

STRUCTURE OF PASSENGER TRAVEL DEMAND MODELS

by

MOSHE EMANUAL BEN-AKIVA

**B.Sc., Technion - Israel Institute of Technology
(1968)**

**S.M., Massachusetts Institute of Technology
(1971)**

**Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy**

at the

Massachusetts Institute of Technology

June, 1973

Signature of Author . . .

Department of Civil Engineering, May 11, 1973

Certified by Thesis Supervisor

**Accepted by Chairman, Departmental Committee on Graduate Students
of the Department of Civil Engineering**

JUN 29 1973

ABSTRACT

STRUCTURE OF PASSENGER TRAVEL DEMAND MODELS

by

MOSHE EMANUEL BEN-AKIVA

Submitted to the Department of Civil Engineering on May 11, 1973 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

This study is concerned with the structure of travel demand models. Two alternative structures are defined: simultaneous and recursive, each based on a different hypothesis about the underlying travel decision making process. The simultaneous structure is very general and does not require any specific assumptions. The recursive structure represents a specific conditional decision structure, i.e., the traveller is assumed to decompose his trip decision into several stages. Thus, it is argued that the simultaneous and the recursive structures represent, respectively, simultaneous and sequential decision making processes.

Theoretical reasoning indicates that the simultaneous structure is more sensible. Moreover, if a sequence assumption is accepted, there are several conceivable sequences, and generally there are no a priori reasons to justify a selection among them. A simultaneous model, however, is very complex due to the large number of alternatives that a traveller is facing in making his trip decision.

An empirical study is conducted to investigate the feasibility of a simultaneous model and to appraise the sensitivity of predictions made by a travel demand model to the structure of the model. The data set for the empirical study was drawn from a conventional urban transportation study data.

Two choices included in a trip decision are considered: destination choice and mode choice. With the same data set, three disaggregate probabilistic models are estimated for the shopping trip purpose: one model with a simultaneous structure, and two recursive models with the two possible sequences. The simultaneous model proved to be feasible in terms of the computational costs and the estimation results. The results of the recursive models showed that estimated model coefficients vary considerably with the different model structures. Thus, it is recommended that the simultaneous model structure should be used.

Thesis Supervisor:

Marvin L. Manheim

Title:

Professor of Civil Engineering

ACKNOWLEDGEMENTS

During the course of this research I have benefitted from the assistance and contributions of many individuals. Professor Marvin Manheim, my thesis supervisor, has contributed the most to the development of this study. His constant encouragement and keen insights were invaluable. Professors Richard de Neufville and Wayne Pecknold, the other members of my doctoral committee, provided many valuable ideas and suggestions. The assistance and advice I received from Robert McGillivray, from the Urban Institute, also contributed greatly to the development and the presentation of the ideas in this research.

I would also like to thank my colleagues and friends at M.I.T., in particular, Earl Ruiter, Frank Koppelman, and Len Sherman, for their constant help along the way. Dan Brand, Gerald Kraft and Thomas Domencich, with whom I have discussed this research, also contributed many ideas through their previous work in the subject area of this research.

Carol Walb and Marianne Koppelman contributed greatly in preparing the final document.

Metropolitan Washington Council of Governments - Department of Transportation Planning, and R.H. Pratt Associates are acknowledged for supplying the data for the empirical analysis. Charles Manski was most helpful in providing the logit estimation program that was used in this study.

This research was supported in part by a grant to M.I.T. from the Ministry of Transport of the State of Israel for transportation planning

research. Early stages of the research were supported in part by a Sloan Research Traineeship from M.I.T.

Finally, I would like to express my gratitude to my wife, Hagit and our son, Ori, whose contributions, of a different nature, were at least as important.

TABLE OF CONTENTS

	<u>Page</u>
TITLE PAGE	1
ABSTRACT	2
ACKNOWLEDGEMENTS	3
TABLE OF CONTENTS	5
<u>CHAPTER I: Introduction and Summary</u>	8
Purpose of the Study	8
Models for Policy Analysis	10
Disaggregate Models	13
Choice Theory	14
The Multinomial Logit Model	16
The Travel Choices	16
The Alternative Structures	17
Alternative Models	20
The Empirical Study	21
Conclusions	27
Outline of the Report	29
<u>CHAPTER II: Nature of Travel Demand and Prediction Models</u>	31
<u>for Transportation Planning</u>	
Introduction	31
The Complexity of Travel Demand	32
The Demand for Travel as a Derived Demand	36
Interaction Between Location and Transport	38
The Importance of the Work Trip	41
Short Run vs. Long Run, Static vs. Dynamic	42
The Overall Structure of Prediction in Transportation Planning	43
Alternative Structures	47
The Structure of the Urban Transportation Model System	52
Modelling Considerations	56
Disaggregate Models	61
A Disaggregate Modelling Framework	67
A Hierarchy of Choices	69
The Travel Demand Function	77
The Behavioral Unit	79
The Alternatives	80
Supply Effects with Disaggregate Demand Models	85

	<u>Page</u>
CHAPTER III: Consumer and Choice Theories	87
Overview	87
Consumer Theory	87
Probabilistic Choice Theory	97
Constant Utility Models	101
The Set of Alternatives	105
Random Utility Models	108
Summary	115
CHAPTER IV: Multi-Dimensional Choice Models: Alternative . . .	117
Structures of Travel Demand Models	
Introduction	117
Dependencies Among Choices	118
The Overall Set of Alternatives	121
Alternative Structures	122
Separability of Choices	125
The Identification Problem and Estimation of Simultaneous Probability Structures	127
Modelling the Travel Choices	131
Alternative Structures of Travel Demand Models	133
The Aggregate Equivalent of Disaggregate Models	143
Direct and Indirect Travel Demand Models	149
The Empirical Problem	155
CHAPTER V: Review of the Structures of Current Approaches . . .	158
to Travel Demand Modelling	
Introduction	158
A Typology of Travel Demand Models	158
Aggregate Recursive Models	160
Aggregate Simultaneous Models	165
Disaggregate Recursive Models	166
Disaggregate Simultaneous Models	168
Summary	169
CHAPTER VI: The Logit Model and Functional Specification of . . .	170
Alternative Models	
Introduction	170
The Multinomial Logit Model	170
The Specification of the Variables	173
Derivations of the Logit Model	177
The Choice Axiom and the Independence from Irrelevant Alternatives Property	178
Elasticities of the Logit Model	183
Aggregate Elasticity	185

	<u>Page</u>
Estimation Technique	187
Application of the Logit Model to a Simultaneous Travel Choice Model	191
Application of the Logit Model to a Recursive Travel Choice Model	204
Previous Applications of a Disaggregate Logit Model to Travel Demand	215
Summary	215
CHAPTER VII: Estimation of Alternative Models	217
Introduction	217
The Data Set	219
The Subsample Used for Estimation	221
Variables Used in the Models	226
Specification of the Variables	228
The Simultaneous Model	231
Alternative Recursive Models	232
Sequence $d \rightarrow m$: The Conditional Probability	233
Sequence $d \rightarrow m$: The Marginal Probability	234
Sequence $m \rightarrow d$: The Conditional Probability	238
Sequence $m \rightarrow d$: The Marginal Probability	239
Sequential Estimation of the Simultaneous Model	242
Comparison of Alternative Models	244
Summary	251
CHAPTER VIII: Conclusions and Recommendations	252
Summary	252
General Conclusions	253
Theoretical Framework	253
Modelling a Simultaneous Structure	254
Choice Models	255
The Choice Set	255
The Aggregation Problem	256
Non-Home-Based Trips	256
Case Studies	256
Data	257
Extensions of this Research	257
BIBLIOGRAPHY	260
BIOGRAPHICAL SUMMARY	267
APPENDIX: List of Figures	268

CHAPTER I

Introduction and SummaryPurpose of the Study

Decision making in transportation planning, as in any other planning activity, requires the prediction of impacts from proposed policies. One of the inputs to the prediction process is the demand function which describes consumers' expected usage of transport services.

The most widely used approach to the prediction of passenger travel demand is the aggregate Urban Transportation Model System (UTMS).* It is characterized by a recursive, or sequential, structure** which represents a conditional decision making process, i.e., the traveller is viewed as decomposing his trip decision into several stages. A trip decision consists of several travel choices, e.g., choice of mode, choice of destination, etc. In a recursive structure the travel choices are determined one at a time, in sequence.

Two recent developments in modelling travel demand have stimulated the present study. The first was the recognition that the representation of the trip decision as a sequential process is not completely realistic. It has been argued that the trip decision should be modelled simultaneously

*The UTMS is described in a large number of references. See, for example, FHWA (1970), Martin et al (1961), and Manheim (1972).

**A model can be expressed mathematically in many different ways. The term "structure" refers to the format of writing a model that has a behavioral interpretation. A model can be used for forecasting in a format which has no behavioral interpretation. The distinction between direct and indirect travel demand model (Manheim, 1972) is made with respect to the format used for forecasting and does not necessarily imply a different behavioral interpretation.

without resorting to an artificial decomposition into sequential stages (Kraft and Wohl, 1967). Attempts to develop simultaneous models were made using the conventional approach of aggregate demand analysis, where the quantity demanded is taken as a continuous variable (e.g., Kraft, 1963; Quandt and Baumol, 1966; Domencich et al, 1968; Plourde, 1968). The second development was the introduction of disaggregate probabilistic demand models that relied on a more realistic theory of choice among qualitative trip alternatives. However, all the disaggregate models that were developed could be used either for a single stage of the UTMS (e.g., Reichman and Stopher, 1971), or, more recently, for all the stages, but again assuming a recursive structure (CRA, 1972).

The common denominator of these two developments is clearly a disaggregate probabilistic simultaneous travel choice model.

However, due to the large number of alternative trips that a traveller is facing, and the large number of attributes that describe an alternative trip, a simultaneous model can become very complex. This raises some important issues concerned with the feasibility of a simultaneous model, and the sensitivity of travel predictions to the simplifying assumption of a recursive structure.

The purpose of this research is to investigate these issues and to recommend a strategy for structuring travel demand models. Specifically, this study first explores the alternative travel demand model structures and their inherent behavioral assumptions. Secondly, an empirical study is conducted to estimate the alternative models and furnish some

evidence with respect to the feasibility and desirability of disaggregate simultaneous travel choice models.

Models for Policy Analysis

We first explain what models are, what their purpose is, and why their behavioral assumptions are important.

In general, models are simplified representations of some objects or phenomena. In this study, we deal with econometric models, i.e., mathematical relationships describing economic phenomena of observed variables and unknown but statistically estimable parameters. We use models to better understand real world phenomena so as to be able to make decisions based on this understanding.

Travel demand models are used to aid in the evaluation of alternative policies. The purpose of the models is to predict the consequences of alternative policies or plans. A model that determines travel consequences independently of the characteristics of various policy options can obviously not be used to evaluate those options (unless policies are, in fact, irrelevant to consequences).

The specification of a travel demand model necessarily embodies some assumptions about the relationships among the variables underlying travel behavior. Predictions made by the model are conditional on the correctness of the behavioral assumptions and, therefore, are no more valid than the behavioral assumptions on which the model is based.

A model can duplicate the data perfectly, but may serve no useful purpose for prediction if it represents erroneous behavioral assumptions.

For example, consider a policy that will drastically change present conditions. In this case the future may not resemble the present, and simple extrapolation from present data can result in significant errors. However, if the behavioral assumptions of the model are well captured, the model is then valid under radically different conditions. It should be noted that this discussion is very general. "Behavioral assumptions" are a matter of degree since there could be many levels of detail in which behavior could be described. (For example, sensitivity to policies could be regarded as a gross level of behavioral assumptions.)

The requirement that models should be policy sensitive is necessary but not sufficient for planning purposes. The additional requirement is that the models should be based on valid behavioral assumptions. A model could be policy sensitive but be useless for policy analysis, even if it fits the data perfectly, if it is not based on valid assumptions.

In general, it is impossible to determine the correct specification of a model from data analysis. It should be determined from theory or a priori knowledge which are based on experience with, and understanding of, the phenomenon to be modelled. The problem is that frequently we do not have a comprehensive theory that will prescribe a specific model. Moreover, important variables are often missing due to lack of data or to measurement problems. There are other potential problems which involve the different kinds of data that could be used to estimate the model (e.g., time series vs. cross-section, attitudinal vs.

engineering, etc.), and the need to use a mathematical form which is amenable to a feasible statistical estimation technique.

The result is that we may have several alternative models to evaluate. Unfortunately, "in statistical inference proper, the model is never questioned... The methods of mathematical statistics do not provide us with a means of specifying the model" (Malinvaud, 1966). In other words, given several alternative models and a data set, statistical inference will not be conclusive as to which model represents the "true" process. This does not say, however, that the data does not play a role in the selection among models. At various stages of an empirical analysis we may revise some aspects of our assumptions that do not agree sufficiently with the findings. More generally, the accumulated past evidence from empirical studies influences the formulation of the assumptions of new efforts.

Suppose that we are faced with a choice among some alternative models that were not discarded in the course of analysis of the data. If these alternative models were based on different sets of assumptions, we should decide which set makes the most sense according to our a priori knowledge about behavior, taken together with "goodness of fit" measures and statistical significance tests.

In modelling passenger travel demand, we are concerned with the trip-making behavior of individuals or households. Hence, a prerequisite to travel demand modelling is a set of assumptions that describe the process of trip-making decisions of these individuals or households.

The basis for comparing different travel demand models should be the reasonableness (or, the correspondence with a priori knowledge) of the behavioral assumptions of each model.

In this study we consider two alternative structures of travel demand models: simultaneous and recursive, each representing a different travel behavior assumption. We assume a priori that a simultaneous structure is appropriate. However, we also consider models with recursive structures, in order to evaluate their significant differences from a simultaneous model.

Disaggregate Models

The behavioral assumptions of a demand model always take the point of view of an individual consumer as he weighs the alternatives and makes a choice. An aggregate model could be constructed based on aggregates of consumers by location or socio-economic category. However, aggregation during the model construction phase of the analysis will only cloud the actual relationships and can cause a significant loss of information (Fleet and Robertson, 1968; McCarthy, 1969). An aggregate model that is constructed based on averages of observations of socio-economic types and geographic location, would not necessarily represent an individual consumer's behavior, and there is no reason to expect that the same relationships would hold in another instance or another location. For planning purposes, we are concerned with the prediction of the behavior of aggregates of people. However, in principle, aggregation to any desired level that is required for forecasting

could always be performed after estimation.

In Urban Transportation Planning (UTP) studies the data are collected on the disaggregate level, and aggregated to a zonal level for use in the conventional UTMS (Martin et al, 1961). Using this disaggregate data directly in disaggregate travel demand models can bring about large savings in data collection and processing costs. Since the data are used in the original disaggregate form, and are not aggregated to the zonal level, a comprehensive home interview survey is not essential as it is for the conventional aggregate models. The experience from previous work with disaggregate travel demand models (e.g., Reichman and Stopher, 1971; CRA, 1972) indicates that it is a feasible modelling approach. Thus, disaggregate travel demand models have several practical advantages over aggregate models: the possible reductions in data collection costs, the transferability of the models from one area to another, and the possibility of using the same set of models for various levels of planning. The problem of aggregating a disaggregate model for forecasting requires more research. However, some simplified methods, such as the use of homogenous market segments (Manheim, 1972; Aldana, 1971) are available and could be used.

Choice Theory

In general, models that describe consumer behavior are based on the principle of utility maximization subject to resource constraints. Conventional consumer theory, however, is not suitable to derive models that describe a probabilistic choice from a qualitative, or discrete,

set of alternatives. Therefore, the travel demand models developed in this study rely on probabilistic choice theories.*

The consumer is visualized as selecting the alternative that maximizes his utility. The probabilistic behavior mechanism is a result of the assumption that the utilities of the alternatives are not certain, but rather random variables determined by a specific distribution.

Denote the utility of alternative i to consumer t as U_{it} . The choice probability of alternative i is therefore:

$$P(i:A_t) = \text{Prob}[U_{it} \geq U_{jt}, \forall j \in A_t]$$

where A_t is the set of alternative choices available to consumer t . The utilities are essentially indirect utility functions which are defined in theory as the maximum level of utility for given prices and income. In other words, the utility U_{it} is a function of the variables that characterize alternative i , denoted as X_i , and of the socio-economic variables describing consumer t , denoted as S_t . Thus, we can write:

$$U_{it} = U_i(X_i, S_t)$$

The set of alternatives A_t is mutually exclusive and exhaustive such that one and only one alternative is chosen. The deterministic equivalent of this theory is simply a comparison of all alternatives available and the selection of the alternative with the highest utility.

*Choice theories are reviewed in Chapter III. See also Luce and Suppes (1965), CRA (1972), and Brand (1972).

The mathematical form of the choice model is determined from the assumption about the distribution of the utility values. The coefficients of the utility functions are estimated with a cross section of consumers using observations of actual choices. Therefore, the observed dependent variable has a value of zero or one. The forecast of the model is a set of probabilities for the set of alternatives.

The Multinomial Logit Model

There are a number of probabilistic choice models that are available: two of the most popular and most useful are the Probit and Logit models. The multinomial Logit model, as described below, appears to be superior to Probit due to the practical considerations of computational time requirements.

We write the Logit model as follows.*

$$P(i:A_t) = \frac{e^{U_i(X_i, S_t)}}{\sum_{j \in A_t} e^{U_j(X_j, S_t)}}$$

With disaggregate cross-sectional data, the logit model is estimated using the maximum likelihood method (McFadden, 1968).

The Travel Choices

A trip decision for a given trip purpose consists of several choices: choice of trip frequency (e.g., how often to go shopping), choice of destination (e.g., where to shop), choice of time of day

*The derivation of the model from a distribution assumption is presented in Chapter III. The Logit model and its properties are described in detail in Chapter VI.

(e.g., when to go), choice of mode of travel, and choice of route. In a probabilistic choice approach we are interested in predicting the following joint probability:

$$P(f, d, h, m, r : \text{FDHMR}_t)$$

which is defined as the probability that individual or household t will make a trip with frequency f , to destination d , during time of day h , using mode m , and via route r . The set of alternatives FDHMR_t consists of all possible combinations of frequencies, destinations, times of day, modes, and routes, available to individual t .

Consider for the purpose of presentation only two travel choices: destination and mode. Denote the set of all alternative combinations of destinations and modes as DM . (For simplicity we drop the subscript t .) We can partition this set according to destination to get the sets of alternative modes to a given destination M_d . If modes and destinations had no common attributes and the two choices were independent then M_d is independent of d and could be written as M . However, this is an unrealistic assumption since there are many attributes, such as travel time, that are in fact characterized by all the travel choices. Therefore, it is assumed that $M_d \neq M_{d'}$. We are interested here in predicting the joint probability: $P(d, m : DM)$.

The Alternative Structures

If we assume that the two choices are independent, we write the

following independent structure*:

$$P(d:D) = \text{Prob}[U_d \geq U_{d'}, \forall d' \in D]$$

$$P(m:M) = \text{Prob}[U_m \geq U_{m'}, \forall m' \in M]$$

and

$$P(d,m:DM) = P(d:D) \cdot P(m:M)$$

where:

D = the set of alternative destinations

M = the set of alternative modes

U_d = the utility from destination d

U_m = the utility from mode m

Consider a conditional decision making process in which, for example, destination is chosen first, and then conditional on the choice of destination a mode is chosen. For this assumption we write the following recursive structure:

$$P(d:D) = \text{Prob}[U_d \geq U_{d'}, \forall d' \in D]$$

$$P(m:M_d) = \text{Prob}[U_{m|d} \geq U_{m'|d}, \forall m' \in M_d]$$

and

$$P(d,m:DM) = P(d:D) \cdot P(m:M_d)$$

where:

M_d = the set of alternative modes to destination d

$U_{m|d}$ = the utility from mode m given that destination d is chosen.

*This is an unrealistic structure for travel demand, but it is presented for the purpose of comparison with other structures.

If we assume that the choice of mode is dependent on the choice of destination and vice versa, we write the following simultaneous structure:

$$P(d:D_m) = \text{Prob}[U_{d|m} \geq U_{d'|m}, \forall d' \in D_m]$$

$$P(m:M_d) = \text{Prob}[U_{m|d} \geq U_{m'|d}, \forall m' \in M_d]$$

where:

D_m = the set of alternative destinations by mode m

In the independent and recursive structures we predict the joint probability by multiplying the structural probabilities. However, in a simultaneous structure, the two conditional probabilities are insufficient information to predict the joint probability. Therefore, we need to estimate either a marginal probability, say $P(d:D)$, or estimate directly the joint probability. The problem with the first approach is that we need to define a utility U_d where we originally specified $U_{d|m}$. The second approach requires a specification of the joint utility U_{dm} , where we consider the combination dm as a single alternative. This approach is more logical since it corresponds with the notion of a simultaneous choice. Hence, in the simultaneous structure, we need to estimate the following choice probability:

$$P(d,m:DM) = \text{Prob}[U_{dm} \geq U_{d'm'}, \forall d'm' \in DM]$$

Given the joint probability we can derive any desired marginal or conditional probability. For example,

$$P(m:M) = \sum_{d \in D_m} P(d,m:DM)$$

and

$$P(d:D_m) = \frac{P(d,m:DM)}{P(m:M)}$$

Alternative Models

For simplicity, we write the probabilities in this section without the notation for the set of alternatives. In other words, we will write $P(d,m:DM_t)$ as $P_t(d,m)$, and $P(m:M_{dt})$ as $P_t(m|d)$.

In predicting the following joint probability:

$$P_t(f,d,m,h,r)$$

the set of alternatives consists of all possible trips, or all possible combinations of frequencies, destinations, modes, times of day and routes, available to individual t. In a simultaneous structure using the logit model, this will be the definition of the set of alternatives, and the choice probability will be for an alternative f,d,m,h,r combination.

The joint probability can be written as a product of marginal and conditional probabilities as follows:

$$P_t(f) \cdot P_t(d|f) \cdot P_t(m|f,d) \cdot P_t(h|f,d,m) \cdot P_t(r|f,d,m,h)$$

We can write this product in many different ways, e.g.,

$$P_t(f) \cdot P_t(h|f) \cdot P_t(m|f,h) \cdot P_t(d|f,h,m) \cdot P_t(r|f,h,m,d)$$

In a recursive structure we will use a logit model for each probability separately, and arrange the set of alternatives for each choice according to the sequence implied by the way we write the above product. For example, the probability $P_t(m|f,d)$ is the probability of choosing mode m , when the set of alternatives consists of the modes available to individual t , to destination d , with trip frequency f .

Estimating a sequential model requires further assumptions beyond the definitions of the relevant sets of alternatives for each choice. Consider, for example, the choice model for the probability $P_t(m|f,h)$. The problem is how to include in the model all the variables which for a given mode vary across destinations, e.g., travel time, travel cost, etc. Clearly, we cannot use all these variables as separate variables with their own coefficients. Therefore, we need to construct composite variables. There are many possible composition schemes. In addition there is the possibility of constructing the composite variables from several original variables together such that the tradeoff among them is kept constant in all choices. For example, for an alternative destination we define a generalized price by each mode which is a function of travel time and travel cost, then we aggregate across destinations to create a composite generalized price which is specific only to mode.

The Empirical Study

The data for this study was taken from a data set prepared by R.H. Pratt Associates (RHP) for the Metropolitan Washington Council of Governments (WCOG). The data set was combined from a home interview

survey conducted in 1968 by WCOG and a network (i.e., level of service) data set assembled by WCOG and RHP.

Due to the scale and the objectives of this empirical study, it was decided to use only a small subsample of the original data set for a single trip purpose: shopping. The data were kept in the disaggregate form where the observation unit is a household. This follows the assumption that the behavioral unit for a shopping trip is also a household.

Hence, the disaggregate data were exclusively drawn from conventional urban transportation study data. Specifically, trip and socio-economic data from a home interview survey, and level of service data from coded networks and other user cost data customarily collected by transportation planning agencies, were used.

Since our purpose is to evaluate the sensitivity of the predictions to the structure of the model we consider in the empirical work only the joint probability of destination and mode (given that a trip is taken) -- $P_t(m,d)$. We model this joint probability with three alternative structures; namely, a simultaneous logit model that estimates this probability directly*, and two possible recursive model sequences:

*The justification of this separation of destination choice and mode choice from other choices is as follows: The choice of time-of-day is assumed to be insignificant since the sample included only off-peak shopping trips. Route choice is not reported in the available data. The actual frequency is also not reported. Trips are reported for a 24 hour period. Therefore, the observed daily frequency is either 0 or 1 (and in a few cases 2). If we use an aggregate of households, this is sufficient information to compute an average
(cont'd on next page)

$$P_t(d) = P_t(m|d)$$

and

$$P_t(m) = P_t(d|m),$$

where a logit model is applied to each probability separately. We also investigate alternative ways of constructing composite variables for the marginal probability.

The sample that was used for estimation consists of 123 household home-shop-home round trips that were selected randomly from the original home interview sample in the northern corridor of Metropolitan Washington. Each household has a choice between two modes - the family car and bus - and several shopping destinations, ranging from one to eight according to the location of the household residence. It is important to note that we need to consider only alternatives that have positive choice probabilities. Therefore, a shopping location that is "too far" or a mode that is "unsafe" and consequently not feasible, or assumed to have negligible choice probability, need not be included in the set of alternatives.

The data consist of level of service variables by mode and destination, shopping opportunities by destination, and socio-economic

*(cont'd)

frequency. For a disaggregate model the actual frequency is not available. We are forced to assume that the choices of mode and destination are independent of the actual frequency and, therefore, can be modelled separately. Note that with 0, 1 daily frequencies: $P_t(f=1|d,m) = 1$ and $P_t(f=0|d,m) = 0$.

characteristics of the household. Each observation included the value of the variables for all the relevant alternatives for this household and the observed choice.

The simultaneous model that included observations with up to 16 alternatives and 7 variables gave reasonable coefficient estimates. The computer cost was only slightly higher ($\approx 20\%$) than the cost of a binary mode choice model with 5 variables. This indicates that a simultaneous model is feasible for the two choices of destination and mode. It also indicates that expanding the set of choices and therefore increasing the number of alternatives and variables may not be an unrealistic objective.

Comparison of the coefficient estimates of the simultaneous model with those of the various recursive models that were estimated suggest that the predictions are sensitive to the structure of the model.

In order to demonstrate this sensitivity here, it is not necessary to present in detail all the models that were estimated in this study.* It will be sufficient to show some examples of the important tradeoffs and elasticities.

The following table shows the values of time implied by the different models:

*The estimation results are described in detail in Chapter VII.

Estimated from a model for:	$P(m,d)$	$P(m d)$	$P(d m)$
Value of Out-of-Vehicle Travel Time	3.02 \$/hr (1.44)	1.36 \$/hr (.98)	4.67 \$/hr (4.36)
Value of In-Vehicle Travel Time	.78 \$/hr (.68)	.28 \$/hr (.66)	2.21 \$/hr (2.01)

The figures are for a household with annual income between \$10,000 and \$12,000. The numbers in parentheses are standard errors.

Although the standard errors are relatively large, this is not atypical for estimates of value of time (e.g., Talvitie, 1972). (The estimated model coefficients that were used to compute the values of time were significantly different from zero.)

Estimated values of time from the simultaneous model are greater than those estimated from a mode choice model (given destination), and smaller than those estimated from a destination choice model (given mode).

The following table shows some direct elasticities of the mode choice probability:

Mode	Bus		Auto	
Estimated From a Model for:	$P(m,d)$	$P(m d)$	$P(m,d)$	$P(m d)$
Out-of-Vehicle Travel Time	-1.01	-.82	-.13	-.10
In-Vehicle Travel Time	-.31	-.26	-.05	-.03
Out-of-Pocket Cost	-.40	-.91	-.13	-.23

The figures are computed for the following case:

- a household with annual income between \$10,000 and \$12,000.
- the probabilities of choosing bus and auto are .2 and .8, respectively,
- out-of-vehicle travel times are 20 minutes by bus, and 10 minutes by auto,
- in-vehicle travel times are 30 minutes by bus and 15 minutes by auto,
- out of pocket costs are 50 cents by bus and 50 cents by auto.

The most striking variation in this table is in the cost elasticity. The mode choice model derived from an estimated joint probability gives cost elasticities which are about half the elasticities computed from a recursive mode choice model.

The differences among the models could be attributed to specification errors which affect differently a mode choice model and a destination choice model. The effects could be in the opposite directions and therefore the joint probability model gave estimates that are in some way between the estimates of the two other models.

The marginal probabilities of the recursive models which were formulated with composite variables also demonstrated significant differences from the corresponding probabilities derived from the simultaneous models.

All the alternative models that were estimated resulted in essentially equal goodness of fit.

Thus, the chosen structure can make a big difference in terms of the values of the estimated coefficients. Since there are *a priori* reasons to assume a simultaneous structure, rather than recursive, we should estimate directly the joint probabilities. Then, if necessary, we can derive any conditional probability.

Conclusions

Models based on disaggregate data and choice theory were estimated in the past either for a single travel choice, primarily mode choice, or for several choices but in a recursive structure. The empirical study that was conducted in this research demonstrated the estimation of a disaggregate simultaneous model. The results from the estimation of a simultaneous destination and mode choice model indicate that this approach is feasible within reasonable computation cost. Moreover, the estimation results of models with recursive structures for the same two choices show that important coefficient estimates vary considerably with the different model structures.

This empirical study was limited in scale, and it is recommended that the evidence should be extended to include: alternative data sets, different trip purpose categories, a complete set of travel choices, and a more extensive set of explanatory variables (in particular, attraction description).

The empirical evidence taken together with the theoretically appealing assumptions of a simultaneous structure and the advantages of disaggregate models suggest that future efforts in travel demand

modelling should be in the direction of simultaneous disaggregate probabilistic models. Given the joint probability (from the simultaneous model), one can derive conditional probabilities and use the model for forecasting in sequential stages, corresponding with the UTMS procedure.

One of the important problems in using disaggregate models for forecasting is the aggregation problem. Future research efforts should investigate this problem. However, for the short run, simplified aggregation procedures, such as market segmentation, are available and can be used.

The use of disaggregate models suggests new emphasis in data collection efforts for transportation planning. It is not clear yet how much data is needed for disaggregate models, but it is clear that a change in the general methodology of travel data collection is appropriate.

The comprehensive home interview survey covering an entire planning region might be replaced by several more descriptive small samples, in selected areas of the region. Thus, the emphasis should be to attempt to represent the full range of socio-economic characteristics affecting travel behavior, rather than on sampling all parts of the region at a uniform rate. Smaller scale surveys will make possible the collection of the detailed information (not conventionally collected) that are important for disaggregate demand models. For example, car pool information, information on how often a trip is made (instead of reporting only the trips made during the last 24 hours), information on institutional constraints such as preferred arrival time, and so forth, would be

obtained. In addition to the travel data requirements, better information is also needed with respect to the attributes of alternative trips. In particular, the attraction data that is available from conventional data sources used in Urban Transportation Planning are not very descriptive. More detailed attraction data is needed in order to achieve better predictions of destination choice.

In depth studies of travel behavior based on detailed interviews and attitudinal data could be fruitful. However, it appears that the most beneficial directions for research toward improvements of transportation planning capabilities are: the aggregation problem, the behavioral modelling of round trips with non-home-based links, and the experimental application of simultaneous disaggregate models to case studies of important transportation issues at different levels of planning.

In conclusion, this research has indicated the desirability and the feasibility of a simultaneous disaggregate travel choice model.

Outline of the Report

In Chapter II, the nature of the demand for travel is discussed in terms of its inherent characteristics and its relation to other economic goods. The system of models used for transportation planning is outlined and the implications for disaggregate modelling of mobility and travel choices are discussed. Chapter III presents a review of some concepts from consumer and choice theories. It indicates that choice theory, rather than demand analysis, is the appropriate approach for the

modelling of mobility and travel choices. Chapter IV establishes alternative model structures, each representing different assumptions with regard to dependencies among choices. Chapter V reviews the structures of current approaches to travel demand modelling. In Chapter VI the multinomial logit model and its application to the alternative travel demand models is described. Chapter VII reports the results of an empirical study in which the alternative models were estimated on a comparable basis. Finally, Chapter VIII presents a discussion of the implications from the theoretical considerations and the empirical evidence toward the modelling of travel demand. It also points the desirable directions for research.

CHAPTER II

Nature of Travel Demand and Prediction Models for Transportation Planning

Introduction

This chapter examines the nature of the travel demand process and the overall structure in which it is embodied. An attempt is made to survey the theoretical considerations on which travel demand models are based.

The overall structure of the system of models used for prediction in transportation planning has implications for the specification, data collection and estimation, and usage of the various planning models. Therefore, we consider the travel demand process and its relations to other processes; i.e., socio-economic activities and transport supply. We establish first an overall structure of the system of models and then focus on the travel demand model.

The models used for planning take a market point of view. However, it is argued that the models that describe consumer behavior should be developed and estimated for the unit that makes the decisions; the individual consumer. Thus, the chapter also establishes the modelling framework to be used in demand modelling from the point of view of an individual consumer.

The Complexity of Travel Demand

The complexity of travel demand is apparent from the way we characterize a trip--origin, destination, time of day, modes of travel, route, and purpose.

Trips are spatially disaggregated; we speak of a trip as made from an origin to a destination. For a work trip the origin and the destination are fixed by the choices of residential location of the household and the tripmaker's job site. However, for most leisure trips one has a choice among several destinations. For example, shopping trips can be made to a nearby grocery store or to a distant shopping center.

Peaking effects bring out the importance of the time of day in which a trip is taken (perhaps also the day of the week if it is not an every day trip). For the majority of workers, working hours are not totally in their control; however, they can still decide to arrive early at their jobs in order to avoid the inconvenience of the peak hour. Staggering working hours in downtown areas is often suggested as a means of relieving congestion during the peak hours. Some leisure activities are less constrained and the consumer has a range of hours in which he can make a trip to those activities.

In most urban areas the traveller has a choice between public transportation modes and private automobile. Even when there is no public transportation available, a traveller can sometimes choose between driving his own car, being driven, or participating in a car pool. In some cases, taxi, walking, and two wheeled vehicles

are also viable options. The situation is more complex when the traveller makes use of several modes for a single trip. For example, a suburban commuter who drives his car to the commuter railroad station in his suburb, parks the car, takes the train to the downtown terminal, and finally rides a bus to his place of work.

In dense networks, the traveller can choose among several alternative routes. A driver of a private automobile may have a choice of a limited access toll road, a less fast free road, and possibly an even slower road offering attractive scenic qualities. Thus, for a given trip the traveller is faced with a set of choices: choice of destination, choice of time of day, choice of mode, and choice of route. In addition, a potential trip maker has to decide whether or not to take a trip, or how often to make a trip for a given trip purpose. For example, families can shop for groceries with different frequencies; every day, twice a week, once a week, etc.

The classification of trips by trip purpose does not represent an inherent characteristic of the trip itself, but rather stems from the recognition that the demand for trips is a derived demand.* A trip is almost always made not for its own sake, but for some purpose that can be served at its destination. Therefore, tripmakers' behavior will be different for different trip purposes. There is no substitution among trips for different purposes unless the corres-

*This point will be discussed further in the following section.

ponding activities are substitutes. However, the practice has been to define broad categories: work, business, recreation, school, social, shopping, and personal business. For this classification, substitution among trip purpose will in general be non-existent. Hence, it is possible to model separately travel demand for different trip purposes.

The usual definition of a linked trip in Urban Transportation Planning (UTP) studies is a one-way travel from origin to destination. However, recognizing that the demand for travel is a derived demand, it is evident that this definition is an inappropriate unit of trips demanded. We have to consider round trips or tours, i.e., single or multiple-stop round trips. To carry out an activity at any destination generally requires at least a stop-over at that destination. Hence, the travel associated with that activity or bundle of activities is the round trip.

The trips to and from each destination are interdependent-- decisions made with regard to the first leg constrain and influence the decisions made with regard to the remaining legs, and the trip alternatives available for the remaining legs influence the decisions made on the first leg. For example, an individual who goes from his home to shop by car will also return by car. If he chooses to make another stop on this tour for shopping or any other activities, his choice of the two destinations will not be independent.

If we consider a round trip or a tour, the complexity of travel demand increases further. A single trip is described by its origin, destination, time of day, mode, route, and purpose. A tour has a single base point, or origin, but a mixture of several different destinations, times of day, modes, routes and purposes.

Simple round trips, e.g., home-work-home, home-shop-home, etc., have in most cases two identical legs, and therefore pose no additional problem for modelling. In a few cases, the mode of travel may be different, e.g., going to the supermarket by bus and returning by taxi, but this could be handled by defining the relatively few possible combinations of modes as separate modes. However, for some complex tours, the dimensions of choice, that is, the number of alternative trip options, can be prohibitively large. If we assume that all tours originate at the home, then all intermediate legs are non-home based trips, which are customarily modelled separately from home-based trips, that is the first and the last leg of the tour.

Other complexities of travel demand stem from the variety of factors that affect the decisions people make. In general, the factors include the characteristics of the areas considered, the socio-economic status of the traveller, and the characteristics of the transportation system. It is not our intention to attempt to enumerate those factors and it will suffice to note that the number of these factors can be enormous and many of them intangible and difficult to quantify. In addition, people's preferences and values

differ considerably. This causes a high level of uncertainty to be present in any attempt to predict the demand for travel.

Thus, travel demand is a complex phenomenon that is dependent upon many factors that are extremely difficult to describe analytically. A difficult analytical problem is to find a way to properly model complex tours. The analytical problem that is the focus of this study arises from the multi-dimensional choice situation, i.e., the set of interdependent choices that a traveller is facing. The modelling of multi-dimensional choice situations will be discussed in Chapter IV. The following sections of this chapter and Chapter III provide the theoretical basis for a close examination of the internal structure of travel demand models.

The Demand for Travel as a Derived Demand

A trip is almost always made for a certain purpose that can be served at its destination. The exception is a pleasure ride which is made solely for its own sake. Pleasure trips are relatively infrequent and probably non-existent in dense urban traffic. Hence, if the traveller does not derive any utility from the trip by itself, it can be regarded as an intermediate good (Oi and Shuldiner, 1962; Kraft and Wohl, 1967). The demand for an intermediate good is derived from the consumption of the final good. In the case of travel the final good is the activity at the destination of the trip, e.g., work, shopping, etc. The demand for this activity is a function of its price which includes the cost of the associated trip. Given

the demand for this activity the demand for trips alone can be derived.* As an example, consider the demand for theater visits. It will depend, among other things, on the price of the ticket and the cost of the trip to the theater, including parking fee if a private automobile is used. The demand for trips to the theater is now determined from the demand for theater visits.

The theory of derived demand as presented in production theory in the micro economic literature (Stigler, 1952; Marshall, 1961), does not have much to offer to the discussion here. Our interest is to explore the implications it has on the modelling of transportation demand. The basic implication stems merely from the recognition of travel demand as derived demand and is quite simple -- in order to model travel demand properly one has to model the demand for the final activities. The other inferences that can be made are more important in the empirical work since they are qualitative statements, or rules, about the derived demand elasticity.

The modelling of the demand for final activities is unfortunately very difficult. It is avoided by classifying trips according to trip purpose and modelling directly the demand for trips, in each trip purpose category. Also, the variables that are used as attraction variables serve as proxy to the intensity of the final activity. It

*It should be mentioned in this context that there are also demands for a variety of services and products which are derived from the demand for travel. One such case is the demand for new planes, which depends on the demand for air travel (Philips, 1971).

will be apparent later that this procedure is particularly unsatisfactory for the work trips.

Interaction Between Location and Transport

The demands for the final activities are among the forces that determine the pattern of development. The location decisions of households are influenced by considerations of accessibility to these activities. Similarly, firms' and services' location decisions are partially based on accessibility to markets for their services and products, and that of the labor force and clientele to their location.

The location decisions determine the pattern of trips for the various activities, and the available supply of transportation influences the location decisions. This observation brings out a rather unique case of a market demand curve.

In general, a demand curve in a static framework is assumed to shift due to income changes and changes in prices of other commodities, and aggregate demand will shift also due to natural growth of the population. However, in this case, the aggregate demand curve for trips will shift also due to changes in the transport supply curve. An improvement in the supply of trips between any two points will increase the number of trips made between those points. In addition, and here we deviate from the conventional static model, the improved accessibility may attract new locators which will further increase the volume of trips, or actually shift the demand curve to the right.*

*Thus, there may be shifts to the left in locations where the supply does not change.

This is illustrated in Figure 1. The demand curve D_1 is the curve that is in effect with the supply curve S_1 . The equilibrium volume of trips is q_1 at a generalized travel price (includes money outlay and travel time) of p_1 . An improvement in the transportation system results in a new supply curve, S_2 . The new equilibrium volume and price are q_2 and p_2 , respectively. However, this is only short run equilibrium. In the longer run the demand curve will shift to the right along the supply curve S_2 and the demand curve D_2 . The shift of the demand curve from D_1 to D_2 is conventionally attributed to changing income, prices of other commodities and population. However, in our case part of the shift is attributed to the fact that the supply curve shifted from S_1 to S_2 . The long run equilibrium is q_3 and p_3 .

It may be argued that the supply of transportation plays only a minor role in determining the general magnitude of activities, and therefore the demand shift attributed to the change in supply is negligible. However, it is reasonable to assume that the spatial distribution of the activities is more directly related to transportation (Mayer, 1968). Since demand curves for trips make no sense unless they are spatially disaggregated (we always consider demand for trips between two points or zones) the shift in the demand curve will be caused primarily by the changing distribution of activities.

The implication of this interaction between location and transport is that if we want to properly evaluate major policy programs in

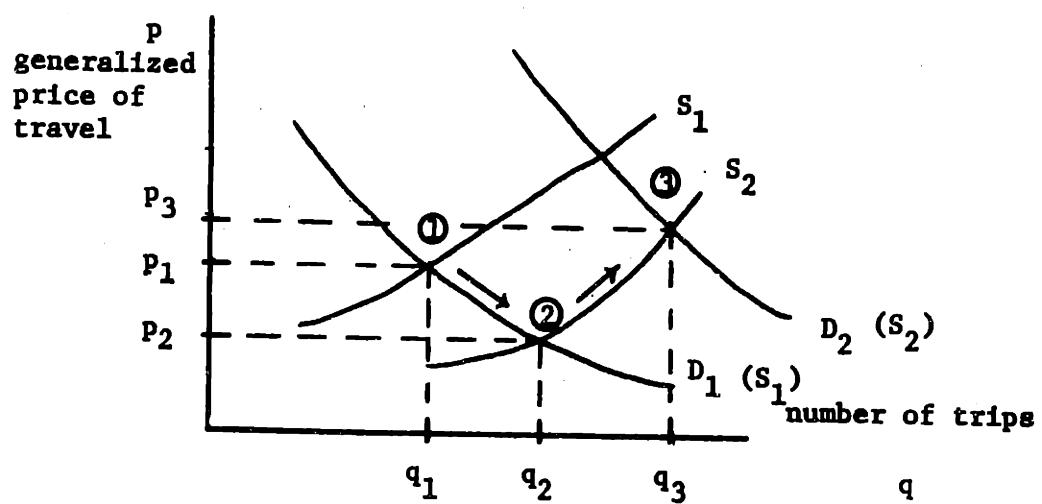


Figure 1.

transportation facilities, we should model the effect of transport on the spatial distribution of activities. In particular, we should model the locations decisions of potential travelers which are closely related to their travel demand. For example, an improved service on a bus line will attract more riders but will also influence people who consider using the bus service to locate near the bus stations.*

The Importance of the Work Trip

By far, the trip from home to work and the trip from work to home are the most important trips made by the household. The work trip has the greatest influence on location decisions amongst all trip purposes. From the transportation system planning point of view, the work trip is also the most important trip purpose. The majority of the work trips are made during the peak hours and constitute the bulk of the trips made during the peak hour. (The only other trip purposes that generate significant volume of trips during the peak hours are school trips during the morning peak and shopping trips during the afternoon peak.) It is the peak hour volume that strains the capacity of the transportation system. In general, transportation systems have more capacity than they need for the off-peak traffic volumes (except for recreation trips on weekends in some areas).

*The location decision and other decisions related to personal mobility that are closely connected with travel decisions will be discussed further later in this chapter.

In the short run when household locations and job sites are fixed, the demand for work trips is very inelastic with respect to the trip cost. This is due to the fact that the price of a trip is generally small compared to the benefit derived — the earned income. An improvement in the transportation system may cause in the short run a person to switch jobs. With the improvement he can now reach a higher paying job without completely compensating for the increased income by increased commuting expenditures. In the long run the problem is more complex; persons can change their place of work, households may change their residence locations, and firms and services can also relocate. This points out that for long range forecasting, we should use location models to predict the demand for work trips.

Short Run vs. Long Run, Static vs. Dynamic

We made a distinction between two types of equilibria: short run and long run. In the short run equilibrium the magnitude and distribution of activities are assumed to be fixed. In the long run equilibrium, however, they change and we assume that the location market also reaches equilibrium.

One can argue that an assumption of a static long run equilibrium is unrealistic. The reaction times in the location markets are very long, and therefore we can never observe, at any point in time, this equilibrium. The importance of the reaction times, or lag effects, points to the necessity of departing from the static framework and introducing some dynamics. We will return to this point later when

we discuss alternative overall structures of the transport prediction problem.

The Overall Structure of Prediction in Transportation Planning*

The foregoing discussion described the interactions of the travel demand function with the location of activities and the supply of transport services. We proceed now to outline a structure for prediction in transportation planning of which the demand function is a part.

The problem of prediction in transportation planning is to anticipate the impacts from a proposed change in the transportation system. We can distinguish between two types of impacts: internal and external. The internal impacts depend on the new equilibrium in the transportation system. The external impacts depend on the effects the change in the transportation system had on the equilibrium of related markets. We will refer to these markets, which constitute the environment of transportation, as the activity system. For example, construction of a new highway will affect the equilibrium in the transportation network. But, it will also affect the activity system, e.g., changes in land values in the vicinity of the highway, changes in the production costs of manufacturing firms located near the highway, etc.

In general, the transport planner cannot directly change the activity system. However, it is evident that changes in the activity system will have a very significant effect on the transportation

*The discussion, terminology, and notation in this section draw heavily from Manheim (1969).

system. For example, consider the effects that a change in the housing market, like housing subsidies or new mortgage policy, will have on the transportation system.

The basic problem for the transport planner is to predict the equilibrium in the transportation system which is described by the pattern of flows in the transportation network. The pattern of flows consists of the volumes of trips V and the service levels experienced by those volumes, L .* The general structure of predicting the flow pattern is the equilibrium between supply and demand. The demand function D describes the demanded volumes of trips V as a function of the levels of service L , and the characteristics of the activity system A .**

$$V = D(L, A) \quad (1)$$

The supply function S describes the levels of service offered L as a function of the volumes of trips V , and the characteristics of the transportation system T .***

$$L = S(V, T) \quad (2)$$

*By service level (L) we refer to all the service characteristics of a trip: travel time, travel cost, safety, comfort, etc. We can regard this service level as a vector of prices that if combined together can be regarded as a generalized price.

**The characteristics of the activity system (A) that determine the demand for travel are the spatial pattern of population and economic activities, or locations and intensities of land uses, and the socio-economic characteristics of the different population segments.

***The transportation system characteristics (T) include a variety of technological factors (e.g., vehicles, network configuration, link characteristics, etc.), and operating policies (e.g., schedules, pricing policy, etc.)

The reduced form* of the system of equations (1) and (2) gives the equilibrium flow pattern in equations (3) and (4).

$$V = V(A, T) \quad (3)$$

$$L = L(A, T) \quad (4)$$

The equilibrium is determined by the characteristics of the activity system A, and those of the transportation system T. However, following our discussion earlier, this is only a partial, or a short run, equilibrium. It does not consider the effects the transportation system has on the activity system. This structure assumes that the characteristics of the activity system are exogenously determined, an assumption that may be justified in some instances, but which is certainly unreasonable for environments of rapid development or for undeveloped transportation systems.

The structure that is needed to determine the more general, or the long run, equilibrium should include the supply and demand functions of all the related markets of the activity system, in addition to those of transport. Only with this general structure can we predict the full set of internal and external impacts caused by changes in both the transportation and the activity systems.

Since the focus of this study is the travel demand function, we will not consider the details of the activity markets. For our purposes here, it will be sufficient to describe the activity markets in terms of their reduced form, as in equation (5).

$$A = A(E, L) \quad (5)$$

*The solution of a system of simultaneous equations is called a reduced form.

The equilibrium pattern of the activity system is a function of a variety of factors and policies that we noted by E,* and the levels of service offered by the transportation system L. The reduced form of the general structure that includes the transport supply and demand, and equation (5) is as follows.

$$A = A'(E, T) \quad (6)$$

$$V = V'(E, T) \quad (7)$$

$$L = L'(E, T) \quad (8)$$

This reduced form describes the long term equilibrium as a function of E and T, which are exogenously determined.

To summarize, the structure that has to be constructed for prediction includes three types of relationships as represented by equations (1), (2), and (5).

Before we continue, it should be noted that there might be some doubt as to the validity of the assumption that all the elements of T are determined outside the system. The characteristics of the transportation system are altered by governments and private firms in response to changes in the equilibrium of the transport markets.

However, government decisions and reactions of the mostly regulated transport firms are in most cases impossible to model, partly due to the political nature of most important decisions, and partly because

*For example, housing characteristics, public services, education, taxes, topography, etc. In addition, this group of variables (E) includes some elements of the transportation system characteristics (T), such as network configuration and capacities.

this process is very ill-defined. A planner who takes the government point of view has no interest in modeling this relationship, since most of the variables in T are his design variables. In a situation where the private transport operators play a significant role, some variables in T could not be assumed to be exogenous. These variables will have to be explained inside the system by relationships describing the investment and pricing behavior of the private firms.

Alternative Structures

In outlining the structure of the prediction problem we used an equation format. It is useful to represent this structure with the aid of a block diagram in which the arrows represent inputs or outputs and the blocks represent processes. The structure of equations (1), (2), and (5) is depicted in Figure 2.a. The equilibrium values of A, V, and L are the long run equilibrium. The system to the right hand side of the dotted line which takes A as exogenous represents the short run equilibrium.

Another useful graphical technique to describe a causal structure is an arrow diagram (de Neufville and Stafford, 1971). In this diagram, Figure 2.b, the arrows indicate the direction of causation between two variables. The short run structure is represented by omitting the two arrows which are crossed by dotted lines.

So far, we used a static framework in which time was not treated explicitly. The usage of the terms short run and long run merely reflects the assumption that the interactions between the transport

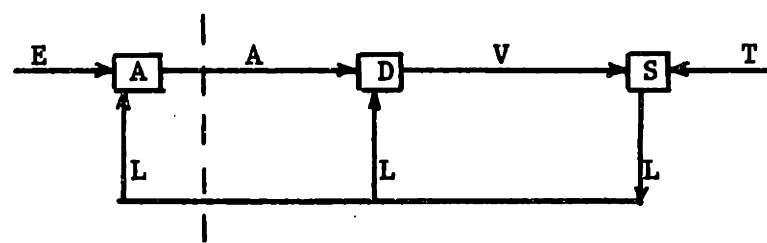


Figure 2.a



Figure 2.b

market and the activity system are slower than the interactions between transport supply and demand. It was argued already that time has to be taken into account explicitly because the assumption of a long run static equilibrium is unrealistic and can never be observed. Transforming the static structure into a dynamic one requires some definite hypothesis about reaction times. For each of the three processes A, D, and S, as shown in Figure 2.a, we need to assume the reaction times each process has in response to its inputs. Consequently, in order to determine the output at the simulated point in time we have to lag each input by the reaction time associated with it. For example, consider the block D that determines V as a function of A and L. It is plausible that the volumes demanded today are a function of present activity system characteristics and the levels of service that were experienced last month.

The determination of these lag periods is a difficult problem. A prerequisite to any empirical work with a dynamic structure is detailed time series data. In addition to theoretical considerations, the length of the time lags will depend on the kind of data available, or on the time period implicit in the definitions of the variables. For example, we can define the volumes as yearly average daily volume (which is equal to total yearly traffic divided by the number of days in a year). This means that the observation period should be one year. Suppose now that the volumes demanded during a particular year depend on the characteristics of the activity system last year. This implies

that in order to model the lag we need a series of observations from consecutive years. If we have an observation only every five years, this lag cannot be modelled. If the period of observation is two years (i.e., instead of yearly volumes, we have data for volumes during periods of two years each), we still cannot model the lag; however, in this case we can assume that the lag effect is less important.

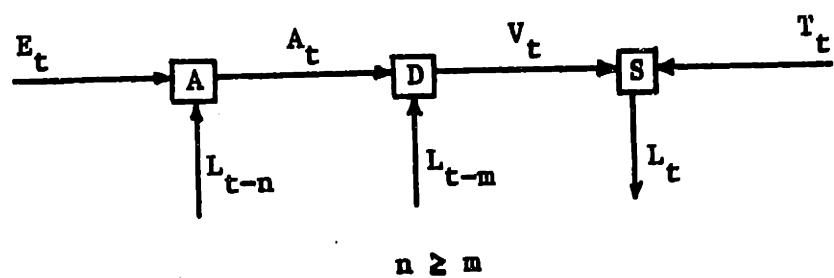
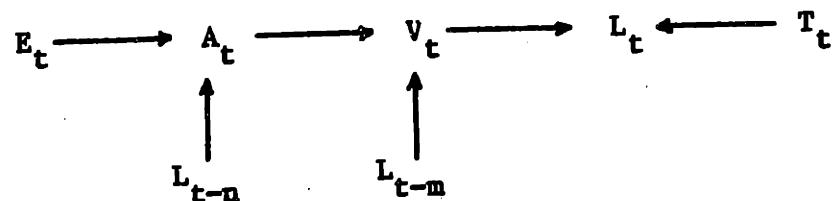
Let us assume the most common case of the variables defined on a yearly basis. Under this circumstance, which of the input flows, in the block diagram of Figure 2.a, can be assumed to have a significant lag effect? The most obvious lag has been discussed already. It is the periods of time it takes the activity system to respond to changes in the transportation system. Another plausible lag, which will most probably be shorter than or equal to the first one, is the period of time it takes the demanded volumes to respond to changes in the levels of service. The dynamic structure that includes these two lags is displayed in the block diagram of Figure 3.a, and the arrow diagram of Figure 3.b.

The system of equations that describe this structure are the following equations (9), (10), and (11).

$$A_t = A(E_t, L_{t-n})^* \quad (9)$$

$$V_t = D(A_t, L_{t-m}) \quad (10)$$

* E_t should also include a lag which represents internal reaction time in the activity system. It could also include lagged dependent variables.

Figure 3.aFigure 3.b

$$L_t = S(T_t, V_t) \quad (11)$$

This system has a recursive structure, i.e., A_t , determined first, influences V_t , which in turn influences L_t . If the lag m is assumed to vanish, we have a structure in which equations (10) and (11) are simultaneous but still recursive with respect to equation (9). This structure is represented in the block and arrow diagrams in Figures 4.a and 4.b.

The empirical problems with a static structure are the solution and estimation of a complex system of simultaneous relationships. With a dynamic structure there are the additional difficulties of obtaining adequate data and convergence properties. However, if adequate data are available, a dynamic structure is much more appealing because of its realistic nature. It has another important advantage which we have not discussed. It allows the explicit treatment of the effects at any point in time of existing stocks and the adjustment costs, e.g., relocation cost.

The Structure of the Urban Transportation Model System (UTMS)*

It will be useful to compare the previously outlined abstract structure with the structure of the Urban Transportation Planning system of models (UTMS) that have been used by metropolitan planning agencies throughout the world (FHWA, 1970).

*Note that we are concerned here only with the overall structure of the system of models and not with the specifics of the models.

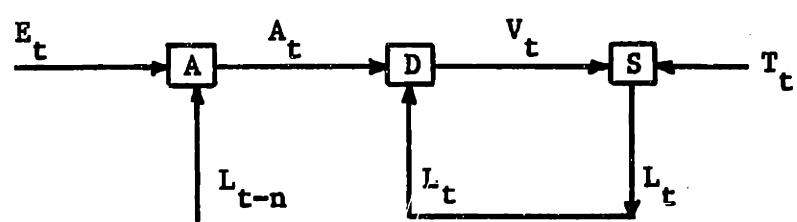


Figure 4.a

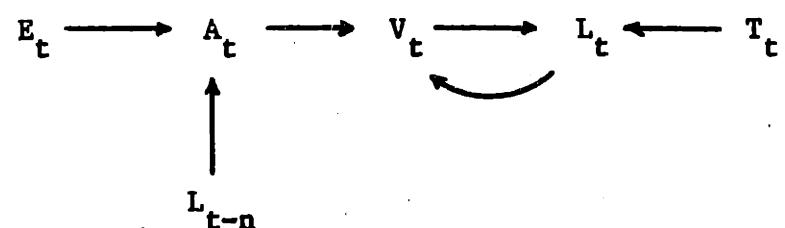


Figure 4.b

The structure of the UTMS is basically recursive and static. It consists of a sequential application of the following models: land use, trip generation, trip distribution, modal split, and network assignment (Martin, Memmott, and Bone, 1961). Using the same notation as before, this structure is shown in Figures 5.a and 5.b, where L_0 is the expected levels of service which are assumed initially and are usually not revised.

The activity system model is usually nothing more sophisticated than a straightforward trend projection of key variables, and in a few cases a somewhat sophisticated land use model that take into account the expected levels of service, e.g., Empiric (Traffic Research Corp., 1967), Lowry (Lowry, 1966). The demand model is broken down into four sequential steps: trip generation, trip distribution, modal split and network assignment. A supply model is used when the network assignment is of the "capacity restraint" type and not "all or nothing." In this case, the supply model is a series of volume-travel time curves. These curves are used to revise the expected values of the levels of service and iterate with the last step in the demand model.* In some cases, iterations were made also with modal split and trip distribution models (Irwin et al, 1961). No iterations are

*When these iterations, or the equilibrium procedure, are not performed in an "all or nothing" assignment, the supply curve is essentially assumed to be infinitely elastic.

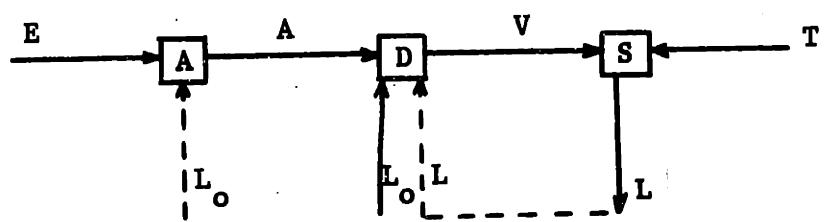


Figure 5.a

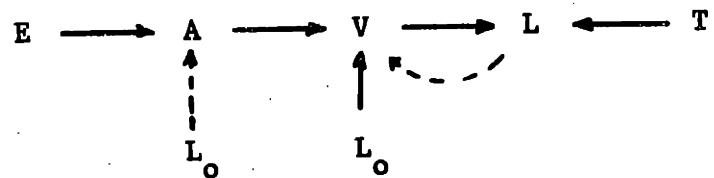


Figure 5.b

possible with trip generation because it is generally assumed to be independent of level of service.*

The structure of the UTMS as in Figure 5 is similar to the dynamic structure of Figure 3 when both n and m are set to zero. However, in such case, the relevant structure is no longer Figure 3, but instead the general static structure of Figure 2. The UTMS can be made to be more in accordance with this structure by making both the land use and trip generation dependent on L and then iterate the whole system of models in solving for equilibrium.

During the last decade, two other major systems of models have been developed: the Northeast Corridor system of models (NECTP 209, 1970), and the Harvard-Brookings simulation model (Roberts and Kresge, 1971) that was implemented for Colombia. Both systems represent considerable theoretical improvement over the UTMS. To an extent, this improvement was partly facilitated by the change of scenario from metropolitan areas to a dense intercity corridor and a less developed country.

Modelling Considerations

Why is it so important that the structure of a system of models conform to the theoretical structure? The answer to this question is

*It could be claimed that level of service does not have a significant effect on trip generation for some trip purposes. However, this does not justify the exclusion of a level of service variable from a trip generation model because it is unreasonable that "price" has no effect at all on "quantity" demanded.

that a wrong structure may lead to distorted and biased prediction. For example, it was argued that even if the incomplete structure of the UTMS produces correct forecasts, it is only because they are "self-fulfilling prophecies" (Meyer and Straszheim, 1971). Travel demand forecasts made independently of the transportation system are biased upward. The sizeable volumes will indicate a need for more highway investment. This great improvement in the transport system will affect the activity system to the extent that the large volumes that were forecasted will now be realized. If this sequence of events is true, we have justified an investment on the basis of benefits that never occur.

The structure determines the way in which the various sub-models should be used and interact with each other. Suppose that all the sub-models in the UTMS are specified correctly and their parameter estimates are unbiased. Unless iterations are made with all phases of the demand model and the land use model, the predictions will be biased. The size of the bias will depend on the difference between the assumed level of service (see Figure 5), and the true equilibrium level of service. At equilibrium the level of service inputs to A and D should be equal to the output level of service from S. Since the initially assumed levels of service are in general based on unloaded networks, presumably the forecasts of demands and changes in land use are biased upwards.

The structure also plays an important role during the estimation (or calibration) of the various parameters of the models. Assume that all the models in the system are specified correctly. An attempt to estimate the parameters of any model with no account of the structure may result in biased parameters. The first issue here is the identification problem (Fisher, 1966). In a system of simultaneous equations, the parameters of any equation can be estimated from the data only if this equation is identified with respect to the others.* If an equation is not identified then there is no way in which its parameters can be estimated (although the parameters of the reduced form can be estimated and used for forecasting). If an attempt is made to estimate an unidentified equation, the parameter estimates are meaningless. To find out whether an equation is identified or not one has to specify explicitly all the equations in the structure. However, with the conventional use of cross section data for area wide demand models, the demand equation will be identified with respect to the supply of transport services. With the assumption of long run equilibrium, it is not possible to make a similar statement about the identification of the travel demand and activity models with respect to each other, unless the models are specified explicitly. For example, assume that car ownership and the choice of mode are each a function of the other plus an identical set of independent variables. For this

*The conditions for an equation in a system of simultaneous equation to be identified are presented in Fisher (1966).

specification each model is not identified with respect to the other and therefore its parameters could not be estimated.

Suppose now that all the equations in a system of simultaneous equations are specified correctly and are identified. Estimating (calibrating) each equation separately using ordinary least squares regression, will result in biased parameters (Malinvaud, 1966). This bias is often referred to as simultaneous equation bias. To get consistent estimators one has to use an estimation method that takes account of the structure, the most commonly used being two stage least squares (Theil, 1971). There is no empirical evidence to determine the importance of this bias in land use and demand models. Therefore, it is not clear whether the additional expense associated with more sophisticated estimation techniques is justified. The implicit assumption in current practices of calibrating each model by itself is that this bias is minimal or non-existent. Thus far, no research has been performed to check whether or not this is a valid assumption.

So far we assumed that all the models are correctly specified, that is, the model includes all the true explanatory variables in their correct mathematical form. The correct specification has to be dictated from theory or a-priori knowledge. It is impossible to determine the correct specification by statistical analysis of the data. The problem is that frequently we do not have a well structured theory in sufficient detail. Moreover, important variables are often missing due to lack of data or measurement problems. Models can be misspecified in two ways--in their mathematical form, and in their

list of explanatory variables. One example of misspecification of the mathematical form is the assumption of constant elasticities in travel demand models. An incomplete list of explanatory variables is the case with trip generation models that omit the level of service variables or trip distribution model with travel time as the only level of service variable. A failure to recognize the dynamic or lag effects in a model -- that is, the erroneous assumption of a static structure -- is also an example of misspecification. If a variable should be lagged and it is not, it is equivalent to an omission of one (the lagged) variable that should be included, and the inclusion of another (the unlagged) variable that does not belong in the model.

Errors due to misspecified models* (specification errors), are probably the most crucial errors in transportation models. The most common misspecified models are trip generation equations which include socio-economic variables, but exclude level of service, or accessibility variables. Assume that the true explanatory variables of trip generation are income and accessibility. What effect will the omission of the accessibility variable have on the model and on the forecasts produced by the model? With cross section data it is likely that the affluent suburban residents will have better accessibility for most trip purposes than central city median and lower income residents. If we assume that this positive correlation between income and

*For a discussion of specification analysis, see Theil (1971), Chapter 11.

accessibility exists in the data then it can be shown that the estimated income coefficient will be upwardly biased. This means that if a forecast is made for a future year, for which income is expected to rise, it will be also biased upward (assuming that accessibility is kept constant).

The structure also has implications on the kind of data that is needed and the way it should be organized. Obviously, for any meaningful empirical work the structure has to be spelled out in a much higher level of detail than the one we used here.

Disaggregate Models

For planning purposes, we are concerned with the prediction of the behavior of aggregates of people. In order to forecast a market equilibrium we need a demand function that represents the aggregate demand of all consumers in the market. The demand models that were considered in the previous sections are aggregate models, and their reduced form with the supply functions was the equilibrium solution. However, an aggregate demand function is simply the sum of individual consumers' demand functions. Demand functions may be developed for different levels of aggregations. The most disaggregate level, or the micro level, is that of the individual consumer, or the single behavioral unit. This is the level at which the behavioral theory on which demand models are based is postulated. Aggregation is achieved by creating groupings or collections of individual consumers. These groupings are often based on geographic location and/or socio-economic characteristics.

Obviously, it is not feasible to compute the demand for all individual consumers and accumulate them in order to obtain an aggregate estimate of demand. Hence, for forecasting, some level of aggregation is necessary. The appropriate level is determined based on considerations of the purpose of the forecasts, the required detail and accuracy of the analysis, and the data processing and computational costs.

If detailed observations are available, aggregation for forecasting could be performed either before or after the model estimation. Aggregation during the model construction phase of the analysis will cloud the underlying behavioral relationships and will result in a loss of information. It is always desirable to estimate a model at the disaggregate level. Sometimes when detailed observations are not available or due to economic reasons (i.e., limited resources for data processing and computational tasks), some level of aggregation is necessary. In principle, aggregation for the purpose of forecasting can always be performed after estimation. Models estimated at the micro level can be aggregated to any level that may be required for forecasting purposes

The theoretical starting point in specifying demand functions is always at the level of the individual consumer. The transformation of individual demand functions to an aggregate function is a difficult problem. Without going into the details of the aggregation problem, a few relevant points will be summarized. If we assume that the

individual functions are identical (i.e., identical specifications and sets of coefficients) and linear, the aggregate function has the same specification with averages of the variables substituted for individual values (Theil, 1971). However, if the individual functions are non-linear then the same functional form with averages of the variables will give biased forecasts. In principle, the transformation could be performed by integrating the relationship over the distributions of the explanatory variables (Kanafani, 1972). This means that for a non-linear relationship the aggregate functional form will not in general be the same as the disaggregate form, and it will be different for different aggregation schemes. In addition, this means that estimation results of a model based on aggregate data will be valid for forecasting only with the same aggregation scheme.

On the other hand, models estimated on the micro level are based on individual observations and represent the average of behavior of individual consumers. Therefore, they are independent of an aggregation method. Furthermore, to the extent that consistent consumer behavior for similar circumstances can be assumed, a model estimated with data on a set of consumers in one area can be used to forecast behavior of other consumers in another area, or at a different point in time.

The transformation of micro non-linear models into aggregate models can be an intractable mathematical problem as was indicated by Kanafani (1972). However, a method of approximation that could be

found to be useful because of its simplicity is the use of relatively homogeneous market segments (Aldana, 1971; Manheim, 1972). All individual consumers within a market segment are assumed to be the same and the forecast of the micro model is therefore simply multiplied by the number of consumers in the segment. Another possible approximation method is to use averages of the independent variables in the disaggregate model and to estimate the aggregation error using a Taylor series expansion.*

In Urban Transportation Planning (UTP) studies, consumer data is conventionally collected at the disaggregate level. A home interview survey is conducted with a sample of households drawn with a uniform sampling rate throughout the planning region. The data is expanded and aggregated geographically using a traffic zone system which is defined largely on the basis of accessibility to various transport facilities. The models are then estimated using the aggregated data with the traffic zone as the observation unit. Measurements of transport level of service with respect to a traffic zone are made to a zone centroid. This means that all consumers within a zone are assumed to be located at one point in space. This is equivalent to the use of the mean of the category in an aggregation by socio-economic categories.

*Talvitie (1972) suggested this procedure and derived the aggregation error for a logit model.

Notwithstanding the aggregation problem, it is argued that there are large potential benefits from using disaggregate data for the model estimation in UTP studies. Fleet and Robertson (1968) and McCarthy (1969) have shown that in a typical partitioning of an area into geographical analysis zones, the variability of behavior among observations within the zones is greater than the variability between zones. Hence, aggregation prior to model estimation carries with it a significant loss of information. In addition, estimation results of aggregate models will have to be carefully interpreted and used for forecasting due to the risk of ecological fallacy (deNeufville and Stafford, 1971) which arises because correlations among aggregate variables do not necessarily reflect actual behavioral associations on the disaggregate level. Using the data in its original disaggregate form has the advantages of better transferability of the models from one situation to another and flexibility of aggregation for forecasting to any desired level.

However, the most important advantage for actual planning studies appear to be the potential for large savings in data collection costs. In the past, data collection and processing have often consumed as much as half the budget of an UTP study (Wickstrom, 1971). If the data is used in its original disaggregate form, and it is not aggregated to a zonal level for estimation, a smaller sample is required to estimate the models. For example, the number of household interviews in the home interview survey conducted in Metropolitan Washington during 1968

was about 30,000. The most detailed analysis was performed with a system of about 1,000 traffic zones. The potential for cost savings is very large because the cost of, say, 10,000 interviews can be as high as \$500,000. It is not clear how many interviews should be conducted for disaggregate modelling. However, the traditional comprehensive home interview survey is not essential as it is for the aggregate models. The large uniform sample might be replaced by several small samples in selected areas of the planning region. The emphasis should be on attempting to represent the full range of socio-economic characteristics affecting travel behavior rather than on sampling all parts of the region at a uniform rate.

Research is needed to fully answer questions of estimation, data, and forecasting related to the use of disaggregate models. The important questions are the following:

- Can we satisfactorily estimate travel behavior models with disaggregate data?
- How can such disaggregate data be obtained efficiently?
- Can we use these models to adequately predict aggregate trip making?

The limited experience from previous work in disaggregate travel demand models, e.g., Reichman and Stopher (1971), CRA (1972), and from disaggregate residential location and auto ownership models, e.g., Aldana (1971), indicates that the models produced satisfactory esti-

mation results. However, more work is needed before disaggregate travel demand models could be extensively relied upon in actual planning studies.

Since the focus of this study is the structure of travel demand models and not the forecasting of market equilibrium, we are concerned with the point of view of the individual consumer. The theoretical discussion in the remainder of this study is at the level of the single behavioral unit. The models that are developed and estimated in later chapters are disaggregate models. They take the point of view of the individual consumer as he weighs the alternatives and makes his choices.

A Disaggregate Modelling Framework

We now turn to consider the aggregate models that were discussed earlier from the point of view of an individual consumer.

Consumer's market behavior is described by a demand function. A consumer acts in many different markets. We are interested here only in the markets that are affected by the supply of transportation services, or the markets that influence the conditions of the transportation system, namely, the transportation and the activity markets. The consumer behavior in the transportation market is in terms of his usage of transportation services, which on the aggregate level was expressed as a volume of trips V . The actions of the consumer in the activity markets are primarily his locational decisions and other mobility related characteristics, such as auto ownership, that on the

aggregate level were expressed as characteristics of the activity system A.

From the point of view of an individual tripmaker, a trip decision for a specific trip purpose consists of the following set of travel choices: choice of trip frequency, choice of destination, choice of time of day, choice of mode, and choice of route. The actions of an individual consumer in the activity markets include the following set of mobility choices: choice of residence location, choice of housing type, choice of auto ownership level, and choice of employment locations.

Thus, on the disaggregate level the travel demand model should explain consumer's travel choices, and the activity system model should explain consumer's mobility choices. The travel choices of an individual tripmaker are interdependent, e.g., the choice of a mode depends on the choice of a destination, and the choice of the destination may also be dependent on the choice of a mode. Similarly, the mobility choices are interdependent, e.g., the choice of residence location influences the choice of auto ownership level, and the choice of auto ownership level may also condition the choice of residence location. In addition, the two sets of choices, mobility and travel, are also dependent on each other as discussed earlier with respect to the interactions between the transport and the activity markets.

The travel demand function explains the trip making behavior of the consumer as a function of his mobility decisions, among other

factors. It was argued that the consumer makes his mobility decisions in a longer time frame than that of his travel decisions. Therefore, in the travel demand function, the mobility decisions could be assumed to be predetermined. The aggregate structures that were outlined in previous sections, i.e., simultaneous, dynamic, and recursive (static), represent alternative assumptions about the interaction between the mobility and the travel choices. Similar assumptions should also be evaluated with respect to the interactions among choices within the same set, e.g., the nature of the interaction between mode and destination choices, or the interaction between auto ownership level and residence location choices.

A Hierarchy of Choices

There is, apparently, a hierarchy to the mobility and travel choices. For example, the choice of mode on any particular trip is certainly a lower level decision than the choice of residence location. A higher level choice seems to condition lower level ones. That is, once the choice of residence location is firmly made, the choice of mode on a particular trip may be limited.

The hierarchy is essentially a result of the different time frames in which the various choices are made. In a long run static model we generally assume that the consumer has the possibility of changing all his mobility and travel decisions. In a short run static model we assume that the mobility decisions are fixed and only travel choices can vary. However, at any point in time a consumer has

different levels of commitment to the different choices previously made. Therefore, given a change in policy it will take each choice a different time period to respond. For example, in response to a change in the transportation level of service, it can be assumed that in the short run only the mode of travel and route can be changed for a work trip. In the medium run, auto ownership level may be altered. In the long run, job and residence locations may also change.

Thus recognizing the different levels of commitments to the various choices, one may conclude that decisions are made at different points in time. Also, since the time periods it takes the various choices to respond to a change in policy or to a change in other factors that influence the choices, are different, one may conclude that there is a sequence, which is in order of decreasing response time. For example, for a choice of mode it can be assumed that the auto ownership level is given, and for the choice of auto ownership level, the residence location is given.

The hierarchy that is suggested is that the mobility choices are on a higher level than the travel choices. In addition, among the mobility choices one can suggest that auto ownership decision is the lowest level, housing type and residence location the next, and employment location finally the highest. Clearly, different hierarchies could be relevant for different types of consumers.

It should be noted that the work trip place in this hierarchy is different from that of trips for other purposes. The mobility choices

determine all the travel choices for the work trip except for mode, route, and also time of day for employees without fixed working hours. The trip from home to work and the trip from work to home are the most important trips made by the household. The work trip generally has the greatest influence on mobility decisions amongst all trip purposes. This means that a household's evaluation of a residential location is not independent of the work place(s) and the evaluation of a work-place is not independent of the residence location.* Clearly, the evaluation of a residential location depends on many other factors; among them some activities that serve as a trip purpose, e.g., shopping opportunities, medical facilities, schools, etc., and other spatial characteristics such as social groupings and housing supply. The open travel choices with respect to the work trip also have significant effect on other mobility choices and in particular, auto ownership. This points out that although the choice of mode to work may be a lower level choice than the choice of auto ownership level, it is a higher level choice with respect to travel choice for other trip purposes. The choices with respect to trips for non-work purposes, with the possible exclusion of school trips, seem to have a shorter response time than the mobility and the work trip choices.

*This points out that a destination choice model for the work trip that assumes the residential location as given is an employment location choice model and may have to be treated simultaneously with a residence location choice model that takes the employment location as given. For some consumers, e.g., low-income black ghetto residents, the choice of residence location can be assumed to be independent of the work trip. For others, e.g., high-income suburban residents, the two choices cannot reasonably be assumed to be independent.

It is also possible to argue for a hierarchy of choices within the travel choices for non-work trips. For example, one can change more rapidly the choice of mode of travel for a trip to the bank than the choice of destination that will require a switch of a bank account from one bank to another.

Response lags, or threshold level changes, complicate the modeling of mobility and travel choices, but are typical of human behavior. Response lags may be different for different consumers, and different consumers may exhibit different behavioral patterns. For example, for a young, white collar household the residence location and auto ownership decisions can be assumed to follow the job decision. However, for a blue collar worker, married with children, the residential location decision may precede the job decision.

Note, however, that in this dynamic scenario, it is realistic that when a higher level, or longer response time, choice is made, a consumer may consider all his lower level choices and alter them at the same time accordingly. For example, a change in residence location may carry with it a change in all lower level decisions. Thus, when a consumer relocates his residence from a central city location to a fringe location, he is forced to reconsider his automobile ownership, his housing type, his mode choice to work, and his pattern of trip making for all other purposes.

The modelling implication of the hierarchy of choices is that we should model the consumer behavior over time. Ideally, a dynamic

model that "simulates" adjustments over time is desirable. Note that the model for any given choice should allow for adjustments of all lower level choices. The model that will describe the highest level choice will include all choices, similar to a static long run model. However, the changes that are precipitated by changes in socio-economic characteristics, such as income, by changes in the spatial opportunities, or by changes in transportation level of service characteristics, take time and in general affect only gradually higher level choices.

For lower level choices, therefore, there are two types of models, one that describes the short run behavior and another that describes long run behavior when a lower level choice is allowed to adjust simultaneously with higher level choices.

Clearly, a dynamic adjustment model will be very complex. It can be argued that our present knowledge of mobility and travel behavior is not detailed enough for the development of operational dynamic models. The aggregate dynamic structures that were considered earlier in this chapter represent assumptions that may not be acceptable in a disaggregate model. In addition, as a practical matter independent of theoretical feasibility or infeasibility, the development of dynamic models requires time series data which are not currently available, and would probably be very expensive to collect. Data must be obtained for each consumer at each significant point in time. In addition to the socio-economic characteristics of each consumer at each point in time, the data should include the spatial opportunity variables, and the transportation level of service variables that existed.

Given that the true behavioral process consists of dynamic adjustments in the framework of a choice hierarchy as described, the issue is what are the appropriate behavioral assumptions in a static model that is based on cross sectional data.* It is clear that the coefficient estimates from a static model should be interpreted to indicate response over a sufficient time period to allow adjustments.

The question is essentially whether the assumption of a hierarchy of choices based on dynamic considerations justifies a static recursive model in which a higher level choice is made independently of lower level choices, and lower level choices are made with the higher level choices assumed as given. In other words, modelling a specific conditional decision structure such that higher level choices are modelled independently of lower level ones, and lower level choices are conditional on higher level choices. Higher level choices could be based on some measures of expectations from lower level choices, but the lower level choices are indeterminate. For example, we can consider destination choice to precede mode choice in a static recursive structure if we assume that the choice of destination is made independently of the actual choice of mode, and the choice of mode is made with the destination assumed as predetermined. However, an assumption of such recursive structure seems to be unrealistic because even in the dynamic

*Note that longitudinal data on two or more points in time, such as data from a before and after study, could be useful but will require a similar modelling approach to that required by cross sectional data.

scenario it was argued that no higher level choice is made independently of other lower level choices. It is more realistic in a static model to assume, for example, that the choices of mode and destination are made together and the choice of one depends on the choice of the other. Therefore, the only reasonable assumption in a static model is that "trip-making decisions are made simultaneously rather than sequentially", as noted by Kraft and Wohl (1967). This argument applies to all the mobility and travel choices, and it implies a long run model in which all choices are allowed to change simultaneously and all choices are interdependent.

It is possible to approach the issue of the appropriate static structure from a different point of view. Given the true dynamic behavioral process and given that all we can observe is a cross section of consumers at one point in time, what are the true behavioral models that can be estimated from the data, or what models provide the best approximation of the true processes? A static long run equilibrium never really exists because of the response of the time lags of the various choices. Changes may have occurred but the consumers may not have brought themselves to the point of adjusting all their choices because of the response lags. Moreover, different consumers at different locations, or with different socio-economic characteristics are responding to different changes and with different lags. Therefore, we not only observe an out-of-equilibrium situation, but we also observe different consumers at different "distances" from their respective

equilibrium conditions. This means that from a cross section we can estimate only models for choices that have instantaneous, or relatively short, response time.

These models for choices with very short response time should assume higher level choices as given. Thus, from a cross section we can estimate only the model for the short run behavior of the lowest choice in the choice hierarchy. If several choices have similar response time and are in the bottom of the choice hierarchy then their short run models could also be estimated. For choices with relatively longer response time we need to resort to the approximation of a long run static equilibrium assumption. This means that modelling a choice which is not at the bottom of the hierarchy requires a simultaneous structure in which lower level choices can vary and higher level choices are assumed as predetermined.

It is reasonable to assume, as already indicated, that travel choices for non-work trips have relatively shorter response times than the mobility choices. Therefore, these choices should be modelled in a simultaneous structure that assumes the mobility choices as predetermined. The modelling of the mobility choices could be done independent of the travel choices for non-work trips based on the assumption that the mobility decisions depend on the overall pattern of these trips and not on specific choices. Thus, a reasonable static approach will be a block recursive structure in which the first block treats simultaneously the mobility choices independent of specific travel choices. The

second block treats simultaneously the travel choices for non-work trips conditional on the mobility choices.

The modelling on the disaggregate level of alternative structures for the interdependencies among choices is discussed in Chapter IV.

The Travel Demand Function

In the remainder of this study we consider a travel demand function for a specific trip purpose that assumes as predetermined the mobility decisions of the consumer and is independent of other trip purposes.

Since the demand for travel is a derived demand, a trip can be considered as an input into the production of the activity at the destination of the trip. Therefore, trips are classified according to the trip purpose which is the activity at the destination. Modelling directly the demand for trips in each trip purpose category independently requires the assumption of no substitution among trips for different purposes. It is assumed that trip purpose categories could be defined such that the assumption of no substitution is justified. This is a strong assumption but it simplifies considerably the modelling of travel demand.

For a specific trip purpose, the trip making behavior of a consumer is assumed to be a simultaneous decision making process. The consumer is making the following choices: choice of trip frequency, choice of destination, choice of time of day, choice of mode, and choice of route. Even if there exists a hierarchy of travel choices,

when we want to predict all travel choices, we should assume simultaneity. The reason is that we are modelling in the time frame of the highest level travel choice. This assumption could also be intuitively justified from the recognition of travel as an intermediate good. For example, a consumer who wants to shop is choosing the frequency, destination, time-of-day, mode of travel and route that best (e.g., minimum disutility) fit his need.

It should be noted that a sequential decision making process which is modelled in a recursive structure is assumed by the conventional aggregate Urban Transportation Model System (UTMS) (Manheim, 1972) and by the recent disaggregate model developed by CRA (1972). The modelling of both a simultaneous and a sequential structure will be discussed in Chapter IV. Briefly, however, the advantage of a sequence assumption is that it reduces the complexity of the estimation problem. The problem is that even if we accept it, (in spite of the previous discussion), there are generally several conceivable sequences due to the similar response times. And there are no a priori reasons to choose among them. For example, for the shopping trip purpose there are no apparent reasons to select between the sequences of destination choice followed by mode choice and vice versa.

Before we continue to consider the alternative modelling approaches of travel demand, we need to address the issues of what unit does an individual consumer represent? and what are the alternatives from which the consumer is choosing?

The Behavioral Unit

There is some question as to exactly what unit the individual consumer represents. Is it an individual person, a household, or some more narrowly defined expenditure unit? The basic difference is between an individual person and a grouping of individuals such as family, household, or expenditure unit. In what follows we will distinguish between an individual and a household. By a household unit we mean a group of individuals (usually related) that live together and make common economic decisions. This is a general definition that takes into account single individuals living alone, groups of single individuals living together, married couples with or without children, families with grown children (married or unmarried) living at home, etc.

There exists some disagreement about what is the behavioral unit for some travel choices. In developing a disaggregate mode choice model for work trips, de Donnea (1971) assumed the behavioral unit to be the individual. However, he incorporated a household effect by using the variable of household income instead of personal income. It might be reasonable to consider the individual commuter's mode choice as an individual decision, but a shopping trip for household goods should clearly be considered as a household decision. Oi and Shuldiner (1962) argue that the household is the pertinent behavioral unit. Although some trips are seemingly initiated by individuals without regard to the rest of the household, even these trips are influenced by characteristics of the household as a whole. de Donnea (1971) labelled these influences as "family constraints".

Some trips may be entirely unrelated to the household, particularly on-the-job or business trips. However, for most trips we may assume either the household or the individual as the behavioral unit depending on the type of trip being considered.

For the mobility choices such as residence location, housing type, and auto ownership, the behavioral unit is clearly the household. For example, the disaggregate model developed by Aldana (1971) for residence location and car ownership assumes the household as the behavioral unit.

The Alternatives

Each choice is made from a set of alternatives, or a choice field. Before we can consider the issue of defining the choice field we need to answer the following fundamental question:

Is there truly a choice, or is travel and mobility behavior completely dictated by constraints?

Considering the travel choices, the answer to this question will depend on the type of trip. For some trips there may be no choice at all and for others the number of alternatives that the traveller can choose from can be very large.

In mode choice analysis some studies have distinguished between "captive transit riders", and "non-captive", or people who have the choice of a private car. Hence, the answer will depend also on the characteristics of the travellers.

It is possible to describe a scenario in which a certain type of traveller for a certain trip purpose has no choice at all; for example, a worker who is assumed not to have an alternative job at the present time, and who cannot use a car, and is required to be punctual.

However, a no-choice scenario will in many cases be of a short run nature. By short run it is meant that we observe a traveller at a certain point in time and consider his possibilities in the very near future. Yet the limited possibilities that we may observe can be a result of long range conscious decisions. For example, an individual may be assumed not to have the choice of using a car because he does not possess a car. However, given that it is economically feasible for him to own a car, the fact that he does not have one available at present may have been the result of the realization that it is more advantageous for him to use public transportation. But then, what about those for whom a car is economically infeasible or who are prevented from driving for other non-economic reasons?

It is important to regard travel decisions as long term decisions. This clearly reduces the number of constraints and opens choices. Still, for a specific trip purpose the choices to some traveller with respect to one or more aspects of travel may be very limited.

In modelling the demand for travel there are two alternative ways of treating travellers with limited choices. One alternative is

to predict outside the demand model the choices available to different groups of people. Then, to apply a demand model that takes into account different choices for different groups. The other alternative is to internalize the determination of the available choices in the demand model. In other words, the demand model is constructed as if the same set of choices is available to everyone, but the variables that describe a traveller in the model will cause the unavailable choices to be infeasible.

For example, consider again a mode choice model. The first alternative is to exclude the alternative of driving a car for all travellers who cannot drive a car due to economic and other reasons. The second alternative is to consider all alternatives but to introduce into the model variables like income and age such that the demand model itself will show the car driving alternative to be infeasible for those who cannot own or drive a car because of their income or age.

The issue is essentially whether car availability is determined within or outside the travel demand function. In any case, some individuals due to health, age, or economic reasons may not have the alternative of using a car. If the only other alternative is a bus, then the choice of mode of these individuals is dictated by constraints, in the long run as well as in the short run. If we include such individuals in a mode choice model then the model does not represent only an actual choice but also a determination of what alternatives are available.

As mentioned already the set of available choices will be different for different trip purposes. The difference is primarily between work and non-work trips. It was assumed that the travel demand model consider the mobility decisions of the consumer as given. The previous discussion of the work trip implies that the choice of destination in a travel demand function for the work trip is completely constrained. It can be argued that the choice of destination for a school trip is similarly constrained.

School and work trips are also constrained with respect to the choices of trip frequency and time-of-day. Except for sick days and vacations an employee, or a student, will go to his workplace, or his school, daily. The choice of departure time may not be completely constrained since the traveller may choose to arrive earlier than required.

Apparently, even if we accept all the constraints that were discussed, some choices are still open in the travel demand function for the work trip, i.e., choice of mode and choice of route.

The source of the above discussion is the separation of the consumer actions into two markets: mobility and travel. The question of the open choices in the travel demand function is essentially dependent on what decisions we assign to what market. Hence we can conclude that there is truly a choice but not all choices are determined by the travel demand function as we have defined it here.

The question that remains is, therefore, how to determine the set of alternatives. Consider, for example, a shopping trip which is

undoubtedly unconstrained with respect to all travel choices. For a shopping trip consider the choice of destination. The question is how to determine the set of destinations that the consumer is choosing from. One possibility is to treat all possible shopping destinations in the entire urban area as the set of relevant choices. The gravity or the opportunity distribution models used in the UTMS (Martin et al., 1961) implicitly assume that the alternative destinations are the largest possible set. Yet, it is likely that the actual choice is made out of only a subset of the above set. The problem is to determine this subset. Any determination of such subset involves an *a priori* arbitrary criterion. For example, we could define the relevant subset to include only shopping destinations within half an hour travel time from the origin of the trip, but this will perhaps eliminate some important specialized shopping trips.

As before when we discussed the different choices available to different travellers, we also have here the possibility of internalizing this problem in the demand model. The characterization of a destination in the model should be such that the prediction of travel to destinations outside the relevant subset will not be possible.

In probabilistic models it is important to include all alternatives that have positive probability of choice. It should not matter whether or not an alternative with zero probability is included in the assumed set of choices. Hence, the *a priori* criterion to include or exclude an alternative will be whether or not we have reason to believe

that the alternative has positive or zero choice probability. Again, if an alternative has zero probability and it is included, the model should be able to predict zero probability of choice for this alternative.

Then why should we want to reduce the larger set into the relevant subset? For at least one reason - it can reduce considerably the complexity of the model and the data requirements. Actually, every existing model explicitly or implicitly makes some a priori assumption that determines the relevant subset of alternatives. For example, mode choice models estimated in U.S. urban areas have not considered a bike as a relevant mode, due to its rare use. However, it may be a viable alternative in other circumstances, e.g., Holland, or Cambridge, Mass. It is justified to exclude the bike as an alternative mode when it is reasonable to expect its probability of use to be nearly zero. Hence the determination of the universe of choices is very important but could be very delicate because it is based on a priori assumptions and observations of current behavior which may not hold in the future.

In conclusion, we consider a travel demand function, that assumes the location decisions of the individual or household as predetermined, that is for a specific trip purpose and is independent of other trip purposes, and that is for a given set of trip alternatives.

Supply Effects with Disaggregate Demand Models

It was argued earlier with respect to the aggregate system of models, that in a static framework the travel demand model is simul-

taneous with the relationships of travel supply and locations decisions. Therefore, the transportation level of service variables should be treated as endogenous variables, and travel demand models should be estimated simultaneously with the transport supply relationships.

However, this argument applies primarily to aggregate models. The supply function is an aggregate relationship and if we estimate the demand function for individual consumers, it can be argued that the transportation level of service variables could be assumed to be exogenous. For the individual consumer, the level of service variables can be assumed to be fixed, since the acts of an individual traveller will not significantly affect their values. Thus, in a disaggregate model these variables could be considered as exogenous. This justifies estimation of the disaggregate travel demand function as a single relationship, without the simultaneous estimation of supply functions.*

Note, however, from the discussion in previous sections, that there are other simultaneous relationships among travel and mobility choices that do play a significant role on the aggregate as well as the disaggregate levels. These interactions among choices will be treated in Chapter IV.

*At least it is more justified for disaggregate models than for aggregate models

CHAPTER III

Consumer and Choice Theories

Overview

Chapter II has laid the overall background for travel demand modelling. The purpose of this chapter is to survey some principles of economic consumer theory and probabilistic choice theory that could be used to formulate demand models. It is concluded that a choice approach is more appropriate than economic demand analysis for the modelling of a disaggregate travel demand function. In particular, choice theories are suitable for the modelling of mobility and travel choices that are qualitative in the sense that the choice on the disaggregate level cannot be expressed as a continuous quantity, e.g., choice of residence location, choice of mode, etc.

Consumer Theory

In general, any travel demand model can be expressed as:

$$V = D(L, A) \quad (1)$$

where volumes of trips V are determined as a function of characteristics of the activity system A , and level of service characteristics L . This expression is similar to demand functions in the micro-economic literature*, where the quantity of a product demanded (V) is a function of its price (L), the prices of related commodities

*See for example Henderson and Quandt (1958)

(substitutes and complements) (L), and income (A). It will be useful to explore briefly the principles of economic consumer theory* and its implications for travel demand models.

An individual consumer is assumed to have preferences over alternative consumption bundles, he is faced with a budget that defines his consumption possibilities, and he is assumed to demand the bundle that maximizes his preferences subject to his budget. The preferences of a "rational" consumer are represented by a utility function. Rational behavior is defined in the sense of a transitive ordering of alternatives (i.e., if X is preferred to Y and Y preferred to Z, then X is preferred to Z), and a capability of comparing all possible alternatives. With a continuity assumption, a mathematical utility function represents an ordering of alternatives. In this general case, the utility function is called ordinal and it is unique only up to an order-preserving transformation.**

The consumer utility function is expressed as follows:

$$U = U(X_1, X_2, \dots, X_N) \quad (2)$$

*Many references are available at different levels of presentation. See for example Henderson and Quandt (1958), Chapter 2, or for a more advanced presentation see McFadden and Winter (1970).

**A distinction is made between ordinal and cardinal utility. Ordinal utility is just a mathematical expression of ranking. Cardinal utility implies a uniqueness of numerical assignment up to positive linear transformation. The assumptions of ordinal utility are less restrictive but sufficient for the development of consumer theory. When the concept of expected utility is employed, a cardinal measure of utility is necessary. (Luce and Suppes, 1965)

where X_1, X_2, \dots, X_N are the quantities of each of the commodities 1 to N. The consumer is assumed to have a fixed income Y and to face fixed prices P_1, P_2, \dots, P_N . Hence, his budget constraint is:

$$P_1 X_1 + P_2 X_2 + \dots + P_N X_N \leq Y \quad (3)$$

The maximization of (2) subject to (3) with respect to X_1, X_2, \dots, X_N , results in a set of demand functions as follows:

$$X_1 = D^1 (P_1, P_2, \dots, P_N, Y)$$

$$X_2 = D^2 (P_1, P_2, \dots, P_N, Y)$$

(4)

.

.

.

.

$$X_N = D^N (P_1, P_2, \dots, P_N, Y)$$

Consumer theory deals with a single consumer. His tastes, or the factors that shape the consumer's preferences, are explicitly contained in the given form of the utility function (2). In empirical work, when observations of different consumers are used to estimate demand functions, we need to specify how tastes and consequently utility functions vary from one consumer to another. This is the reason why socio-economic characteristics are introduced into utility and demand functions, or why demand functions are sometimes separately estimated for different socio-economic groups.

Consumer theory is developed without any assumptions about the nature of the alternatives from which the consumer is making a choice.

As a result, the demand functions that are derived, equation (4), have very limited empirical content. What (4) says is that we should include the prices of all commodities. However, this is never done; instead, empirical demand functions include only prices of related commodities. The problem is how to define related commodities.

Strictly speaking, the demand for any commodity is affected by the price of all commodities that get a part from the consumer's budget. In principle therefore, all commodities are related. However, it is possible to hypothesize that the relationships between some commodities reflect only income effects and not substitution effects.* This imposes some restrictions on the utility function. Strotz proposed to formulate the utility function as a "utility tree" (Strotz, 1957 and 1959). The utility function is additively separable in several utility functions, each defined on a group of commodities, or a branch. Hence, the utility function can be written as follows:

$$U(X_1, X_2, \dots, X_N) = \quad (5)$$

$$F[U^1(X_1^1, X_2^1, \dots, X_{N_1}^1) + U^2(X_1^2, X_2^2, \dots, X_{N_2}^2) + \\ + \dots + U^S(X_1^S, X_2^S, \dots, X_{N_S}^S)]$$

and

$$N_1 + N_2 + \dots + N_S = N$$

*The total effect of price change on demand is the sum of the substitution and income effects. The substitution effect is the change in demand with utility held constant. The income effect is the change in demand with prices held constant (Henderson and Quandt, 1958).

where F is a monotone-increasing function and U^1, U^2, \dots, U^S are functions of different groups of commodities.

The consumer is visualized to follow a two stages procedure: "first, to decide how much expenditure to allocate to each branch, and then, secondly, to decide how best to spend each allocation on commodities within each branch" (Strotz, 1957). The branches correspond to groups of commodities such as: food, clothing, vacation, housing, transportation, etc.

The problems are: how to define "natural" branches that would be the same for all individuals, and how to define composite prices, such as the "price of food", or the "price of transportation", on the basis of which the consumer makes his interbranch allocation (i.e., the first stage).

The concept of utility tree, or separable utility function, suggests one alternative to the definition of related commodities. Another hypothesis, suggested by Muth, is to view the commodities purchased by the consumer as inputs into a production process of final goods (Muth, 1966). The utility function is defined over the final goods. The final goods are defined in the same manner as the branches of the utility tree. For each final good the consumer possesses a production function in terms of actual commodities. Substituting these production functions into the utility function we get

again a separable utility function.* The marginal rate of substitution in consumption between two related market goods is the marginal rate of substitution in the production of a final good, and consequently independent of all commodities not used in the production of that final good.

Muth's hypothesis corresponds directly with the observation that trips are not consumed for their own sake but rather to accomplish an activity at the trip destination. (See the section "The Demand For Travel as A Derived Demand" in Chapter II.) Thus, the utility function is defined in terms of the final activities and for each final activity the consumer possesses a production function in which trips are the main factors.

The new approach to consumer theory suggested by Lancaster (1966) is essentially a more fundamental level of the hypothesis suggested by Muth. Lancaster defined the utility function not over the final goods, but over attributes of the goods. Hence, the output from a production process is an attribute of the final good, instead of the final good itself. The most important advantage of Lancaster's approach is that it can accommodate with relative ease the introduction

*Muth and Strotz get similar results. Strotz assumes an additively separable utility function. Muth assumes weakly separable utility function, but in addition assumes the production functions to be homogenous of degree one. The notion of separability was introduced by Leontief (1947). The differences between additively separable, or strongly separable, and weakly separable are explained by Goldman and Uzawa (1964).

of a new commodity. With the conventional utility function, equation (2), this involves the undefined transformation of the function from N commodities to $N + 1$ commodities.

Thus, the assumption of a separable utility function justifies a demand function with only the prices of commodities in the same branch. The income variable is the income allocated to this branch, which is assumed to be predetermined. The problem is that a model is needed to explain the allocation of income among branches. This function, however, depends on the prices of all commodities and total income.

The theoretical demand functions can be further simplified by assuming that the proportion of income spent on a particular branch of related commodities is relatively small. Thus, the interbranch income effect on the commodities in this particular branch is assumed negligible. Recall the discussion of the utility tree. A demand function for a particular commodity that describes the consumer decisions in the second stage, is a function of the prices of all the commodities in its branch and the income allocated to this branch. The income allocated to a branch is in turn a function of total income and the composite prices of all branches, as in the first stage. However, if the proportion of income allocated to a particular branch is very small, we can assume that the effects of the composite prices of other branches on this allocated income is negligible. Therefore, we can write demand functions for commodities in this branch which are completely independent of prices of commodities in other branches. In essence, we

assume that the income allocated to this branch is a fixed proportion of the total income.

This assumption will justify the modelling of travel demand independently of all other goods and services. The same exact considerations also apply to the separation of the travel demand function into independent functions for different trip purposes. As noted earlier, this is somewhat a stronger assumption, but it can also be justified based on the broad definitions of trip purpose categories (and hence separable utility), and the small portion of income spent on a trip for a given purpose (and hence negligible income effects).

A stronger assumption about the unimportance of income effects is to assume that if the proportion of income spent on a particular group of commodities is very small, the consumer makes his decisions on this group of commodities independent of a budget constraint. In other words, the consumer maximizes a net utility function without a constraint. However, this involves a different definition of the utility function from the one used here. Essentially it requires the introduction of the expenditures into the utility function. This approach was used by Golob and Beckmann (1971), in deriving travel demand models from different utility functions.

Another assumption particularly relevant to travel demand models, will be explored in more detail in the following section. Briefly, it applies to a case of a group of commodities that are assumed to be perfect substitutes, and a given achievable level of utility. The

consumer is then selecting the commodity from this group that minimizes his expenditures. Then, if he decides to obtain any commodities from this group it will be only from this selected commodity. For example, different trips to a given destination for a specific trip purpose are assumed to be perfect substitutes. The traveller is choosing the mode and route that has the least cost in terms of money, travel time, etc. (McGillivray, 1972).

Another approach is to assume a selection among a group of perfect substitute commodities that results, however, in different levels of utility. In this case, the utility again will be a function of only one commodity. Therefore, the consumer will obtain, if at all, only from that commodity that for a given expenditure level and price, maximizes his utility. An example could be alternative destinations for a specific trip purpose.

These examples deal with a finite set of discrete choices which is considered in more detail in the next section, vs. the continuous commodity quantities and consequently the infinite set of alternative commodity bundles that were assumed in this section.

These examples indicate the weaknesses of consumer theory in its application to travel demand. One is the fact that on the individual level, travel choices represent a finite set of alternatives. Another is the conventional assumption in utility function analysis that the solution is at an "interior" point, or a point where there is a "tangency" between the budget line and the maximum utility surface. In the case of discrete travel alternatives which are perfect substi-

tutes we have a "corner" solution, or a point where the usual first order conditions for utility maximization (i.e., price ratio equal, the marginal rate of substitution) do not hold (Henderson and Quandt, 1958). For these reasons, probabilistic choice theories that consider a finite set of mutually exclusive alternatives from which one element must be chosen, can be more readily adapted to formulate travel demand models.

To end this discussion of consumer theory it will be useful to return to the travel demand function of equation (1) and relate it to the demand functions that were discussed in this section. In the first place, a demand function derived from consumer theory is for a single consumer, while a demand function that is used to determine equilibrium in markets, as equation (1), represents the behavior of a group of consumers. Thus, equation (1) represents an aggregate demand function which represents the sum of individual consumers' demand functions. The set of variables denoted by A for activity system characteristics consists of income and socio-economic variables, as discussed in this section, and aggregation variables, such as population or employment, that indicate the number of consumers whose demand functions were aggregated. Aggregate demand functions are often formulated on a per capita basis so that it represents a sort of average demand function for a group of consumers.

The set of variables A includes also some variables that are specific to travel demand models (disaggregate and aggregate). Namely,

attraction variables which describe the spatial opportunities to achieve a specific trip purpose activity at various destinations. Fundamentally, these variables are proxies for the prices of performing the activity which is the purpose of the trip. These prices could be positive to represent an outlay or negative to represent a reward. These prices are different in different locations and, since we cannot measure them, we replace them by variables that describe the locations in terms of opportunities for a specific trip purpose.

Probabilistic Choice Theory

The development of probabilistic choice theories has been in the field of psychology.* It is attributed to the heavy reliance on experimentation in the development of psychological theories (Luce and Suppes, 1965). The introduction of a probability choice mechanism was required to explain behavioral inconsistencies that were observed.** Essentially, it amounts to an admission of lack of knowledge, because if it was possible to specify the causes of these inconsistencies, then a deterministic model could have been used. The causes could be completely unknown, or known but unmeasurable.

It should be noted that consumer theory (discussed in the previous section) is a deterministic theory. However, in applications, when

*Basic references for this subject are Luce and Suppes (1965), Bock and Jones (1968), Luce (1959), and CRA (1972).

**Alternatively, probabilistic choice theories can be viewed as a result of logical arguments that human choice is essentially probabilistic.

the parameters of a demand model are estimated by a regression technique, the existence of a random unobserved disturbance is assumed to account for all the factors that were not included in the model.

Recall from the previous section the examples of choice among perfect substitutes. In a deterministic model, the consumer is assumed to choose only one alternative (except for the case of indifference). However, we may observe two kinds of inconsistencies. The first, which is what was observed in psychological experiments, is to observe an individual not always selecting the same alternative in repetitions of the same choice, under the same conditions. The second kind of inconsistency is relevant when we do not follow a single individual, but rather observe once the choices made by each of several individuals. In this case, we explain the differences among individuals by using income and other socio-economic variables. An inconsistency here will be when two different individuals who are identical, in terms of the variables describing them to the model, are observed to select different alternatives.

It is the latter kind of inconsistencies which require the use of probabilistic choice models to explain the behavior of a cross-section of travellers. An example would be two individuals, whose socio-economic characteristics are nominally identical, who take two different modes for exactly the same trip. When we are dealing with an aggregate group of travellers, we do not expect a unique choice, but rather a split among different alternatives. A random unobserved

disturbance is assumed to account for different splits among groups with identical characteristics. Therefore, in travel demand models that were developed with an aggregate observation unit, the inconsistencies were not recognized. As a result, the deterministic consumer theory was sufficient to formulate the model. However, if we want to model the choice processes inherent in travel demand on the disaggregate level, we need to apply a probabilistic choice theory.

Suppose that there is a finite set of possible alternatives that we denote as A, and suppose that an element must be chosen from A. If i is an element in A we denote the probability of choosing i from the set A as:

$$P(i:A) \quad (6)$$

Clearly, we require that:

$$0 < P(i:A) < 1 \quad (7)$$

and

$$\sum_{i \in A} P(i:A) = 1$$

The problem of the choice theory is to relate the individual preferences to these probabilities. It should be noted that the usual theorems from probability theory are assumed to hold. In particular, since only one alternative is chosen the following holds.

$$P(i \cap i':A) = 0 \quad (8)$$

and

$$P(i \cup i':A) = P(i:A) + P(i':A)$$

where i and i' are two different alternatives. In general, if B is a subset of A ($B \subseteq A$), we can write:

$$P(B:A) = \sum_{i \in B} P(i:A) \quad (9)$$

where $P(B:A)$ is the probability that the choice lies in the subset B . Clearly, equation (8) applies also to two non-overlapping subsets, i.e., $P(B \cap B':A) = 0$.

As in consumer theory the notion of utility is used to express the consumer preferences. However, the utility function here is not defined over a space of N commodities, or attributes. Since only one alternative is chosen, a utility measure is identified for each alternative. The utility function is therefore defined as a function of the attributes of a single alternative.* When we consider different consumers we define the utility function also as a function of socio-economic variables to account for different tastes.

Luce and Suppes (1965) distinguish between two approaches to relate the choice probabilities and the utility measures. In one approach, the utility function is a fixed function over the set of alternatives, and the choice probabilities are determined as a specific assumed function of the utility measures. Therefore, Luce and

*It is assumed that the utility from a given alternative is only a function of its own attributes because it makes sense that the final utility is not a function of attributes of non-chosen alternatives. However, this assumption is not very restrictive. If it is assumed that the choice environment (i.e., what else is available?) influences the final utility, then it implies the existence of attributes for each alternative that express the attributes of other available alternatives.

Suppes label models resulting from this approach as "constant utility models". In the other approach that results in "random utility models", the utility measures are assumed to be random variables. However, given the utilities the choice is deterministic as in consumer theory. Thus, the probabilistic behavior arises either from a probabilistic choice behavior as in the first approach, or from the randomness of the utility function as in the second approach.

Constant Utility Models

Two concepts which result in the same model are of interest here. The first is the "strict utility model" which is defined as follows (Luce and Suppes, 1965):

$$P(i:A) = \frac{v_i}{\sum_{j \in A} v_j} \quad (10)$$

where the v_i 's are positive values that determine a utility measure which is unique up to a multiplication by a positive constant, or a ratio scale.*

The second concept which results in the same model, is the "choice axiom" which was first suggested by Luce (1959). A revised definition of the choice axiom appears in Luce and Suppes (1965) and it is written as follows:

*If we substitute $v_i = e^{w_i}$ in equation (10), we get the multinomial logit model, where the w_i 's form an interval scale, unique up to an addition or subtraction of a constant. Both scales are specific cases of cardinal utility.

- A set of choice probabilities defined for all the subsets of a finite set A satisfy the choice axiom provided that for all i, B, and A such that $i \in B \subseteq A$

$$P(i|B) = P(i|B:A) \quad (11)$$

whenever the conditional probability exists. The conditional probability $P(i|B:A)$ is the probability that i will be chosen, given that the choice (from A) lies in the subset B.

In other words, if some alternatives are removed from the set of alternative choices, the relative probabilities of choice among the remaining alternatives are unchanged. The choice from a subset is dependent only on the alternatives included in this subset and is independent of what other alternatives exist. (Note that dependency here is meant in the probability sense which is not the same as functional dependency.)

The qualification that the conditional probability exists means essentially that:

$$P(B:A) \neq 0 \quad (12)$$

This is because the conditional probability of equation (11), by the definition of conditional probabilities, can be written as follows:

$$\begin{aligned} P(i|B:A) &= \frac{P(i \cap B:A)}{P(B:A)} \\ &= \frac{P(i:A)}{P(B:A)} \end{aligned} \quad (13)$$

Equation (12) together with equation (9) imply that at least one alternative in the subset B has a positive choice probability.

Equations (11) and (13) imply that:

$$P(i:A) = P(i:B) \cdot P(B:A) \quad (14)$$

which means that the choice probability of i from the set A can be written as a product of the choice probability of i from a subset B and the probability that the choice lies in B. Hence the product on the right hand side is independent of the definition of B.

Luce proceeds to prove that if the choice axiom holds, and given a utility measure directly proportional to the choice probability, then there exists a ratio scale such as in equation (10). Clearly, equation (10) applies also to any subset of A. Thus, (10) can either be assumed, or can be derived as a consequence of the choice axiom.

Luce (1959) argues that the choice axiom can be viewed as a probabilistic version of both concepts of transitivity and independence from irrelevant alternatives, from decision theory. The assumption of independence from irrelevant alternatives means that the decision maker's preference between two alternatives is independent of the existence of other alternatives. (Hence if X is preferred to Y in the presence of Z it is also preferred to Y in the absence of Z). The probabilistic analog is the result of the choice axiom that the ratio of the probabilities of choice of two alternatives is fixed and does not vary by changing the set of alternatives.

The empirical validity of the choice axiom depends on the specification of the structure of the set of alternatives. Consider the fre-

quently used example of red and blue buses as two of the available travel modes (Mayberry, 1970).* Suppose that a traveller who is making a choice between a private car and the available bus service is just indifferent between the two. Hence, his choice probabilities are $P_{car} = 1/2$ and $P_{bus} = 1/2$. Now suppose that another bus service is introduced which is equal in all aspects to the old bus service except that its buses are painted differently. Since the ratio of choice probabilities is assumed not to vary, the new choice probabilities will be: $P_{car} = 1/3$, $P_{bus_1} = 1/3$, and $P_{bus_2} = 1/3$. This result is unrealistic because the traveller will treat the two bus modes as a single mode and only if car is rejected will he consider a choice among them.

(This result can also be explained by the properties of the elasticities in the strict utility specification. This will be discussed in Chapter VI when we derive the elasticities of the multinomial logit model. Briefly, cross elasticities are uniform, such as in a choice among three modes -- say car, bus and rapid transit -- the cross elasticity of the car and bus choice probabilities with respect to any attribute of rapid transit are equal.)

The problem is with how the alternatives are defined. All this argues that an additional assumption is required, indicating "hierarchical structure" of the alternatives, i.e., how they are

*Luce and Suppes (1965) present an equivalent example of a boy who must select between a pony and two nearly identical bicycles.

perceived. For this example, we can represent the perceived structure using the "choice tree" in Figure 1.

This example considered two identical alternatives, but the fundamental problem of the choice axiom and the strict utility model is the consequences of similar or related alternatives which need not be identical. Bock and Jones (1968) noted that "it is doubtful that the principle is strictly true in practice, since personal preferences for given objects usually extend to similar objects". However, they conclude from their experiments that the model "appears to be well enough approximated in many applications to allow reasonably accurate predictions of choice". The best asset of the strict utility model is its simplicity and ease of application relative to the alternative models.* Therefore, there is a tradeoff between the accuracy of alternative models and the ease of application of the strict utility model.

The Set of Alternatives

The problem here can also be traced to the development of a theory without any assumptions about the special character of the alternatives. In essence, the choice axiom could be expected to hold when the set of alternatives is relatively "homogenous". Thus, in order to apply the choice axiom it is necessary to properly define the alternatives. In the case of red bus and blue bus this redefinition was simple. But what about the choice among car, bus and street car? Are bus and street car two independent alternatives?

All this is not a criticism of only the choice axiom or the strict utility model, but of all theories. It amounts to an observation that

*For a discussion of alternative choice models see Bock and Jones (1968), Luce and Suppes (1965), and CRA (1972).

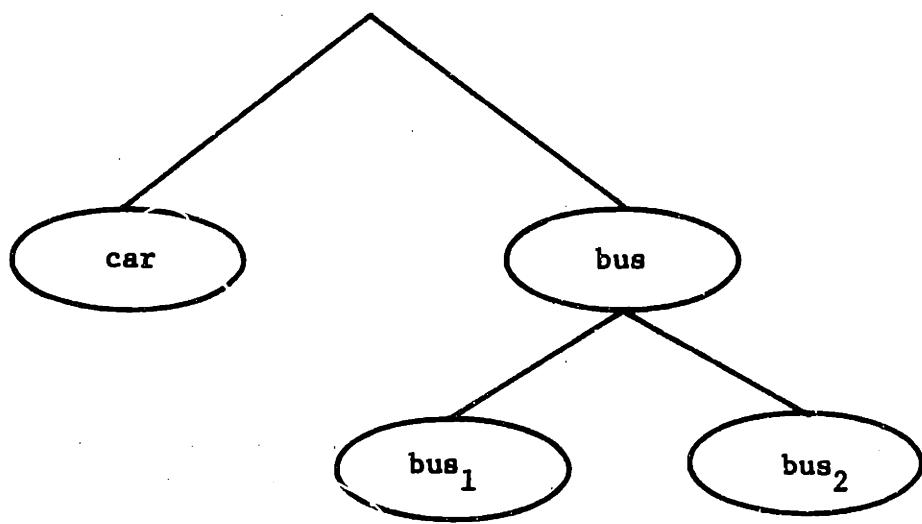


Figure 1

we actually do not know what set of alternatives it is from which the consumer is choosing. This problem is more serious for complex decisions, of which the trip choices are a relevant example. A common practice is to decompose a complex decision into two or more stages, by subdividing the complex decision into component choices. In other words, a complex decision is replaced by a conditional decision structure. Recall equation (14); we can define such subsets B that the consumer is assumed to first select a subset and then, conditional on the selection of the subset, choose an alternative within this subset. Luce (1959) noted that "it is commonly accepted, and it is probably true, that when such a multi-stage process is needed, the over-all result depends significantly upon which intermediate partitionings are employed". The problem is therefore to know whether the consumer does indeed decompose a certain decision and, if so, what are the partitionings that he considers. In some cases, it may appear as if there are "natural" subdivisions; however, it is not clear whether this is the case in the trip decision. We will return to this point when we discuss alternative structures of the travel demand model in Chapter IV.

An important point that is brought out by Luce (1959) is the fact that if we assume a certain choice mechanism for the choice within a subset, it does not imply that the same choice mechanism applies for the choice from the whole set. Therefore, in a conditional decision structure, given the assumption of the choice axiom for the choice

within a subset, it does not imply that the choice axiom applies for the whole set of alternative. On the other hand, in the discussion of the choice axiom, it was shown that if the axiom applies to a set of alternatives, it also applies to any subset thereof.

Random Utility Models

Random utility models are more in line with consumer theory than constant utility models. The consumer is visualized as selecting the alternative that maximizes his utility. The probabilistic behavior mechanism is a result of the assumption that the utility which a consumer will place on an alternative is not certain, but is rather a random variable.

Denote the utility of alternative i as U_i . The choice probability of alternative i is therefore:

$$P(i:A) = \text{Prob } [U_i \geq U_j, \forall j \in A] \quad (15)$$

The deterministic equivalent of equation (15) is simply a comparison of all alternatives available and the selection of the alternative with the highest utility. It should be noted that this maximization process is not subject to an income constraint. (Nor did we consider income constraint in the discussion of constant utility models). The source of this omission is basically the application of these theories to sets of alternatives that are all feasible. In terms of the application of these theories to travel demand, it also does not introduce any problem as the discussion at the end of the section of consumer

theory would indicate.* The assumption that the consumer is minimizing a certain disutility, or expenditures, function that is often assumed in travel demand is equivalent to changing the direction of the greater than sign in equation (15).

Equation (15) can also be written as follows:

$$P(i:A) = \int_{-\infty}^{\infty} \text{Prob}[U_i = t, U_j \leq t, \forall j \in A] dt \quad (16)$$

If we assume that the utilities of the alternatives are independent random variables, the model is then called an independent random utility model. In this case equation (16) can be written as:

$$P(i:A) = \int_{-\infty}^{\infty} \text{Prob}[U_i = t] \prod_{j \in A} \text{Prob}[U_j \leq t] dt \quad (17)$$

In order to obtain a specific functional form for the choice probabilities, we have to assume a specific joint distribution for the utilities and then solve the integral. Luce and Suppes (1965) and also Beilner and Jacobs (1971) show a family of independent distributions of the utilities that if substituted in equation (17) result in a strict utility model as in equation (10). An example would be the following independent distribution:

*The set of alternatives A includes only alternatives that are within the budget set. The utilities of the various alternatives are essentially indirect utility functions which are defined as the maximum level of utility for given prices and income (McFadden and Winter, 1970). It is also possible to include in A all alternatives and require the choice model to predict zero (or very small) choice probabilities for infeasible alternatives.

$$\text{Prob}[U_i \leq t] = \begin{cases} 1 - e^{-v_i \cdot t^n}, & \text{if } t \geq 0 \\ 0, & \text{if } t < 0 \end{cases} \quad (18)$$

where n is any positive constant, and the v_i 's will be the utility scales in the strict utility model.

Another way to introduce the randomness into the utility measures is to assume that the utility consists of two parts: systematic and random. Thus we can write the utility as:

$$\begin{aligned} U_i &= v_i \cdot e_i, \\ \text{or} \\ U_i &= v_i + e_i \end{aligned} \quad (19)$$

Equation (19) is similar to the usual regression models where an additive disturbance term is used to account for unobserved variables that are not included in the model. However, in the case here the distribution assumptions of the disturbance terms will determine the functional form of the choice probabilities.

Substituting the additive disturbance formulation from equation (19) in equation (15) we get:

$$P(i:A) = \text{Prob} [v_i + e_i \geq v_j + e_j, \forall j \in A] \quad (20)$$

or,

$$P(i:A) = \text{Prob} [e_j - e_i \leq v_i - v_j, \forall j \in A] \quad (21)$$

Similar to equation (16) we can write:

$$P(i:A) = \int_{-\infty}^{\infty} \text{Prob}[e_i = t, e_j \leq t + v_i - v_j, \forall j \in A] dt \quad (22)$$

Define:

$$w_j = t + v_i - v_j \quad (23)$$

Substituting in equation (22) it becomes:

$$P(i:A) = \int_{-\infty}^{\infty} \text{Prob}[e_i = w_i, e_j \leq w_j, j \in A] dt \quad (24)$$

Again, if we assume the disturbances to be distributed independently across alternatives, we get an independent random utility model as follows:

$$P(i:A) = \int_{-\infty}^{\infty} \text{Prob}[e_i = w_i] \cdot \prod_{j \in A} \text{Prob}[e_j \leq w_j] dt \quad (25)$$

Assume that the disturbances are independently distributed with the reciprocal exponential distribution as follows:

$$\text{Prob}[e_j \leq w_j] = e^{-nw_j} \quad (26)$$

where n is any positive constant. This distribution results in the multinomial logit model:

$$P(i:A) = \frac{e^{v_i}}{\sum_{j \in A} e^{v_j}} \quad (27)$$

The logit model is essentially equivalent to the strict utility model by defining the exponent of the utility as the ratio scale in the strict utility model. We can get directly the strict utility model by assuming multiplicative disturbance and defining:

$$w_j = t \cdot \frac{v_i}{v_j} \quad (28)$$

and assuming the distribution:

$$\text{Prob}[e_j \leq w_j] = e^{-1/t} \quad (t \geq 0) \quad (29)$$

Thus, the strict utility model is an independent random utility model. However, it is only a very special case of the family of possible independent random utility models. Any joint distribution of the utilities, or the disturbances, will result in a random utility model. The strict utility or the logit models are a result of specific distribution function assumptions. As was pointed out by CRA (1972), the models that result from other distributions are very complicated and virtually impractical. They mention several examples including the multiple probit model which results from a multivariate normal distribution. A binary version of the probit model was used by several researchers in a mode choice model (Lisco, 1967; Lave, 1969).

The empirical violations of the choice axiom or the lack of realism of the strict utility model can be traced to the assumption of independent random utilities. This was noted by Bock and Jones (1968) and it is evident from the observation that the strict utility model is only a special case which generalizes to the independent random utility models. The relationships among the different models is shown in Figure 2. If two alternatives are similar or close in some respect, then we expect their utilities to be correlated. If two alternatives are essentially identical, e.g., blue and red buses, then we expect the utilities not only to be correlated but equal.

Again, we should mention the tradeoff between accuracy and cost. A multiple probit model that assumes independence is more expensive to apply than the logit, but as we have seen above, since it is also an independent random utility model, it does not constitute a substantial

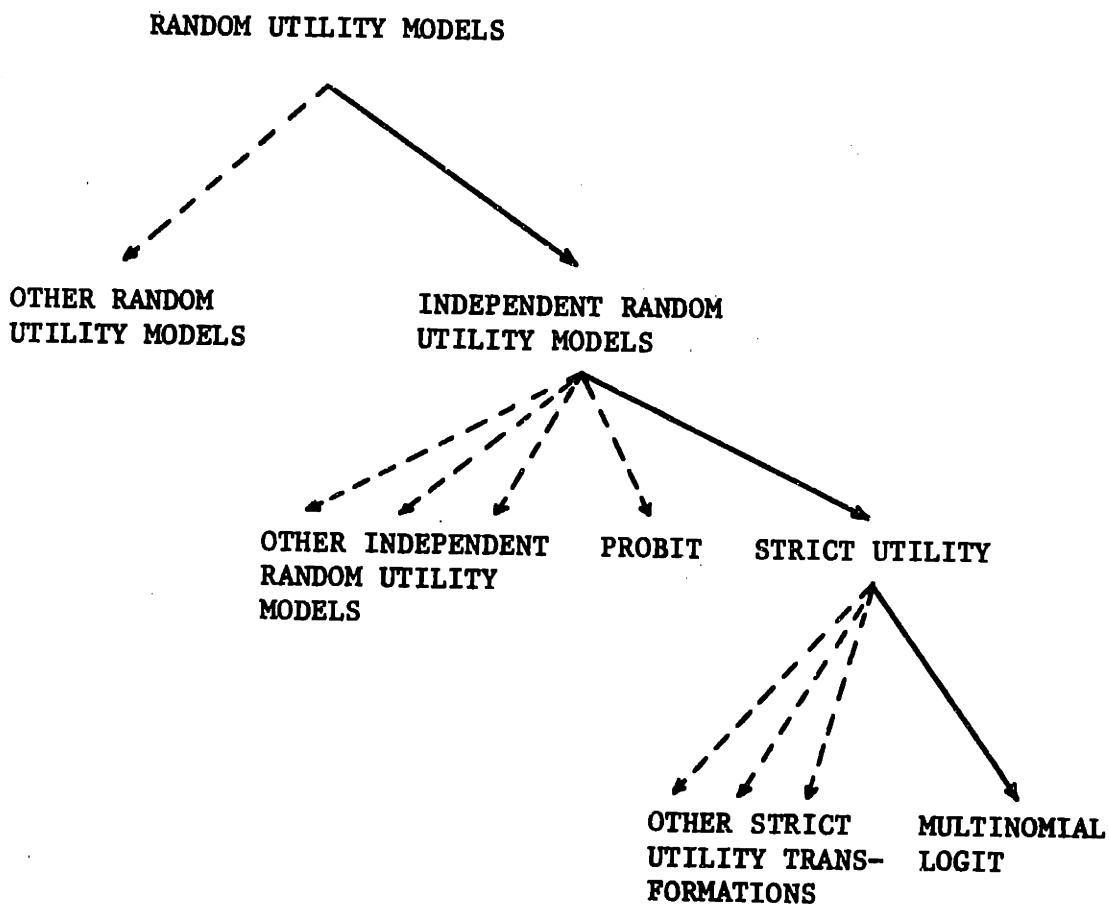


Figure 2

improvement over the logit model. A more realistic utility model could be devised. The problem is that we do not know if the additional resources are warranted in the case of travel choice models. The limited evidence from Bock and Jones (1968) and the state of the art of estimation programs suggest that at the present we should rely exclusively for multiple choice travel demand model on a strict utility specification.

The use of a strict utility specification, or more generally, an independent random utility model, will require a careful examination of the structure of the set of alternatives. The model should be applied only to sets of alternatives for which the assumption of independent (or, uncorrelated) utilities can be accepted as a good approximation. Therefore, a crucial point is the definitions of the alternatives. For example, for a choice of shopping destination, we should probably define alternative shopping centers, rather than alternative shops within centers, as the set of alternatives. If all shops were taken as independent alternatives, then there would be "clusters" of alternatives -- shops within each center -- that have similar, or close to one another, attributes (observed and unobserved). Again, this argues for a hierarchical structure of the field of alternatives where within each hierarchy the alternatives can be assumed to reasonably approximate independent utilities.

Summary

The demand approach of economic consumer theory assumes a selection of quantities from a set of commodities and the quantities demanded are treated as continuous variables. On the other hand, choice theories assume a selection from a finite set of mutually exclusive and exhaustive alternatives. In consumer theory, with a continuum of alternatives, we assume a deterministic behavior, except for a random disturbance added in estimation to explain variations among individuals and account for "optimization" errors and unobserved variables. On the other hand, in choice theory, with qualitative or discrete alternatives, we assume a probabilistic behavior that explains observations of different choices for the same set of observed independent variables.

Travel demand on the disaggregate level is more appropriately viewed in a choice context rather than in the traditional demand analysis framework. Travel and mobility choices can best be described as a choice from a finite set of mutually exclusive and exhaustive alternatives, rather than as a selection of quantity of commodity. Almost all the mobility and travel choices that were considered in Chapter II could only be described on the disaggregate level as choices from sets of qualitative alternatives. The exceptions are the choices of automobile ownership level and trip frequency that could be described either as continuous quantities or as qualitative alternatives. (Auto ownership levels are discrete but a continuous variable can be inter-

preted to express expected values.) The alternatives for the auto ownership and frequency choices are different from those of other choices because they are sequentially ordered rather than unordered. However, there is no problem in treating these alternatives as sets of discrete and qualitative alternatives.

The definition of the set of alternatives in choice models allows for the evaluation of non-price rationing, or non-pecuniary constraints. This is an important advantage for the purposes of policy evaluation.

It should be noted, that in a choice model the definition of the set of alternatives plays an important role. It brings out an issue which was not apparent in consumer theory which is how do we know from what alternatives the consumer is choosing. Predictions of a choice model will be biased if we include in the set of alternatives an alternative which seems desirable but is actually not available, or if we exclude a desirable alternative which is available. In general, and in particular in a strict utility specification, the exclusion or inclusion of an alternative that has a zero, or negligible, choice probability will not bias predictions. Possibly, this discussion points out the need for models that determine the set of alternatives for inclusion in a choice model. In this study, however, lacking such models, we will rely on a priori logical arguments and observations of current behavior in determining the set of alternatives.

CHAPTER IV

Multi-Dimensional Choice Models:
Alternative Structures of Travel Demand Models

Introduction

The choice theories that were described in the previous chapter considered a uni-dimensional choice situation. A consumer was assumed to select an alternative i out of a set of alternative choices A . If the set A includes the alternative choices of a single commodity, then the choice probability $P(i:A)$ is the choice analog of a demand function for a given commodity. A consumer is faced with a multi-dimensional choice situation in determining his consumption pattern.* For example, a consumer who is selecting a residence location within the metropolitan area is choosing also among alternative housing types, alternative auto ownership levels, etc.

The total number of choices that a consumer makes is very large. The assumptions of a "utility tree", or a separable utility function, and negligible income effects, that were discussed in the previous chapter, permit the independent modelling of demand for a subset of commodities. That is, the demand functions for a subset of commodities are independent of the prices of all other commodities.

*The term multiple choice refers to a choice from a set of more than two alternatives. A choice from two alternatives is termed binary choice. The term multi-dimensional choice is used for a set of choices where each choice can be either multiple or binary.

We assume here that mobility and travel choices are such an independent "branch" or subset of the consumer's utility function. Choices within this subset are interdependent. In Chapter II assumptions with respect to the dependencies among the mobility and the travel choices were discussed. It was concluded that a block recursive system is appropriate. That is, the first block consists of the mobility choices, and the second block consists of the travel choices (assuming the mobility choices as fixed). It was further assumed that travel choices with respect to different trip purpose categories could also be considered independently of each other. Thus we can model separately the set of mobility choices, and the sets of travel choices for different trip purposes (assuming the mobility choices to be predetermined). Yet, each of the above sets of choices represent a multi-dimensional choice situation.

The purpose of this chapter is to extend the choice theories from uni-dimensional to multi-dimensional choice situations. In a multi-dimensional choice situation different assumptions about the dependencies among choices result in models with different structures. The alternative structures are identified and their applicability to travel demand models is discussed.

Dependencies Among Choices

In order to simplify the discussion we will rely on an example of two choices. We consider a consumer who is making a trip for a given

trip purpose, say shopping, and is faced with the choices of destination d and mode of travel m .

We distinguish between two types of dependencies among choices: dependency in the structural sense, and dependency of the sets of alternative choices in a physical sense.

Dependency in the structural sense arises from substitution and complementary relationships among choices, and when different choices are made with respect to the same final commodity, i.e., the utilities from different choices are not independent. For example, the choices of auto ownership level and residence location are dependent on each other because a downtown location could be a substitute for a high auto ownership level. The utility from an alternative location will therefore depend on the chosen car ownership level, and vice versa.

The choices of mode and destination are made with respect to the same final commodity - a trip. Some of the attributes of a mode, such as travel time by bus, will be different for different destinations. Therefore, mode m to destination d is a different alternative from the same mode to destination d' ($d' \neq d$). Similarly, some of the attributes of destination d depend on the chosen mode. Therefore, destination d reached by mode m is a different alternative from the same destination reached by mode m' ($m' \neq m$). In other words, the utility from an alternative mode is dependent on the destination, and vice versa.

Thus, the dependency among travel choices can be attributed to the commonality of the attributes. In other words, some attributes

of a trip are specific to all travel choices. For example, the travel cost for shopping at a certain frequency depends on where one shops, what mode he is using, etc. Similarly, the travel cost of shopping at a given destination depends on how often one shops, what mode he is using, etc. Therefore, a traveller can trade off between different choices. For example, he can shop frequently at a nearby grocery store, or less frequently at a distant shopping center.

The dependency, or the causality, can be assumed either in one direction, e.g., the utility from a mode depends on the chosen destination, but the utility from an alternative destination is independent of the chosen mode; or in two directions, e.g., the utility from an alternative mode depends on the chosen destination and the utility from an alternative destination depends on the chosen mode. It was concluded previously that all travel choices are interdependent. However, we consider here also alternative assumptions that result in models with different structures, as will be shown in the following sections.

If the choices of mode and destination depend on each other, then the set of alternative modes is different for different destinations, and the set of alternative destinations is different for different modes. We denote the set of alternative modes for a given destination as M_d , and the set of alternative destinations for a given mode as D_m .

In addition, the set of alternative modes can be physically dependent on the chosen destination, and vice versa. For example, a bus service may be available to one destination but not to the other.

Therefore, the sets of alternative modes M_d can have different numbers of alternatives for different destinations.

If two choices are independent, then their alternative sets will also be independent. If the choice of mode and destination are assumed to be independent we denote the set of alternative modes as M and the set of alternative destinations as D .

The Overall Set of Alternatives

The consumer can be viewed as selecting an alternative destination and mode combination dm from an overall set of alternatives DM that includes all possible destination and mode combinations. For example, if the number of alternative modes available to every destination is identical and equal to M , and the number of alternative destinations is D , then the total number of alternatives in the overall set will be $D \times M$.

The overall set of alternatives DM can be partitioned according to modes or according to destinations. If we partition according to destination, then we can write the overall set of alternatives as follows:

$$DM = [M_1, M_2, \dots, M_d, \dots, M_D] \quad (1)$$

In this scheme we denote the set of destinations used for partitioning as D . Partitioning according to modes, we write:

$$DM = [D_1, D_2, \dots, D_m, \dots, D_M] \quad (2)$$

The set of modes used for partitioning is denoted as M . If the alternative sets are independent, then:

$$\begin{aligned} M_d &= M, \forall d \in D \\ \text{and} \\ D_m &= D, \forall m \in M \end{aligned} \tag{3}$$

Alternative Structures

If we assume that the choices are independent, then we can write the following structural choice probabilities:*

$$\begin{aligned} P(d:D) &= \text{Prob}[U_d \geq U_{d'}, \forall d' \in D] \\ \text{and} \\ P(m:M) &= \text{Prob}[U_m \geq U_{m'}, \forall m' \in M] \end{aligned} \tag{4}$$

where U_d and U_m are the utilities from destination d and mode m , respectively. In essence, the independence assumption implies an additive utility function:**

$$U_{dm} = U_d + U_m \tag{5}$$

In words, the total utility from a destination and mode combination is equal to the utility from the destination plus the utility from the mode. Since the choices are independent, we can write the joint probability of d and m as follows:

$$P(d, m:DM) = P(d:D) \cdot P(m:M) \tag{6}$$

The structure that represents independent choices, or an independent structure, consists of marginal probabilities of the different choices.

If the choice of mode and destination are dependent on each other then we can write the following conditional choice probabilities.

$$\begin{aligned} P(d:D_m) &= \text{Prob}[U_d|m \geq U_{d'}|m', \forall d' \in D_m] \\ \text{and} \\ P(m:M_d) &= \text{Prob}[U_m|d \geq U_{m'}|d', \forall m' \in M_d] \end{aligned} \tag{7}$$

*The probabilities that have direct behavioral interpretation and are originally written to describe a structure are called structural probabilities.

**See the discussion of separable utility function in Chapter III.

where $U_{d|m}$ is the utility from destination d given that mode m is chosen, $U_{m|d}$ is the utility from mode m given that destination d is chosen. The conditional probability $P(d:D_m)$ is the choice probability of destination d given that mode m is chosen, and similarly $P(m:M_d)$ is the choice probability of mode m given that destination d is chosen.

For forecasting, however, the two conditional probabilities are insufficient information to compute the joint probability of destination and mode. In this case, as opposed to independent choices, the joint probability is not a product of two marginal probabilities, since $P(m:M_d)$ is functionally dependent on d, i.e., $P(m:M_d) \neq P(m:M)$. If we had $P(m:M)$ then the joint probability is equal to its product with $P(d:D_m)$. However, in order to model the marginal probability, $P(m:M)$ we need to identify a utility function for an alternative mode which is independent of what destination is actually chosen. Therefore, for such a simultaneous structure, in which the choice of destination depends on the choice of mode and vice versa, we must model explicitly the joint probability $P(d,m:DM)$. Given the joint probability, we can derive the marginal probabilities, and the structural probabilities as follows:

$$\begin{aligned}
 P(m:M) &= \sum_{d \in D_m} P(d,m:DM) \\
 P(d:D) &= \sum_{m \in M_d} P(d,m:DM) \\
 P(d:D_m) &= \frac{P(d,m:DM)}{P(m:M)} \\
 P(m:M_d) &= \frac{P(d,m:DM)}{P(d:D)}
 \end{aligned} \tag{8}$$

A dependency that goes only in one direction results in a recursive structure. If we assume that the choice of destination is independent of what mode is actually chosen, and that the choice of mode is dependent on the chosen destination, we write the following probabilities:

$$P(d:D) = \text{Prob}[U_d \geq U_{d'}, \forall d' \in D]$$

and

(9)

$$P(m:M_d) = \text{Prob}[U_{m|d} \geq U_{m'|d}, \forall m' \in M_d]$$

This recursive structure implies the following additive utility function:

$$U_{dm} = U_d + U_{m|d}$$
(10)

The utility for a destination and mode combination is equal to utility from the destination plus a utility from the mode which is dependent on the destination. Note that in a recursive structure the joint probability is the product of the structural probabilities.

Since we assume in this recursive structure that $P(m:M_d) \neq P(m:M)$, it is possible to derive from the joint probability a conditional $P(d:D_m)$ which is not equal to $P(d:D)$. However, this conditional probability is not causal but simply a mathematical relationship derived from the model with no behavioral interpretation.

A recursive structure represents a hierarchical conditional decision structure. It is a common practice to replace a complex decision with a large number of alternatives by a recursive structure. The decision is decomposed into stages by successive partitions of the

overall set of alternatives. Luce (1959) noted that different partitions give different results. Therefore, a recursive structure can be viewed either as a simplifying assumption (this will require a sensitivity analysis of the partitioning scheme to determine how the results are affected), or as truly representing a sequential, or conditional, decision making process.

Separability of Choices

Implicit in the discussion of the alternative structures was a separability of choices assumption. The conditional choice probability of mode given a destination was written as $P(m:M_d)$. This implies that the choice of m given d is independent of alternative modes to all other destinations d' ($d' \neq d$), and is dependent only on the alternative modes for the given destination.

This is a reasonable assumption. It is required in order to be able to model choices separately. If we model directly a joint probability and assume a simultaneous dependency, then it appears that this assumption is not necessary. However, the interpretation of the derived conditional probabilities will not be the same as the one used here. It was also impossible to find an example of a model that does not make this assumption.

If we partition the set DM according to destinations we can write the joint probability as follows:

$$P(d,m:DM) = P(m:M_d) \cdot P(d:D) \quad (11)$$

This equation is similar to the way in which the choice axiom was expressed in the previous chapter:

$$P(i:A) = P(i:B) \cdot P(B:A) \quad (12)$$

for $i \in B \subseteq A$. The subset B corresponds to the subset of alternative modes to a given destination. However, the choice axiom is more general than the separability of choices assumption. It applies to any partitions of A to non-overlapping subsets B. The separability of choices assumption applies only to partitions according to choices.

There is some similarity between the concept of functional separability that was considered for utility functions in Chapter III and the separability of choices assumption. Functional separability is based on the idea that the marginal rate of substitution among a set of variables is independent of other variables. Separability of choices implies that a conditional probability for a given choice depends only on a part of the total utility function. The choice of mode given a destination is assumed to be dependent on $U_{m|d}$, which is the part of the utility function that for a given d varies across modes.

Hence, a separability assumption implies that from the utility function for a destination and mode combination U_{dm} , we can identify the utility from a mode given a chosen destination $U_{m|d}$, and the utility from a destination given a chosen mode $U_{d|m}$. Clearly, their sum is not equal to U_{dm} . The separability assumption in an independent structure implies the additive utility function of equation (5):

$$U_{dm} = U_d + U_m \quad (5)$$

The separability assumption in a recursive structure where m depends on d implies the additive utility of equation (10):

$$U_{dm} = U_d + U_m|_d \quad (10)$$

Another assumption that is implied here is that the coefficients of the utility function U_{dm} are invariant with d and m . It is assumed that the specification of the variables in the utility function explains all the differences among utility values for different alternatives except for an unobserved random disturbance. This is not a restrictive assumption, since we can always use alternative specific variables as will be described in Chapter VI. Examples of violation of this assumption can be traced to specification errors.* For a mode choice model it may sometimes be found that travel time appears to be more onerous to one destination than another. A mode choice model may be quite sensitive to trip length in a way not specified in the utility formulation. Another source of the difference could be the omission of some other variable which described the way in which travel time is spent and which is different between the two destinations.

The Identification Problem and Estimation of
Simultaneous Probability Structures

In a structure of simultaneous equations the coefficients of a structural equation can be estimated from the data only if the equation is identified with respect to all other equations (Fisher, 1966). It appears that in a simultaneous structure of conditional probabilities this problem does not exist. First, it will be useful to draw the

*Specification errors are the consequences of an incorrect set of explanatory variables and/or incorrect mathematical form.

analogy between simultaneous equations and conditional probability structures.

The endogenous variables in the probabilistic models are the choices and the exogenous variables are those included in the utility functions.

The joint probability plays the role of the reduced form of a system of simultaneous equations. The reduced form can always be directly estimated and used for forecasting. Similarly, the joint probability can be directly estimated and used for forecasting. However, a structural equation can be estimated only if it is identified, while the coefficients of a structural conditional probability can always be derived once the joint probability has been estimated.

In a system of independent simultaneous equations, one can always compute the reduced form solution. However, it is impossible to derive the joint probability from the structural conditional probabilities in a simultaneous probabilistic model. Therefore, for forecasting, it is always necessary in a simultaneous probabilistic structure to estimate the joint probability.

This apparent contradiction becomes clear when one recognizes that the conditional probabilities are not the exact analog of the structural equations in a system of simultaneous equations. A conditional probability gives the choice probability for one choice, when all other choices are fixed. This is equivalent to a structural equation with all the right hand side endogenous variables fixed at given values. The variability of the left hand side endogenous variable is attributed to a random disturbance. This means that a conditional probability supplies a differ-

ent type of information and is therefore not subjected to the same kind of identification issues inherent to structural equations.

It is possible to estimate directly the conditional probabilities, or to derive their estimates from the estimated joint probability.* If the purpose of the analysis is to make only conditional predictions of one choice, given that all other choices remain constant, then the conditional probabilities are all that is needed and one can estimate them directly. However, the coefficient estimates of the conditional probabilities will not necessarily be equal whether they were estimated directly or indirectly through the estimation of the joint probability.

It appears that the coefficient estimates of the conditional probabilities can gain in statistical efficiency and can be less sensitive to specification errors if estimated through the joint probability. The basis for this statement is the possibility of incorporating restrictions across conditional probabilities, and thereby using more information to estimate some coefficients, in estimating the joint probability. As an example consider the simultaneous structure of destination choice and mode choice described previously. It is possible that $U_{d|m}$ and $U_{m|d}$ have common coefficients. By estimating directly U_{dm} we constrain them to be equal and we use simultaneously all the information from the choice among alternative modes as well as from the choice among alternative destinations. If we estimate directly $U_{d|m}$ we can only use information on alternative destinations for the chosen mode, i.e., the alternatives

*Estimating the joint probability and then deriving the conditional probabilities is analogous to the method of indirect least squares (Malinvaud, 1966).

in D_m . For direct estimation of $U_{m|d}$ we can only use information on alternative modes for the chosen destination, i.e., the alternatives in M_d . In estimating U_{dm} we use information on all the alternatives in the overall set DM.

It is only under very restrictive conditions that direct estimate of, say, $U_{m|d}$ will result in the same coefficient estimates as indirect estimation through U_{dm} . This happens when the alternatives in DM that are not in M_d do not provide additional information to that obtained from M_d alone. In other words, this happens when the variability of modal attributes for destinations d' ($d' \neq d$) is the same as for the chosen destination d . The exact conditions that have to be fulfilled by the data for this to occur depend on the exact specification of the choice model. However, it seems that knowledge of the exact conditions is unimportant because as a practical matter it never occurs. Furthermore, even if it occurs there is no reason not to estimate U_{dm} if it can only be more efficient and it is needed for forecasting anyway.

It should be noted here that in a recursive probabilistic structure there is no reason to estimate directly the joint probability. Therefore, it could be estimated in its structural form, as was done by CRA (1972).*

A simultaneous structure could also be estimated as a recursive structure, as follows: first, estimate one conditional, say $P(m|M_d)$; secondly, derive from the analytical form of the joint probability the marginal

*The model developed by CRA (1972) is discussed in Chapter V.

$P(d:D)$; and finally estimate the marginal with the coefficients that are included in $P(m:M_d)$ constrained to their estimates from $P(m:M_d)$. This estimation procedure is suggested only when for some reason it is computationally difficult to estimate directly the joint probability.

Modelling the Travel Choices

From the discussion in this chapter and in Chapter II it is evident that the appropriate structure for the travel choices is a simultaneous one. In the remainder of this chapter we discuss alternative structures of travel demand models in more detail.

A trip taken for a specific purpose is characterized by its origin, destination, time of the day, mode of travel, and route. We are interested in predicting the volume of trips V_{idhmr} , from an origin i , to a destination d , during a time of day h , using a mode m , and via a route r . From the point of view of the individual trip-maker or the household, we consider the probability of a trip instead of a quantity or volume of trips. A trip decision consists of several choices: choice of trip frequency f , (e.g., how often to go shopping), choice of destination d (e.g., where to shop), choice of time of day h (e.g., when to go), choice of mode m , and choice of route r . Hence, for an individual traveller we are interested in predicting the joint probability:

$$P(f, d, h, m, r : FDHMR_t) \quad (13)$$

where t denotes an individual or a household in origin i ,* and $FDHMR_t$ is the overall set of alternative trips that consists of all possible combinations of frequencies, destinations, modes, times of day, and routes available to individual t .** The alternatives in this set are exhaustive and mutually exclusive. The individual t is always selecting one and only one alternative from this set.

For simplicity we will write the probabilities in the remainder of this chapter without denoting the set of alternatives. We write the above probability (13) as:

$$P_t(f, d, h, m, r) \quad (14)$$

A conditional probability previously denoted as $P(m|M_d)$ will now be written as $P(m|d)$. The joint probability previously written as $P(d, m: DM)$ will now be written as $P(d, m)$.

On the disaggregate level the travel demand function for a given trip purpose predicts the joint probability $P_t(f, d, h, m, r)$. On the aggregate level the demand function predicts the volume V_{idhmr} . In either case, we have a complex product — a trip — with an enormous number of substitutes. Micro-economic consumer theory tells us that a demand function expresses the quantity of a product demanded as a

*The choice of residence location is assumed as given. Travel demand models assume that mobility decisions are fixed. (See the discussion in Chapter II.)

**In the following sections a notation for different subsets of $FDHMR$ is used. This notation follows the same logic that was used to define subsets of DM . Therefore, it is not explained in text.

function of its price, the prices of related commodities (substitutes and complements), and income. The complexities stem from the large number of relevant "prices" (i.e., price and many price-like attributes), for all the alternative trips.

Alternative Structures of Travel Demand Models

With no further assumption, the travel demand model predicts the probability $P(f,d,m,h,r)*$, or the volume V_{fdmhr} , as a function of the attributes of all the alternative combinations of $fdmhr$. Denote the explanatory variables as $x_{fdmhr}^1, x_{fdmhr}^2, \dots, x_{fdmhr}^k, \dots, x_{fdmhr}^K$, or as a vector X_{fdmhr}^{**} . Hence, we can write the travel demand model as follows:

$$P(f,d,m,h,r) = F([x_{fdmhr}, \forall fdmhr \in FDMHR]) \quad (15)$$

where $[x_{fdmhr}, \forall fdmhr \in FDMHR]$ is a vector that includes all the variables X for all relevant combinations of the subscripts (f,d,m,h,r) , and F is the demand function. Alternatively, we can write the utility function for an alternative trip as:

$$U_{fdmhr} = U(X_{fdmhr}) \quad (16)$$

* For additional simplicity, we drop the subscript t in writing the probabilities in this section.

**The explanatory variables include all the level of service, the spatial opportunities, and the socio-economic variables. The socio-economic variables are specific to an individual and not to a trip alternative. However, we assume here that they are introduced into the model as having alternative specific values. The ways in which this is done are described in Chapter VI.

Clearly, this results in a very complex demand model. Without further assumptions, for a simultaneous structure this is the type of travel demand model which must be estimated.

If, however, we make some assumptions about the travel decision making process we can divide the overall travel demand function into several less complex functions, each including only a subset of all the explanatory variables. That is, under some assumptions we can formulate the travel demand function as a recursive or as an independent structure.

The first assumption that is required is the separability of choices assumption that was described earlier, and is usually made also with respect to a simultaneous model. The separability assumption with respect to a certain choice says that the conditional probability of this choice given other choices is a function of only a specific subset of the explanatory variables, as depicted in the following example for route choice:

$$U_{r|fdmh} = U^r(X_{fdmhr}) \quad (17)$$

and

$$P(r|f,d,m,h) = F^r([X_{fdmhr}, \forall r \in R_{fdmh}])$$

In words, the conditional probability of choosing a route given other choices is a function only of the explanatory variables for all routes for given fdmh. If we considered only two choices, say mode and destination, then the separability assumption with respect to mode choice says that the conditional probabilities of choosing a mode given a destination is a function of the variables for all modes but for only one specific destination. For this example we write:

$$\begin{aligned}
 P(d, m) &= F^{dm}([X_{dm}, \forall dm \in DM]) \\
 U_{dm} &= U(X_{dm}) \\
 P(d|m) &= F^d([X_{dm}, \forall d \in D_m]) \\
 U_{d|m} &= U^d(X_{dm}) \\
 P(m|d) &= F^m([X_{dm}, \forall m \in M_d]) \\
 U_{m|d} &= U^m(X_{dm})
 \end{aligned} \tag{18}$$

Note that if we calculate the marginal probabilities $P(d)$ and $P(m)$, they will be a function of the vector $[X_{dm}, \forall dm \in DM]$.

An independent structure is possible only if the set of attributes is separable. That is,

$$[X_{fdmhr}] = [X_f, X_d, X_m, X_h, X_r] \tag{19}$$

where we can identify only attributes that vary only across a single choice. The independent utility function can be written as:

$$U_{fdmhr} = U^f(X_f) + U^d(X_d) + U^m(X_m) + U^h(X_h) + U^r(X_r) \tag{20}$$

The independent travel demand model can be written as:

$$\begin{aligned}
 P(f) &= F^f([X_f, \forall f \in F]) \\
 P(d) &= F^d([X_d, \forall d \in D]) \\
 P(m) &= F^m([X_m, \forall m \in M]) \\
 P(h) &= F^h([X_h, \forall h \in H]) \\
 P(r) &= F^r([X_r, \forall r \in R])
 \end{aligned} \tag{21}$$

and

$$P(f, d, m, h, r) = P(f) \cdot P(d) \cdot P(m) \cdot P(h) \cdot P(r) \quad (22)$$

It should be clear that this is an unrealistic structure for a travel demand model.

A recursive structure requires the assumption of a sequential decision making process, or a hierarchy of conditional decisions. The sequence is expressed in a recursive travel demand model in two ways. The first is the manner in which the set of all trip alternatives is partitioned. In a recursive model of mode and destination choices where mode choice is conditional on the chosen destination, the set of all alternative combinations of mode and destination is partitioned according to destination. The second way is the "composition" of explanatory variables. For the same example, the problem is how do we include in a model of the marginal probability of destination choice the variables that are defined by destination and mode, such as travel time, fare, etc. The way this is handled is to construct a composite variable that combines the above variables across modes to create a variable which is specific only to a destination.

Consider for example the following recursive structure:

$$U_{fdmhr} = U_f + U_{d|f} + U_{m|fd} + U_{h|fdm} + U_{r|fdmh} \quad (23)$$

$$= U^f(x_f) + U^d(x_{fd}) + U^m(x_{fdm}) + U^h(x_{fdmh}) + U^r(x_{fdmhr})$$

$$P(f) = F^f([x_f, \forall f \in F]) \quad (24)$$

$$P(d|f) = F^d([x_{fd}, \forall d \in D_f])$$

$$P(m|f,d) = F^m([X_{fdm}, \forall m \in M_{fd}]) \quad (24)$$

$$P(h|f,d,m) = F^h([X_{fdmh}, \forall h \in H_{fdm}])$$

$$P(r|f,d,m,h) = F^r([X_{fdmhr}, \forall r \in R_{fdmh}])$$

where each variable is defined as follows:

$$X_{fdmh} = [X_{fdmhr}, \forall r \in R_{fdmh}]$$

$$X_{fdm} = [X_{fdmh}, \forall h \in H_{fdm}]$$

$$X_{fd} = [X_{fdm}, \forall m \in M_{fd}] \quad (25)$$

$$X_f = [X_{fd}, \forall d \in D_f]$$

If we keep the variables in their original form then the model for $P(f)$ will include all the explanatory variables $[X_{fdmhr}, \forall fdmhr \in FDMHR]$.

The definition of composite variables allows the treatment of X_{fdmh} , X_{fdm} , X_{fd} , and X_f as single variables. In other words, these variables are expressed as a specific function of their elements. For example, we express:

$$X_{fdmh} = g([X_{fdmhr}, \forall r \in R_{fdmh}]) \quad (26)$$

where g is the composition function. The functional form of the composition rule requires further assumptions.

There are a variety of possible composition schemes. One such scheme which was derived from an assumption of additive utility function by CRA (1972) is as follows:

$$X_{fdmh} = \sum_{r \in R_{fdmh}} X_{fdmhr} \cdot P(r|f,d,m,h) \quad (27)$$

This composition scheme is essentially a computation of the expected value of the original variable. Another way to observe this is to rewrite equation (27) using the definition of conditional probability as follows:

$$X_{fdmh} \cdot P(f,d,m,h) = \sum_{r \in R_{fdmh}} X_{fdmhr} \cdot P(f,d,m,h,r) \quad (28)$$

Thus the composite variable as defined by equation (27) is in accordance with a consistency requirement that the expected value of a variable is maintained. If X is a price variable then equation (28) says that the expected expenditure is consistent in the different stages of a recursive model.

Clearly, there are many other schemes of creating the composite variables, among them: a simple sum;

$$X_{fdmh} = \sum_{r \in R_{fdmh}} X_{fdmhr} \quad (29)$$

or the value for the "best" route (Quandt and Baumol, 1966);

$$X_{fdmh} = X_{fdmhb} \quad (30)$$

where $r = b$ is the "best" route according to some criterion.

Often, several price variables are combined together to form a "generalized price". Then the composite variable is a composition of the generalized price instead of each variable separately (Manheim, 1972; CRA, 1972).

Constructing a composite variable from several explanatory variables together amounts to maintaining equal "marginal rates of

"substitution" among those variables in the different probabilities of a recursive structure.

Thus, given a separability assumption, a specific sequence assumption, and an assumption on the mathematical form of the composite variables, the overall travel demand model can be formulated as a recursive structure.

A simultaneous structure requires the estimation of an equation that includes a very large number of explanatory variables. On the other hand, each equation in a recursive structure includes only a subset of the explanatory variables that are included in a simultaneous model. In addition, the number of variables is reduced by constructing composite variables. Therefore, a recursive model can be easier to implement, computationally and analytically, than a simultaneous model.

The separability and the sequence assumptions required by a recursive travel demand model are equivalent to an assumption of a conditional decision structure. The choice of a particular fdmhr combination is made from a relatively large set of alternatives. It makes "sense" to partition the set of all alternatives into collections of non-overlapping subsets. Consider, for example, two choices: destination and mode. The set of all alternative combinations of d and m, DM, is very large. We can partition DM into the subsets $M_1, M_2, \dots, M_d, \dots, M_D$, where each subset includes all the alternative modes to a specific destination. The assumption is that the traveller is first choosing among these subsets, or choosing a destination, and

then, secondly, choosing within the chosen subset, or choosing a mode. The choice of mode is now a function of only the characteristics of available modes to a given destination. The choice of destination depends on some measure of the expected attributes of all modes to a given destination. The utility function of a dm combination is assumed to consist of two parts; one for each choice. The choice of destination is based on the utility of the destination, which is also dependent on the expected attributes from the modes available to this destination.

Note, however, that we can also partition the set DM according to modes as follows: $D_1, D_2, \dots, D_m, \dots, D_M$. When we apply choice models to this or the previous sequence we do not expect the predictions to be the same. The problem is, therefore, to know when does the consumer decompose his decision into stages and what partitions are used.

If we modelled the choice of an fmdhr combination as a deterministic optimization problem, it would not be important what partitions were used. The reason that we expect different partitions to give different results is due to the probabilistic choice mechanism and the computation of expected attributes from lower stages.

The problem with travel decisions is that we cannot find a unique "natural" sequence of partitions that will be generally applicable. Therefore, a simultaneous structure is superior to a recursive structure. In general, the simultaneous structure of a travel demand model consists of the following conditional Probabilities:

$$\begin{aligned}
 & P(f|d,m,h,r) \\
 & P(d|f,m,h,r) \\
 & P(m|f,d,h,r) \\
 & P(h|f,d,m,r) \\
 & P(r|f,d,m,h)
 \end{aligned} \tag{31}$$

Under particular behavioral assumptions we can place restrictions on this general structure and obtain alternative simultaneous structural forms.

Consider the following simultaneous structure:

$$\begin{aligned}
 & P(f|d,m,h,r) \\
 & P(d|f,m,h) \\
 & P(m|f,d,h) \\
 & P(h|f,d,m,r) \\
 & P(r|f,d,m,h)
 \end{aligned} \tag{32}$$

Note that the conditional probabilities of mode choice and destination choice are not conditional on the chosen route. This is due to the fact that we cannot generally identify alternative modes or destinations for a given route.

The choices that are conditional on f in either a simultaneous or a recursive structure are defined only for $f>0$, because it does not make sense to define alternative trips when no trip is taken. It may be argued that for some trip purposes the choice of trip frequency is based on some measure of expected accessibility and is not dependent on

the actual values of d , m , h and r . Therefore, it is "natural" to partition according to f , and for each f have all possible combinations of $mdhr$.

If for some trip purpose the choice of time of day is constrained, or limited to alternative times for which the traveller can be assumed to be indifferent, then it is possible to partition according to f , and for each f have all possible combinations of m and d . Then, partition according to dm combinations to create the sets of alternative routes for a given trip. This decomposition implies the following structural probabilities:

$$P(f)$$

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

(33)

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

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

The choices of mode and destination are simultaneous, but recursive with respect to f . The choice of route is recursive with respect to f , d and m . This is essentially the structure that is assumed in the empirical study reported in Chapter VII.*

It should be clear that in simultaneous and recursive structures we can derive any conditional or marginal probabilities. (However, only the structural probabilities are causal.) Therefore, for forecasting, it is possible to use the joint probability directly, or any

*Time of day was excluded because the sample included only off-peak shopping trips.

combination of marginal and conditional probabilities provided that their product is equal to the joint probability. For example,

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

$$\begin{aligned} &= P(f) \cdot (P_d | f) \cdot P(m | f, d) \cdot P(h | f, d, m) \cdot P(r | f, d, m, h) \\ &= P(f) \cdot P(h | f) \cdot P(m | f, h) \cdot P(d | f, h, m) \cdot P(r | f, h, m, d) \\ &= P(f) \cdot P(h, m, d | f) \cdot P(r | f, h, m, d) \end{aligned}$$

etc.

The Aggregate Equivalent of Disaggregate Models

In order to compute a volume of trips we need to multiply the joint probability $P(f, d, m, h, r)$ by the actual values of the alternative trip frequencies. The definition of alternative trip frequencies f is based on a time period. If the total number of trips is expressed for a period of twenty-four hours then f is a daily trip frequency. If we assume that the alternative daily trip frequencies take either the value of 1 for one trip or the value of 0 for no trip, then:

$$V_{tdhmr} = P_t (f=1, d, m, h, r) \quad (34)$$

where V_{tdhmr} is the expected number of trips that individual t will make to destination d , during time-of-day h , by mode m and via route

r. $P_t(f=1, d, m, h, r)$ is the probability that individual t will make a trip to destination d, during time-of-day h, using mode m, and via route r.

In general, we must write:^{*}

$$V_{tdhmr} = \sum_{f \in F_{dhmr}} f \cdot P_t(f, d, m, h, r) \quad (35)$$

The aggregate volume is computed by summing over individuals, as follows:

$$V_{idhmr} = \sum_{t \in T_i} V_{tdhmr}$$

where T_i is the set of all potential trip makers in origin i, e.g., the resident population, and V_{idhmr} is the expected volume of trips from origin i to destination d, during time-of-day h, using mode m, and via route r.

Note from (36) that the "per capita" volume $\frac{V_{idhmr}}{T_i}$ is an average of V_{tdhmr} . It corresponds to the average of the joint probability $P_t(f=1, d, m, h, r)$ in the case of 0, 1 alternative daily trip frequencies.

The expected number of trips that individual t will make daily can be computed as follows:

*It is possible to define the following volume of trips:

$$V_{tfdhmr} = f \cdot P_t(f, d, m, h, r)$$

However, all existing aggregate travel demand models predict volumes of trips for all frequencies. Therefore, we use here V_{tdhmr} .

If $f = 0,1$:

$$\begin{aligned} V_t &= P_t(f=1) \\ &= \sum_{dmhr \in DMHR} P_t(f=1, d, m, h, r) \end{aligned} \quad (37)$$

In general, however,

$$\begin{aligned} V_t &= \sum_{f \in F} f \cdot P_t(f) \\ &= \sum_{f \in F} f \cdot \sum_{dmhr \in DMHR_f} P_t(f, d, m, h, r) \end{aligned} \quad (38)$$

"Trip generation", or the total number of trips originating from i , can be computed as follows:

$$V_i = \sum_{t \in T_i} V_t \quad (39)$$

Again, the average trip generation $\frac{V_i}{T_i}$ corresponds to the average of V_t , and to the average of $P_t(f=1)$ in the case of $f=0,1$. This indicates that alternative trip frequencies are not explicitly represented in any travel demand model that explains a volume of trips as the dependent variable.

Using the same procedure we can compute any element of volume, such as V_{id} , V_{idm} , V_{im} , V_{idmh} , V_{idmr} , V_{ih} , etc. For example, if $f=0,1$:

$$\begin{aligned} V_{tdm} &= P(f=1, d, m) \\ &= \sum_{hr \in HR_{dm}} P_t(f=1, d, m, h, r) \end{aligned} \quad (40)$$

or in general,

$$\begin{aligned} V_{tdm} &= \sum_{f \in F_{dm}} f \cdot P_t(f, d, m) \\ &= \sum_{f \in F_{dm}} f \cdot \sum_{hr \in HR_{fdm}} P_t(f, d, m, h, r) \end{aligned} \quad (41)$$

and

$$V_{idm} = \sum_{t \in T_i} V_{tdm} \quad (42)$$

Note that V_{tdm} can also be computed as follows:

$$\begin{aligned} V_{tdm} &= \sum_{hr \in HR_{dm}} \sum_{f \in F_{dmhr}} f \cdot P_t(f, d, m, h, r) \\ &= \sum_{hr \in HR_{dm}} V_{tdhm} \end{aligned} \quad (43)$$

therefore, we can compute V_{idm} as follows:

$$V_{idm} = \sum_{hr \in HR_{dm}} V_{idhm} \quad (44)$$

All these volumes satisfy the condition of "internal consistency" (Manheim, 1972). For example, it can be shown that:

$$V_{id} = \sum_{m \in M_d} V_{idm} \quad (45)$$

$$V_i = \sum_{d \in D} V_{id}$$

etc. This indicates that the condition of "internal consistency" is equivalent to the assumption of a mutually exclusive and exhaustive set of alternatives. This assumption is necessary to compute the marginal probabilities by summation of the joint probability.*

Again, it should be noted that all these volumes are equal to the sum of the marginal probabilities only in the case of 0,1 daily trip frequencies. For example,

$$V_{idm} = \sum_{t \in T_i} P(f=1, d, m) \quad (46)$$

*The probability functions are discrete.

and the average probability for the group T_i is:

$$\frac{V_{idm}}{T_i} \quad (47)$$

"Modal split" is defined sometimes in aggregate travel demand models as follows:

$$S_{idm} = \frac{V_{idm}}{V_{id}} \quad (48)$$

In words, the fraction, or the share, of the volume from i to d that goes by mode m. Computing this share for individual t directly from the probabilities we get:

$$S_{tdm} = \frac{\sum_{\substack{hr \in HR \\ dm}} V_{tdhmr}}{\sum_{\substack{mhreMHR \\ fd}} V_{tdhmr}} \quad (49)$$

$$= \frac{\sum_{\substack{f \in F \\ dm}} P_t(f, d, m)}{\sum_{\substack{f \in F \\ d}} P_t(f, d)}$$

For the case of 0,1 alternative trip frequencies, this share becomes:

$$S_{tdm} = \frac{P_t(f=1, d, m)}{P_t(f=1, d)} \quad (50)$$

$$= P_t(m | f=1, d)$$

If we assume that the choices of m and d are independent of f^* , we get:

*This is not a realistic assumption, as was already indicated.

$$\begin{aligned}
 S_{tdm} &= \frac{P(d, m | f>0) \cdot \sum_{f \in F} f \cdot P_t(f)}{P(d | f>0) \cdot \sum_{f \in F} f \cdot P_t(f)} \\
 &= \frac{P(d, m | f>0)}{P(d | f>0)} \\
 &= P(m | d, f>0)
 \end{aligned} \tag{51}$$

We find that shares, or fractions, of volumes correspond to the conditional probabilities only in the case of 0,1 frequencies, and when we assume that the choice of f is independent (in the probability sense) of all other choices.

Note that the average share is not equal to the ratio of the volumes:

$$\frac{\sum_{t \in T_i} S_{tdm}}{T_i} \neq \frac{V_{idm}}{V_{id}} \tag{52}$$

However, the average share is equal to the average of the conditional probabilities for 0,1, or independent, frequencies. This indicates that an aggregate share model, that predicts an aggregate share of volume, does not predict exactly the corresponding average conditional probability. It predicts the ratio of the averages of two probabilities. For example,

$$S_{idm} = \frac{\sum_{t \in T_i} P_t(f=1, d, m)}{\sum_{t \in T_i} P_t(f=1, d)} \neq \frac{\sum_{t \in T_i} P_t(m | f=1, d)}{T_i} \tag{53}$$

We conclude that the structural conditional probabilities, with the assumption of 0,1, or independent, frequencies, are equivalent to

shares of individual volumes. For example,

$$P_t(d|f,h,m) = \frac{V_{tdhm}}{V_{thm}} \\ = S_{thmd} \quad (54)$$

where S_{tdmh} is the share of trips by individual t during time-of-day h by mode m that are going to destination d .

Direct and Indirect Travel Demand Models

A distinction was made between simultaneous, recursive, and independent travel demand models. It was based on the behavioral assumptions of the model. Another distinction that is often made is between direct and indirect travel demand models (Manheim, 1972). This distinction, however, is based on the way that the travel demand model is used for forecasting.

A direct demand model predicts directly the joint probability $P_t(f,d,m,h,r)$, or the volume V_{idmhr} , as a function of all the explanatory variables. In an indirect travel demand model the joint probability, or the volume, is predicted with several intermediate steps. Each step corresponds to a single choice, or to a single subscript of the volume. For example, one equation can predict the number of trips taken by the household, another equation will distribute trips among the various destinations, and so forth. Hence, in a direct model a forecast is made with a single equation, while in an indirect model a forecast is made using a multi-equation model.

Note that the sequence used for forecasting does not necessarily have a behavioral interpretation. Even a recursive model could in principle be used for forecasting in an indirect fashion that does not correspond to the structural sequence.

An example of a widely used indirect travel demand model with a recursive structure would be the conventional Urban Transportation Model System (UTMS).* The aggregate demand model of the UTMS is structured into a sequence of four steps, commonly called: trip generation, trip distribution, modal split, and route assignment** (Manheim, 1972, Martin et al., 1961). The trip generation model predicts the total volume V_i from origin i. The trip distribution model distributes the volume V_i among the various destinations to predict the total volume V_{id} from origin i to destination d. In a similar way the modal split model predicts V_{idm} , and finally the route assignment model predicts V_{idmr} . The UTMS demand model that was described represents a particular sequence. Clearly, other possible sequences of the intermediate steps are possible. In particular, some UTMS demand models placed the modal split model after the trip generation model and before the trip distribution model. (Weiner, 1969).

* The behavioral structure of the UTMS variants will be discussed in the next chapter.

**The volume of trips is generally for a single time period of 24 hours. A peak hour volume was sometimes estimated as a constant proportion of the 24 hours volume.

Note that each step of an indirect aggregate travel demand model actually predicts a share, or a fraction of volume. In the UTMS variant that was described the following shares are predicted:

$$\begin{aligned}
 \text{"trip generation": } \quad s_i &= \frac{v_i}{T_i} \\
 \text{"trip distribution": } \quad s_{id} &= \frac{v_{id}}{v_i} \\
 \text{"modal split": } \quad s_{idm} &= \frac{v_{idm}}{v_{id}} \\
 \text{"route assignment": } \quad s_{idmr} &= \frac{v_{idmr}}{v_{idm}}
 \end{aligned} \tag{55}$$

The product of all these shares gives:

$$\frac{v_{idmr}}{T_i} \tag{56}$$

which is what we want to predict.

Assume, for example, the following sequence:

$$f \rightarrow d \rightarrow m \rightarrow h \rightarrow r$$

where the arrows indicate the sequence of forecasting the various choices. The following list includes the disaggregate probabilities for the case of 0,1 alternative frequencies, and the equivalent aggregate volumes that are predicted in each intermediate step of an indirect travel demand model:

Disaggregate

$$P(f=1)$$

Aggregate

$$\frac{V_i}{T_i} \quad (57.1)$$

$$P(f=1, d)$$

$$\frac{V_{id}}{T_i} \quad (57.2)$$

$$P(f=1, d, m)$$

$$\frac{V_{idm}}{T_i} \quad (57.3)$$

$$P(f=1, d, m, h)$$

$$\frac{V_{idmh}}{T_i} \quad (57.4)$$

$$P(f=1, d, m, h, r)$$

$$\frac{V_{idmhr}}{T_i} \quad (57.5)$$

The intermediate steps actually predict shares of volume or conditional probabilities as in the following list:

Disaggregate

$$P(f=1)$$

Aggregate

$$s_i = \frac{V_i}{T_i} \quad (58.1)$$

$$P(d|f=1)$$

$$s_{id} = \frac{V_{id}}{V_i} \quad (58.2)$$

$$P(m|f=1, d)$$

$$s_{idm} = \frac{V_{idm}}{V_{id}} \quad (58.3)$$

$$P(h|f=1, d, m)$$

$$s_{idmh} = \frac{V_{idmh}}{V_{idm}} \quad (58.4)$$

$$P(r|f=1, d, m, h)$$

$$s_{idmhr} = \frac{V_{idmhr}}{V_{idmh}} \quad (58.5)$$

The product of the marginal probability (58.1) and the conditional probabilities (58.2 - 58.5) gives the probability $P(f=1, d, h, m, r)$, and the product of the aggregate shares (58.1 - 58.5) gives the volume $\frac{V_{idmhr}}{T_1}$ as follows.

$$\begin{aligned} P(f=1, d, m, h, r) &= P(f=1) \cdot P(d|f=1) \cdot P(m|f=1, d) \\ &\quad \cdot P(h|f=1, d, m) \cdot P(r|f=1, d, m, h) \end{aligned} \quad (59)$$

and,

$$\frac{V_{idmhr}}{T_1} = s_i \cdot s_{id} \cdot s_{idm} \cdot s_{idmh} \cdot s_{idmhr} \quad (60)$$

Manheim (1972) introduced the "general share model" which is expressed by equation (60). The disaggregate equivalent of the "general share model" is equation (59). Clearly, we can write equations (59) and (60) for any possible sequence. For example, assume the following sequence:

$$f \rightarrow h \rightarrow m \rightarrow d \rightarrow r$$

For this sequence we can write:

$$\begin{aligned} P(f=1, d, m, h, r) &= P(f=1) \cdot P(h|f=1) \cdot P(m|f=1, h) \\ &\quad \cdot P(d|f=1, h, m) \cdot P(r|f=1, h, m, d) \end{aligned} \quad (61)$$

and,

$$\frac{V_{idmhr}}{T_1} = s_i \cdot s_{ih} \cdot s_{ihm} \cdot s_{ihmd} \cdot s_{ihmhr} \quad (62)$$

Thus, any indirect travel demand model, disaggregate or aggregate, could be written as a single equation which would be the product of the different functions for all the intermediate steps. The reverse is also possible. Given a single function in a direct travel demand model

we can derive functions for all the intermediate steps and for any sequence of an indirect travel demand model. This is shown in the following set of equations for the sequence of equations (59) and (60):*

Disaggregate

$$P(f=1, d, m, h) = \sum_r P(f=1, d, m, h, r)$$

$$P(f=1, d, m) = \sum_h P(f=1, d, m, h)$$

$$P(f=1, d) = \sum_m P(f=1, d, m)$$

$$P(f=1) = \sum_d P(f=1, d)$$

$$P(r | f=1, d, m, h) = \frac{P(f=1, d, m, h, r)}{P(f=1, d, m, h)}$$

$$P(h | f=1, d, m) = \frac{P(f=1, d, m, h)}{P(f=1, d, m)}$$

$$P(m | f=1, d) = \frac{P(f=1, d, m)}{P(f=1, d)}$$

$$P(d | f=1) = \frac{P(f=1, d)}{P(f=1)}$$

Aggregate

$$v_{idmh} = \sum_r v_{idmhr} \quad (63.1)$$

$$v_{idm} = \sum_h v_{idmh} \quad (63.2)$$

$$v_{id} = \sum_m v_{idm} \quad (63.3)$$

$$v_i = \sum_d v_{id} \quad (63.4)$$

$$s_{idmhr} = \frac{v_{idmhr}}{v_{idmh}} \quad (63.5)$$

$$s_{idmh} = \frac{v_{idmh}}{v_{idm}} \quad (63.6)$$

$$s_{idm} = \frac{v_{idm}}{v_{id}} \quad (63.7)$$

$$s_{id} = \frac{v_{id}}{v_i} \quad (63.8)$$

The direct and the indirect models that were described are two extremes. There are a variety of possible indirect models in which an intermediate step may predict directly more than one choice. For example, one equation can predict the number of trips taken by the

*We are assuming throughout that we have a set of exhaustive and mutually exclusive alternatives or that the condition of "internal consistency" (Manheim, 1972) holds. This is necessary to carry out the summations in equations (63.1 - 63.4).

household to a certain destination, another equation will split these trips among the various modes of travel, and so forth.

Thus, from the forecasting point of view it does not make any difference whether we use a model as direct or as indirect. The way a model is used for forecasting should be determined only on the basis of computational efficiency and convenience considerations.

In this study we are concerned with the behavioral structure of travel demand models. However, as was shown, we can express any given model in many different ways. Therefore, an obvious question to ask is: how can the behavioral structure of a given model be recognized?

In general, the answer to this question is that the behavioral structure cannot be determined unless the model is written in its structural form. This answer could be explained by using the analogy of a structure of simultaneous equations. Given a reduced form, which is used for forecasting, it is impossible to determine the original structure. However, in travel demand models that were structured with composite variables, the structure may be discerned. It is possible to recognize the sequence through the order of composition (e.g., order of summation) that is maintained in a composite variable no matter how the model is expressed.

The Empirical Problem

As mentioned earlier, the complexity of the overall travel demand function stems primarily from the large number of alternatives and attributes which call for a large number of variables. In order to

appreciate the dimensions of the overall travel demand function, consider the following example of travel choices.

Suppose that for a certain trip purpose a person has the following options:

- two daily trip frequencies (one trip or no trip),
- four destinations,
- two modes of travel, and
- two times of day (peak or off-peak).

The total number of alternatives facing the decision maker is 17 (16 one-trip alternatives and one no-trip alternative). Suppose that for each one-trip alternative there are only two price variables, travel time and travel cost. (The price of no trip is zero.) The total number of price variables is 32. If we increment each choice by one additional option we have a total of 91 alternatives and 180 price variables.

It appears that the joint probability may be too complex and the number of variables too large to be condensed into a single relationship. The most important question is whether we can estimate a choice model with such large numbers of alternatives and variables. Using a recursive structure, we will have to estimate four choice models but with the number of alternatives in each model equal to the number of options for the corresponding choice. It should be pointed out that the data requirements are identical for both structures unless further assumptions are made.

It is not clear whether it is less expensive to estimate four models each with a small number of alternatives rather than one model with many alternatives. (Assuming, of course, that estimation of a joint probability is feasible.)

Under the presumption that the implementation of a recursive model is easier and less expensive, is the additional expense to implement a simultaneous model justified? The answer is unclear. Costs can be compared only together with the benefits. Therefore, we need to know how the simplifying assumptions of a recursive model affect the results of the prediction process.

These are critical issues that can only be addressed by an empirical study. The evidence from the estimation of alternative structures that was conducted in this study is reported in Chapter VII.* Briefly, it indicates that a simultaneous model is feasible and that the results are highly sensitive to the assumed structure. This empirical evidence is not absolutely conclusive, however, because it is based on one small sample and only on a subset of the travel choices for a single trip purpose. Future research is needed to extend the empirical evidence to different data sets, larger samples and a complete set of travel choices for all trip purpose categories.

*Note that the purpose of the empirical study is not to determine which structure is the "true" one.

CHAPTER V

Review of the Structures of Current Approaches to Travel Demand Modelling

Introduction

The purpose of this chapter is to provide a review of the structures of current approaches to travel demand modelling. Specifically, this chapter will distinguish among current approaches according to their structure: either simultaneous or recursive, and according to their use of data: either disaggregate or aggregate. No attempt will be made in this chapter, however, to survey the detailed specifications of existing models.

A Typology of Travel Demand Models

A recursive travel demand model can be recognized from the definitions of the composite variables which are always based on specific sequence assumptions.*

Conversely, simultaneous models are recognized by their lack of composite variables (i.e., no explicit sequence of travel choices is assumed). There have been no simultaneous travel demand models that treated simultaneously all the travel choices. Thus, in this review, a model that has a simultaneous structure for any subset of travel choices is considered to be a simultaneous model.

The dependent variable of an aggregate travel demand function is a volume of trips, V . A disaggregate travel demand function predicts

*The use of composite variables in recursive travel demand models is discussed in Chapter IV.

an individual's probability of travel, P. When a travel demand function is factored into its component choices then the intermediate stages predict either shares of volume in an aggregate model or conditional probabilities in a disaggregate model.

The independent variables are of two distinct types: characteristics of the activity system, A, and the level of service characteristics, L. The set of variables, A, includes three categories of variables: socio-economic variables, S; spatial opportunities (also land use or attraction) variables, E; and, in aggregate models, aggregation variables, T.

The travel choices (except frequency) that are included in a trip probability are also used as subscripts for volume and explanatory variables. These travel choices are as follows:

f = trip frequency (in disaggregate models only)

d = destination zone

m = mode of travel

h = hour of day

r = route

In addition, in aggregate models, trips are aggregated by the origin zone, i.

Travel demand models for different trip purpose categories and for different market segments (or person types) can generally be assumed to be independent. Therefore, the models that are presented in the following sections are for a given trip purpose, and for a given market seg-

ment (if such stratification is performed). Composite variables will be defined in the text where applicable.

Aggregate Recursive Models

The Urban Transportation Model System (UTMS) is the acronym given to a variety of aggregate recursive travel demand models that have been commonly used in urban transportation planning studies. There are many references describing these models: Martin et al (1961), FHWA (1970), Fertal et al (1966), Weiner (1969), and others. In addition, there are several critical appraisals of the UTMS including: Kain (1967), Stopher and Lisco (1970), Meyer and Straszheim (1971), Manheim (1970) and (1972), Brand (1972), and CRA (1972).

The UTMS has many variants differing in the functional specification of the models and in the set of explanatory variables employed. Most important (for the purposes of this review) are the differences in their structural forms. All the UTMS variants are recursive, however different sequences of choices were (implicitly) assumed. All the alternative sequences of the UTMS assumed route choice to be the last in the sequence. Time of day choice was usually represented as a constant proportion of the 24-hour volume of trips.

Three basic UTMS variants were classified according to the place of mode choice in the overall sequence by Weiner (1969):

1. Direct generation of transit trips.
2. Trip-end modal split models.
3. Trip-interchange modal split models.

The first of these model sequences has been used primarily in small urban areas, and/or in areas where it was considered that public transportation ridership is largely independent of transport level of service. This structure essentially represents a simultaneous choice of frequency and mode followed by the choice of destination as expressed below:

Trip generation and modal split:

$$\text{Production: } v_{im} = g_1(s_i, t_i) \quad (1)$$

$$\text{Attraction: } v_{dm} = g_2(e_d) \quad (2)$$

Trip distribution:

$$\frac{v_{imd}}{v_{im}} = g_3(\tilde{v}_{dm}, L_{imd}, \tau_d) \quad (3)$$

where \tilde{v}_{dm} are the revised attractions in a trial and error procedure that is performed to satisfy the attractions constraints:

$$\sum_i v_{imd} = v_{dm}$$

This model is clearly misspecified. Most obviously, equations (1) and (2) should include level of service measures as follows:

$$v_{im} = g_4(s_i, t_i, [L_{im}, \tau_m]) \quad (4)$$

$$v_{dm} = g_5(e_d, [L_{dm}, \tau_m]) \quad (5)$$

where L_{im} and L_{dm} are composite variables (or accessibility measures). (The composition is done over destination zones and origin zones, respectively.)

In the trip end sequence, modal split is performed after trip generation but before trip distribution as expressed below:

Trip generation:

$$\text{production: } v_i = g_6(s_i, t_i) \quad (6)$$

$$\text{attraction: } v_d = g_7(e_d) \quad (7)$$

Modal split:

$$\frac{v_{im}}{v_i} = g_8(s_i, [L_{im}, \#m]) \quad (8)$$

$$\frac{v_{dm}}{v_d} = g_9(e_d, [L_{dm}, \#m]) \quad (9)$$

Trip distribution:

$$\frac{v_{imd}}{v_{im}} = g_3([\tilde{v}_{dm}, L_{imd}, \#d]) \quad (\text{Same as equation (3)})$$

Again, the trip generation models, Equations (6) and (7), are misspecified. An improved specification of these models (incorporating level of service measures) could be expressed as follows:

$$v_i = g_{10}(s_i, t_i, L_i) \quad (10)$$

$$v_d = g_{11}(e_d, L_d) \quad (11)$$

where L_i and L_d are composite (or accessibility) variables defined over all modes, and destinations and origins, respectively.

The most common sequence in UTMS applications, especially in large metropolitan areas, is the trip interchange modal split sequence as presented below:

Trip generation:

Same as equations (6) and (7).

Trip distribution:

$$\frac{v_{id}}{v_i} = g_{12}(\tilde{v}_d, L_{id}, \tau_d) \quad (12)$$

Modal split:

$$\frac{v_{idm}}{v_{id}} = g_{13}(S_i, E_d, L_{idm}, \tau_m) \quad (13)$$

where \tilde{v}_d are revised attractions to satisfy the constraint:

$$\sum_i v_{id} = v_d$$

All the models presented above suffer in practice from numerous specification errors that are unnoticeable from the way we expressed them. For example, the composite level of service variables L_{id} in Equation (12) were in many applications taken as only travel time by highway. Furthermore, many of the models, such as the attraction equations, have no behavioral justification, and in reality represent a process of data fitting.

There have been numerous attempts to improve the submodels of the UTMS and in particular to make them policy sensitive. The work of Wilson (1969) and Wilson et al (1969) develops a consistent formulation of level of service variables in trip distribution and modal split models. Gur (1970) estimated a special type of trip generation equations with accessibility variables.

Given that the UTMS submodels are properly specified, they still do not constitute a pure travel demand model due to the constraint on attractions. If it is assumed that the number of trips arriving at a destination affects its attractiveness, then there is clearly a need for a supply type relationship to express this phenomenon. Thus, the demand model that is represented by the above procedure consists of the production part of trip generation, the distribution model without the attraction constraint, and the modal split model.

The travel demand model that was used in the Northeast Corridor Transportation Project (NECTP) is also an aggregate recursive model (McLynn and Woronka, 1969). This model can be written as follows:

$$V_{id} = g_{14}(S_i, T_i, E_d, L_{id}) \quad (14)$$

$$\frac{V_{idm}}{V_{id}} = g_{15}([L_{idm}, V_m]) \quad (15)$$

where L_{id} is a composite variable of L_{idm} . Note that this model suffers from a specification error not present in the UTMS, that is V_{id} is determined independent of attraction and level of service characteristics of destinations other than d. This implies that destination choice is not represented, a characteristic of almost all demand models that were developed for the NECTP (Crow et al, 1971).

Manheim (1972) specified theoretically (without estimation) a class of recursive models that he termed Special Product Models. These models represent a consistent specification of a recursive structure with a specific method of forming the composite variables.

Aggregate Simultaneous Models

Several aggregate simultaneous models have been estimated which predict v_{idm} directly. This type of model was first developed for intercity travel in connection with the NECTP by Kraft (1963), Quandt and Baumol (1966), Quandt and Young (1969), and others.* Later, the same type of models were applied to urban travel demand by CRA (1967),** Plourde (1968) and Talvitie (1971). A common characteristic of these models is that they do not take account of alternative destinations, similar to the recursive NECTP model. Therefore, it can be said that only the choices of frequency*** and mode of travel have been modelled simultaneously. These models can be expressed in general form, as follows:

$$v_{idm} = g_{16}(T_1, S_1, E_d [L_{idm}, \nabla_m]) \quad (16)$$

A correct specification of equation (16) will include also the E_d and L_{idm} variables for all alternative destinations.

An attempt to deal with the problem of predicting ridership on a new mode led to the development of the abstract mode models by Quandt and Baumol (1966), and others (Crow et al, 1971). Several versions of

*For a survey of the models developed by Quandt, Young, Blackburn, Monsod and others for the NECTP see Crow et al (1971) and Quandt (1970).

**Also reported by Domencich et al (1968).

***Note from the discussion in Chapter IV, that all existing aggregate models by their nature cannot represent properly frequency choice. They implicitly assume a choice between 0 and 1 alternative daily frequencies, or independent frequency choice.

the abstract mode model were formulated in terms of the attributes of the "best" mode. As a result cross elasticities were defined only with respect to the attributes of one mode.*

These aggregate simultaneous models have some inherent shortcomings which can be traced to the reliance on aggregate data and on linear regression estimation techniques. The use of aggregate data resulted in severe problems of multicollinearity. As a result some models were estimated using a constrained regression technique (Kraft, 1963; CRA, 1967; and Talvitie, 1971). The abstract mode models included smaller numbers of coefficients and therefore reduced the multicollinearity problem. The use of linear regression restricted the functional forms that could be used. In particular it resulted in the inability to model a large number of alternatives, such as alternative destinations.

Disaggregate Recursive Models

Until recently the use of disaggregate data in travel demand models has been limited to scattered estimation of some submodels in a framework of a recursive structure. The sequence that was implied was that of the trip-interchange version of the UTMS. Disaggregate models were estimated for trip generation, mode choice, and route choice models.

The trip generation models that were estimated with household data rather than zonal averages were deterministic (i.e., predicted volume rather than probability). Therefore, they have the same specification as the aggregate models, and also suffer from the same specification errors.**

*Other versions of the abstract mode model were formulated with the attributes of the "best" mode replaced by averages or weighted averages of the attributes of all available modes.

**See, for example, Fleet and Robertson (1968), FHWA (1970), and Kannel and Heathington (1973).

Disaggregate probabilistic mode choice models, primarily for work trips, were initially developed for binary choice, between auto and transit, similar to the UTMS modal split models. These models were briefly reviewed by Reichman and Stopher (1971).* Rassam et al (1971) estimated a disaggregate mode choice model with four alternative modes, but with a relatively small set of explanatory variables.

A disaggregate probabilistic model of a binary choice of route, between free and toll roads, was estimated by Thomas (1968).**

The theoretical framework for the application of disaggregate probabilistic models to the entire travel demand function was outlined by Stopher and Lisco (1970). They described a recursive structure that followed the UTMS trip-interchange sequence. Recently, a complete set of disaggregate models was developed by CRA (1972) for the shopping trip purpose. Their model can be expressed as follows:

$$P_t(f) = g_{17}(S_t, E, L) \quad (17)$$

$$P_t(d | f = 1) = g_{18}(S_t [E_d, L_d, \not d]) \quad (18)$$

$$P_t(h | f = 1, d) = g_{19}(S_t, [L_{dh}, \not h]) \quad (19)$$

$$P_t(m | f = 1, d, h) = g_{20}(S_t, [L_{dhm}, \not m]) \quad (20)$$

where t denotes an individual traveller. The sequence implied by the order of the conditional probabilities was used to formulate the composite variables for L and E .

*See, for example, the binary mode choice models developed by Warner (1962), Quarby (1967), Lisco (1967), Lave (1969), McGillivray (1969), Stopher (1969), Plourde (1971), de Donnea (1971), Stopher and Lavender (1972) and Talvitie (1972).

**See also Thomas and Thompson (1970).

The innovations of the CRA model are numerous. Disaggregate data was used for the destination choice model and choice of time of day was explicitly modelled. A consistent formulation of composite variables, that was based on the concept of expected generalized or inclusive price, was used throughout the model.

Disaggregate Simultaneous Models

It appears that no disaggregate simultaneous travel demand model has been developed so far. CRA (1972) recognized the desirability of a simultaneous structure but used a recursive structure in order to reduce the complexity of the estimation problem.* However, they did not test a simultaneous model and did not justify the particular sequence that was used.

This review does not include models of mobility choices, however, the model that was developed by Aldana (1971)** is relevant since it was based on disaggregate probabilistic models. The model included three choices: residence location, auto ownership level, and choice of mode. The choice of mode was with respect to the frequency of use of public transport.

Aldana hypothesized a simultaneous structure but he did not estimate directly the joint probability. Rather, he used a sequential estimation procedure in which the following probabilities are estimated:

*See in CRA (1972) the section "Factoring the Demand Model" in Chapter 4.

**See also Aldana et al (1973).

the marginal probability of residence location, the conditional probability of auto ownership level given residence location, and the conditional probability of transit use given auto ownership and residence location.

Summary

The review of travel demand approaches reported in this chapter indicates that current approaches to travel demand modelling fall into the categories of: aggregate recursive, aggregate simultaneous (frequency and mode), and disaggregate recursive. It appears that there has been no attempt to develop a disaggregate simultaneous travel demand model.

CHAPTER VI

The Logit Model and Functional Specification of Alternative Models

Introduction

To this point, we have presented the alternative structures of travel demand models without specifying functional forms. The purpose of this chapter is to describe the multinomial logit model as used for the models estimated in this study. First, the logit model and its properties are discussed in general terms and then, secondly, its applications to simultaneous and recursive travel demand models are described.

The Multinomial Logit Model

We characterized the demand for travel as a choice among a set of discrete, or qualitative, alternatives, just as the total amount of travel can be regarded as a choice among different trip making frequencies. A potential traveller can decide not to take a trip at all; therefore, a no-trip alternative is also included in the set of alternatives that he is facing. Consequently, one, and only one, alternative is always chosen from this set. We assume that the set of alternatives can be different for different travellers.

We are interested in the probability of each alternative being chosen. (In the aggregate, we are interested in the share of each alternative.) Clearly, the probability (share) of any alternative must be between zero and one and the sum of the probabilities (shares) of all alternatives must equal unity. The model that we use should

be able to handle any number of alternatives and different sets of alternatives for different individuals.*

We turn now to describe the choice model used in this study and its important properties. The basic choice model that is used in this study for all alternative models is the multinomial logit model. Other choice models that might be considered to be superior from a theoretical point of view, such as a multiple probit model, are more complicated. It is not evident, however, that the added expense for more sophisticated choice models is worthwhile.**

We write the logit model as follows:

$$P(i:A_t) = \frac{e^{v_{it}}}{\sum_{j \in A_t} e^{v_{jt}}} \quad (1)$$

where:

t = a behavioral unit,

= 1, 2, ..., T,

A_t = the set of relevant alternatives for behavioral unit t ,

$P(i:A_t)$ = the probability that behavioral unit t will choose alternative i out of the set A_t ,

v_{it} = the utility of alternative i to behavioral unit t .

*This is a requirement that is not met by the technique of discriminant analysis which was often used for mode choice models, e.g., Quarmby (1967), McGillivray (1969), Plourde (1971).

**See the discussion of choice models in Chapter III.

The utility V_{it} is a function of the characteristics of alternative i (or "prices") and the socio-economic characteristics of behavioral unit t (or "income").* This is in essence an indirect utility function which is defined as a function of prices and income. (Alternatively, it can be viewed as an expenditure function which is a function of prices.) Hence, the function V_{it} can be expressed as:

$$V_{it} = V_i (X_i, S_t) \quad (2)$$

where:

X_i = a vector of characteristics of alternative i

S_t = a vector of socio-economic characteristics of behavioral unit t

Assume the function V_{it} to be a linear function in the parameters:**

$$V_{it} = X_{it}'\theta = \sum_{k=1}^K X_{itk} \theta_k \quad (3)$$

where:

X_{it} = a $K \times 1$ vector of finite functions which are constructed from the various X_i and S_t variables and are different from one alternative to the other.***

*The socio-economic characteristics include income and other variables, such as life cycle, education, etc., that account for differences in "tastes".

**This assumption can be justified if we consider the V 's to be the expenditure functions for the different alternatives. This assumption is required, however, since the available estimation program requires linearity in the parameters.

***This definition is explained in detail in the following section.

$$= (X_{it_1}, X_{it_2}, \dots, X_{itK})$$

θ = a Kx1 vector of parameters

$$= (\theta_1, \theta_2, \dots, \theta_K)$$

Equation (1) can now be expressed as follows:

$$P(i:A_t) = \frac{e^{X_{it}^{\top} \theta}}{\sum_{j \in A_t} e^{X_{jt}^{\top} \theta}} \quad (4)$$

The Specification of the Variables

A variable in X_{it} , say X_{itk} , can be specified either as a generic variable or as an alternative specific variable. If the variable X_{itk} appears only in the utility function of alternative i , then it has a value of 0 for all other alternatives. That is,

$$X_{jtk} = 0, \forall j \neq i \in A_t$$

and it is termed an alternative i specific variable. If the variable X_{itk} appears in the utility functions of all the alternatives, it is termed a generic variable. Consider, for example, the variable of travel time in a mode choice model. If the travel times by the different modes are assumed to have a common coefficient in their respective utility functions, then travel time is specified as a generic variable. The variable X_{itk} will take the value of travel time by mode i . On the other hand, if we assume mode specific coefficients of travel time, we will introduce a series of mode specific variables. For example, mode m travel time variable will be specified as follows:

$$x_{itk} = \begin{cases} \text{travel time by mode } m, & \text{for } i = m \\ 0 & \text{, for } i \neq m \end{cases}$$

If there are M alternative modes, we will have a total of M travel time variables. If all the variables in the model are generic, then it is an "abstract alternative model". This is an essential property when the set of alternatives used for estimation is not identical with the one used for forecasting, for example, the forecasting of mode choice probability for a "new mode".

Alternative specific variables are relevant only when there is some natural correspondence between the alternatives available to different individuals, e.g., alternative modes. CRA (1972) termed these kinds of alternatives ranked alternatives. In this case, if all the variables in the model are specified as alternative specific, this implies that the utility functions are alternative specific as well. When there is no correspondence between individual's sets of alternatives, we can use only generic variables. For example, the set of alternative shopping destinations at one origin may be entirely different from the destination set at a different origin. Another example will be alternative routes which are different for different trips. CRA (1972) referred to these kinds of alternatives as unranked alternatives. The set of alternatives can consist of both ranked and unranked alternatives, for example, alternative destinations that include the CBD as one of the alternatives. For this example it is possible to specify CBD specific variables.

If a variable takes the same value for all alternatives by all individuals, i.e., for some k

$$x_{itk} = x_{jtk}, \forall t \in T, \forall j \in A_t$$

it will have no effect on the model, due to the linear specification.

In other words, the coefficient θ_k is not identified. For example, if the variable x_{itk} is income, then we have the same term multiplying the numerator and each member of the sum in the denominator, and it will cancel out. This means that all the variables in x_{it} must have alternative specific values.*

Hence, the socio-economic characteristics s_t can enter the model only when they are transformed to have alternative specific values. This could be done in two ways:

- (1) Combining them with the x_i variables:

$$x_{itk} = g^k(x_i, s_t)$$

where g^k is a finite function (e.g., income multiplied by price.) The function $g^k(x_i, s_t)$ has an alternative specific value, and can be used to define either a generic or an alternative specific variable.

- (2) Introducing a series of alternative specific dummy variables, each one of them takes the value of the socio-economic variable for a certain alternative and is zero

*Not to be confused with alternative specific variables. The reference here is to the value a variable assumes.

otherwise (e.g., alternative one income variable which equals income for alternative one and zero for all other alternatives.) This specification is relevant only for ranked alternatives. For a given socio-economic variable, such as income, we can introduce only N-1 alternative specific dummy variables, where N is the total number of ranked alternatives.* If we introduce such variables for the first N-1 alternatives, then the coefficient estimate of alternative i dummy variable is the difference between the actual coefficient of alternative i variable and that of alternative N variable. This can be observed from rewriting equation (4) as follows:

$$P(i:A_t) = \frac{1}{\sum_{j \in A_t} e^{(X_{jt} - X_{it})' \theta}} \quad (5)$$

If the model was originally written with N dummy variables with the same value, the exponents in the sum will include differences of the original coefficients. Therefore, only N-1 differences can be identified. It should be clear that the above also applies to dummy variables that take the value of 1 and represent "pure alternative effect". For example, a mode specific constant.

*An attempt to estimate a model with N alternative specific dummy variables with the same value (e.g. income) will fail due to perfect "collinearity".

Derivations of the Logit Model

The logit model has been used in several previous studies of travel demand models. It has been derived from a variety of theoretical considerations. In Chapter III we have already discussed two approaches to the derivation of the logit model: from the choice axiom as a constant utility model, and from a specific distribution as a random utility model. Alternative derivations of the logit model are presented among others by McFadden (1968), McLynn et al (1968), Rassam et al (1971), and Wilson (1972). McFadden derives the logit model starting from the assumption of independence from irrelevant alternatives. This assumption was mentioned as a property of the choice axiom in Chapter III. (It is presented for the logit model in the following section.) McLynn et al and Rassam et al derive the logit model from assumptions about the elasticities of the model. (The elasticities of the logit model will be derived in a following section.) The derivation of the logit model by Wilson is made from an entirely different perspective. It is based on the entropy maximizing method and it can be regarded as a maximization of a specific distribution function subject to constraints.

The formulation employed in the McLynn and Woronka (1969) model is equivalent to equation (4) if all the variables x_{itk} are replaced by their logs. This formulation can be written as follows:

$$P(i:A_t) = \frac{\prod_{k=1}^K \pi_{itk}^{\theta_k}}{\sum_{j \in A_t} \prod_{k=1}^K \pi_{jtk}^{\theta_k}} \quad (6)$$

Many other models that have been used for travel demand models have essentially a strict utility specification of which the logit model, equation (4), is only one possible transformation.

Some of the properties of the logit model were already discussed in Chapter III. In the following sections we consider again some properties and derive the elasticities of the model.

The Choice Axiom and the Independence from Irrelevant Alternatives Property

We defined A_t as the set of all relevant alternatives for behavioral unit t . Therefore,

$$\sum_{i \in A_t} P(i:A_t) = 1 \quad (7)$$

Suppose that we are interested in a subset of the relevant alternatives, B_t , and we want to predict the conditional probability of choosing i given that the choice lies within this subset:

$$P(i|B_t:A_t) \text{ for } i \in B_t \subset A_t \quad (8)$$

assuming that $P(i:A_t) > 0$.

From the definition of conditional probability, we can write the following:

$$P(i:A_t) = P(i|B_t:A_t) \cdot P(B_t:A_t) \quad (9)$$

This means that the probability of choosing i is equal to the probability of choosing i from the subset times the total probability of the subset, i.e., the probability that the choice lies in B_t .

From (9) we can write:*

$$\begin{aligned} P(i|B_t : A_t) &= \frac{P(i:A_t)}{P(B_t : A_t)} \\ &= \frac{P(i:A_t)}{\sum_{k \in B_t} P(k:A_t)} \end{aligned} \quad (10)$$

Hence, given the probability function for $P(i:A_t)$, we can always derive the conditional probability of equation (8).

The Choice Axiom as introduced by Luce (1959), and discussed in Chapter III, says that a choice from a set of alternatives is independent of what else may have been available. This property means that the following holds:

$$P(i|B_t : A_t) = P(i:B_t) \quad (11)$$

It says that the choice of i from the subset B_t is as if the alternatives in B_t are the only alternatives considered.

To see that this property holds for the Logit model, we compute the conditional probability as in equation (10) using equation (1) as follows:

$$P(i|B_t : A_t) = \frac{\frac{e^{V_{it}}}{\sum_{j \in A_t} e^{V_{jt}}}}{\sum_{k \in B_t} \frac{e^{V_{kt}}}{\sum_{j \in A_t} e^{V_{jt}}}} \quad (12)$$

*We assumed that only one alternative is chosen. Therefore, we can write:

$$P(B_t : A_t) = \sum_{k \in B_t} P(k:A_t)$$

$$\begin{aligned}
 &= \frac{e^{v_{it}}}{\sum_{k \in B_t} e^{v_{kt}}} \\
 &= P(i:B_t)
 \end{aligned}$$

The last equality follows from the definition of the logit model, equation (1).

Another way of observing this property for the logit model is to rewrite equation (1) and divide and multiply by the sum of the utility exponents over the subset of alternatives B_t :

$$\begin{aligned}
 P(i:A_t) &= \frac{e^{v_{it}}}{\sum_{k \in B_t} e^{v_{kt}}} \cdot \frac{\sum_{k \in B_t} e^{v_{kt}}}{\sum_{j \in A_t} e^{v_{jt}}} \\
 &= P(i:B_t) \cdot \sum_{k \in B_t} P(k:A_t) \\
 &= P(i:B_t) \cdot P(B_t:A_t)
 \end{aligned} \tag{13}$$

By equation (9), we get:

$$P(i|P_t:A_t) = P(i:B_t)$$

Thus, equation (13) again says that the probability of choosing i from the set A_t is equal to the probability of choosing i from the subset B_t , given that only B_t is considered, times the total probability of any of the alternatives in B_t being chosen out of A_t .

This property of the logit model holds for any subset of alternatives. Consider the following choice probability:

$$P(f,d,m,h,r:FDMHR_t) \tag{14}$$

In words, the joint probability of choosing trip frequency f , destination d , mode m , time-of-day h , and route r , from the set of all trip alternatives. Consequently, if we used the logit formulation to model a joint probability such as equation (14), then equation (11) implies that the separability of choices assumption discussed in Chapter IV holds for all choices. For example, if the subset B_t corresponds to all relevant modes of travel to a given destination, then given a logit model for the joint probability of mode and destination we can derive a conditional probability for the choice of mode given a destination which is a function of only the alternatives in B_t . Therefore, a simultaneous model is separable in this case.

Using the logit model for a joint probability implies, however, an assumption stronger than the separability assumption. This is due to the fact that the property of equation (11), or the Choice Axiom, holds for any subset of alternatives and not only for subsets created by partitions that correspond to the various travel choices. It applies to any subset of alternatives created across and within the various travel choices. For example, all modes to a group of destinations, and a group of modes to a given destination. This stronger "separability property" may not be as easily justifiable as the separability assumption. However, it is the price that one has to pay for the relative simplicity of the logit model.

The Choice Axiom is sometimes referred to as the "independence from irrelevant alternatives" axiom, or property. The idea is that

the comparison between two alternatives is not influenced by what other alternatives are available. In probability terms the idea is that the ratio of the probability of choosing one alternative to the probability of choosing the other is independent of the set of available alternatives. To show how it is implied by the Choice Axiom, consider the subset A_t to consist of two alternatives i and j . Then, using equations (9) and (11), we get:

$$\frac{P(i:A_t)}{P(j:A_t)} = \frac{P(i:[i,j])}{P(j:[i,j])} \quad (15)$$

It can also be observed directly as a property of the logit model.

From (1) we can write the ratio in (15) as follows:

$$\begin{aligned} \frac{P(i:A_t)}{P(j:A_t)} &= \frac{e^{V_{it}}}{e^{V_{jt}}} \\ &= e^{V_{it} - V_{jt}} \end{aligned} \quad (16)$$

Hence the ratio of the probabilities is only a function of the characteristics of the two alternatives. This implies, for example, that if a logit model is used for a mode choice model then the ratio of the choice probabilities of, say auto and bus, is independent of whether or not rapid transit is also available.

Define Q_{ij} as the odds of choosing i over j :

$$Q_{ij} = \frac{P(i:A_t)}{P(j:A_t)} \quad (17)$$

Then, from (16) we get:

$$\ln Q_{ij} = v_{it} - v_{jt}$$

If the utility functions are linear as in (4), we get the following linear function:

$$\ln Q_{ij} = (x_{it} - x_{jt})'\theta \quad (18)$$

Thus, the log of the odds is expressed as a linear function of the differences of the variables. For the formulation of equation (6) we get:

$$Q_{ij} = \prod_{k=1}^K \left(\frac{x_{itk}}{x_{jtk}} \right)^{\theta_k} \quad (19)$$

and the odds are expressed as a product function of the ratios of the variables. Note that if the k^{th} variable is alternative i specific, then in equation (18) we get,

$$x_{itk} - x_{jtk} = x_{itk}$$

and in (19) we get,

$$\frac{x_{itk}}{x_{jtk}} = x_{itk}$$

Elasticities of the Logit Model

A direct elasticity is defined as:

$$E_{x_{itk}}^{P(i:A_t)} = \frac{\partial P(i:A_t)}{\partial x_{itk}} \cdot \frac{x_{itk}}{P(i:A_t)} \quad (20)$$

A cross-elasticity is defined as:

$$E_{x_{jtk}}^{P(i:A_t)} = \frac{\partial P(i:A_t)}{\partial x_{jtk}} \cdot \frac{x_{jtk}}{P(i:A_t)} \quad (21)$$

For the logit model as in equation (4), equations (20) and (21) result in the following relationships:

$$E_{X_{itk}}^{P(i:A_t)} = [1 - P(i:A_t)] \cdot \theta_k \cdot X_{itk} \quad (22)$$

$$E_{X_{jtk}}^{P(i:A_t)} = - P(j:A_t) \cdot \theta_k \cdot X_{jtk} \quad (23)$$

In general, we can write:

$$E_{X_{jtk}}^{P(i:A_t)} = [\delta_{ij} - P(j:A_t)] \cdot \theta_k \cdot X_{jtk} \quad (24)$$

where:

$$\delta_{ij} \sim \begin{cases} 1 & , \text{ for } i=j \\ 0 & , \text{ for } i \neq j \end{cases}$$

If variable X_{itk} enters the model as $\log X_{itk}$ then the elasticities are as given above divided by X_{itk} .

The direct elasticity, equation (22), varies from zero when the choice probability is 1, to $\theta_k X_{itk}$ when the choice probability is 0. (If a variable enters the model in a log form, the upper limit of the elasticity will be a constant equal to the coefficient θ_k .) If X_{itk} is a price variable we expect a negative direct elasticity and therefore θ_k must be negative.

The cross-elasticity, equation (23), is dependent only on values related to alternative j and not to alternative i for which the cross-elasticity is computed. This means that the cross-elasticities of all alternatives with respect to an attribute of alternative j are equal. This pattern of equal substitutability is equivalent to the independence

from irrelevant alternatives property. However, it provides another viewpoint to observe again the basic problem that results from the assumption of independent utilities as discussed in Chapter III. The cross-elasticity varies from zero when the choice probability of alternative j is 0, to $-\theta_k X_{jtk}$ when the same choice probability equals 1. Since θ_k must be negative, as noted before, the cross-elasticity will be positive as expected.

Aggregate Elasticity

For a ranked alternative it is possible to calculate an aggregate elasticity. It is defined with respect to the average choice probability as follows:

$$\bar{P}_i = \frac{\sum_{t=1}^T P(i:A_t)}{T} \quad (25)$$

Assume that,

$$X_{jtk} = X_{jk} \quad \text{for } t=1, \dots, T$$

Then,

$$\bar{E}_{X_{jk}} = \frac{\partial \bar{P}_i}{\partial X_{jk}} \cdot \frac{X_{jk}}{\bar{P}_i}$$

$$= \frac{\sum_{t=1}^T P(i:A_t) \cdot E_{jk}^{P(i:A_t)}}{\sum_{t=1}^T P(i:A_t)} \quad (26)$$

Note that the aggregate elasticity is a weighted sum of the individual elasticities using the individual probabilities as the weights. As a general rule, it can be shown that the elasticity of a sum of functions is the weighted sum of the elasticities of the functions with the functions as the weights. It can also be shown that the elasticity of a product of functions is the (unweighted) sum of the elasticities of the functions.*

*These two rules can be used to compute the elasticity of the joint probability in a recursive structure (for an example, see Manheim, 1970), and the elasticities for the volumes of trips discussed in Chapter IV.

Estimation Technique

For a cross section of behavioral units that are making a choice we do not observe the probabilities but only the actual choices. Hence, the observed dependent variable with disaggregate data takes a value of either 0 or 1. The independent variables are continuous and/or discrete. With observed 0,1 dependent variables, we use the maximum likelihood method to estimate the model coefficients. The estimator of the θ 's of the logit model as in equation (4) and its properties are described in McFadden (1968).*

The likelihood function of a disaggregate sample is written as follows:

$$L = \prod_{t=1}^T \prod_{i \in A_t} P(i:A_t)^{g_{it}} \quad (27)$$

where T is the number of observations and g_{it} equals 1 if alternative i was chosen in observation t , and 0 otherwise. Taking the log of both sides we get:

*If we aggregate groups of behavioral units according to socio-economic categories or geographic location and use the groups as the observation units, then the observed dependent variable is a share with a value between 0 and 1. With aggregated data both a maximum likelihood and ordinary least squares techniques can be used. The ordinary least squares technique is based on the linearization of the model by dividing the probability of each alternative by a "base" alternative, as in equation (18). The estimated coefficients are, therefore, sensitive to the choice of a "base" alternative, except for binary choice. See for example McLynn and Woronka (1969). A maximum likelihood estimator with aggregate data has not been used but in principle it is a simple extension of the disaggregate estimator.

$$L^* = \sum_{t=1}^T \sum_{i \in A_t} g_{it} \cdot \log P(i:A_t) \quad (28)$$

Substituting equation (4) for $P(i:A_t)$, we get for the first order conditions*:

$$\frac{\partial L^*}{\partial \theta_k} = \sum_{t=1}^T \sum_{i \in A_t} [g_{it} - P(i:A_t)] \cdot x_{itk} = 0 \quad (29)$$

for $k = 1, \dots, K$

McFadden (1968) showed that only under very specific conditions in the data, the maximum of L^* obtained from (29) is not unique. This estimator has optimal asymptotic properties. Minus the inverse of the matrix of second derivatives of L^* is the asymptotic variance-covariance matrix (Theil, 1971).

It should be noted that since we do not observe probabilities, it will be misleading to compare the computed probabilities with the g_{it} variables if we assume that the actual choice is made with a probability and not with certainty as the g_{it} variables would indicate. Therefore, a "goodness of fit" measure such as R^2 in ordinary least squares which is based on estimated residuals does not exist. In addition, a comparison of a sum of probabilities for a given ranked alternative with the total number of observations that chose this

*For more detail see McFadden (1968). The equations in (29) are nonlinear and their solution requires an iterative procedure. In this study we use the Newton-Raphson method. The estimation computer program that was used was developed by C. F. Manski at the Department of Economics at M.I.T.

alternative, is also misleading under the following conditions. If the set of variables includes an alternative specific dummy variable as follows:

$$x_{jtk} = \begin{cases} \text{Const., for } j = i \\ 0 \quad , \text{ for } j \neq i \end{cases} \quad (30)$$

then from the first order conditions, equation (29), the following always hold:

$$\sum_{t=1}^T g_{it} = \sum_{t=1}^T P(i:A_t) \quad (31)$$

Thus, for a disaggregate logit model because we cannot estimate residuals there is no statistic such as R^2 that will indicate how well the model "fits" the data. However, it is possible to define a measure analogous to R^2 , which is based on the value of the log likelihood function, and can be used to compare alternative models, as follows:

$$f^2 = 1 - \frac{L^*(\hat{\theta})}{L^*(0)} \quad (32)$$

where $L^*(\hat{\theta})$ is the value of L^* for the vector of estimated coefficients, and $L^*(0)$ is the value of L^* for $\theta=0$. Thus, f^2 is equal to the ratio of the explained log likelihood over the total log likelihood, and it lies between 0 and 1. This measure suffers from the same deficiencies as R^2 . In particular, it does not take account of the degrees of freedom. Adjusting for the degrees of freedom, we get:

$$\bar{f}^2 = 1 - \frac{\frac{L^*(\hat{\theta})}{T} / \sum_{t=1}^T (J_t - 1) - K}{\frac{L^*(0)}{T} / \sum_{t=1}^T (J_t - 1)} \quad (33)$$

where J_t is the number of alternatives in A_t . Setting $\theta=0$ amounts to an assumption of equally likely alternatives. However, for ranked alternatives, we can compute $L^*(0)$ that takes into account pure alternative effects.

Consider the simultaneous model in which we estimate directly $P(m,d:DM)$, and a recursive structure in which we estimate separately two probabilities, say $P(d:D)$ and $P(m:M_d)$. The problem is how to compare the goodness of fit between the recursive and the simultaneous structures. For the simultaneous model we can compute ϕ^2 directly. For the recursive model we have to compute the joint likelihood of $P(d:D) \cdot P(m:M_d)$, in order to be able to calculate a comparable ϕ^2 . The joint log likelihood of the recursive model can be written as follows:

$$\begin{aligned}
 L^*_{dm} &= \sum_{t=1}^T \sum_{d \in DM_t} g_{dmt} \cdot \log [P(d:D_t) \cdot P(m:M_{dt})] \\
 &= \sum_{t=1}^T \sum_{d \in D_t} \sum_{m \in M_{dt}} g_{dmt} \cdot [\log P(d:D_t) + \log P(m:M_{dt})] \\
 &= \sum_{t=1}^T \sum_{d \in D_t} g_{dt} \cdot \log P(d:D_t) \\
 &\quad + \sum_{t=1}^T \sum_{m \in M_{d't}} g_{d'mt} \cdot \log P(m:M_{d't}) \\
 &= L^*_d + L^*_{m|d}
 \end{aligned} \tag{34}$$

where d' is the chosen destination in observation t , and L^*_d and $L^*_{m|d}$ are the log likelihood values from the two separate models. Thus, the

joint log likelihood function of a recursive model is equal to the sum of the log likelihood functions from the structural probabilities.

An asymptotic t test can be used to test the significance of a single coefficient. The significance of a group of coefficients can be tested using the statistic of minus twice the log of a likelihood ratio that is asymptotically distributed as Chi-Square (Theil, 1971).

Application of the Logit Model
to a Simultaneous Travel Choice Model

The overall travel demand model includes several choices: trip frequency, destination, mode of travel, time of day, and route. In what follows we consider for simplicity only the first three choices. We assume the choices are made simultaneously. Hence, the set of alternatives includes all the possible combinations of the various trip frequencies, destinations and modes. For example, assume that there are two trip frequencies (one trip and no trip), four relevant destinations, and two possible modes of travel. The total number of alternatives will be nine (eight possible combinations of modes and destinations if a trip is made and the additional alternative of no trip). If we double the number of destinations and modes we have a total of 33 alternatives.

Denote frequency - f, destination - d, mode - m. The choice model should give the probability of choosing any FDM combinations out of the set of all possible combinations. Denoting this set as FDM, we can write the logit model, equation (4), as follows:

$$P(fdm:FDM_t) = \frac{e^{x_{fdmt}'\theta}}{\sum_{ijk \in FDM_t} e^{x_{ijk}'\theta}} \quad (35)$$

To simplify, we can drop the subscript t which represents a behavioral unit or an observation and write the model as:

$$P(fdm:FDM) = \frac{e^{x_{fdm}'\theta}}{\sum_{ijk \in FDM} e^{x_{ijk}'\theta}} \quad (36)$$

This choice model is very complex and the number of alternatives in FDM can be very large. However, in principle, it could be estimated directly and applied directly. In what follows, the indirect structure of the simultaneous model will be derived. It will be shown that the model of equation (36) can be expressed as a sequence

of three logit models that could also be estimated and applied in a sequence. However, the sequence assumption that is involved is not behavioral and therefore not reflected in the structure of the model. It is only needed in order to derive an indirect structure. Note that a sequential estimation procedure for a simultaneous model is suggested only if it is infeasible to estimate directly the joint probability.

We can separate this simultaneous model into its three component choices. The problem here is to decide on the sequence in which we create the various subsets of FDM. Since the number of choices is three, the total number of possible sequences is $6(3! = 6)$. However, it is "natural" to start by grouping the alternatives according to the trip frequency - f.* This limits the number of sequences to only two: group by f and then by d, or group by f and then by m. Therefore, we consider the following subsets of FMD: (We assume that the relevant

*It is "natural" because the alternatives in f include a "null" alternative which is not to make a trip at all. For this alternative we cannot associate alternative destinations or modes. In addition, the number of alternative frequencies that are observed during the conventional sampling period of 24 hours for a specific trip purpose is generally only zero, one, and in limited cases two. If we consider only zero and one then for each combination of mode and destination there is only one possible f alternative.

modes are not the same for all destinations or that the relevant destinations are not the same for all modes.)

For both sequences,

$F = \text{the set of all relevant trip frequencies}$
 $= [1, \dots, f, \dots, i, \dots, F]$

For the first sequence,

$D_f = \text{the set of all relevant destinations for trip frequency } f$
 $= [1, \dots, d, \dots, j, \dots, D_f]$

$M_{df} = \text{the set of all relevant travel modes for destination } d \text{ and trip frequency } f$
 $= [1, \dots, m, \dots, k, \dots, M_{df}]$

and for the second sequence,

$M_f = \text{the set of all relevant modes for trip frequency } f$
 $= [1, \dots, m, \dots, k, \dots, M_f]$
 $D_{mf} = \text{the set of all relevant destinations for travel mode } m \text{ and trip frequency } f$
 $= [1, \dots, d, \dots, j, \dots, D_{mf}]$

For the first sequence we can now separate equation (36) in the following ways:^{*}

*See the discussion of the choice axiom for the logit model in a previous section.

$$P(fdm: FDM) =$$

$$\frac{e^{X_{fdm}'\theta}}{\sum_{k \in M_{df}} e^{X_{fdk}'\theta}} \cdot \frac{\sum_{k \in M_{df}} e^{X_{fdk}'\theta}}{\sum_{j \in D_f} \sum_{k \in M_{jf}} e^{X_{fjk}'\theta}} \cdot \frac{\sum_{j \in D_f} \sum_{k \in M_{jf}} e^{X_{fjk}'\theta}}{\sum_{i \in F} \sum_{j \in D_i} \sum_{k \in M_{ji}} e^{X_{ijk}'\theta}}$$

$$= P(m:M_{df}) \cdot P(d:D_f) \cdot P(f:F) \quad (37)$$

For the second sequence, in a similar manner, we write:

$$P(fdm: FDM) =$$

$$\frac{e^{X_{fdm}'\theta}}{\sum_{j \in D_{mf}} e^{X_{fjm}'\theta}} \cdot \frac{\sum_{j \in D_{mf}} e^{X_{fjm}'\theta}}{\sum_{k \in M_f} \sum_{j \in D_{kf}} e^{X_{fjk}'\theta}} \cdot \frac{\sum_{k \in M_f} \sum_{j \in D_{kf}} e^{X_{fjk}'\theta}}{\sum_{i \in F} \sum_{k \in M_i} \sum_{j \in D_{ki}} e^{X_{ijk}'\theta}}$$

$$= P(d:D_{mf}) \cdot P(m:M_f) \cdot P(f:F) \quad (38)$$

So far, we have shown how a simultaneous demand model in the form of the multinomial logit model can be separated into its component choices. We have separated the choice model into three separate choices: trip frequency, destination, and travel mode. The first choice in the sequence, i.e., mode choice in equation (37) and destination choice in equation (38), is expressed in the form of the multinomial logit model. The two other choices, for both sequences, are expressed using terms

that have already appeared in the first probability. This observation may lead to a wrong conclusion that if one estimates a mode choice model, he can then derive from it the frequency and destination choice models. It will be apparent from what follows that the expression of equation (36) in the form of equation (37) or (38) does not provide any useful information unless we are willing to specify in more detail the explanatory variables, i.e., the vector X_{fdm} .

When we defined the vector X_{it} in equation (4) we assumed that the variables are different from one alternative to another. This means that the variables X_{fdm} have to be different for all the alternatives in FDM. However, this assumption does not hold with respect to all the variables in X_{fdm} . A variable like price may indeed be different among all travel alternatives, but a variable which describes a given destination will be equal for all modes of travel. The following is an exhaustive list of possible variations among the explanatory variables:

X_{fdm} - variables which differ among all alternatives

X_f - variables which differ only among trip frequencies

X_d - variables which differ only among destinations

X_m - variables which differ only among travel modes

X_{fd} - variables which differ only among trip frequencies and destinations.

X_{fm} - variables which differ only among trip frequencies and travel modes.

X_{dm} - variables which differ only among destinations and travel modes.

An example of a variable that could be subscripted as X_{fdm} is travel time. It varies among modes and destinations and also among frequencies. (If a trip is not taken, the travel time is zero.) A variable which could be characterized as X_f is a frequency specific income variable. Most destinations are unranked alternatives, but if the CBD is one of the destinations, it is possible to have a CBD dummy variable which will be characterized as X_d . On the other hand, travel modes (and also trip frequencies) are ranked alternatives. Therefore, any mode specific variable which takes a certain value for a given mode and zero otherwise is an X_m variable. X_{fd} variables are all those that describe the opportunities for a given trip purpose at the destinations. (If a trip is not taken, the opportunities are zero.) For the last two possibilities, X_{fm} and X_{dm} , there does not seem to exist any example. We will keep them in what follows for completeness.

Using the above classification of the explanatory variables, we can now write equation (36) as follows:

$$P(fdःFDM) = \frac{\sum e^{X_{fdm}'\theta^{fdm} + X_f'\theta^f + X_d'\theta^d + X_m'\theta^m + X_{fd}'\theta^{fd} + X_{fm}'\theta^{fm} + X_{dm}'\theta^{dm}}}{\sum e^{X_{ijk}'\theta^{fdm} + X_i'\theta^f + X_j'\theta^d + X_k'\theta^m + X_{ij}'\theta^{fd} + X_{ik}'\theta^{fm} + X_{jk}'\theta^{dm}}} \quad (39)$$

$$\sum_{ijk \in FDM} e^{X_{ijk}'\theta^{fdm} + X_i'\theta^f + X_j'\theta^d + X_k'\theta^m + X_{ij}'\theta^{fd} + X_{ik}'\theta^{fm} + X_{jk}'\theta^{dm}}$$

where the vectors X_{fdm} and θ of equation (36) are now divided into seven vectors according to the classification of the explanatory variables.

Given this model, the first probability in equation (37) is reduced to the following:

$$P(m:M_{df}) =$$

$$\frac{e^{X_{fdm}'\theta^{fdm} + X_m'\theta^m + X_{fm}'\theta^{fm} + X_{dm}'\theta^{dm}}}{\sum_{k \in M_{df}} e^{X_{fdk}'\theta^{fdm} + X_k'\theta^m + X_{fk}'\theta^{fm} + X_{dk}'\theta^{dm}}} \quad (40)$$

$$\sum_{k \in M_{df}} e^{X_{fdk}'\theta^{fdm} + X_k'\theta^m + X_{fk}'\theta^{fm} + X_{dk}'\theta^{dm}}$$

The terms that appear in the total demand model but do not appear in the above mode choice model were cancelled out because they are equal for all the alternatives in M_{df} . Now it is also clear why one cannot recover all the parameters of the total demand model by estimating only the mode choice model. One can only estimate the parameters of those variables that vary across modes.

The second probability in the first sequence, equation (37), is the destination choice model, which can now be written as:

$$P(d:D_f) =$$

(41)

$$\frac{\sum_{k \in M_{df}} e^{X_{fdk}'\theta^{fdm} + X_d'\theta^d + X_k'\theta^m + X_{fd}'\theta^{fd} + X_{fk}'\theta^{fm} + X_{dk}'\theta^{dm}}}{\sum_{j \in D_f} \sum_{k \in M_{jf}} e^{X_{fjk}'\theta^{fdm} + X_j'\theta^d + X_k'\theta^m + X_{fj}'\theta^{fd} + X_{fk}'\theta^{fm} + X_{jk}'\theta^{dm}}}$$

The only term in the total demand that does not appear here is the term of X_f which is equal for all the alternatives in D_f . Since not all the terms in equation (41) are mode specific, we can take them outside the sum over modes and rewrite equation (41) as follows:

$$P(d:D_f) = \frac{e^{X_d' \theta^d + X_{fd}' \theta^{fd}}}{\sum_{k \in M_{df}} e^{X_{fdk}' \theta^{fdm} + X_k' \theta^m + X_{fk}' \theta^{fm} + X_{dk}' \theta^{dm}}} \quad (42)$$

$$\frac{\sum_{j \in D_f} e^{X_j' \theta^d + X_{fj}' \theta^{fd}}}{\sum_{k \in M_{jf}} e^{X_{fjk}' \theta^{fdm} + X_k' \theta^m + X_{fk}' \theta^{fm} + X_{jk}' \theta^{dm}}}$$

This equation is still not in the form of the multinomial logit model.

Define,

$$P_{fjk}^M = e^{X_{fjk}' \theta^{fdm} + X_k' \theta^m + X_{fk}' \theta^{fm} + X_{jk}' \theta^{dm}} \quad (43)$$

and,

$$P_{fj}^M = \sum_{k \in M_{jf}} P_{fjk}^M \quad (44)$$

If we have estimated the mode choice model of equation (40), we already have all the parameters of P_{fjk}^M . P_{fd}^M is equal to the denominator of equation (40), and P_{fdm}^M is the numerator. Hence, equation (40) can

now be expressed as:

$$P(m:M_{df}) = \frac{P_{fdm}^M}{P_{fd}^M} \quad (45)$$

Substituting the P terms in equation (42) we get:

$$P(d:D_f) = \frac{\frac{x_d' \theta^d + x_{fd}' \theta^{fd}}{e} \cdot P_{fd}^M}{\sum_{j \in D_f} \frac{x_j' \theta^d + x_{fj}' \theta^{fd}}{e} \cdot P_{fj}^M} \quad (46)$$

or, alternatively,

$$P(d:D_f) = \frac{\frac{x_d' \theta^d + x_{fd}' \theta^{fd} + \ln P_{fd}^M}{e}}{\sum_{j \in D_f} \frac{x_j' \theta^d + x_{fj}' \theta^{fd} + \ln P_{fj}^M}{e}} \quad (47)$$

Equation (47) is in the form of the multinomial logit model and could be estimated. However, we have one variable $\ln P_{fj}^M$ with a coefficient that should be constrained to equal one.* Estimating this

*The relaxation of this constraint implies a departure from the multinomial logit form of the total demand model. Later we will analyze a model in which this coefficient is allowed to vary from unity.

model is still not sufficient to recover all the parameters of the demand model. The only parameters that have not been estimated yet are θ^f . We can estimate θ^f only with a frequency choice model. But before we write the frequency choice model, we define some more variables to simplify the equations.

Define,

$$P_{ij}^D = e^{x_j' \theta^d + x_{ij}' \theta^{fd} + \ln P_{ij}^M} \quad (48)$$

and,

$$P_i^D = \sum_{j \in D_i} P_{ij}^D \quad (49)$$

Again, all the parameters of P_{fj}^D are known from the destination choice model.

Equation (47) can be written in a simplified form as:

$$P(d:D_f) = \frac{P_{fd}^D}{\sum_{f \in F} P_f^D} \quad (50)$$

Using the variables that we have defined, the frequency choice model of equation (37) can now be reduced to the following:

$$P(f:F) = \frac{e^{x_f' \theta^f + \ln P_f^D}}{\sum_{i \in F} e^{x_i' \theta^f + \ln P_i^D}} \quad (51)$$

Equation (51) is a trip frequency choice model which is in the form of the multinomial logit model. Again the coefficient of one variable is constrained to be one.

Using the same notation as before, we can now define,

$$P_i^F = e^{X_i' \theta^F + \ln P_i^D} \quad (52)$$

and,

$$P^F = \sum_{i \in F} P_i^F \quad (53)$$

With these definitions, equation (51) can now be written as,

$$P(f:F) = \frac{P_f^F}{P^F} \quad (54)$$

To summarize, we have shown that the simultaneous model of equation (39) can be expressed as a multiplication of three separate choice models, each having the form of the multinomial logit model.

Equation (37) can now be written in the following form:

$$P(fdm:FDM) = \frac{P_{fdm}^M}{\sum_{k \in M_{df}} P_{fdk}^M} \cdot \frac{P_{fd}^D}{\sum_{j \in D_f} P_{fj}^D} \cdot \frac{P_f^F}{\sum_{i \in F} P_i^F} \quad (55)$$

where:

$$P_{fdm}^M = e^{X_{fdm}'\theta^{fdm}} + X_m'\theta^m + X_{fm}'\theta^{fm} + X_{dm}'\theta^{dm}$$

$$P_{fd}^D = e^{X_d'\theta^d} + X_{fd}'\theta^{fd} + \ln \sum_{k \in M_{df}} P_{fdk}^M$$

$$P_f^F = e^{X_f'\theta^f} + \ln \sum_{j \in D_f} P_{fj}^D$$

We can estimate each model separately by starting with the mode choice model. Using the estimates of the parameters obtained in this stage we create the variables P_{fdm}^M and estimate the destination choice model. The estimates of the parameters of this model are used to create the variables P_{fd}^D which are used to estimate the trip frequency choice model. The procedure for the second sequence, equation (38), is essentially the same and there is no need to repeat the derivation. Mathematically the results are identical.

Equation (55) is mathematically equivalent to equation (39) from which we have started. However, we expect that the coefficient estimates will not be identical whether we estimate directly equation (39) or in sequence equation (55). Furthermore, we expect that the different possible sequences of estimation will also result in different coefficient estimates. If due to practical limitations a simultaneous model could not be estimated directly, a sequential estimation is feasible. But the coefficient estimates from a sequential estimation will be useful only if it is shown that the estimated coefficients are not highly sensitive to the structure used for estimation.

Application of the Logit Model to a Recursive Travel Choice Model

In this section we assume that travel choices are made sequentially.

Assume, for the purpose of presentation, the following sequence:

$$f \longrightarrow d \longrightarrow m \quad (56)$$

Hence, the potential traveller is assumed to consider first how often to make a trip for a specific trip purpose, then, secondly, he decides where to go given that he takes a trip, and then, finally he selects a mode given that he makes a trip to a specific destination.

Since the mode choice is the last in the sequence there is no problem in writing the model for this step. It will be identical to equation (40), but now the coefficients will be only specific to this model and will be written with a subscript denoting the given choice model. Hence, we write the mode choice model as follows:

$$P(m:M_{df}) = \frac{\sum_e x_{fdm}' \theta_M^{fdm} + x_m' \theta_M^m + x_{fm}' \theta_M^{fm} + x_{dm}' \theta_M^{dm}}{\sum_{k \in M_{df}} x_{fdk}' \theta_M^{fdm} + x_k' \theta_M^m + x_{fk}' \theta_M^{fm} + x_{dk}' \theta_M^{dm}} \quad (57)$$

However, formulating the destination choice model requires a further assumption about the specification of the X variables with mode as one of the subscripts. Several alternative ways of creating these composite variables were already discussed in Chapter IV.

Assume for the moment that the composition rule is known and write any composite variable with an asterisk replacing the subscript over which the composition was performed. For example, X_{fd^*} will be the composite variable of X_{fdm} aggregated over modes. Hence we can write the destination choice model as follows:

$$P(d:D_f) = \frac{\sum_e X_{fd^*} \theta_D^{fdm} + X_d \theta_D^d + X_{fd} \theta_D^{fd} + X_{d^*} \theta_D^{dm}}{\sum_{j \in D_f} e X_{fj^*} \theta_D^{fdm} + X_j \theta_D^d + X_{fj} \theta_D^{fd} + X_{j^*} \theta_D^{dm}} \quad (58)$$

The frequency choice model is:

$$P(f:F) = \frac{\sum_e X_{f^**} \theta_F^{fdm} + X_f \theta_F^f + X_{f^*} \theta_F^{fd} + X_{f^*} \theta_F^{fm}}{\sum_{i \in F} e X_{i^**} \theta_F^{fdm} + X_i \theta_F^f + X_{i^*} \theta_F^{fd} + X_{i^*} \theta_F^{fm}} \quad (59)$$

Note that the destination and the frequency choice models do not include all the variables that were included in the same models that were derived from the simultaneous model. Specifically, all the variables that are specific only to one choice appear only in that choice model, while before a variable such as X_m appeared in all choice models. On the other hand, variables specific to more than one choice appear in all their corresponding choices, but with possibly a

different set of parameters. For example, the variables denoted as

x_{fdm} appear in each choice model with a different set of parameters:

θ_M^{fdm} , θ_D^{fdm} , and θ_F^{fdm} . This means that the marginal rate of substitution (or trade off) between any two variables, say travel time and travel cost, may be different in the various choice models. In other words, it implies, for example, that the traveller possesses implicitly a different value of time for different choices.

This property of different marginal rates of substitution among the same variables for different choices appears unreasonable from a theoretical point of view. A recursive model such as the one presented above implies an assumption about the additivity of the utility function, where each additive term corresponds to a single choice.* The different "prices" of travel enter each separate utility function.** It might be reasonable to assume that the "price effects" are different for the different choices. However, since we are dealing with several "prices" for each alternative, such as travel time and travel cost, the relative "price effects" will also be different in the recursive model that was written above. From a theoretical point of view we regard a trip as a factor input into the production of the final activity at the destination. The utility is derived only from the final activity. The above prices, say travel time and cost, are resources required

*See the discussion of the alternative structures in Chapter IV.

**Note that the exponent function on the logit model is interpreted as an indirect utility function.

to produce a trip. The total expenditure on a trip is therefore the travel cost plus travel time, multiplied by a coefficient that was termed "value of time". By dividing the utility function into additive terms for each component choice and maintaining a single expenditure function the marginal rate of substitution is entirely determined by the expenditure function and is therefore unique.

This leads to the creation of the concept of "generalized price" (or, "inclusive price" by CRA, 1972). The various price variables, primarily those denoted as X_{fdm} , are combined together with a unique set of coefficients. Consequently, the vector of variables X_{fd*} and X_{f**} are only single variables in the destination and frequency choice models, respectively. Thus, when we create the composite variables from any vector of X variables we maintain their coefficients from the more detailed choice model and aggregate together the "generalized" variable. With this procedure all the vectors of composite variables that include an asterisk in their subscripts are now only a single variable instead of a vector.

We can now write the destination choice model, equation (58), as follows:

$$P(d:P_f) = \frac{e^{\tilde{X}_{fd*} \cdot \tilde{\theta}_D^{fdm} + X_d' \theta_D^d + X_{fd}' \theta_D^{fd} + \tilde{X}_{d*} \cdot \tilde{\theta}_D^{dm}}}{\sum_{j \in D_f} e^{\tilde{X}_{fj*} \cdot \tilde{\theta}_D^{fdm} + X_j' \theta_D^d + X_{fj}' \theta_D^{fd} + \tilde{X}_{j*} \cdot \tilde{\theta}_D^{dm}}} \quad (60)$$

where \sim denotes a single variable or coefficient. The frequency choice model become:

$$P(f:F) = \frac{e^{\sum_{i \in F} \tilde{x}_{i*} \cdot \tilde{\theta}_F^{fdm} + x_i' \theta_F^f + \tilde{x}_{i*} \cdot \tilde{\theta}_F^{fd} + \tilde{x}_{i*} \cdot \tilde{\theta}_F^{fm}}}{\sum_{i \in F} e^{\sum_{i \in F} \tilde{x}_{i*} \cdot \tilde{\theta}_F^{fdm} + x_i' \theta_F^f + \tilde{x}_{i*} \cdot \tilde{\theta}_F^{fd} + \tilde{x}_{i*} \cdot \tilde{\theta}_F^{fm}}} \quad (61)$$

In these models the marginal rates of substitution are constant but the relative effect of a group of variables that appear in more than one model can be different. This procedure was used by CRA (1972).

If we regard the exponent in the mode choice model, equation (57), as a utility, or expenditure, function associated with a mode, we can create a composite variable by aggregating this function as a single variable. In this case, we interpret the composite variable as an expected utility, or expenditure, from all modes to a given destination. We can do the same in creating the composite variables for the frequency choice model. In this case, for any group of X variables with the same subscripts we associate only one set of parameters. Hence, we can drop the subscript for θ that denotes the choice model.

Define,

$$GP_{fdm}^M = x_{fdm}' \theta^{fdm} + x_m' \theta^m + x_{fm}' \theta^f + x_{dm}' \theta^{dm} \quad (62)$$

where GP_{fdm}^M is the utility, or expenditure, of mode m given fd. Each composite variable is now formed from several groups of X variables.

In writing the models for this procedure we also present a specific

composition rule - expected value - which is consistent with the interpretation of the composite variable as expected utility. The expected utility denoted as GP_{fd*}^M is computed as follows:^{*}

$$GP_{fd*}^M = \sum_{k \in M_{df}} P(k:M_{df}) \cdot GP_{fdk}^M \quad (63)$$

The mode choice model, equation (57), can now be written as:

$$P(m:M_{df}) = \frac{\sum_e GP_{fdm}^M}{\sum_{k \in M_{df}} GP_{fdk}^M} \quad (64)$$

The destination choice model is written as:

$$P(d:D_f) = \frac{\sum_e x_d' \theta^d + x_{fd}' \theta^{fd} + GP_{fd*}^M \cdot \delta_D}{\sum_{j \in D_f} x_j' \theta^d + x_{fj}' \theta^{fd} + GP_{fj*}^M \cdot \delta_D} \quad (65)$$

where δ_D is the coefficient of the composite variable.

Define,

$$GP_{fd}^D = x_d' \theta^d + x_{fd}' \theta^{fd} + GP_{fd*}^M \cdot \delta_D \quad (66)$$

and,

$$GP_{f*}^D = \sum_{j \in D_f} P(j:D_f) \cdot GP_{fj}^P$$

^{*}Note that this is one of the alternative ways to create the composite variables discussed in Chapter IV. This composition rule was used by CRA (1972) to create the composite variables in the procedure of equations (60) and (61).

Equation (65) can be written as:

$$P(d:D_f) = \frac{e^{GP_{fd}^D}}{\sum_{j \in D_f} e^{GP_{fj}^P}} \quad (67)$$

The frequency choice model is:

$$P(f:F) = \frac{e^{X_f \theta^f + GP_{f*}^D \cdot \delta_F}}{\sum_{i \in F} e^{X_i \theta^f + GP_{i*}^D \cdot \delta_F}} \quad (68)$$

where δ_F is the coefficient of the composite variable.

Define,

$$GP_f^F = X_f \theta^f + GP_{f*}^D \cdot \delta_F \quad (69)$$

and,

$$GP_*^F = \sum_{i \in F} P(i:F) \cdot GP_i^F \quad (70)$$

Equation (68) can be written as:

$$P(f:F) = \frac{e^{GP_f^F}}{\sum_{i \in F} e^{GP_i^F}} \quad (71)$$

Note that this recursive model is very similar to the indirect structure of the simultaneous model described in the previous section, equation (55). GP_{fdm}^M corresponds with P_{fdm}^M , GP_{fd}^D with P_{fd}^D , and GP_f^F with P_f^F . (P_{fdm}^M is an exponential transformation of GP_{fdm}^M .) The aggregation of the P variables corresponds to a different composition rule, a simple sum vs. the weighted sum that was used here. Thus, writing the indirect structure of the simultaneous model without the unity constraints corresponds to an alternative formulation of a recursive model. However, the formulation used in this section appears to be more sensible because of the computations of expected values, rather than simple sum.*

The relaxation of the unity constraints for the coefficients of $\ln P_{fd}^M$ in the destination choice model and $\ln P_f^D$ in the frequency choice model is equivalent to rewriting equation (55) as follows:

$$P(fdm:FDM) = \frac{P_{fdm}^M}{\sum_{k \in M_{df}} P_{fdk}^M} \quad \frac{(P_{fd}^D)^{\hat{\delta}_D}}{\sum_{j \in D_f} (P_{fj}^D)^{\hat{\delta}_D}} \quad \frac{(P_f^F)^{\hat{\delta}_F}}{\sum_{i \in F} (P_i^F)^{\hat{\delta}_F}} \quad (72)$$

where we introduced two new parameters $\hat{\delta}_D$ and $\hat{\delta}_F$, which will also have to be estimated. Equation (72) is equivalent to one of the Special Product Models suggested by Manheim (1972).

The mode choice model is unchanged. The destination choice model is now,

*A simple sum can be taken to represent equal weights, which are in order only under the assumption that for a higher level decision (e.g., destination choice) lower level alternatives (e.g., modes) are considered equally likely to be chosen.

$$P(d:D_f) = \frac{\hat{\delta}_D \cdot x_d' \theta^d + \hat{\delta}_D \cdot x_{fd}' \theta^{fd} + \hat{\delta}_D \cdot \ln P_{fd}^M}{\sum_{j \in D_f} e^{\hat{\delta}_D \cdot x_j' \theta^d + \hat{\delta}_D \cdot x_{fj}' \theta^{fd} + \hat{\delta}_D \cdot \ln P_{fj}^M}} \quad (73)$$

The destination choice model has now also a multinomial logit specification but with no constraints on the coefficient of $\ln P_{fj}^M$. The coefficient estimates of all the X variables will have to be divided by the estimate of $\hat{\delta}_D$ to get the parameters θ .

The trip frequency model is:

$$P(f:F) = \frac{e^{\hat{\delta}_F \cdot x_f' \theta^f + \hat{\delta}_F \cdot \ln P_f^D}}{\sum_{i \in F} e^{\hat{\delta}_F \cdot x_i' \theta^f + \hat{\delta}_F \cdot \ln P_i^D}} \quad (74)$$

The sequence that we assumed to derive equation (55) from equation (39) had no behavioral interpretation; it was just a way to group all the travel alternatives so that the total demand model could be factored to its component decisions. All possible sequences were mathematically equivalent because they could all be expressed as equation (39). However, from equation (72) we can no longer return to equation (39). The sequence in equation (72) has a behavioral significance because if we assume a different sequence, we will get another model which is no longer mathematically equivalent to the first one.

The total demand model of equation (72) (with $\hat{\delta}_D$ and $\hat{\delta}_F \neq 1$), assumes that the choices of trip frequency, destination and mode of travel are made sequentially and in that order. No behavioral sequence assumption was involved in the formulation of the model in equation (55).

An empirical "test of simultaneity" can be made by estimating the unconstrained model, equation (72), and performing a statistical test to determine whether $\hat{\delta}_D$ and $\hat{\delta}_F$ are each significantly different from 1. However, it only compares a recursive model with a sequential estimation of a simultaneous model.

It should be noted that in the previous section the definition of the P variables was dictated from the initial assumption of the mathematical form of the simultaneous demand model. In this section, we have started with a recursive structure, and assumed that each choice has a multinomial logit form but together they do not necessarily collapse to a multinomial logit total demand. In this case the composite variables were defined differently. They were based partly on parameter estimates from another choice model. Or alternatively, each choice model could be estimated independently.

It is interesting to note that the two composition rules used with the GP variables and with the P variables have equal derivatives with respect to the original linear function. This can be seen in the following example:

We defined:

$$GP_{fd*}^M = \sum_{k \in M_{df}} P(k:M_{df}) \cdot GP_{fdk}^M \quad (75)$$

The derivative is:

$$\frac{\partial GP_{fd*}^M}{\partial GP_{fdm}^M} = P(m:M_{df}) \quad (76)$$

We defined:

$$P_{fd}^M = \sum_{k \in M_{df}} P_{fdk}^M = \sum_{k \in M_{df}} e^{GP_{fdk}^M} \quad (77)$$

However, P_{fd}^M was used in the destination choice model in its log form.

Therefore, the corresponding derivative will be:

$$\begin{aligned} \frac{\partial \ln P_{fd}^M}{\partial GP_{fdk}^M} &= \frac{e^{GP_{fdm}^M}}{\sum_{k \in M_{df}} e^{GP_{fdk}^M}} \\ &= P(m:M_{df}) \end{aligned} \quad (78)$$

Thus, the log of the sum of e^{GP} or the weighted sum of GP are two composition rules that should give similar results. From a practical point of view the composition rule of the P variables is therefore superior because it does not necessitate the computation of all the probabilities.

Previous Applications of a Disaggregate Logit Model to Travel Demand

A binary mode choice or route choice model using the logit formulation has been estimated by Warner (1962), Stopher (1969), Thomas and Thompson (1971), and others. However, only recently has a multinomial logit model been applied. The first study was that of Rassam et al (1971), that estimated a mode choice model for more than two modes. However, the logit model that was used included the same set of alternatives for each behavioral unit. The first application of the logit model as expressed in equation (4) to travel demand was by CRA (1972).

CRA were the first to apply disaggregate models to the entire set of choices included in the travel demand function. However, all previous disaggregate models have been an application of the logit, or other choice model, to a single travel choice, or to a series of travel choices with each choice modelled separately. CRA used a logit model separately for mode, time of day, destination, and frequency of travel choices. Their model is similar to the recursive model described in the previous section with the weighted generalized price composite variables and it was formulated and estimated for the following sequence:

$$f \rightarrow d \rightarrow h \rightarrow m \quad (79)$$

where h denotes time-of-day choice.

Summary

This chapter has presented functional specifications of a simultaneous and of several recursive structures using the multinomial logit choice model. For the simultaneous model a sequential estimation

procedure was also described. The different recursive models represent alternative composition schemes. The alternative models that were presented are listed below:

Simultaneous:

Direct --- equation (39)

Indirect (for a particular sequence) -- equation (55)

Recursive (for a particular sequence):

Composition of single variables -- equations (57), (58) and (59)

Composition of groups of variables -- equations (57), (60) and (61)

Composition of utility functions -- equations (57), (65) and (68).

In addition, two composition rules were presented:

Weighted sum -- equations (63), (66) and (70)

Log of the denominator -- equation (55).

For the two rules we presented the models with composite utility functions:

Weighted sum -- equations (64), (67) and (71)

Log of the denominator -- equation (72).

The following chapter reports the results of an empirical study in which models were estimated for the various structures described in this chapter (with the exception that we model only the choices of mode and destination given that a trip is made).

CHAPTER VII

Estimation of Alternative ModelsIntroduction

This chapter describes an empirical study of alternative structures of travel demand models. The objectives of the empirical study are twofold. First, to investigate the feasibility of estimating a disaggregate simultaneous travel demand model, and second, to furnish some evidence on the sensitivity of travel predictions with respect to the structural specification of the demand model.* A priori considerations indicate that a simultaneous structure, rather than a recursive one, is appropriate for modelling travel choices with cross-sectional data.** It is not expected that this empirical study can conclusively prove or disprove this assertion.

It is not the objective of this study to develop a full set of models to be used in an actual planning context. Therefore, we will estimate the models only for a single trip purpose and only for a sub-set of the choices included in the travel demand function. Specifically, the choices of mode and destination for a shopping trip purpose were selected. The reasoning for this selection was the relative clarity

*See the discussion of the empirical problem in Chapter IV.

**See the discussion in Chapters II and IV.

of the decision making process with respect to these two choices. Other trip purposes, such as work trips, are characterized by more constraints on travellers' choices.* We assume a simultaneous choice of destination and mode given that a trip is made. The justification of this structure was discussed in detail in Chapter IV. In addition, the dependency on frequency cannot be modelled properly because the available data for this study consists of reported trips for a 24 hour period. The actual frequency is not reported.

Based on the objective of this study, which is to investigate alternative structures, and the available resources, it was decided to use only a small sample and to estimate models with different structures without detailed investigations of alternative variable specifications.

In the following section the data set that was available for Metropolitan Washington is described. Following this is a description of the selection of the subsample used in this study. We did not use all the available data, but required only a relatively small subsample which represents a suitable cross-section of household and trip characteristics. The data were kept in the disaggregate form where the observation unit is a household. This follows the assumption that the behavioral unit for a shopping trip is also a household.** The

*See the discussion in Chapter II.

**See the discussion in Chapter II and also Oi and Shuldiner (1962).

final data section describes the actual variables to be used in the empirical estimation of the models.

Then, estimation results of the alternative models that were described in the previous chapter are presented. With the same data set we estimate a simultaneous model, and several recursive models for the two possible sequences. All the models are based on an identical set of variables.

The Data Set

The basic source for this study was a data set prepared by R. H. Pratt Associates (RHP) for the Metropolitan Washington Council of Governments (WCOG). The data set was combined from a home interview survey conducted in 1968 by WCOG and level of service (or network) data assembled by WCOG and RHP.

The home interview survey followed the conventional practice of comprehensive data collection in urban transportation planning studies (Martin et al, 1961). This type of survey is based on field interviews of a random sample of households. In the WCOG survey the sampling rate was 3% inside the capital beltway, or one out of every 33 households, and 5% outside the beltway, or one out of every 20 households. In total about 30,000 households were interviewed. The survey is designed to collect factual information rather than opinions or attitudes. The information included in each interview questionnaire can be classified into three basic parts: socio-economic characteristics of the household, socio-economic characteristics of persons, and trip

information. The first two parts contain information on the household composition, income, car ownership, occupation, education, sex, age, race, and several other data items. The trip information consists of a complete listing, in sequence, of every trip made by every person in the household during a 24 hour period. A total of about 250,000 one-way trips were reported. Related to each trip, information is obtained on its characteristics, such as: trip purpose, mode of travel, origin and destination, time of day, etc. Information on the travel route is usually not obtained.

The level of service data were obtained primarily from a comprehensive inventory of existing transportation, highway and transit, facilities and services performed by WCOG. Additional level of service data were assembled by RHP.

The data set includes a record for each one-way trip. Each record consists of a set of variables describing: the trip maker, the household, the origin and destination zones, the trip and its characteristics and the characteristics of the same trip if the competing mode were to be used.

It should be noted that although the data are available on a disaggregated basis, some level of aggregation is present. It is due to the fact that some of the level of service data were derived from coded networks based on a traffic zone system. (The WCOG study area, 2,350 square miles and population of about 2,750,000, was divided into 1065 traffic zones varying in area according to population and employment densities.)

No data collection effort was undertaken as part of the present study. All the data used were taken from the data set described above and from tabulations from the home interview survey that were obtained from WCOG. As yet, the home interview survey and the supply data are not fully documented. However, later in this chapter, when the specific variables are presented, it will be briefly indicated how each data item was obtained.

The Subsample Used for Estimation

Based on the purposes of the empirical work in this study and the available resources, it was judged that a small subsample will be sufficient. In selecting the subsample it was decided to sample residents of a single sector rather than the entire metropolitan area. This ensures that the subsample represents a range of socio-economic characteristics and travel patterns.

The sector to the north of the Metropolitan Washington Central Business District (CBD) was selected.* Examination of data contained in a series of information reports issued by the Transportation Planning Board of WCOG, indicated that this sector contains a wide range of socio-economic characteristics and has a relatively high proportion of transit use.

*The boundaries of the sector were defined as follows: 16th St. and the continuation of Georgia Ave. beyond the district line on the west, and Georgia Ave., New Hampshire Ave. and the boundary of Montgomery County on the east.

This sector still included a very large number of households. Therefore, within this sector 20 traffic zones were randomly selected.* This represents a cross section of central city, inner suburbs, and outer suburbs zones.

It was decided to model round trips: home-shop-home.** Since the data set consists of records for one way (linked) trips, it was necessary to chain the individuals' trips and retain only the home-shop-home round trips. A total of 195 such person trips were reported for the 20 selected traffic zones. However, since some of these trips were taken together by members of the same household, this number represents only 160 household trips. This number was further reduced to 123 by excluding the following trips: 11 auto passenger trips, 5 taxi passenger trips, 6 trips with more than one mode (except walk access and egress), 2 trips by households that did not report income, and 13 trips made within the residential traffic zone. The auto passenger trips were excluded because they represent a shared ride and there was no information on their composition. The sample for taxi trips was judged to be too small. The same applies to the sample of the various mode combination trips. The within traffic zone trips (all by auto) were excluded because some level of service variables were

*Each zone was assigned a weight according to the number of households residing in the zone. Then, using a random number generator, 20 traffic zones were selected.

**See the discussion in Chapter II.

missing for these observations. Thus, the final sample consists of 123 household trips out of which 102 used a car and 21 were by bus.

The set of alternative choices for each observation is not reported in a conventional home interview survey. The determination of the set of alternatives to be included in a choice model is a difficult problem.* If an alternative has zero or very close to zero choice probability, its inclusion or exclusion from the set of alternatives will have negligible effects on the estimation and the forecasting results of the model. This means that we can attribute to any observation any non-relevant alternatives which will be predicted by the model to have zero or nearly zero probability of choice. Alternatively, we can delete it as an alternative. A practical reason to reduce the set of alternatives to only the relevant set is the savings in data collection and computation costs. It is difficult, if not impossible, to formulate criteria that will determine the set of relevant alternatives. We assume that there are no alternatives that are explicitly not available to a household due to a factor that is not expressed in the model. The problem, then, is that we do not know a priori which alternatives have zero choice probability. (In essence, from a behavioral point of view, we do not know the set of alternatives that a consumer is evaluating in making his decisions.) Therefore, trip alternatives for estimation are customarily determined from the observed trip making pattern. For example, if a taxi is rarely used,

*See the discussion in Chapters II and III.

then it is not considered a relevant alternative mode, or if a shopping center at location d is not visited by travellers from area i then destination d is not considered a relevant destination for households at i. It should be noted too that the set of alternatives used for forecasting does not have to be absolutely identical with the ones used in estimation.*

Hence, the next task was to identify the alternative shopping destinations for each household. The available information was the destination traffic zones of all the shopping trips. The procedure that was followed is in principle equivalent to that used by CRA(1972). The alternative destinations for a given household were determined as the set of all destinations to which trips were made by any of the households residing in the same area. This procedure excludes destinations which could be assumed to have a small choice probability.**

The identification of a traffic zone as an alternative shopping destination is difficult to accept. Primarily, in downtown and other central city shopping locations it is not clear how to identify an alternative shopping destination. Suburban shopping locations can be more easily identified as regional shopping centers. In downtown areas, where the land area of each traffic zone is very small, two department

*See the discussion of generic variables in Chapter VI.

**Since the logit model is also valid for any subset of alternative, the exclusion of alternatives with positive probabilities should not significantly affect the estimation results.

stores which are within walking distance from each other could be in two different traffic zones. However, the traveller may identify the two stores as a single destination alternative. Based on these considerations it was decided to identify an alternative shopping destination with more than one traffic zone when necessary. WCOG divided the metropolitan Washington area into about 150 districts, each one including from one to twenty traffic zones. Examination of the data indicated that the district system is more reasonable as a means of identifying shopping destinations in the central city, while in the suburban areas each district did not include more than one regional shopping center. The number of destinations that were identified for each household varied from one to eight. One alternative indicates that the sample was too small in the given area to infer the alternative destinations and therefore the trip is essentially excluded from the destination choice model. In total there were 114 observations which could be used for a destination choice model with the number of alternatives between two and eight. The total number of alternative destinations for all observations was 679, or an average of about six alternative destinations per observation.

The data consist of level of service variables by mode and destination, shopping opportunities by destination, and socio-economic characteristics of the household. Each observation includes the value of the variables for all the relevant alternative destination and mode combinations for this household and the observed choice. The observa-

tions included from two alternatives (in nine observations for which alternative destinations could not be inferred), up to 16 alternatives (in observations with eight alternative destinations). The total number of alternative destination and mode combinations for all observations was 1,376, or an average of about eleven alternatives per observation.

Variables Used in the Models

This section describes the set of explanatory variables that were used in the empirical estimation of the models.

Level of service variables:

Out-of-vehicle travel time -- Out-of-vehicle travel time for bus consists of the walk time to and from bus stations and waiting time in stations. For auto it is defined as the walk time to and from car, and park/unpark time. The data for this variable were developed by WCOG and are based on measures of expected walking distance.

In-vehicle travel time -- In-vehicle travel time is defined for auto as the time spent in the line-haul portion of the trip. For bus it is the time spent inside the bus. For both modes this variable was computed by WCOG using coded networks.

Out-of-pocket cost -- For bus: fare multiplied by the number of household members who make the trip together. For auto: parking charges plus operating cost of driving an automobile. The fares were determined for each trip from 1968 fare tables by RHP. The number of

persons, 5 years and older*, making the trip together was determined from the trip data in this study. The parking charges were developed in a parking study conducted by WCOG. They represent perceived cost that was computed using the probability of finding a free parking spot. The auto operating costs were computed with a formula developed by RHP. Variables included in this formula are speed and distance. The computed cost does not include ownership expenditures.

Socio-economic variables:

Household income--Household yearly income classified into ten ranges, as reported in the home interview survey. Other household and personal data were available. Number of cars was excluded a priori to avoid the problem of simultaneous dependency between auto ownership and mode choice. Based on experience from previous studies (e.g., McGillivray, 1969; CRA, 1972), it was determined that other variables that could be obtained play only a marginal role in the choice of mode and destination for shopping trips.

Attraction variables:

Employment--As an attraction variable to indicate the shopping opportunities in an alternative destination, the variable of employment in retail and wholesale was used. It was available from tabulations made from the home interview survey data by WCOG. The only other

*Children under five years do not pay fare.

attraction variable that was available was land area, but since we deal with shopping locations with different densities it was not used. The attractiveness of alternative shopping destinations can be represented by a variety of other variables, factual quantities such as retail floor space, gross revenue of shops, variety of goods and services available, etc.; as well as subjectively or attitudinally derived values, such as an evaluation of shopping centers on a quality scale. Unfortunately, the above data were not available. The shopping destination choice model estimated by CRA (1972) with data from Pittsburgh employed a single attraction variable: employment.

Other variables:

A few dummy variables that were created and the way in which the different variables were introduced in the models are described in the next section.

Specification of the Variables

The following list describes the definitions of the explanatory variables:*

$T_{O_{dm}}$ = Out-of-vehicle travel time to destination d by mode m (in minutes),

$T_{I_{dm}}$ = In-vehicle travel time to destination d by mode m (in minutes),

*The alternative specifications of the explanatory variables in the logit model are discussed in Chapter VI.

- $\frac{C_{dm}}{INC}$ = Out of pocket cost to destination d by mode m
 (in cents), divided by household income (code*),
- E_d = Wholesale-retail employment (# of jobs),
- $DCBD_d$ = CBD specific dummy variable for destination d
 = $\begin{cases} 1, & \text{for } d = \text{CBD} \\ 0, & \text{otherwise} \end{cases}$
- DA_m = Auto specific dummy variable for mode m
 = $\begin{cases} 1, & \text{for } m = \text{auto} \\ 0, & \text{for } m = \text{bus} \end{cases}$
- $DINC_m$ = Auto specific income variable for mode m
 = $\begin{cases} INC, & \text{for } m = \text{auto} \\ 0, & \text{for } m = \text{bus} \end{cases}$

Note that the level of service variables are specified as generic variables. An alternative would be to specify them as mode specific variables. (This would increase the number of level of service variables from 3 to 6.) In this case, the marginal rates of substitution among

*Household income was reported using the following code:

Code	Value (in \$1000 per year)
1	0 - <3
2	3 - <4
3	4 - <6
4	6 - <8
5	8 - <10
6	10 - <12
7	12 - <15
8	15 - <20
9	20 - <25
10	25+

level of service variables will differ for alternative modes. From a theoretical point of view it makes more sense to have equal marginal rates of substitution. The differences among modes that are not explained by the included level of service variables, such as differences in comfort, safety, etc., are accounted for by the mode specific dummy variables. From an empirical point of view this assumption was tested by CRA (1972). They estimated a mode choice model with mode specific level of service variables and found that the modal coefficients were not significantly different. This evidence sustains the a priori assumption of equal marginal rates of substitution.

We turn now to present the estimation results of the alternative models. The models are presented in terms of the log of the odds of choosing one alternative over another. That is, the models are expressed as follows:

$$\log \frac{P(i)}{P(j)} = \sum_{k=1}^K (X_{ik} - X_{jk}) \hat{\theta}_k$$

where

$P(i)$ = the choice probability of alternative i

X_{ik} = the k^{th} explanatory variable for alternative i

$\hat{\theta}_k$ = the coefficient estimate of the k^{th} explanatory variable

*See the discussion of the independence from irrelevant alternatives property in Chapter VI.

The Simultaneous Model

In the simultaneous model presented below, the joint probability of destination and mode (given that a trip is made) was directly estimated. The sets of alternatives consist of combinations of mode and destination. The number of alternatives varies among observations, ranging from 2 to 16. The results that were obtained are as follows:

$$\begin{aligned}
 \log \frac{P(d, m)}{P(d', m')} &= -1.36 (DA_m - DA_{m'}) \\
 &\quad (.970) \\
 &\quad -.0633 (TO_{dm} - TO_{d'm'}) \\
 &\quad (.0202) \\
 &\quad -.0164 (TI_{dm} - TI_{d'm'}) \\
 &\quad (.0116) \\
 &\quad -.0757 (C_{dm}/INC - C_{d'm'}/INC) \quad (1) \\
 &\quad (.0216) \\
 &\quad +.114 (DINC_m - DINC_{m'}) \\
 &\quad (.158) \\
 &\quad +.000171 (EMP_d - EMP_{d'}) \\
 &\quad (.0000875) \\
 &\quad +.316 (DCBD_d - DCBD_{d'}) \\
 &\quad (.554)
 \end{aligned}$$

$L^*(0) = -277.678$
 $L^*(\hat{\theta}) = -207.380$
 $\hat{\sigma}^2 = .25$
 $\bar{\sigma}^2 = .25$

where

$P(d, m)$ = the joint probability of choosing destination d
 and mode m

$L^*(0)$ = the log likelihood for zero coefficients*

$\hat{L}^*(\hat{\theta})$ = the log likelihood for the estimated coefficients*

r^2 = coefficient of determination*

\bar{r}^2 = r^2 adjusted for degrees of freedom*

and the numbers in parentheses below the model coefficients are standard errors.

All the signs and the relative values of the coefficient estimates are as expected. The pure auto effect, θ_{DA} , gave a minus sign; however, it should be interpreted as a "transit bias" only together with the coefficient of the auto specific income variable which is positive. Out-of-vehicle travel time is on the order of four times more onerous than in-vehicle travel time. The standard errors of the coefficients of the auto specific income variable and of the CBD dummy are relatively large; however, they have the expected signs.

Alternative Recursive Models

Three alternative composition rules were employed: weighted prices, weighted generalized price, and log of the denominator.** We define the composite variables when we present the estimation results. In addition, there are two possible sequences;

(i) $d + m$: d followed by m,

and

(ii) $m + d$: m followed by d.

*See Chapter VI for the definitions.

**These composition rules are explained in Chapter VI.

Hence, we estimated a total of six recursive models, three for each sequence. The estimation starts with the conditional probability, i.e., $P(m|d)$ in sequence (i) and $P(d|m)$ in sequence (ii). Then, the marginal probability is estimated using the composite variables that are calculated with results from the conditional probabilities. Note that for each sequence there is only one conditional probability and three marginal probabilities for the alternative composition rules.

Sequence d + m: The Conditional Probability

The conditional probability presented below is the equivalent of a trip interchange mode split model (Weiner, 1969). The model predicts the probability of mode choice for a given destination (and given that a trip is made). The sets of alternatives consist of the modes of bus and auto for the chosen destination. The estimation results are presented below:

$$\begin{aligned}
 \log \frac{P(m|d)}{P(m'|d)} &= -.639 (DA_m - DA_{m'}) \\
 &\quad (.133) \\
 &\quad -.0515 (TO_{dm} - TO_{dm'}) \\
 &\quad (.0237) \\
 &\quad -.0108 (TI_{dm} - TI_{dm'}) \\
 &\quad (.0261) \\
 &\quad -.137 (C_{dm}/INC - C_{dm'}/INC) \\
 &\quad (.0530) \\
 &\quad +.0490 (DINC_m - DINC_{m'}) \\
 &\quad (.199)
 \end{aligned} \tag{2}$$

$$L^*(0) = -85.257$$

$$L^*(DA) = -56.216$$

$$L^*(\hat{\theta}) = -23.033$$

$$\rho^z = .73 \quad \rho_{DA}^z = .59$$

$$\bar{\rho}^z = .72 \quad \bar{\rho}_{DA}^z = .58$$

where

$P(m|d)$ = the conditional probability of choosing mode m given that destination d is chosen,

$L^*(DA)$ = the log likelihood for zero coefficients except for pure auto effect DA,

ρ_{DA}^z = coefficient of determination in addition to the pure auto effect.

It can be seen that all the coefficients have their expected signs. Out-of-vehicle travel time is almost five times more onerous than in-vehicle travel time. The standard errors of the coefficients of in-vehicle travel time and income are relatively large; however, the coefficients have their expected signs.

Sequence d → m: The Marginal Probability

The marginal probability of destination choice is the equivalent of a pre-modal split distribution model. This model predicts the probability of destination choice with the mode choice indeterminate. The sets of alternatives consist of the alternative shopping destinations. Three models with the alternative composition rules were estimated for this probability. The estimation results are presented below.

Weighted prices:

$$\begin{aligned}
 \log \frac{P(d)}{P(d')} &= -.0227 \frac{(TO_d^M - TO_{d'}^M)}{(.0523)} \\
 &\quad -.0374 \frac{(TI_d^M - TI_{d'}^M)}{(.0173)} \\
 &\quad -.0269 \frac{(C_d^M / INC - C_{d'}^M / INC)}{(.0327)} \\
 &\quad + .000130 \frac{(EMP_d - EMP_{d'})}{(.0000910)} \\
 &\quad + .638 \frac{(DCBD_d - DCBD_{d'})}{(.595)} \tag{3}
 \end{aligned}$$

$$L^*(0) = -192.421$$

$$L^*(\hat{\theta}) = -182.485$$

$$L^*_{dm}(\hat{\theta}) = -205.518$$

$$\hat{f}^z = .05$$

$$\bar{f}^z = .04$$

$$f_{dm}^z = .26$$

where

$P(d)$ = the marginal probability of choosing destination d ,

$$TO_d^M = \sum_m TO_{dm} \cdot P(m|d)$$

$$TI_d^M = \sum_m TI_{dm} \cdot P(m|d)$$

$$C_d^M = \sum_m C_{dm} \cdot P(m|d)$$

$L^*_{dm}(\hat{\theta})$ = the log likelihood for the joint probability*

*See the explanation in Chapter VI.

f_{dm}^2 = coefficient of determination for the joint probability.*

Weighted generalized price:

$$\begin{aligned}
 \log \frac{P(d)}{P(d')} &= .000149 (\text{EMP}_d - \text{EMP}_{d'}) \\
 &\quad (.0000867) \\
 &\quad + .353 (\text{DCBD}_d - \text{DCBD}_{d'}) \\
 &\quad (.510) \\
 &\quad + .507 (\text{GP}_d^M - \text{GP}_{d'}^M) \\
 &\quad (.141) \\
 L^*(0) &= -192.421 \tag{4} \\
 L^*(\hat{\theta}) &= -184.866 \\
 L^*_{dm}(\hat{\theta}) &= -207.899 \\
 f^2 &= .04 \\
 \bar{f}^2 &= .04 \\
 f_{dm}^2 &= .25
 \end{aligned}$$

where

$$\begin{aligned}
 \text{GP}_d^M &= \sum_m (-.0515 \cdot T_0_{dm} - .0108 \cdot T_I_{dm} - .137 \cdot C_{dm}/\text{INC}) \\
 &\quad \cdot P(m|d)
 \end{aligned}$$

Log of the denominator:

$$\begin{aligned}
 \log \frac{P(d)}{P(d')} &= .000149 (\text{EMP}_d - \text{EMP}_{d'}) \\
 &\quad (.0000862) \\
 &\quad + .295 (\text{DCBD}_d - \text{DCBD}_{d'}) \\
 &\quad (.510)
 \end{aligned} \tag{5}$$

*Note that \bar{f}_{dm}^2 is not computed. The reason is that the two separate models have different numbers of degrees of freedom.

$$+.549 (\log P_d^M - \log P_{d1}^M) \\ (.147)$$

$$L^*(0) = -192.421$$

$$L^*(\hat{\theta}) = -184.068$$

$$L_{dm}^*(\hat{\theta}) = -207.101$$

$$\sigma^L = .04$$

$$\bar{\sigma}^2 = .04$$

$$\sigma_{dm}^L = .25$$

where

$$P_d^M = \sum_m e^{- .639 \cdot DA_m - .0515 \cdot TO_{dm} - .0108 \cdot TI_{dm}}$$

$$-.137 \cdot C_{dm} / INC + .0490 \cdot DINC_m$$

All the models have relatively low coefficients of determination which is attributed to the lack of more descriptive attraction data. All the three models gave coefficient estimates with the expected signs. However, in the model with weighted prices, equation (3), the coefficient of out-of-vehicle travel time is smaller than the coefficient of in-vehicle travel time, in contrast to what we would expect. The standard errors in this model are relatively large, however, it fits the data as well as the two other models.

The model with weighted prices represents the assumption that the marginal rates of substitution among level of service attributes are different for different choices. The two other models assume equal rates for different choices. From a theoretical point of view, the

latter assumption seems more reasonable.* It is more likely that a traveller will have an identical trade-off between travel time and money cost for different travel choices rather than several of them, each being used for a different choice. The poor results from the weighted prices model support this assumption. It appears that all previous travel demand models that are reported in the literature have made the assumption of equal marginal rates of substitution for different choices.

Comparing the two other alternative formulations, equations (4) and (5), it can be observed that there are no significant differences.** The coefficient estimates of the CBD dummy variable have relatively large standard errors in the two models. However, the coefficients have the expected sign. The model with weighted generalized price, equation (4), is equivalent to the model developed by CRA(1972).

Sequence m + d: The Conditional Probability

The conditional probability in this sequence is the equivalent of a post-modal split trip distribution model. The model predicts the probability of destination choice for a given mode. The sets of alternatives consist of the alternative shopping destinations for the chosen mode. The estimation results of this model are as follows:

*See the discussion in Chapter VI.

**It was noted in Chapter VI that the two composite variables, GP and log P, are very similar since they both have equal derivatives with respect to their variables.

$$\begin{aligned}
 \log \frac{P(d|m)}{P(d'|m)} &= -.0610 \frac{(T_{0dm} - T_{0d'm})}{(.0380)} \\
 &\quad -.0287 \frac{(TI_{dm} - TI_{d'm})}{(.0136)} \\
 &\quad -.0470 \frac{(C_{dm}/INC - C_{d'm}/INC)}{(.0263)} \\
 &\quad + .000148 \frac{(EMP_d - EMP_{d'})}{(.0000899)} \\
 &\quad + .330 \frac{(DCBD_d - DCBD_{d'})}{(.548)} \tag{6}
 \end{aligned}$$

$$L^*(0) = -192.421$$

$$L^*(\hat{\theta}) = -179.680$$

$$\hat{f}^1 = .07$$

$$\bar{f}^1 = .06$$

where

$P(d|m)$ = the conditional probability of choosing destination
d given that mode m is chosen.

The signs of the coefficient estimates are as expected. Out-of-vehicle travel time is more than two times more onerous than in-vehicle travel time. The coefficient of the CBD dummy variable has the expected sign, but with a relatively large standard error. The goodness of fit of this model is relatively low due to the large number of alternatives and the lack of better attraction description.

Sequence m → d: The Marginal Probability

The marginal probability of mode choice is the equivalent of a trip-end modal split model (Weiner, 1969). This model predicts the

probability of mode choice with indeterminate destination choice.

The sets of alternatives include the modes of bus and auto. Again, we model this probability with the three alternative composition rules. The results that were obtained are presented below.

Weighted prices:

$$\begin{aligned}
 \log \frac{P(m)}{P(m')} &= -.952 (DA_m - DA_{m'}) \\
 &\quad (.1.27) \\
 &\quad -.0509 (TO_m^D - TO_{m'}^D) \\
 &\quad (.0204) \\
 &\quad +.109 (TI_m^D - TI_{m'}^D) \\
 &\quad (.0429) \\
 &\quad -.183 (C_m^D / INC - C_{m'}^D / INC) \\
 &\quad (.0725) \\
 &\quad +.293 (DINC_m - DINC_{m'}) \\
 &\quad (.225) \tag{7}
 \end{aligned}$$

$$L^*(0) = -85.257$$

$$L^*(DA) = -56.216$$

$$L^*(\hat{\theta}) = -24.596$$

$$L^*_{dm}(\hat{\theta}) = -204.276$$

$$\mathcal{J}^t = .71 \quad \mathcal{J}_{DA}^t = .56$$

$$\bar{\mathcal{J}}^t = .70 \quad \bar{\mathcal{J}}_{DA}^t = .55$$

$$\mathcal{J}_{dm}^t = .26$$

where

$P(m)$ = the marginal probability of choosing mode m

$$TO_m^D = \sum_d TO_{dm} \cdot P(d|m)$$

$$TI_m^D = \sum_d TI_{dm} \cdot P(d|m)$$

$$C_m^D = \sum_d C_{dm} \cdot P(d|m)$$

Weighted generalized price:

$$\begin{aligned} \log \frac{P(m)}{P(m')} &= -2.07 \frac{(DA_m - DA_{m'})}{(.959)} \\ &\quad + .117 \frac{(DINC_m - DINC_{m'})}{(.157)} \\ &\quad + 1.62 \frac{(GP_m^D - GP_{m'}^D)}{(.371)} \end{aligned} \tag{8}$$

$$L^*(0) = -85.257$$

$$L^*(DA) = -56.216$$

$$L^*(\hat{\theta}) = -31.039$$

$$L^*_{dm}(\hat{\theta}) = -210.719$$

$$\hat{\sigma}^2 = .64 \quad \hat{\sigma}_{DA}^2 = .45$$

$$\bar{\sigma}^2 = .63 \quad \bar{\sigma}_{DA}^2 = .44$$

$$\hat{\sigma}_{dm}^2 = .24$$

where

$$GP_m^D = \sum_d (-.0610 \cdot TO_{dm} - .0287 \cdot TI_{dm} - .0470 \cdot C_{dm}/INC) \cdot P(d|m)$$

Log of the denominator:

$$\begin{aligned} \log \frac{P(m)}{P(m')} &= -1.74 \frac{(DA_m - DA_{m'})}{(.955)} \\ &\quad + .0489 \frac{(DINC_m - DINC_{m'})}{(.168)} \\ &\quad + 1.42 \frac{(\log P_m^D - \log P_{m'}^D)}{(.303)} \end{aligned} \tag{9}$$

$$L^*(0) = -85.257$$

$$L^*(DA) = -56.216$$

$$L^*(\hat{\theta}) = -27.832$$

$$L^*_{dm}(\hat{\theta}) = -207.512$$

$$\beta^z = .67 \quad \beta^z_{DA} = .50$$

$$\bar{\beta}^z = .67 \quad \bar{\beta}^z_{DA} = .50$$

$$\beta^z_{dm} = .25$$

where

$$P_m^D = \sum_d e^{-0.0610 \cdot T0_{dm} - 0.0287 \cdot TI_{dm} - 0.0470 \cdot C_{dm} / INC + 0.000148 \cdot EMP_d + 0.330 \cdot DCBD_d}$$

Again, the model with weighted prices, equation (7), gave unreasonable coefficient estimates, similar to equation (3). The two other models, equations (8) and (9), gave better results. The coefficients of the income variable have the expected signs but with relatively large standard errors. The model of equation (8) uses the same composition scheme as the model developed by CRA (1972); however, this model assumes a different sequence.

Sequential Estimation of the Simultaneous Model

The simultaneous model can also be estimated indirectly. This procedure requires the estimation of a conditional probability, i.e., equation (2) or equation (6), and the estimation of a marginal probability, equation (5) or equation (9), with the coefficient of the log

of the denominator, $\log P_d^M$ or $\log P_m^D$, constrained to unity.* The results from the constrained estimation of equations (5) and (9) are presented below.

For the sequence $d \rightarrow m$, equation (5), the following results were obtained:

$$\begin{aligned}
 \log \frac{P(d)}{P(d')} &= .000232 (\text{EMP}_d - \text{EMP}_{d'}) \\
 &\quad (.0000890) \\
 &\quad + .566 (\text{DCBD}_d - \text{DCBD}_{d'}) \\
 &\quad (.561) \\
 &\quad + 1. (\log P_d^M - \log P_{d'}^M) \\
 L^*(0) &= -192.421 \tag{10} \\
 L^*(\hat{\theta}) &= -188.278 \\
 L^*_{dm}(\hat{\theta}) &= -211.311 \\
 \beta^1 &= .02 \\
 \bar{\beta}^1 &= .02 \\
 \beta^2_{dm} &= .24
 \end{aligned}$$

Equation (10) together with equation (2) supply alternative coefficient estimates for the simultaneous model.

For the sequence $m \rightarrow d$, equation (9), the estimation results were as follows:

*This procedure was explained in detail in Chapter VI. Note that the sequence used for estimation has no behavioral interpretation.

$$\begin{aligned}
 \log \frac{P(m)}{P(m')} &= -1.48 \frac{(DA_m - DA_{m'})}{(.810)} \\
 &\quad + .133 \frac{(DINC_m - DINC_{m'})}{(.140)} \\
 &\quad + 1. (\log P_m^D - \log P_{m'}^D) \tag{11}
 \end{aligned}$$

$$L^*(0) = -85.257$$

$$L^*(DA) = -56.216$$

$$L^*(\hat{\theta}) = -28.964$$

$$L^*_{dm}(\hat{\theta}) = -208.644$$

$$\beta^1 = .66 \quad \beta_{DA}^1 = .48$$

$$\bar{\beta}^1 = .65 \quad \bar{\beta}_{DA}^1 = .48$$

$$\beta_{dm}^1 = .25$$

Equation (11) together with equation (6) furnish another alternative set of coefficients for the simultaneous model.

The coefficient estimates of the two constrained models, equations (10) and (11), have the expected signs.

Comparison of Alternative Models

In what follows we compare the estimation results of the alternative models on the basis of coefficient values (including values of time and elasticities), goodness of fit, and estimation costs. We exclude from this discussion the recursive models with weighted prices, equations (3) and (7), that gave unreasonable coefficient estimates. The alternative models that we consider in this comparison are the following:

Simultaneous:

Direct estimation -- equation (1)

Indirect estimation $d \rightarrow m$ -- equations (2) and (10)

Indirect estimation $m \rightarrow d$ -- equations (6) and (11)

Recursive $d \rightarrow m$:

Weighted generalized price -- equations (2) and (4)

Log of the denominator -- equations (2) and (5)

Recursive $m \rightarrow d$:

Weighted generalized price -- equations (6) and (8)

Log of the denominator -- equations (6) and (9)

It was not expected from this empirical study to accept or reject the a priori assumption of a simultaneous decision making process. As expected, the empirical evidence does not show which one of the alternative structures, a simultaneous and two recursive, is more likely to be correct. All the models gave reasonable coefficient estimates.

Furthermore, all the models gave an essentially equal goodness of fit:

$\rho^2 = .25$. The simultaneous model includes seven coefficients compared with eight coefficients in the recursive models. This implies that the simultaneous model has a slight edge in this category, but it is certainly not a conclusive difference.

The coefficient estimates of the simultaneous model can be directly compared with the coefficients of the conditional probabilities in the recursive structures. For the conditional probability $P(m|d)$, we obtain the following comparison:

<u>Variable</u>	<u>Simultaneous Equation (1)</u>	<u>Recursive Equation (2)</u>
DA	-1.36	-.639
TO	-.0633	-.0515
TI	-.0164	-.0108
C/INC	-.0757	-.137
DINC	.114	.0490

For the conditional probability $P(d|m)$ we get:

	<u>Equation (1)</u>	<u>Equation (6)</u>
TO	-.0633	-.0610
TI	-.0164	-.0287
C/INC	-.0757	-.0470
EMP	.000171	.000148
DCBD	.316	.330

The above comparisons indicate that the indirect estimation of the conditional probabilities through the joint probability gave different results from the direct estimation of the conditionals, as expected.

In describing the results of the various models, we noted the relative weights of out-of-vehicle and in-vehicle travel times. In equation (1) the ratio of the coefficients is 3.9, in equation (2) the ratio is 4.8, and in equation (6) it is 2.1. Although the values of the coefficients are not widely apart, the above ratios indicate that the implied rates of substitution can be significantly different. This can also be observed by comparing the rates of substitution among travel times and out-of-pocket cost, or values of time, as follows:

	<u>Equation (1)</u>	<u>Equation (2)</u>	<u>Equation (6)</u>
Value of out-of-vehicle travel time	3.02 (1.44)	1.36 (.98)	4.67 (4.36)
Value of in-vehicle travel time	.78 (.68)	.28 (.66)	2.21 (2.01)

The figures are in \$/hour and for a household with annual income between \$10,000 and \$12,000. The numbers in parentheses are the standard errors.* Again, the ratios of two coefficients demonstrate a considerable variability. It is interesting to note that the ratios between the two travel time coefficients in the three equations follow an opposite pattern to that of the values of time. The ratio takes its highest value in equation (2) and its lowest value in equation (6). The values of time take their highest values in equation (6) and the lowest values in equation (2). This can be attributed to specification errors, such as the omission of a comfort variable and important

*The values of time are computed as ratios of the coefficient estimates of travel time and travel cost. The estimated coefficients were all significantly different from zero. However, the standard errors of the ratios are quite large. The distribution of a ratio of two coefficient estimates is skewed and therefore a t-test cannot be used (Lianos and Rausser, 1972). In addition, the standard errors of values of time obtained in previous studies were of the same order of magnitude, e.g., Talvitie (1972).

attraction variables, which affect differently the two conditional probabilities. Comparison of the level of service coefficients in equations (2) and (6) shows that the coefficients of travel times are greater in equation (6), while the cost coefficient is greater in equation (2). This could be attributed to greater variability of out-of-pocket costs than travel times among modes to a given destination, and greater variability of travel times than out-of-pocket costs among alternative destinations for a given mode.

The significance of the differences among the level of service coefficients can be observed by computing the corresponding elasticities. For example, the elasticities of the conditional probability $P(m|d)$ derived from equations (1) and (2) compare as follows:

	<u>$P(\text{bus} d)$</u>		<u>$P(\text{auto} d)$</u>	
Elasticities with respect to:	<u>Eq (1)</u>	<u>Eq (2)</u>	<u>Eq (1)</u>	<u>Eq (2)</u>
Out-of-vehicle travel time	-1.01	-.82	-.13	-.10
In-vehicle travel time	-.31	-.26	-.05	-.03
Out-of-pocket cost	-.40	-.91	-.13	-.23

The figures are computed for the following case:

- a household with annual income between \$10,000 and \$12,000,
- the probabilities of choosing bus and auto are .2 and .8, respectively,
- out-of-vehicle travel times are 20 minutes by bus, and 10 minutes by auto,
- in-vehicle travel times are 30 minutes by bus and 15 minutes by auto,
- out-of-pocket costs are 50 cents by bus and 50 cents by auto.

The most striking difference is in the cost elasticity. The cost elasticities estimated from the mode choice model (given destination) are twice the estimates from the simultaneous model.

In comparing the marginal probabilities of the recursive models with the simultaneous model, it is important to note the values of the coefficients of the composite variables. In all the marginal probability models for the two sequences, equations (4), (5), (8), and (9), the coefficients are significantly different from unity. Their values indicate that the weights of the level of service variables are about 50% smaller for destination choice than for mode choice. In the models for $P(d)$ the coefficients are smaller than unity, about .5, and in the models for $P(m)$ the coefficients are greater than unity, about 1.5.

The above comparison indicates that although the various structures result in an almost equal goodness of fit, they result in policy implications which are significantly different.

The comparison of the conditional probabilities together with the constrained marginals, equations (10) and (11), with the directly estimated simultaneous model indicates that although a sequential estimation is preferred to a recursive structure, it is not desirable due to the significantly different rates of substitution among level of service variables.

Another important criterion in comparing the alternative models is the computational cost. (Note that the data requirements of all models

were identical). In general, experience with the logit estimation program* indicates that the computer time required increased fairly linearly with the number of observations, the average number of alternatives, and the number of iterations required to achieve convergence. (The estimation method requires the solution of a system of non-linear equations.) Time requirements increase more than linearly with the number of coefficients in the logit model. As a rule of thumb, each iteration on a binary logit model requires twice the time needed for a linear regression having the same number of observations and variables.

In this empirical study the estimation of the simultaneous model, equation (1), required .53 minutes of CPU time compared with .14 minutes for the binary mode choice model of equation (2).** However, comparing the total charge that includes also the input and output operations, the difference is less significant, about 20%. This indicates that a direct estimation of a simultaneous model is computationally feasible. Furthermore, the estimation of any of the recursive models in this study was more expensive than the estimation of the simultaneous model.

*Based primarily on experiments conducted by C. F. Manski at the Department of Economics at M.I.T.

**An IBM 370-155 computer with a FORTRAN G compiler was used. Core requirements depend in general on the number of coefficients. In this study we used 110K which was the core required by the compiler.

Summary

This empirical study has demonstrated the feasibility of a simultaneous choice model with regard to the computational costs, as well as the actual results. The simultaneous model was less expensive to estimate than the recursive models, and it gave the best overall results in terms of goodness of fit and significance of coefficients. Moreover, this empirical study has demonstrated that different structures can result in significantly different estimates of important parameters. The use of a recursive structure instead of an a priori hypothesized simultaneous one may result in different policy implications. Furthermore, different recursive structures for different sequences may also result in different predictions. Given the lack of a priori reasons and empirical evidence to support a choice among sequences, the results demonstrate a practical argument for the direct estimation of simultaneous models.

In spite of the small sample size and the simple specification that was used, the results of the simultaneous model are highly encouraging. All the coefficients have the correct sign, plausible magnitudes, and reasonably standard errors. The successful estimation of the simultaneous model indicates that a disaggregate probabilistic simultaneous travel demand model is a viable modelling approach that should be further investigated.

CHAPTER VIII

Conclusions and RecommendationsSummary

The motivation for this study was the development of disaggregate probabilistic simultaneous travel demand models. First, we examined the framework of modelling in transportation planning and established the overall structure for mobility and travel choice models. This led to an evaluation of alternative demand theories--the abstract consumer and choice theories, from which it was concluded that a choice approach is more appropriate than a conventional demand analysis. However, it was necessary to extend the choice theories to account for dependencies among choices. This provided the framework for the specification of simultaneous and recursive structures of travel demand models. It was noted that in a recursive structure fewer variables have to be included in any one choice model and this can considerably simplify the estimation problem.

Using the concepts developed earlier, the structures of current approaches to travel demand modelling were reviewed. It was apparent that none of the existing models have a simultaneous probabilistic structure. The next task was to specify and estimate alternative travel demand models with different structures. The purpose of the empirical study was to test the feasibility of a simultaneous model, and to investigate the sensitivity of travel predictions to the simplifying assumption of a recursive structure.

It was decided that the multinomial logit model is the most satisfactory choice model that is available. Using the logit model, the functional forms of the alternative models were specified. To simplify the empirical analysis we selected only a single trip purpose, a subset of the travel choices, and a relatively small sample. However, the estimation results indicated that a probabilistic simultaneous structure is a desirable and promising approach to travel demand modelling.

General Conclusions

The overall conclusion of this study is that travel demand models should be developed and estimated with a simultaneous probabilistic structure. The theoretical discussion indicated that a recursive structure that represents a sequential decision making process is not completely realistic. With our present knowledge of travel behavior it appears that the only reasonable description of the trip decision process is a simultaneous structure. The empirical study indicated that the estimation of disaggregate simultaneous travel demand models is feasible in terms of the actual results and the computational costs. Moreover, the empirical study demonstrated that the use of a recursive structure, due to its practical advantages, rather than a simultaneous structure, may result in considerably different policy implications.

The following sections present some recommendations based on this research for future modelling efforts.

Theoretical Framework

It was argued that the mobility and travel choices should be modelled

as a block recursive structure, where each block has a simultaneous structure. The justification for this structure was the assumption that mobility choices depend on the overall pattern of travel for non-work purposes rather than on specific trips. This implies that in modelling the mobility choices we need to define composite variables that represent the spatial opportunities and level of service characteristics for non-work trips.* In modelling the travel choices it implies that the mobility choices are taken as predetermined.

Modelling a Simultaneous Structure

It was argued that a probabilistic choice approach with disaggregate data is superior to an aggregate demand analysis. In the probabilistic choice framework it was shown that for a simultaneous structure we need to model directly the joint probability. This implies the definition of joint utility functions for the mobility choices, and for the travel choices for a given trip purpose.

Given the joint probability, it is possible to use the model for forecasting directly, or to derive any set of marginal and conditional probabilities and use an indirect forecasting procedure. The format in which the model is used for forecasting depends on the purpose of the analysis. For example, if it is assumed that a subset of the choices

*It appears that the use of the expected utility composition rule, discussed in Chapters IV and VI is preferable to the various types of accessibility measures that have been conventionally used in auto ownership and location models (e.g., Dunphy, 1973; Sherman, 1972; Aldana, 1971; TRC, 1967)

is unchanged then only a conditional probability is required. An indirect forecasting procedure may also be useful in order to check less detailed forecasts for reasonableness, and for comparison with other sources of information.

The application of this modelling framework appears to be feasible. However, the following sections outline research that is needed on some important problems.

Choice Models

In this study we used a strict utility specification* of the choice model. This specification represents some restrictive and unrealistic assumptions. More realistic specifications of choice models are not yet practical, and there is some evidence from psychological choice experiments (Bock and Jones, 1968) that the strict utility specification provides a reasonable practical approximation. Therefore, it appears that, in the short run, model development efforts will rely on this specification, while more long term research is conducted toward the development of improved choice models.

The Choice Set

The determination of the alternative destinations in the empirical study was based on an ad-hoc procedure. Even if people were asked to

*The multinomial logit model that was used in a special case of the strict utility specification that was described in Chapter III.

report their perceived alternatives, we still need to determine a set of alternatives for forecasting. The determination of the actual set of alternatives that an individual perceives can result in more reliable models and savings in data collection and processing. It seems desirable to investigate alternative procedures and to research the feasibility of modelling this aspect of choice behavior.

The Aggregation Problem

The use of disaggregate models in planning requires an aggregation procedure. A limited amount of research has been conducted to develop appropriate procedures. Some simplified procedures that have been suggested, e.g., market segmentation, need to be evaluated empirically.

Non-Home-Based Trips

Customarily non-home-based trips are modelled separately from home-based trips. It was argued that this is an unrealistic procedure because of the interdependencies among the different legs of a tour. In the empirical study only simple round trips--home-shop-home--were modelled. Research is needed to develop travel demand models for more complex trips that includes non-home-based legs.

Case Studies

The use of disaggregate probabilistic simultaneous travel demand models in planning should be evaluated in the context of case studies. This seems to be the appropriate strategy to extend the empirical experience, to test aggregation procedures, and to investigate alternative

methods of modelling non-home-based trips. Experimental application of the recommended modelling approach appears to be essential for providing guidelines to make this approach operational in actual planning studies.

Data

The recommended modelling approach suggests some changes in the data collection procedures that are commonly used by Urban Transportation Planning Studies. Research is needed to determine how much data is needed for disaggregate models. The data is used in its original disaggregate form and is not aggregated for estimation. Therefore, for estimation less observations are needed. The amount of data that needs to be collected also depends on the data requirement of the aggregation procedure.

There is clearly a need for more detailed travel information such as: actual trip frequencies, composition of shared rides, desired arrival times, car availability during the different times of day, etc. In addition it seems desirable to collect more detailed level of service characteristics and spatial opportunities variables. The potentially large savings in data collections that could be realized from the reduced sample size could be devoted to the collection of more detailed data and the further development of behavioral models.

Extensions of This Research

The empirical study that was conducted in this research could be extended to include a complete set of travel choices, and different trip purpose categories.

It is apparent that with more detailed data the specification of the models that were estimated could be improved. The set of explanatory variables could be extended to include more attraction variables, more socio-economic characteristics, and level of service measures such as comfort and convenience that were not available. The assumption of 0, 1 daily trip frequencies which is dictated by the available UTP data is also an example of misspecification.

It should be noted that every model developed in an environment of non-experimental research is misspecified to a certain degree due to unquantifiable variables and lack of data. In addition, if data for the omitted variables is not available we cannot evaluate the severity of the specification errors. Therefore, the need for more data that was noted is based primarily on logical a priori arguments with respect to the importance of the omitted variables. However, this only implies that the desirability of the additional data should at least be tested on a limited scale.

The empirical study could also be extended to include different data sets from different urban areas. This would indicate the stability of the model coefficients. It can also serve as a procedure to detect specification errors. For example, Rogers et al (1970) estimated a disaggregate mode choice model with data from different urban areas in the U.K. Comparison of the models from the different urban areas led to a revised specification that resulted in a model that was more generally applicable. In this context estimation of the model with data from different time periods would serve as a test of the validity of the proposed modelling approach.

The theoretical development and the empirical study of this research could also be readily applied to the modelling of mobility choices. The state of the art of activity system models* and the work of Aldana (1971) indicate the desirability of the application of simultaneous probabilistic models to mobility choices.

*For a review see Lee (1968).

BIBLIOGRAPHY

Aldana, E., 1971, TOWARDS MICROANALYTIC MODELS OF URBAN TRANSPORTATION DEMAND, PhD Thesis, Department of Civil Engineering, M.I.T., Cambridge, Massachusetts.

Aldana, E., R. de Neufville, and J.H. Stafford, 1973, "Microanalysis of Urban Transportation Demand," Paper prepared for presentation at the Annual Highway Research Board Meeting, Washington, D.C.

Beilner, H., and F. Jacobs, 1971, "Probabilistic Aspects of Traffic Assignment," in Newell, G.F., ed., TRAFFIC FLOW AND TRANSPORTATION, American Elsevier, New York.

Bock, R.D., and L.V. Jones, 1968., THE MEASUREMENT AND PREDICTION OF JUDGMENT AND CHOICE, Holden-Day, San Francisco, California.

Brand, D., 1972, "Theory and Method in Land Use and Travel Forecasting," Paper prepared for presentation at the Annual Highway Research Board Meeting, Washington, D.C.

Brand, D., 1972, "Travel Demand Forecasting: Some Foundations and a Review," Background paper for HRB/U.S. Department of Transportation Conference on Travel Demand Forecasting, Williamsburg, Virginia.

Charles River Associates, Inc. (CRA), 1967, A MODEL OF URBAN PASSENGER TRAVEL DEMAND IN THE SAN FRANCISCO METROPOLITAN AREA, Prepared for the California Division of Bay Toll Crossings, Cambridge, Massachusetts.

Charles River Associates, Inc. (CRA), 1972., A DISAGGREGATE BEHAVIORAL MODEL OF URBAN TRAVEL DEMAND, Federal Highway Administration, U.S. Department of Transportation, Washington, D.C.

Crow, R.T., K.H. Young, and T. Cooley, 1971, MODELS OF INTERCITY TRAVEL DEMAND, PART I THEORETICAL ASPECTS, Northeast Corridor Transportation Project, U.S. Department of Transportation, Washington, D.C.

de Donnea, F.X., 1971, THE DETERMINANTS OF TRANSPORT MODE CHOICE IN DUTCH CITIES, Rotterdam University Press, Louvain, Belgium.

de Neufville, R., and J.H. Stafford, 1971, SYSTEMS ANALYSIS FOR ENGINEERS AND MANAGERS, McGraw-Hill, New York.

Domencich, T.A., G. Kraft, and J.P. Valette, 1968, "Estimation of Urban Passenger Travel Behavior: An Economic Demand Model," HIGHWAY RESEARCH RECORD #238, Highway Research Board, Washington, D.C.

Dunphy, R.T., 1973, "Transit Accessibility as a Determinant of Automobile Ownership," Paper prepared for presentation at the Annual Highway Research Board Meeting, Washington, D.C.

Federal Highway Administration (FHWA), 1970, URBAN TRANSPORTATION PLANNING, GENERAL INFORMATION AND INTRODUCTION TO SYSTEM 360, U.S. Department of Transportation, Washington, D.C.

Fertal, M.J., E. Weiner, A.J. Balek, and A.F. Sevin, 1966, MODAL SPLIT, Bureau of Public Roads, U.S. Department of Commerce, Washington, D.C.

Fisher, F.M., 1966, THE IDENTIFICATION PROBLEM IN ECONOMETRICS, McGraw-Hill, New York.

Fleet, C.R., and S.R. Robertson, 1968, "Trip Generation in the Transportation Planning Process," HIGHWAY RESEARCH RECORD #240, Highway Research Board, Washington, D.C.

Goldman, S.M., and H. Uzawa, 1964, "A Note on Separability in Demand Analysis," ECONOMETRICA, Volume 32, No. 3, July.

Golob, T.F., and M.J. Beckmann, 1971, "A Utility Model for Travel Forecasting," TRANSPORTATION SCIENCE, Volume 5, No. 1, February.

Gur, Y., 1970, AN ACCESSIBILITY SENSITIVE TRIP GENERATION MODEL, Chicago Area Transportation Study, Chicago, Illinois.

Henderson, J.M., and R.E. Quandt, 1958, MICROECONOMIC THEORY, McGraw-Hill, New York.

Irwin, N.A., N. Dodd, and H.G. von Cube, 1961, "Capacity Restraint in Assignment Programs," HIGHWAY RESEARCH BOARD BULLETIN 297, Washington, D.C.

Kain, J.F., 1967, "Urban Travel Behavior," in Schnose, L.F., and H. Fagin, ed., URBAN RESEARCH AND POLICY PLANNING, VOLUME I, URBAN AFFAIRS ANNUAL REVIEW, Sage Publications, Beverly Hills, California.

Kanafani, A., 1972, "An Aggregate Model of Trip Making," TRANSPORTATION RESEARCH, Volume 8.

Kannel, E.J., and K.W. Heathington, 1973, "The Temporal Stability of Trip Generation Relationships," Paper prepared for presentation at the Annual Highway Research Board Meeting, Washington, D.C.

Kraft, G., 1963, Part V in: Systems Analysis and Research Corp., DEMAND FOR INTERCITY PASSENGER TRAVEL IN THE WASHINGTON-BOSTON CORRIDOR, U.S. Department of Commerce, Washington, D.C.

Kraft, G., and M. Wohl, 1967, "New Directions for Passenger Demand Analysis and Forecasting," TRANSPORTATION RESEARCH, Volume 1, No. 3, November.

Lancaster, K.J., 1966, "A New Approach to Consumer Theory," JOURNAL OF POLITICAL ECONOMY, Volume 74, April.

Lave, C.H., 1969, "A Behavioral Approach to Modal Split Forecasting," TRANSPORTATION RESEARCH, Volume 3, No. 4, December.

Lee, D.B., 1968, MODELS AND TECHNIQUES FOR URBAN PLANNING, Cornell Aeronautical Laboratory, Inc., Buffalo, New York.

Leontief, W., 1947., "Introduction to a Theory of the Internal Structure of Functional Relationships," ECONOMETRICA, Volume 15, No. 4, October.

Lianos, T.P., and G.C. Rausser, 1972, "Approximate Distribution of Parameters in a Distributed Lag Model," JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION, Volume 67, March.

Lisco, T.E., 1967, THE VALUE OF COMMUTERS' TRAVEL TIME, A STUDY IN URBAN TRANSPORTATION, PhD Thesis, Department of Economics, University of Chicago, Chicago, Illinois.

Lowry, I.S., 1966, MODEL OF METROPOLIS, Rand Corp., Santa Monica, California.

Luce, R.D., 1959, INDIVIDUAL CHOICE BEHAVIOR, John Wiley, New York.

Luce, R.D., and P. Suppes, 1965, "Preference, Utility and Subjective Probability," in Luce, R.D., R.R. Bush, and E. Galanter, ed., HANDBOOK OF MATHEMATICAL PSYCHOLOGY, Volume III, John Wiley, New York.

Malinvaud, E., 1966, STATISTICAL METHODS OF ECONOMETRICS, Rand McNally, Chicago, Illinois.

Manheim, M.L., 1969, "Search and Choice in Transport Systems Analysis," HIGHWAY RESEARCH RECORD #293, Highway Research Board, Washington, D.C.

Manheim, M.L., 1970, "Fundamental Properties of Systems of Demand Models," Discussion paper T70-1, Department of Civil Engineering, M.I.T., Cambridge, Massachusetts.

Manheim, M.L., 1972, "Practical Implications of Some Fundamental Properties of Travel Demand Models," Paper prepared for presentation at the Annual Highway Research Board Meeting, Washington, D.C.

Marshall, A., 1961, PRINCIPLES OF ECONOMICS, Ninth Edition, MacMillan, London.

Martin, B.V., F.W. Memmott, and A.J. Bone, 1961, PRINCIPLES AND TECHNIQUES OF PREDICTING FUTURE URBAN AREA TRANSPORTATION, M.I.T. Press, Cambridge, Massachusetts.

Mayberry, J.P., 1970, "Structural Requirements for Abstract-Mode Models of Passenger Transportation," Chapter 5 in: Quandt, R.E., ed., THE DEMAND FOR TRAVEL: THEORY AND MEASUREMENT, Heath Lexington Books, Lexington, Massachusetts.

McCarthy, G.M., 1969, "Multiple Regression Analysis of Household Trip Generation - A Critique," HIGHWAY RESEARCH RECORD #297, Highway Research Board, Washington, D.C.

McFadden, D., 1968, "The Revealed Preferences of a Government Bureaucracy," Technical Report No. 17, Institute of International Studies, University of California, Berkeley.

McFadden, D., and S.G. Winter, 1970, Chapter B-1: Consumer Theory, in THEORY OF RESOURCE ALLOCATION AND PRICES, Unpublished Lecture Notes, Department of Economics, University of California, Berkeley.

McGillivray, R.G., 1969, BINARY CHOICE OF TRANSPORT MODE IN THE SAN FRANCISCO BAY AREA, PhD Thesis, Department of Economics, University of California, Berkeley.

McGillivray, R.G., 1972, "Mode Split and the Value of Travel Time," TRANSPORTATION RESEARCH, Volume 6.

McLynn, J.M., A.J. Goldman, P.R. Meyers, and R.H. Watkins, 1968, "Analysis of a Market Split Model," JOURNAL OF RESEARCH, National Bureau of Standards, Mathematical Sciences, Volume 72B, No. 1.

McLynn, J.M., and T. Woronka, 1969, PASSENGER DEMAND AND MODAL SPLIT MODELS, Northeast Corridor Transportation Project, U.S. Department of Transportation, Washington, D.C.

- Meyer, J.R., 1968, "Urban Transportation," in Wilson, J., ed., THE METROPOLITAN ENIGMA, Harvard University Press, Cambridge, Massachusetts.
- Meyer, J.R., and M.R. Straszheim, 1971, PRICING AND PROJECT EVALUATION, The Brookings Institution, Washington, D.C.
- Muth, R.F., 1966, "Household Production and Consumer Demand Functions," ECONOMETRICA, Volume 34, No. 3, July.
- Northeast Corridor Transportation Project (NECTP), 1970, NORTHEAST CORRIDOR TRANSPORTATION PROJECT REPORT, NECTP 209, Office of High Speed Ground Transportation, U.S. Department of Transportation, Washington, D.C.
- Oi, W.Y., and P.W. Shuldiner, 1962, AN ANALYSIS OF URBAN TRAVEL DEMANDS, Northwestern University Press, Evanston, Illinois.
- Philips, A., 1971, TECHNOLOGY AND MARKET STRUCTURE, A STUDY OF THE AIR-CRAFT INDUSTRY, Heath Lexington Books, Lexington, Massachusetts.
- Plourde, R.P., 1968, CONSUMER PREFERENCE AND THE ABSTRACT MODE MODEL: BOSTON METROPOLITAN AREA, Research Report R68-51, Department of Civil Engineering, M.I.T., Cambridge, Massachusetts.
- Plourde, R.P., 1971, DEVELOPMENT OF A BEHAVIORAL MODEL OF TRAVEL MODE CHOICE, PhD Thesis, Department of Civil Engineering, M.I.T., Cambridge, Massachusetts.
- Quandt, R.E., ed., 1970, THE DEMAND FOR TRAVEL: THEORY AND MEASUREMENT, Heath Lexington Books, Lexington, Massachusetts.
- Quandt, R.E., and W.J. Baumol, 1966, "Abstract Mode Model: Theory and Measurement," JOURNAL OF REGIONAL SCIENCE, Volume 6, No. 2.
- Quandt, R.E., and K.H. Young, 1969, "Cross-Sectional Travel Demand Models: Estimation and Tests," JOURNAL OF REGIONAL SCIENCE, Volume 9, No. 2.
- Quarmby, G.A., 1967, "Choice of Travel Mode for the Journey to Work: Some Findings," JOURNAL OF TRANSPORT ECONOMICS AND POLICY, Volume 1, No. 3.
- Rassam, P.R., R.H. Ellis, and J.C. Bennett, 1971, "The n-Dimensional Logit Model: Development and Application," HIGHWAY RESEARCH RECORD #369, Highway Research Board, Washington, D.C.
- Reichman, S., and P.R. Stopher, 1971, "Disaggregate Stochastic Models of Travel-Mode Choice," HIGHWAY RESEARCH RECORD #369, Highway Research Board, Washington, D.C.

- Roberts, P.O., and D.T. Kresge, 1971, SYSTEMS ANALYSIS AND SIMULATION MODELS, The Brookings Institution, Washington, D.C.
- Rogers, K.G., G.M. Townsend, and A.E. Metcalf, 1970, PLANNING FOR THE WORK JOURNEY, Local Government Operational Research Unit, Royal Institute of Public Administration, Reading, U.K.
- Sherman, L., 1972, "Notes on Accessibility Analysis," Unpublished paper, U.S. Department of Transportation, Washington, D.C.
- Stigler, G.J., 1952, THE THEORY OF PRICE, MacMillan, New York.
- Stopher, P.R., 1969, "A Probability Model of Travel Mode Choice for the Work Journey," HIGHWAY RESEARCH RECORD #283, Highway Research Board, Washington, D.C.
- Stopher, P.R., and J.O. Lavender, 1972, "Disaggregate, Behavioral Travel Demand Models: Empirical Tests of Three Hypotheses," TRANSPORTATION RESEARCH FORUM PROCEEDINGS.
- Stopher, P.R., and T.E. Lisco, 1970, "Modelling Travel Demand: A Disaggregate Behavioral Approach, Issues and Applications," TRANSPORTATION RESEARCH FORUM PROCEEDINGS.
- Strutz, R.H., 1957, "The Empirical Implications of a Utility Tree," ECONOMETRICA, Volume 25, No. 2, April.
- Strutz, R.H., 1959, "The Utility Tree - A Correction and Further Appraisal," ECONOMETRICA, Volume 27, No. 3, July.
- Talvitie, A.P., 1971, AN ECONOMETRIC MODEL FOR DOWNTOWN WORK TRIPS, Chicago Area Transportation Study, Chicago, Illinois.
- Talvitie, A., 1972, "Comparison of Probabilistic Modal-Choice Models: Estimation Methods and System Inputs," HIGHWAY RESEARCH RECORD #392, Highway Research Board, Washington, D.C.
- Talvitie, A., 1972, "Aggregate Travel Demand Analysis with Disaggregate or Aggregate Travel Demand Models," Unpublished paper, Department of Civil Engineering and Environmental Science, University of Oklahoma, Norman, Oklahoma.
- Theil, H., 1971, PRINCIPLES OF ECONOMETRICS, John Wiley, New York.
- Thomas, T.C., 1968, "Value of Time for Commuting Motorists," HIGHWAY RESEARCH RECORD #245, Highway Research Board, Washington, D.C.

Thomas, T.C., and G.I. Thompson, 1970, "The Value of Time for Commuting Motorists as a Function of Their Income Level and Amount of Time Saved," HIGHWAY RESEARCH RECORD #314, Highway Research Board, Washington, D.C.

Thomas, T.C., and G.I. Thompson, 1971, "Value of Time Saved by Trip Purpose," HIGHWAY RESEARCH RECORD #369, Highway Research Board, Washington, D.C.

Traffic Research Corp. (TRC), 1967, EMPIRIC LAND USE FORECASTING MODEL, FINAL REPORT, Eastern Massachusetts Regional Planning Project.

Warner, S.L., 1962, STOCHASTIC CHOICE OF MODE IN URBAN TRAVEL: A STUDY IN BINARY CHOICE, Northwestern University Press, Evanston, Illinois.

Weiner, E., 1969, "Modal Split Revisited," TRAFFIC QUARTERLY, January.

Wickstrom, G.V., 1971, "Are Large-Scale Home-Interview Origin-Destination Surveys Still Desirable?" Paper presented at the Transportation Forecasting Committee Meeting at the Annual Highway Research Board Meeting, Washington, D.C.

Wilson, A.G., 1969, "Entropy Maximizing Models in the Theory of Trip Distribution," JOURNAL OF TRANSPORT ECONOMICS AND POLICY, Volume 3.

Wilson, A.G., 1972, "Travel Demand Forecasting, Achievements, and Problems," Working Paper 28, Department of Geography, University of Leeds, Leeds, U.K.

Wilson, A.G., A.F. Hawkins, G.J. Hill, and D.J. Wagon, 1969, "Calibrating and Testing the SELNEC Transport Model," REGIONAL STUDIES, Volume 3.

BIOGRAPHICAL SUMMARY

Moshe E. Ben-Akiva was born in Tel-Aviv, Israel on June 11, 1944. In 1968, he received a Bachelor of Science degree in Civil Engineering from the Technion-Israel Institute of Technology, Haifa, Israel. Subsequently, he pursued graduate studies in the Transportation Systems Division of the Department of Civil Engineering at M.I.T. There he received the degrees of Master of Science in 1971, and Doctor of Philosophy in 1973. His S.M. thesis was entitled "Public and Private Transportation in an Urban Corridor - The Southeast Corridor of Boston". His PhD thesis was entitled "Structure of Passenger Travel Demand Models".

The professional interests of Moshe Ben-Akiva are primarily in the areas of Transportation Systems Analysis, Transportation Economics, and Transportation and Urban Planning Models. His professional experience includes research and teaching at M.I.T. since 1968, and several summer and part time jobs with the following organizations: Cambridge Systematics, Inc. (since 1973), the Israel Ministry of Transport (1970), Charles River Associates, Inc. (1969), the Haifa Area Transportation Planning Group (1967-8), and the Tel-Aviv Master Plan Group (1966-7).

APPENDIX**List of Figures**

	<u>Page</u>
CHAPTER II:	
Figure 1	40
Figure 2	48
Figure 3	51
Figure 4	53
Figure 5	55
CHAPTER III:	
Figure 1	106
Figure 2	113