline{U}_2-\overline{U}_1$.

In the conventional derivation of random utility models we identify with each individual $I_i \varepsilon S$ utility functions $U_1^i$ and $U_2^i$ and decree that alternative $A_1$ will be selected if

$$U_1^i > U_2^i \tag{6.1}$$

Let us assume that this inequality is satisfied and $A_1$ is selected. A stimulus of magnitude $\delta U_{12}^i$ in favour of the rejected option $A_2$ must be applied before a response – that is, a change of option – will be observed. The minimum size of $\delta U_{12}^i$ is $|U_1^i - U_2^i|$.

In the presence of habit, which will be interpreted as an equivalent utility value $h^i$, this minimum required stimulus is raised to

$$\delta U_{12}^i = U_1^i - U_2^i + h^i$$

and the total population response to change is given by

$$R = \sum_{I_i \varepsilon S} \eta(\delta U_{12}^i) \tag{6.2}$$

in which

$$\begin{aligned} \eta(x) &= 1 \text{ if } \quad x>0 \\ &= 0 \text{ otherwise} \end{aligned} \tag{6.3}$$

The existence of the habit term will clearly accompany those members of S who are currently associated with an alternative(s) experiencing a stimulus to the relative advantage of another option(s). This feature introduces a basic assymetry into response behaviour and gives rise to the phenomenon of hysteresis (Goodwin, 1977; Wilson, 1976). Now, the present state of the population S, identified in terms of the proportions on each alternative, is dependent on not only the utility values $\overline{U}_1$ and $\overline{U}_2$, but on how these utility variables attained their current value. Formally, the state of the system $P_\rho$ may be expressed as a path integral in the space of utility components $(\overline{U}_1, \overline{U}_2, \ldots \overline{U}_N)$. The value of this integral is path independent when habit is absent, but path dependent when it is present.

The features of these hysteresis curves have been reproduced by Monte Carlo Simulation and the results shown in Figure 6.1(c). As Goodwin (1977) has noted, when there exists a distribution of habit $h^i$ over the population, the response behaviour of S, which is dependent on the prior state of the system, can be complicated. In the figure we have identified the system 'history' by means of arrows. (Utilities were sampled from Weibull distributions and habit values from a negative exponential function with mean $\overline{H}$). Now it can be seen that for a given utility difference $\overline{U}_2 - \overline{U}_1$, a multiplicity of $P_2$ values can exist, each of which is dependent on the path of utility values by which $\overline{U}_2 - \overline{U}_1$ was achieved. Now the response to change is likewise history dependent.

It is important to remember that these curves correspond to a single population S responding to changes in the utility differences. We now return to the interpretation of the S-shaped curve in Figure 6.1(a) which is assumed to provide a good statistical fit to underlying observations on dichotomous choice. Admitting the possibility of habit, we must consider the history of cost changes which accompanies each population S $(\overline{U}_2 - \overline{U}_1)$ at the cross-section. One such history might be a continually increasing difference of $\overline{U}_2 - \overline{U}_1$ so that all populations simply move up a curve identified on the cross-section. Another history is that shown in Figure 6.1(b) which accompanies an initial rise and subsequent fall of utility differences experienced by all populations S. We cannot be sure of the response of each population identified on an observed S-curve unless we know something of its history. The traditional assumption of zero habit, and the movement of all populations S along a single curve, will in general introduce a response error - but of what size? To examine this theoretically, we introduce the perspectives $\Lambda^*$ and $\Lambda$ characterised as follows:

| | |
|---|---|
| $\Lambda^*$: Distribution of habit; $D^*$: utility maximisation from $W(o,\sigma^*)$ | $\Lambda$: No habit; D: utility maximisation from $W(o,\sigma)$ |

and look for the response error under a given 'path' of utility changes. We have assumed that a given population $S(\overline{U}_2 - \overline{U}_1 = 10)$ is identified on the right-hand curve in Figure 6.1(c) which is appropriate to a continually increasing utility difference. The stimulus is now identified with three stages: an initial increase of $\overline{U}_2 - \overline{U}_1$ of 10 units, a subsequent reduction of 40 units and a final increase of 30 units. (This kind of situation might correspond to an initial increase in the relative advantage of car travel (1:bus; 2:car) which is in line with previous cost movements, followed by a sudden large rise in car costs, accompanying say a series of petrol rises. Finally, the cost difference is eroded when public transport operators increase charges to further increase revenue).

The response error corresponding to this situation is mapped in Figure 6.1(d). The simulated response to cost changes with $\overline{H} = 15$ units are plotted against the modelled response derived from the single logit curve under the assumption of zero habit. On the first stage no error is involved as the past trend is reinforced. A significant over-estimate of the modelled response takes place on the second stage, and this decreases on the third.

While the existence of habit will influence the response of a population and indeed the interpretation of data in the context of model development, the <u>detailed</u> implications for location and mode switching in the presence of, for example, petrol price increases are unclear. The extent of the response error will, of course, be dependent on the size <u>and</u> distribution of the habit effect over the population. We must not forget also that changes of say mode or location may be triggered by a series of other stimuli (life cycle effects, etc.) and the extent to which the system is considered to be in 'disequilibrium' because of inertia will be dependent on these other stimuli.

We shall discuss the more general implications of habit for response forecasting, together with the other implications of the paper in the next section.

## 7. CONCLUSION

It is perhaps both obvious and particularly unsympathetic to reassert that the validity of a travel response forecasting model cannot be assessed with reference to the cross section alone. Indeed we are not yet able to discern the contributions to behavioural variability at the cross section from the multitude of sources: preference dispersion; the heterogeneous disposition of a wide variety of constraints; the role of information; the 'inertia' or 'disequilibrium' associated with habit effects; sub-optimal behaviour by individuals because they "... have not the wits to maximise ..." (Simon, 1955); or the measurement, aggregation and representational error on the part of an observer in the process of providing the framework within which dispersion is recorded! From a theoretical viewpoint it is not possible to discriminate between hypotheses relating to behavioural responses.

The examples addressed in this paper reaffirm the potential danger of the correlation /causation syndrome often associated with cross-sectional examinations. From the point of view of predicting response, it is not sufficient simply to know that a good fit to base data has been obtained.

While the analysis of variability at the cross section may well provide clues to the behaviour of individuals including their response to various policy stimuli, the examples provided in Section 1, involving before and after studies, do not provide encouragement. Of course, it could be argued that in any particular case discussed in the introduction, a state-of-the-art model or method of estimation was not adopted. Equally it could be proposed that we should appeal directly to the perspective A* instead of A to construct appropriate models - after all were not the Multinomial Probit, GEV model, Hierarchical Logit, DOGIT and other models inspired by the simple defects of the Multinomial Logit?

In one sense the GEV and Multinomial Probit models represent the end of a road in the range of utility functions they can address. The important question is: should the framework provided by random utility theory be generalised to consider further: habit; information contents; and elements of satisficing behaviour? Their inclusion, of course, increases the number of parameters to be determined and adds to the dilemma of how they should be estimated. While transfer pricing methods have their appeal (Daly, 1978; Bonsall, 1979) expected preference methods are not without their limitations.

> "... The most we can assert is that any individual will exhibit some tendency to act in accordance with his stated preferences..." (Brown, 1972).

It is inevitable that the now popular constraint-based studies will allow a greater degree of variation to be accounted for in observed travel behaviour. While we have much sympathy with the view that

> "... the model should not be calibrated on the same travel components that it was called to reproduce. For instance, a truly behavioural model, based on the constraints under which travel choices are made, should produce such choices independently; that would then be compared with the observed choices for its validity, and not calibrated to them ..." (Zahavi, 1978).

it is unlikely that the totality of variation can be accommodated through the incorporation of constraints alone (if this is the implication of the above comment), with no free parameters. What methods and what data should be used to assess these? What range of policies could be accommodated by their absence?

As more information on behavioural response becomes available - and it would be useful if this exercise were co-ordinated - it will be possible to increase the effectiveness of forecasting methods. The new movement case will always prove somewhat problematic for the obvious reason that attributes may not be satisfactorily matched to the existing system (Train, 1978).

In spite of, or perhaps more appropriately because of, the difficulties associated with conventional travel forecasting methods, new approaches are appearing and new methods tested which are geared towards the response context and the very detailed analysis of variability in behaviour. (See, for example, the papers by Kreibich, 1979; Bonsall, 1979; and Jones, 1979). In the HATS framework (Jones, 1979) it is attempted through loose structuring of interviews to decrease the likelihood that an observer (in this case the interviewer) will impose an inappropriate framework for behavioural analysis. At present such studies are in their infancy and must be broadened if preferences are to be accommodated. Traditional problems may then be met.

From a theoretical viewpoint it is necessary to design appropriate frameworks for the analysis of choice contexts and with the aid of response data, to allow the means for both the testing and refutation of hypotheses underpinning the models to be established. Until this is achieved, the correlation / causation syndrome, and more generally the problems of misrepresentation, will, in our opinion, continue to plague cross-sectional studies.

REFERENCES


ALBRIGHT, R.L., LERMAN, S.R., and MANSKI, C.F. (1977) "Report on the development of an estimation program for the multinomial probit model", Prepared for the Federal Highway Administration, *U.S. Department of Transportation*, by Cambridge Systematics Inc.

BANISTER, D.J. (1978) "Decision-making, habit formation and a heuristic modal split model based on these concepts", *Transportation, 7*, 5-18.

BEN-AKIVA, M.E. (1974) "Structure of passenger travel demand models", *Transportation Research Record, 526*, 26-42.

BEN-AKIVA, M.E. (1977a) "Choice models with simple choice set generating processes", Working Paper, Department of Civil Engineering, Massachusetts Institute of Technology, U.S.A.

BEN-AKIVA, M.E. (1977b) "Passenger travel demand forecasting: applications of disaggregate models and directions for research", *Proceedings*, World Conference on Transportation Research, Rotterdam.

BLASE, J.H. (1979) "Hysteresis and catastrophe theory: a demonstration of habit and threshold effects in travel behaviour", PTRC Proceedings, Summer Annual Meeting, University of Warwick.

BONSALL, P.W. (1979) "Microsimulation of mode choice: a model of organised car sharing", PTRC Proceedings, Summer Annual Meeting, University of Warwick.

BRAND, D. (1973) "Travel demand forecasting, some foundations and a review", In *Urban Travel Demand Forecasting*, Special Report 143, Highway Research Board, Washington DC.

BROWN, G.R. (1972) "Analysis of user preferences for systems characteristics to cause a modal shift", *Highway Research Record 417*, 25-36.

BURNETT, P., and HANSON, S. (1979) "A rationale for an alternative mathematical approach to movement as complex human behaviour", PTRC Proceedings, Summer Annual Meeting, University of Warwick.

BURNETT, K.P., and THRIFT, N.J. (1979) "New approaches to understanding traveller behaviour", In Hensher, D.A., and Stopher, P.R. (Eds) *Behavioural Travel Modelling*, Croom Helm, London.

COCHRANE, R.A. (1975) "A possible economic basis for the gravity model", *Journal of Transport Economics and Policy, 9*, 34-49.

DAGANZO, C.F., BOUTHELIER, F., and SHEFFI, Y. (1977) "Multinomial probit and qualitative choice: a computationally efficient algorithm", *Transportation Science, 11*, 338-58.

DALY, A.J. (1978) "Issues in the estimation of journey attribute values", In Hensher, D.A., and Dalvi, M.Q. (Eds.) *Determinants of Travel Choice*, Saxon House, Sussex.

DALY, A.J., and ZACHARY, S. (1978) "Improved multiple choice models", In Hensher, D.A., and Dalvi, M.Q. (Eds.) *Determinants of Travel Choice*, Saxon House, Sussex.

DOMENCICH, T., and McFADDEN, D. (1975) *Urban Travel Demand - a behavioural analysis*, North Holland, Amsterdam.

EILON, S. (1972) "Goals and constraints in decision-making", *Operational Research Quarterly, 23 (1)*, 3-15.

GAUDRY, M.J.I., and DAGENAIS, M.G. (1979) "The DOGIT model", *Transportation Research, 13B(2)*, 105-12.

GENSCH, D.H., and SVESTKA, J.A. (1978) "An exact hierarchical algorithm for determining aggregate statistics from individual choice data", Working Paper, University of Wisconsin at Milwaukee.

GOODWIN, P.B. (1977) "Habit and hysteresis in mode choice", *Urban Studies, 14*, 95-8.

GOODWIN, P.B. (1979) "Two problems of travellers' behaviour in mathematical models", *Bulletin of the Institute of Mathematics and its Applications* (in press).

GOULD, P., and WHITE, R. (1974) *Mental Maps*, Pelican Books, Harmonsworth, U.K.

HAUSMAN, J.A., and WISE, D.A. (1978) "A conditional probit model for qualitative choice: discrete decisions recognising interdependence and heterogeneous preferences", *Econometrica, 46*,403-26.

HAWKINS, A.F. (1978) "Some anomalies due to misspecification in transport models", PTRC Proceedings: Summer Annual Meeting, University of Warwick.

HEGGIE, I.G. (1978) "Putting behaviour into behavioural models of travel choice", *Journal of the Operational Research Society, 29*, 541-50.

HEGGIE, I.G., and JONES, P.M. (1978) "Defining domains for models of demand", *Transportation, 7*, 119-25.

HENSHER, D.A. (1975) "Perception and commuter modal choice: an hypothesis", *Urban Studies, 12*, 101-4.

HOROWITZ, J. (1979) "The accuracy of the multinomial logit model as an approximation to the multinomial probit model of travel demand", Working Paper, Department of Civil Engineering, Massachusetts Institute of Technology.

JONES, P.M. (1979) "'HATS': a technique for investigating household decisions", *Environment and Planning A, 11*, 59-70.

KENT COUNTY COUNCIL (1976) "Report on experimental park-and-ride services", *Maidstone Transportation Study*, Maidstone, U.K.

KIRBY, H.R. (1979) "Is distance no object? Location behaviour and the journey to work", Presented to the 11th Annual Conference, University Transport Study Group, University of Southampton. (unpublished).

KREIBICH, V. (1979) "Modelling car availability modal split and trip distribution by Monte Carlo simulation: a short way to integrated models", *Transportation, 8,* 153-66.

LANGDON, M. (1976) "Modal split models for more than two modes", In *Urban Traffic Models,* PTRC Summer Meeting, University of Warwick.

MAKOWSKY, G.G., BAKR, M.M., and SINHA, K.C. (1977) "Estimating transportation mode utilities", *Transportation Engineering Journal, of ASCE, 103* (TE6), 807-23.

MANSKI, C.F. (1977) "The structure of random utility models", *Theory and Decisions, 8,*229-54.

MANSKI, C.F., and LERMAN, S.R. (1978) "On the use of simulated frequencies to approximate choice probabilities", Cambridge Systematics Inc.

McFADDEN, D. (1973) "Conditional logit analysis of qualitative choice behaviour", In Zarembka, P. (Ed.) *Frontiers in Econometrics,* Academic Press, New York.

McFADDEN, D. (1975) "Economic applications of psychological choice models", Working Paper 7519, Travel Demand Forecasting Project, Institute of Transportation and Traffic Engineering, University of California at Berkeley.

McFADDEN, D. (1979) "Quantitative methods for analysing travel behaviour of individuals: some recent developments", In Hensher, D.A., and Stopher, P.R. (eds.) *Behavioural Travel Modelling,* Croom Helm, London.

MURCHLAND, J.D. (1976) "Decision making when we cannot reduce everything to money terms", PTRC Proceedings, Summer Annual Meeting, University of Warwick.

ORTUZAR, J.D. (1978) "Mixed mode demand forecasting techniques: an assessment of current practice", PTRC Proceedings, Summer Annual Meeting, University of Warwick.

ORTUZAR, J.D. (1979) "Testing the theoretical accuracy of travel choice models using Monte Carlo simulation", Working Paper (forthcoming), Institute for Transport Studies, University of Leeds.

PAPAOULIAS, D.B., and HEGGIE, I.G. (1976) "A comparative evaluation of forecast and use of Park-and-Ride in Oxford", *Traffic Engineering and Control, 17(4),* 144-9.

RICHARDSON, A.J. (1978) "A comparative analysis of transport choice models", Working Paper 78/11, Civil Engineering, Monash University, Australia.

ROBERTSON, D.I. (1977) "A deterministic method of assigning traffic to multiple routes of known cost", Department of the Environment, Department of Transport, TRRL Report 757, Crowthorne.

ROBERTSON, D.I., and KENNEDY, J.V. (1979) "The choice of route mode, origin and destination by calculation and simulation", Department of the Environment, Department of Transport, TRRL Report 877, Crowthorne.

SIMON, H.A. (1955) "A behavioural model of rational choice", *Quarterly Journal of Economics, 69,* 99-118.

SPEAR, B.D. (1977) "Applications of new travel demand forecasting techniques to transportation planning: a study of individual choice models", Federal Highway Administration, *U.S. Department of Transport,* Office of Highway Planning, Washington, D.C.

STOPHER, P.R., and MEYBURG, A.H. (Eds.) (1976) *Behavioural Travel Demand Models,* Lexington Books, D.C. Heath & Co., Lexington, U.S.A.

THRIFT, N.J., and WILLIAMS, H.C.W.L. (1979) "On the development of behavioural location models", *Economic Geography* (in press).

TRAIN, K. (1978) "A validation test of a disaggregate mode choice model", *Transportation Research, 12 (3),* 167-74.

TVERSKY, A. (1972) "Elimination by aspects: a theory of choice". *Psychological Review, 79(4),* 281-99.

VINCENT, R.A., and LAYFIELD, R.E. (1977) "Nottingham zones and collar study-overall assessment", Department of the Environment, Department of Transport, TRRL Report LR805, Crowthorne.

WEIBULL, J.W. (1978) "A search model for microeconomic analysis with spatial applications", In Kalqvist, A., Lundqvist, L., Snickars, F., and Weibull, J.W. (Eds.) *Spatial Interaction Theory and Planning Models,* North Holland, Amsterdam.

WILLIAMS, H.C.W.L. (1977) "On the formation of travel demand models and economic evaluation measures of user benefit", *Environment and Planning A, 9,* 285-344.

WILLIAMS, H.C.W.L. (1979) "Travel demand forecasting - an overview of theoretical developments", Presented at Research Seminar on "Transport and Public Policy Planning", University of Reading.

WILLIAMS, H.C.W.L., and SENIOR, M.L. (1977) "Model based transport policy assessment, 2: Removing fundamental inconsistencies from the models", *Traffic Engineering and Control, 18,* 464-9.

WILSON, A.G. (1969) "Entropy maximising model in the theory of trip distribution, mode split and route split", *Journal of Transport Economics and Policy, 3,* 108-26.

WILSON, A.G. (1976) "Catastrophe theory and urban modelling: an application to modal choice", *Environment and Planning A, 8,* 351-6.

YOUNG, W., and RICHARDSON, A.J. (1978) "Residential preference: a perceptual view of Melbourne", *Proceedings, 4th Australian Research Forum.*

ZAHAVI, Y. (1978) "Can transport policy decisions change urban structure?", PTRC Proceedings, Summer Annual Meeting, University of Warwick.

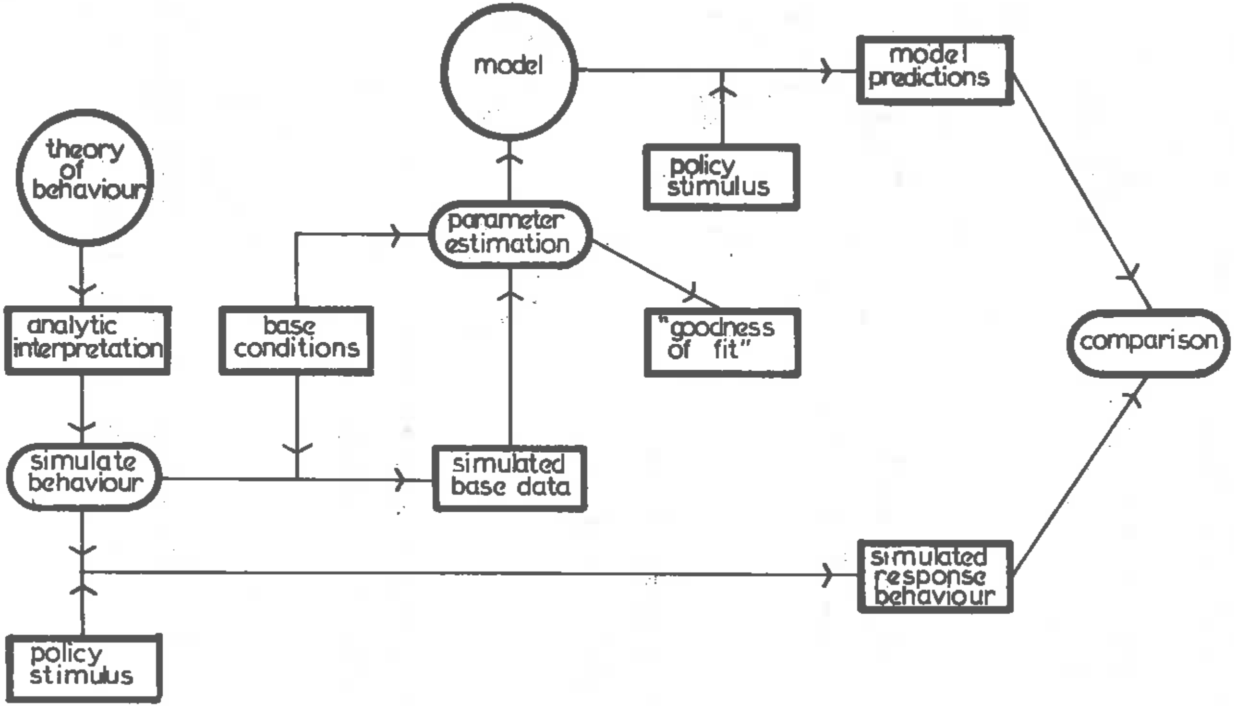



FIGURE 2.1: Testing models of behaviour using simulation.

| STRUCTURE OF THE VARIANCE - COVARIANCE MATRIX | PICTORIAL REPRESENTATION |
|---|---|

a.) Uncorrelated Structure (X-Y)

$$\Sigma_{\mu\nu,\mu'\nu'} = \sigma^2_{XY} \cdot \delta_{\mu\mu'} \cdot \delta_{\nu\nu'}$$

b.) Nested Structure (X/Y)

$$\Sigma_{\mu\nu,\mu'\nu'} = \sigma^2_{X}\, \delta_{\mu\mu'} + \sigma^2_{XY} \cdot \delta_{\mu\mu'} \cdot \delta_{\nu\nu'}$$

c.) Nested Structure (Y/X)

$$\Sigma_{\mu\nu,\mu'\nu'} = \sigma^2_{Y}\, \delta_{\nu\nu'} + \sigma^2_{XY} \cdot \delta_{\mu\mu'} \cdot \delta_{\nu\nu'}$$

d.) Cross- Correlated Structure (X+Y)

$$\Sigma_{\mu\nu,\mu'\nu'} = \sigma^2_{X}\, \delta_{\mu\mu'} + \sigma^2_{Y}\, \delta_{\nu\nu'} + \sigma^2_{XY} \cdot \delta_{\mu\mu'} \cdot \delta_{\nu\nu'}$$

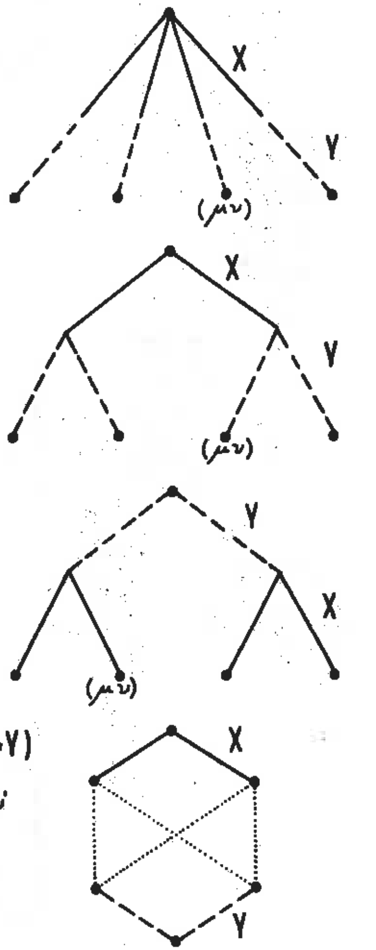
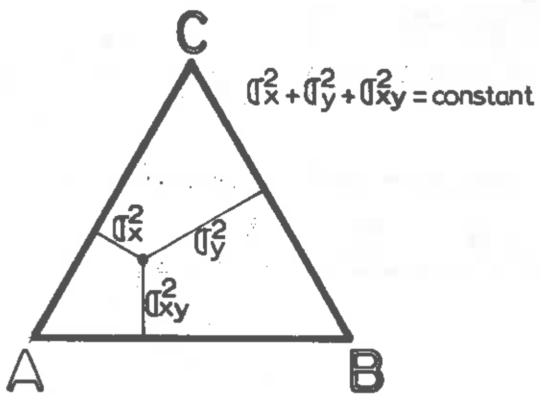
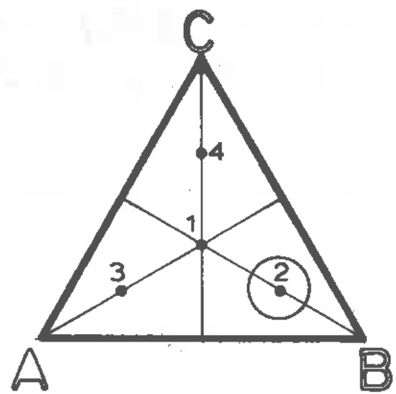



FIGURE 3.1: Representation of the structure of choice models.



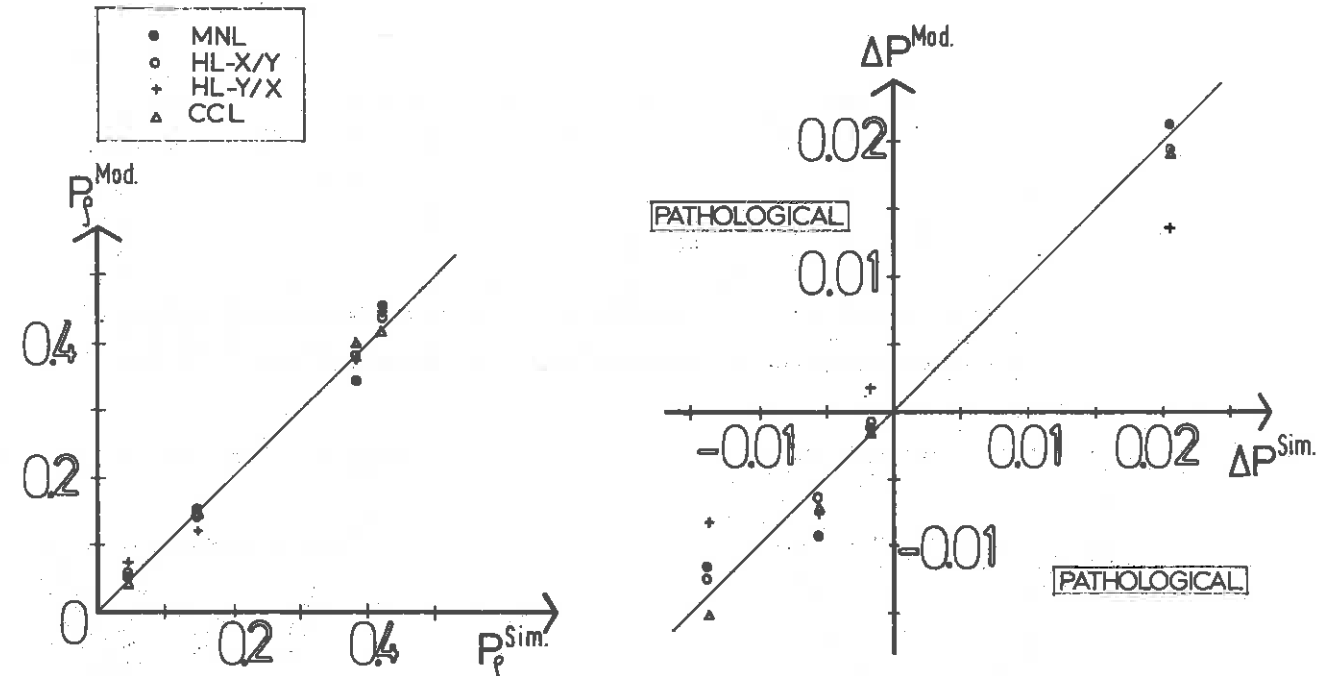

a.) The normalization condition.




b.) The test point coordinates $(\sigma_x, \sigma_y, \sigma_{xy})$

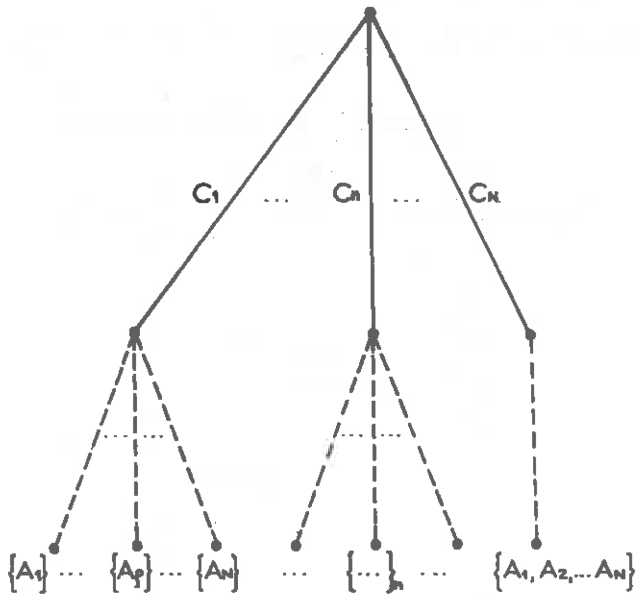
| POINT | MNL | HL-X/Y | HL-Y/X | CCL |
|---|---|---|---|---|
| 1 | good | good $\Delta \sim \beta$ | good $\Delta \sim \beta$ | very good |
| 2 | regular | very good | pathological $\Delta < \beta$ | good |
| 3 | regular | pathological $\Delta < \beta$ | very good | good |
| 4 | very good | very good $\Delta = \beta$ | very good $\Delta = \beta$ | very good |

c.) Model performance at different test points.

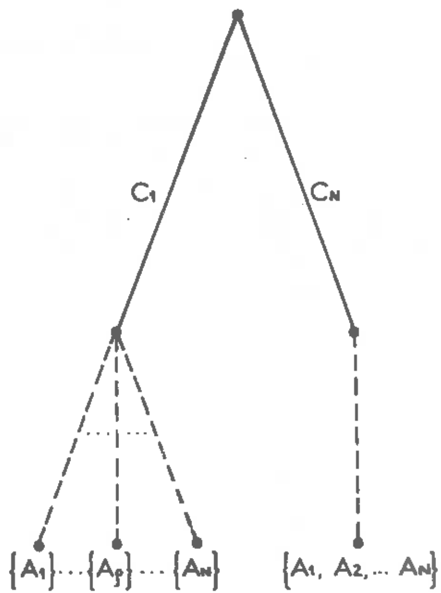



d.) An illustration of base performance and response errors at test point 2.

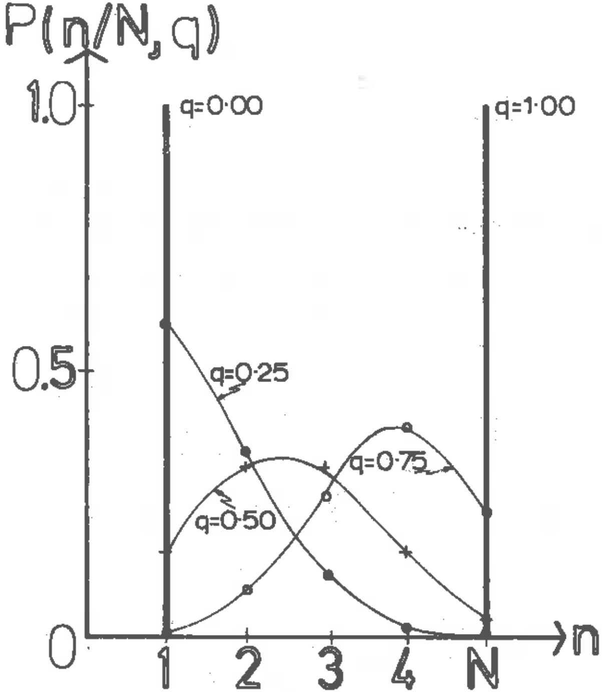
FIGURE 3.2 : The design and results of simulation tests to investigate model structure variation.



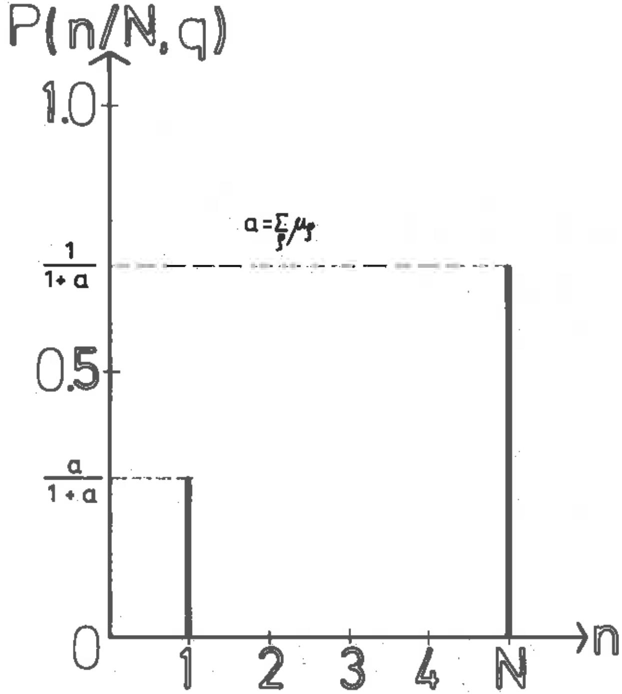

a.) Choice set possibilities in the general case.




b.) Choice set possibilities appropiate to the Dogit model.

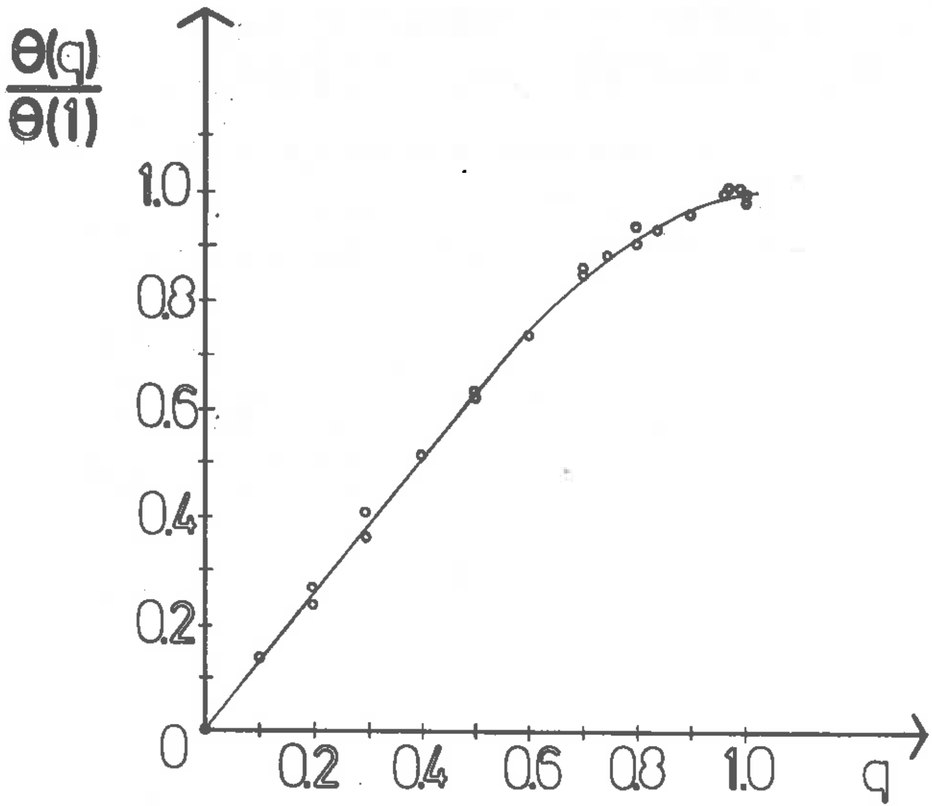



c.) Binomial choice set generation.

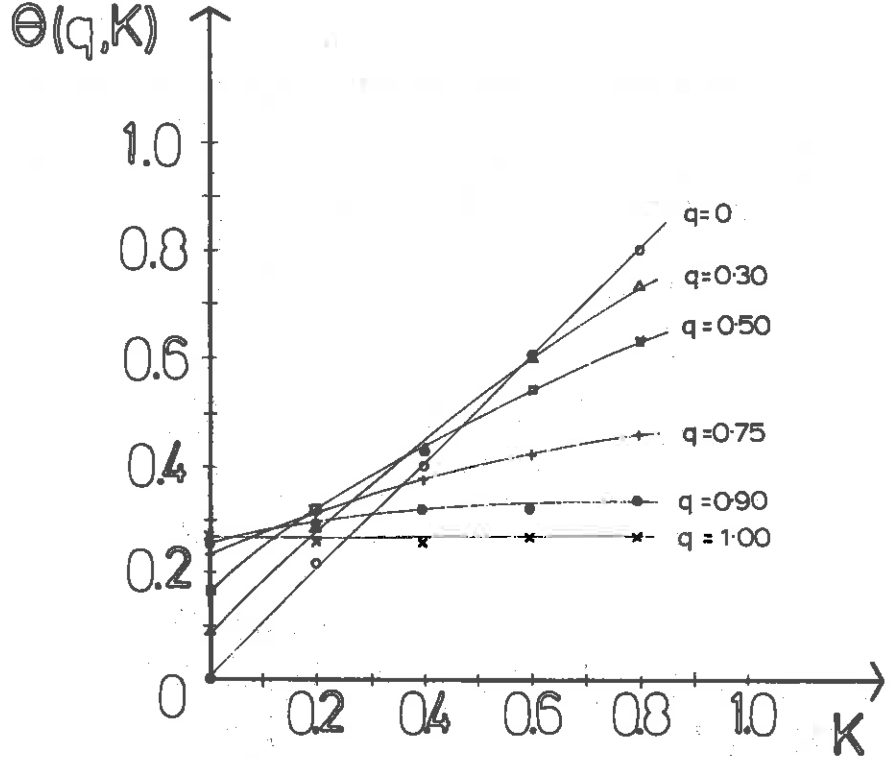



d.) Dogit choice set generation.

**FIGURE 5.1: Choice set generation partitioned by size of choice sets.** (see text for notation)




a.) Variation of $\theta(q)$ with the binomial parameter q.




b.) Variation of θ(q,K) with K for different values of q.

**FIGURE 5.2: Variation of the dispersion parameter θ with the choice set generation parameters.**

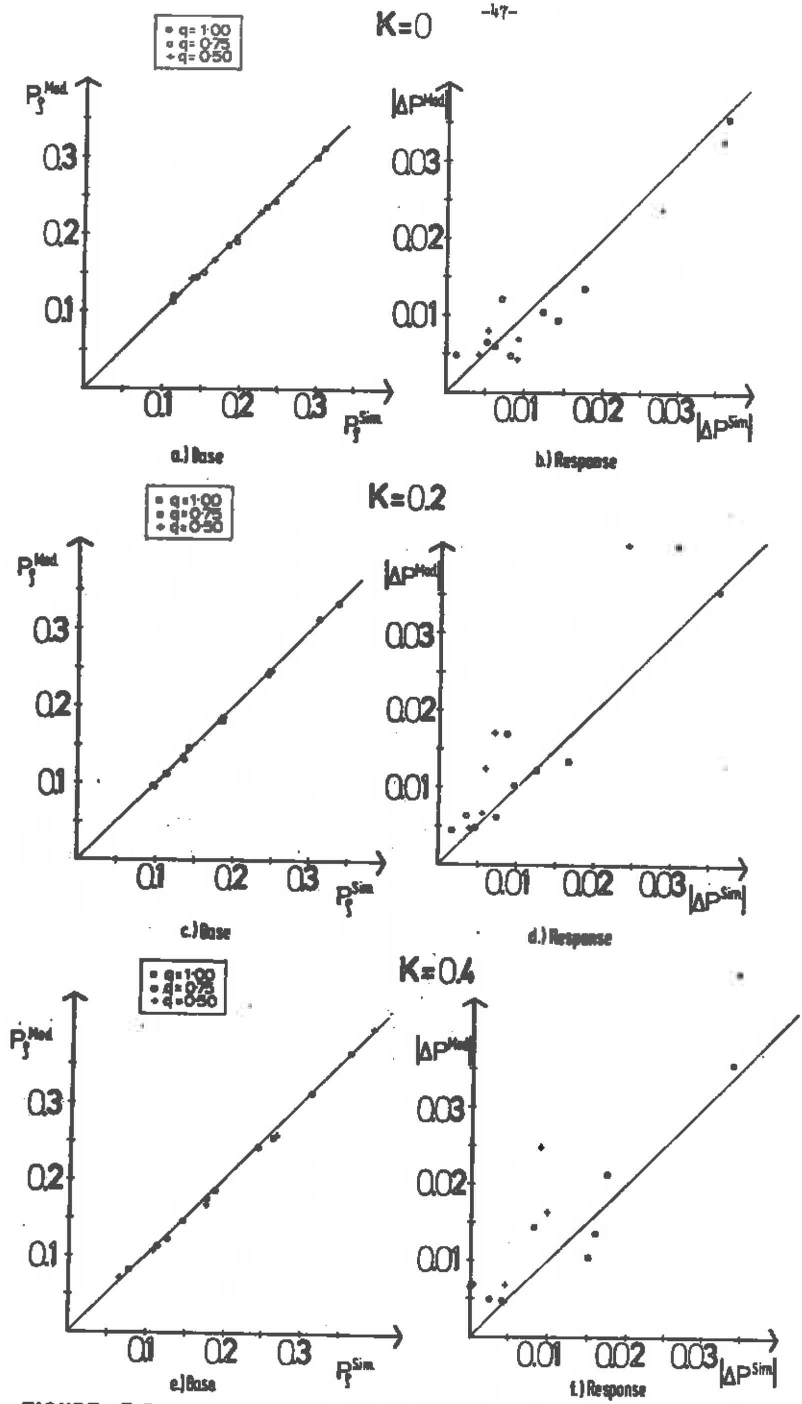
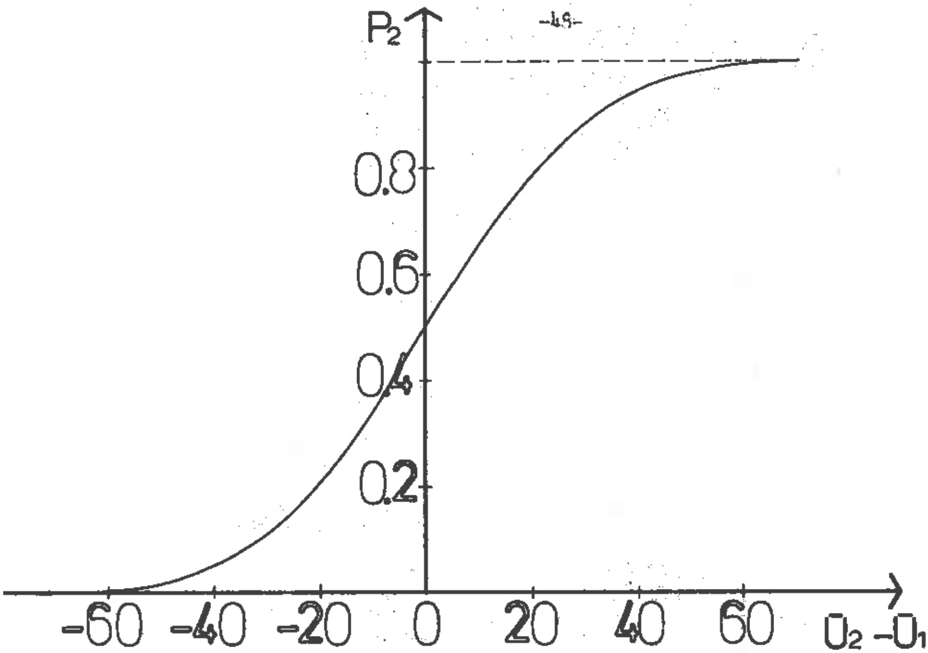
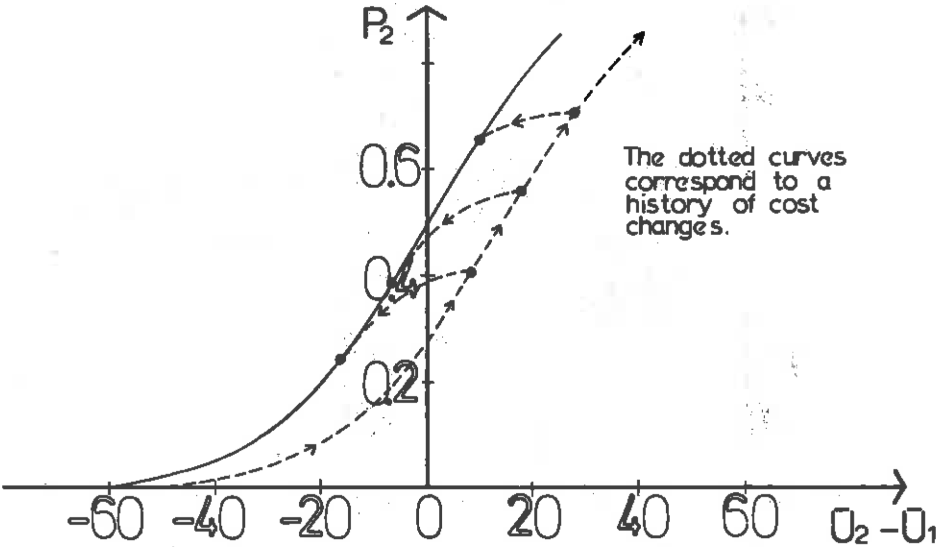



**FIGURE 5.3:** A comparison of simulated and modelled results in the base and response contexts for various combinations of K and q.

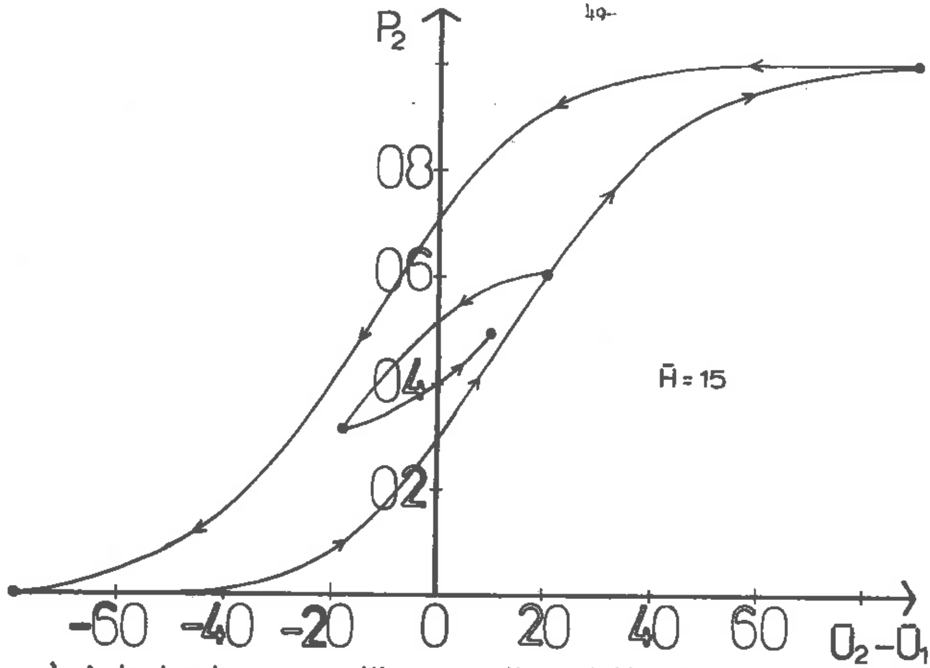



a.) The logit model fitted to cross-sectional binary data.

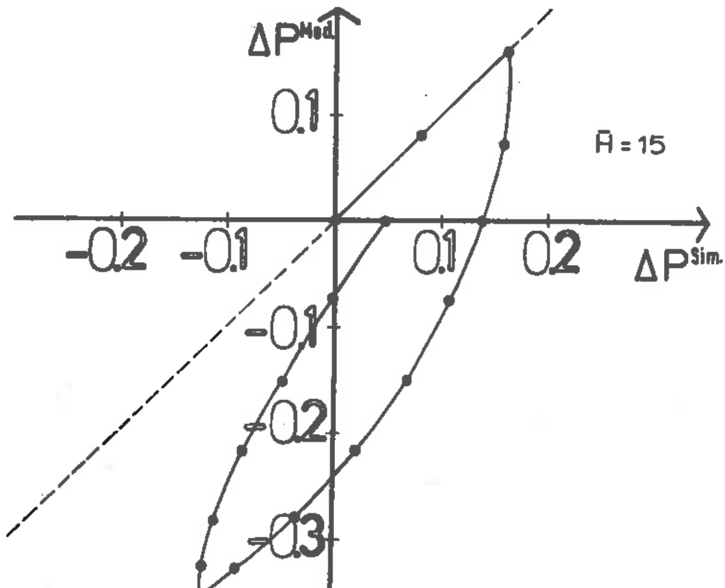



b.) The evolution of cross-sectional data points.

FIGURE 6.1: The influence of habit on cross-sectional models.




c.) A hysteresis curve with non-uniform habit effect.




d.) Response error in the presence of habit.

the interpretation and response of