Disaggregate behavioural airport choice models

PUBLISHER

Loughborough University

This book was bound by
Badminton Press
18 Half Croft, Syston, Leicester, LE7 8LD
Telephone: Leicester (0533) 602918.

This book was bound by
Badminton Press
18 Half Croft, Syston, Leicester, LE7 8LD
Telephone: Leicester (0533) 602918.

Disaggregate Behavioural Airport Choice Models

by

M. BENCHEMAM

A Doctoral thesis
submitted in partial fulfilment of the
requirements of the award of
Doctor of Philosophy
of the Loughborough University of Technology

Supervisor: Pr. N.J.ASHFORD

To my dear wife NADJIA who had admirably supported this work although it caused her much deprivation

To my daughter AMINA

CONTENTS

SUMMARY

The identification of the distribution of air passengers among airports is an important task of the airport planner. It would be useful to understand how trip makers choose among competing airports.

The ultimate purpose of this study is to research into passengers' choice of airport so that the airport system can be planned on a more reliable basis. The choice of airport of passengers originating from Central England in 1975 is explained by constructing multinomial disaggregate behavioural models of logit form. The data used for model calibration, were collected during two Civil Aviation Authority surveys.

This work makes contribution to:

- The definition of the major determinants of airport choice
- The responsiveness of passengers' choice to changes in these determinants
- The policy implications for the regional airports
- The transferability of the model in time and space

The method of analysis has been selected after outlining the potential advantages and shortcomings of logit and probit models and after a test on the validity of the Independence from Irrelevant Alternatives (I.I.A.) property has been carried out.

The results show that the multinomial logit model used for the airport choice is good in terms of its explanatory ability and successful in predicting the choices actually made. Travel time to the airport, frequency of flights and air fare are found to be decisive factors for a passenger to select a given airport but are not of equal importance. By influencing these factors, it appears that there exists room for the transport planner to shift traffic from one airport to another to have an economically and/or environmentally efficient airport system.

In their original form, the models have been tested and found not to be transferable to the London area in 1978. However, after a Bayesian updating procedure was applied, the business and inclusive tours models were transferable. The leisure model was not statistically transferable but had a good predictive ability while the domestic model was not transferable.

Finally, subsequent directions for further research are outlined.

ACKNOWLEDGEMENTS

I am grateful to Professor F.D.Hales, my director of research for his support and for making the facilities of the Department of Transport Technology available for this work.

I would like to express my most sincere appreciation and profound gratitude to my supervisor, Professor N.J.Ashford who provided an invaluable stimulus through his suggestions and his encouraging approach.

I am greatly indebted to Dr.D.E.Pitfield for his help and advice in the field of discrete choice modelling.

Thanks are also due to all the staff from the Department of Transport Technology for their active cooperation, particularly to Mr.R.E.Caves for the useful discussions in the initial stages of this work.

In addition, I wish to thank E.N.E.S.A. for the financial support provided in order to pursue the work in this thesis.

I would like to do credit to the memory of my father-in-law who encouraged me to undertake postgraduation studies.

Finally, I would like to thank my parents whose sacrifices over many years cannot be gauged.

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

I N T R O D U C T I O N

1.1 Objectives of the study:

An important task of the planning effort of air transportation is to identify the distribution of air passengers among airports in order to provide the right facilities at the right time. For example, in the United Kingdom, one airport planning issue is how the future growth of air traffic outside the London area should be handled, i.e. to determine locations and levels of service of alternative regional airports. Failure of a regional airport to attract trip makers from the region it intends to serve would result in inefficient investments.

There is a general belief that each airport serves a paricular territory or "catchment area". However, as de Neufville (1976) pointed out: "The expression catchment area indeed conjures up a totally inaccurate mental image of how people choose transport services. Rainwater flows down a catchment area to a sewer according to physical laws; it has no choice as to the direction it will go. People, on the other hand, do have a choice as to which airport they use, and differ from water in that they can and do make a choice."

Indeed, even in regions where all airports are operated by the same agency, they compete, as a market phenomenon.

Table 1.1 shows the percentage of the total international terminating passengers at selected airports (those to be considered in this research) originating from selected planning regions (the study area to be considered in this research) in 1978 in the United Kingdom.

Regions \ Airports	Manchester	Birmingham	East Midlands	Heathrow	Luton
Yorkshire/ Humberside	33	5	7	22	11
North West	73	1	*	13	3
West Midlands	6	41	4	30	8
East Midlands	4	9	18	37	18

* less than half per cent

Table 1.1 Percentage of the total international terminating passengers at selected airports from selected planning regions. 1978. (Source: CAA, 1980).

This table shows the extent to which airports in a region handle the traffic from that region varies considerably. In North West and West Midlands regions, a local airport (Manchester and Birmingham respectively) caters for the highest proportion of regional traffic but Manchester airport handles the largest proportion of traffic from Yorkshire and Humberside (Leeds/ Bradford, the region's local airport handling only 8%) while Heathrow is the principal airport for the East Midlands region (37% versus 18% to East Midlands airport).

It is therefore important to obtain a better understanding of how trip makers choose from among competing airports. The ultimate purpose of this study is to research into passengers choice of airport to help:

- Airports and airlines managements take the appropriate decisions to increase their market share,

- Airports planners to decide on investment priorities in order to have a more efficient airport system.

The main aim, after ascertaining the factors which determine the choice, is to construct a model which shall be based on the microeconomic theory of consumer choice rather than on mere correlation between observations. It was felt that, if such a tool could be developed, it would be useful in forecasting the redistribution of passenger traffic among airports if new facilities were to be added to the system. It could also be helpful in determining the optimum location of such new facilities or in forecasting a redistribution of traffic which might result from improvements in airport ground access such as high speed rail or motorway, or from the effects of shifting airline flights from one airport to another.

A second motivation for carrying out such a study is that the literature available leads to the conclusion that very little work has been attempted in this field, despite an increased awareness of the important implications of passenger behaviour on airport traffic. In the course of this research, only one piece of work using disaggregate behavioural model was discovered; this was being carried out concurrently by the Civil Aviation Authority (CAA, 1984) with respect to passengers' choice between the two major London airports, Heathrow and Gatwick. In August 1984, however, the modelling work was stopped and a report presented the results of the research to date. Previous works in this field are based on aggregate data, not on the behaviour of the passenger. The statistical modelling techniques traditionally used such as multiple regression, are not conceptually suitable for building behavioural choice models and the results of previous works suggested that aggregate models cannot adequately explain airport choice. However, in other fields, several studies have used behavioural models. The number of applications have been increasing significantly over the last years for example in areas of transport (e.g. mode choice), education (e.g. college choice), recreation (e.g. choice of leisure activities) etc. This large range of meaningful applications and the inadequacies identified in aggregate models lead this study to investigate the feasibility of logit analysis, using data collected at the most disaggregate level, as an analytical technique for explaining airport choice. This method of analysis has been selected after outlining the potential advantages and shortcomings of logit and probit models and after a test on the validity of the Independence from Irrelevant Alternatives (I.I.A.) property has been carried out. For reasons advocated in Chapter 5, it was decided to analyse multiple choice decisions (choice in a set of more than two airports) and so to use multinomial models.

The main constraint to disaggregate behavioural airport choice model has been the weakness of available data which lack the desired detail. The information required could not be collected on a routine basis from the available air transport industry data and it is essential to design and carry out a survey to collect these data. Often, the considerable costs involved will not allow the fulfilment of this task. For the purpose of this study, the data needed were available from two origins/destinations surveys (CAA, 1976, 1980) carried out for the Civil Aviation Authority by National Opinion Polls Ltd.

This research is intended to determine the importance of various factors which influence passenger's choice of airport. The responsiveness of airport choice to changes in its determinants will also be assessed, using elasticity measures. The policy implications, particularly for the regional airports, will be discussed.

This work also aims to study whether the model developed for an area in a given time can be transferred in another area in a different time period. Transferring existing models to new situations and areas provides a low cost alternative to the development of new models which is relatively expensive in data requirements and staff analysis time.

1.2 Scope of the Study

In Chapter 2, a review of previous work is presented. These published works have been broken down into two distinct categories reflecting different approaches: behavioural and non-behavioural models. Chapter 3 defines the determinants of airport choice and selects those which will be included as explanatory variables in the model. Chapter 4 describes the surveys with the definition of the study area, the airports considered, the passengers' origins and destinations, the form in which each variable will be introduced in the model and the data preparation. Chapter 5 outlines the general methodology with emphasis on the theory of random utility maximization, the structure of the decision-making process and the advantages of disaggregate and multinomial models. Subsequently, the analytical form of the models is derived. The methods of evaluation of the model are also presented. Chapter 6 discusses the potential advantages and shortcomings of the multinomial logit model and the multinomial probit model and the more appropriate model is selected after carrying out a test based on conditional choice. Chapter 7 presents and analyses the results of the model calibration for the four samples of passengers (domestic, international business, international leisure and international inclusive tours). An evaluation of the responsiveness of choice to changes in the significant explanatory variables is also presented. Chapter 8 tests the transferability of the model in time and space; from 1975 Central England to 1978 London area. Chapter 9 outlines the general conclusions of this study and makes recommendations for further research.

CHAPTER 2

REVIEW OF PREVIOUS WORK

2.1 Introduction

Relatively little work has been published concerning the trade-offs air passengers make when choosing between alternative airports in comparision to the considerable work that has been done elsewhere in the field of modal choice modelling. Research into airport choice modelling can be separated into two distinct approaches : behavioural and non-behavioural models. The term behavioural is interpreted differently by different investigators. Some persons classify a model as behavioural if a given statistical technique is used in estimating its parameters. Others imply that only a model which is based on attitudinal survey data is behavioural. Domencich and McFadden (1975), define a behavioural model as one which represents the decisions that consumers make when confronted with alternative choices. For this study a behavioural model is a model which attempts to describe the causal relationships between socioeconomic and transport system characteristics on the one hand, and travel decisions on the other.

2.2 Non behavioural models

The distribution of air passengers among airports has usually been identified, in past studies, using a technique which distributes traffic as a function of relative attractiveness of the airports.

The method of assigning the traveller to the minimum generalised cost mode and route, was used by the UK Civil Aviation Authority (CAA) to allocate traffic to regional airports in 1976 (Department of Trade,1976).

The traffic allocation model handles international terminating passengers only, and excludes domestic traffic and passengers interlining from one international flight to another. The forecasts of total traffic in the system for a particular year are prepared on the basis of six types of passengers: business, charter and other leisure traffic each sub-divided into UK and foreign. This traffic is then allocated between 260 separate geographical areas covering the whole of Great Britain.

The model simulates, in respect of each block of passengers (for example, a United Kingdom businessman travelling from Leeds to Paris) the total journey costs and total journey times for the various ways of getting from the journey's origin to its overseas destination.

Each block of traffic is assumed to travel by the route offering the most attractive combination of cash cost and journey time.

In its basic form, the model reproduces a "free market" situation in which the travel costs facing the air travellers are those which would be expected in the normal course of events, and in the absence of any constraints on the expansion of any airport in the system.

The basic output of the model is the distribution of total input traffic among airports broken down into the categories in which the input traffic was defined.

The basic problem of dividing low traffic density between too many airports requiring a graded hierarchy of facilities, is crucially dependent on the allocation model. Any lack of accuracy in the model is not too important to the London airports, but small errors in total U.K. traffic could cause serious discrepancies at some smaller regional airports. This method could also be criticised on the ground that it took no account of the airports' level of service and that the possibility of error was large because the airport access cost was only a small part of total trip cost.

The US Federal Aviation Authority (1979) traditional approach of forecasting the distribution of air passengers among a group of airports was to allocate national forecast to individual airports according to their market-share. For example, the domestic passenger emplanements at airport i were given by

$$E_i = M_{i/j} \cdot M_{j/s} \cdot M_{s/us} \cdot E_{us}$$

where

$M_{i/j}$ = % market share for airport i of scheduled domestic total enplanement in region j

$M_{j/s}$ = % market share for region j of total for state s

$M_{s/us}$ = % market share of state s of total for US

E_{us} = total scheduled domestic enplanements in US.

The proportions were considered constant to a time horizon equal to the time-span used for calibration. This gives very poor results for smaller airports. This approach is not useful for examining the impacts of changes in the availability and level of service of competing airports.

As a partial remedy, Ruben and Fagan (1976) applied a procedure similar to that used in ground transportation to forecast air passengers in the Washington-Baltimore region which has three competitive air carriers airports. A 2-pronged approach was used. Macroforecasts were developed based on national forecasts and historic market shares and were distributed to destination cities. At the same time, microforecasts were used in which forecasts were developed for 72 aviation analysis zones based on the relationship between trip generation and socioeconomic factors in the region. The model, in the form of a set of regression equations, does not incorporate airport attributes as variables and, thus, still lacks the ability to deal with changes in the make-up of competing airports.

Augustinus,J.C. (1973) considered that a major question in analysing and solving such problems as how existing facilities in the New York - New Jersey metropolitan area, can be utilized most efficiently, how the system might possibly be expanded etc..., was to determine on what basis an air passenger makes his choice of airport in a multi-airport system. To do this use was made of the Rand model which is a mathematical formulation of the passenger's choice of airport as a function of ground access time. The formulation, a variant of the well-known gravity model often used in transportation studies, was as follows:

$$W_{ijk} = \frac{\left\{\frac{1}{T_{ijk}}\right\}^{\alpha}}{\sum_{\ell=1}^{P} \sum_{m=1}^{Q} \left\{\frac{1}{T_{ilm}}\right\}^{\alpha}}$$

in which:

W_{ijk} = fraction of passengers from centroid i travelling to destination k which will select airport j;

i = area centroid = 1,.....A

j = airport, = 1,....P

k = destination = 1,....Q

T_{ijk} = travel time from centroid i to airport j (roadway time, process time, waiting time, etc...) as related to destination k.

This model says in essence that the fraction (W) of the total passenger volume originating in a particular centroid (i), and terminating in (k) which is to select airport (j) is a function of the access time from this centroid to the particular airport versus the access time to all other competing airports. The value of $\alpha = 2.4$ produced the best fit to the data.

However, Augustinus recognized that the possibility existed that this value was biased because passenger distributions are affected by other factors besides access time, such as access cost, schedule availability and possibly psychological factors, such as crossing the Hudson River, etc. when more factors have to be entered into the analysis and not one but more α's have to be estimated, iterative methods become involved. Therefore, further attempts at measurement have been based on a somewhat simpler estimating model by means of linear regression. This model reads as follows:

$$W_{ijk} = b_1 \frac{T_{ij}}{\sum_{j=1}^{p} T_{ij/P}} + b_2 \frac{C_{ij}}{\sum_{j=1}^{p} C_{ij/P}} + \ldots$$

where the b's are regression coefficients, representing passenger sensitivities with respect to access time, access cost, etc.

T_{ij} = travel time from centroid i to airport j

C_{ij} = access cost from centroid i to airport j

The effect of length of haul of the passenger's air trip has been considered by stratifying the passengers into five groups by mileage range.

Basically, the second model says that the air passenger, confronted with a choice of airports representing different convenience levels in terms of access time, cost and possibly other factors, will make his choice by weighting the convenience characteristics of each particular airport in relation to those of the other choices available to him.

However, the equations leave much to be desired (the best model has an $R^2 = 0.72$); the standard errors of estimate are still large. Augustinus pointed out that the schedule variable should be more precisely defined in terms of destinations and probably also with respect to departure time rather than in terms of the fairly imprecise criterion of "length of haul".

In a case study of Texas intrastate air markets, de Neufville and King (1979) considered the effects of airport access and of airline fares and frequency on airport traffic using the quasi-experimental method. Indeed, a series of important quasi-experimental situations arose in this market between 1970 and 1975 that enabled this research to examine the importance of these variables in a relatively controlled environment. Specifically, almost all the Dallas air passengers had to shift from Dallas/Love Field to the new Dallas/Fort Worth airport in 1974, and southwest Airlines engaged its competitors in a series of price and frequency "wars" that led to a series of sudden changes in fares and schedules in narrow markets. The research design comprised a series of before-and-after analyses through which was estimated in turn and then was compared the effect of airport access and of airline fares and frequency on air traffic. It was concluded that:

- Airport access considerations appear to have an important impact on short haul travel. This effect occurs both when travellers have a choice of airports, showing a strong preference for convenient access and when the only available airport is made more remote, there is a resulting decrease in overall demand.

- The fare elasticity is approximately -1.

- Relative frequency is an important determinant of market share, according to an S-shaped relationship, when no other significant factors intervene. When significant price and locational considerations enter, relative frequency appears to become a secondary factor.

However, the quasi-experimental method has a particular limitation. The changes of interest almost without exception come about to resolve some operational problem (e.g. to raise profits or to accommodate the use of larger aircraft, etc...) rather than to satisfy research purposes. The changes are therefore not modulated over a range for the convenience of the observer. Thus statistical analyses are often difficult to apply if not meaningless.

These studies have shown that non-behavioural models have particular limitations in the identification of the distribution of air passengers among competing airports. Most of the approaches used are not useful for examining the impacts of changes in the availability and level of service of competing airports because they do not incorporate airport attributes as variables. These models also ignore the important fact that the passenger can and does make a choice as to which airport to use; the passenger does not necessarily use the airport offering the least expensive or most convenient access. The model formulation often represents nothing more than curve fitting with an a priori assumption between airport choice and system characteristics. Besides, a model that results in a good fit to existing data cannot guarantee a reliable estimate for the future because the contribution of passengers' behaviour has not been considered.

These reasons suggest that non-behavioural models cannot adequately explain airport choice.

2.3 Behavioural models

The studies reviewed in the preceding section have suggested that non-behavioural models perform poorly even when they include the most important a priori factors affecting the airport choice or that they have particular limitations. This does not, however, seem as yet engendered the use of behavioural models to explain airport choice.

A behavioural model attempts to explain the choice behaviour of the individual. The axioms of behavioural models are that individuals represent the basic decision making unit and that each individual will choose an alternative among those available that he or she finds most desirable or useful. The model seeks to explain the probability that an individual will choose a given alternative in terms of a number of system and user characteristics. Such models have been tested using a number of statistical techniques most notably discriminant analysis, probit analysis and logit analysis.

As stated in Chapter 1, the passenger can and does make a choice as to which airport he uses. It is likely that the factors influencing the choice outcome will differ in volume and range across the population. This argues for the relevance of studying the passenger's behaviour in the prediction of airport traffic. Thus, a behavioural model is expected to provide a better approach to the problem under study.

In the course of this research, only one such piece of work was discovered; this was being carried out concurrently by the Civil Aviation Authority (CAA, 1984) with respect to passengers' choice between the two major London airports, Heathrow and Gatwick on the basis of the 1978 survey (CAA, 1980). The stated aims of this research were:

- To ascertain the key factors which determine the choice of London airport for passengers on scheduled services who travel to those airports by surface mode.

- To construct a model or models which could be used to predict the changes in passenger flows resulting from different assumptions about the factors affecting those flows.

In August 1984, the modelling work was terminated because it was felt that "the particular combination of software and computer bureau being used, while producing the required results, was not providing the best value for money"; a report presented the results to date.

The approach to the data collection and preparation was similar to that in this research. The relevant observations were extracted from the 1978 survey where 91,086 passengers were interviewed. As with this research, only the complete interviews were considered. A number of assumptions were made concerning the surface journey characteristics, for example, the car times were based on a speed of 60 mph on motorways, 40 mph on rural roads and 20 mph on urban roads. These speed values are reasonable and similar to those considered in this work. The international passengers were stratified by length of haul and into business and leisure. The inclusive tours passengers were not considered as there were no charter flights from Heathrow. However, this category of passengers had a choice, in the London area, between Gatwick and Luton airports and this choice could also have been modelled. In this research, the distinction between inclusive tours passengers (package holiday) and leisure passengers is made. The multinomial logit model (which is presented in Chapter 5) was used by the CAA and calibrated, for the business passengers only, using the Quail computer package with the assumption that the traveller makes a simultaneous choice of access mode and departure airport. This structure of decision making is believable for the business passengers particularly as a large range of access modes to the London airports (car, train, tube, bus, BR bus) is available with no predominant mode. This structure of decision making, however, has not been considered in this research for reasons stated in Chapter 3. In the CAA study, the set of alternatives (nine in total) as distinguished by the mode of arrival to the airport (e.g. Gatwick by car, Heathrow by tube etc.) were determined by a priori considerations. The "Independence from Irrelevant Alternatives" difficulties, which are discussed in Chapter 5, were not mentioned and the multinomial logit model was chosen for its computational convenience.

The estimation for the CAA was carried out for a sample of 319 business passengers starting their journey in London and travelling to a short or medium haul destination (mainly Paris, Amsterdam and Brussels). The effect of a multi-airport situation at Paris on the choice outcome was not considered. Indeed, the choice of London airport is inextricably linked with the choice at the other end between Orly and Charles de Gaulle airports. Thus, in this research, Paris will be excluded from the set of selected destinations. For the CAA, the passenger was assumed to be influenced by:

- surface access time and cost
- changes between modes of transport
- air journey characteristics: frequency of flights and the earliest arrival time at the destination.

It was reported that the access fare was highly correlated with taxi distance. The number of changes between modes of access was found not to be significant. The variable earliest arrival time at the destination was found not to be so effective and the vast proportion of airport specific effects could be explained solely in terms of the number of flights available on a given day at the airport. It was, thus, concluded for the business sample that:

- Travellers were influenced both by surface access characteristics and by frequency of flights available to their chosen destination.

- These two influences provided a model which predicted satisfactorily the aggregate proportions of travellers choosing the various options. Table 2.1 compares the model predicted shares with those observed.

The air fare variable was, rightfully, not considered as there were no fare differences between Heathrow and Gatwick for the selected destinations but this variable will be investigated in this research as fare differences do in this case exist among the selected airports.

	Heathrow		Gatwick	
	Obs %	Pred %	Obs %	Pred %
by car	55.6	53.3	5.6	6.2
by taxi	15.4	14.3	0.7	1.0
by train	-	-	2.3	1.0
by bus	4.6	5.7	0.0	0.5
by tube	15.0	17.1	-	-
by BR bus	0.8	0.9	-	-

Table 2.1 Observed and predicted shares - London area 1978 (Source: CAA, 1984)

However, in other fields, several studies have used behavioural models, for example in areas of transport (particularly in modal choice modelling), education, recreation, communications etc.

The number of studies using the logit model (binary or multinomial) has mushroomed in recent years. Benabi (1983) developed a disaggregate binary logit model to explain and predict freight mode choice, from the shipper's point of view, in the manufacturing industry in Leicestershire (U.K.). The alternatives considered were road and rail. A "consignment-shipper" oriented survey stratified according to industrial affiliation was designed and carried out. Different mode choice variables were determined for each sub-group of industry. The results have shown that the models are good in terms of their explanatory ability and highly successful in predicting the actual choice of mode (the overall percentage correctly predicted ranges from 81.3 to 97.6).

The study by Rassam and al (1971) was the first to extend the logit model to analyse the choice between multiple transportation alternatives. The specific application deals with the choice of access mode to airports in the Washington Baltimore area where four modes were available. The model was estimated both by maximum likelihood technique and by constrained least squares regression. The authors report that the estimated parameters had the expected signs and relative magnitudes as well as low standard errors of estimate.

Another major area of development has been that of destination choice with the work of Adler and Ben-Akiva (1976) who used simple measures of destination attractiveness such as employment and used the destination within a specified cost and time range as the choice set. Other successful efforts have been concerned with choices of residential location (Lerman (1975)) and auto-ownership (Atherton and Ben-Akiva (1977)).

Even though multinomial probit models (which are presented in Chapter 5) have many attractive theoretical features, they have not been used in practice due to the lack of an adequate numerical technique for their application until the late 1970s when the work of Albright, Lerman and Manski (1977), Hausman and Wise (1978) and Daganzo, Bouthelier and Sheffi (1977) made significant progress in developing practical estimation procedures. Binary probit, the estimation of which was computationally feasible, was used for a wide range of choice modelling problems. We can mention the work of Lisco (1967) and Lave (1968) who both used probit analysis to calibrate a modal choice model. K.H.White (1975) identified and analysed some of the factors influencing the consumer's decision to use bank credit card versus checking accounts.

The CAA attempt to develop a behavioural airport choice model which had encouraging results and the wide range of successful applications in other fields provide evidence of the practicability of behavioural models.

2.4 Conclusion

In this chapter it is shown that:

- Non-behavioural models do not adequately explain passengers' choice of airport

- Only one attempt to explain airport choice has been made using a behavioural model

- The use of behavioural models to reflect choice behaviour, has provided, in other fields, evidence of the practicability of such models.

Thus, a behavioural model is expected to provide a better approach to the problem under study.

CHAPTER 3

DETERMINANTS OF AIRPORT CHOICE

3.1 Introduction

Having decided to fly, the traveller should answer the question of the airport from which to fly. The aim of this Chapter is to identify and select the determinants of this choice to be included as explanatory variables in an airport choice model.

3.2 Identification of variables

The central problem facing airport designers is that airport traffic is highly erratic. The number of passengers using an airport may easily double in 3 to 4 years, or stagnate or there may be a shift from one airport to another. Only part of this kind of variability in airport traffic can be attributed to overall economic patterns that would affect all modes of transport (i.e. the recession of the early 1970's), or to factors peculiar to the air transport industry which would impact all airports (i.e. deregulation in the U.S.A.).

Much of the variability in the traffic at an airport may be due to how well the airport competes with other airports (de Neufville,1976).

The variables which are likely to be determinants of airport choice can be divided into three categories:

- transport system variables
- socioeconomic variables
- non-quantitative variables

3.2.1 Transport system variables

In choosing between airports, a passenger attempts to minimise its travel times and travel costs. Specifically, the distance between an airport and its market is likely to affect the number of passengers that use this airport in preference to another airport. Likewise, the frequency of air service from an airport may well be a deciding factor for or against the use of an airport. In addition, the general level of fare is also clearly important in this regard. Other variables such as changes between modes of access, length of haul or airline nationality could also play a role in the passenger's choice of airport.

3.2.2 Socioeconomic variables

In many studies in mode choice modelling, income was assumed to be an important variable, to which age and sex were often added. The effect of socioeconomic variables on the choice process can be reflected by two possible strategies. First, socioeconomic variables can enter the model as interactive variables. Three problems arise in this method. The socioeconomic variables so used may enter one or two terms of a multivariate utility function, or else severe multicollinearity problems arise. The number of socioeconomic variables that may enter is likewise very restricted and the choice of which variables they should interact with becomes crucial to model performance. Further, many socioeconomic variables are measured using grouped information, e.g., age group, income group, etc. These groups provide, at best, only interval measures and may often provide nonmetric information. The use of such variables as quantitative interaction terms is clearly ill-advised (Stopher and Meyburg, 1976). The second method of reflecting the effect of socioeconomic variables on the choice process is to use these variables as the basis for stratifying the population. This method, however, requires substantially larger data sets for calibration, because of the need to have significant subpopulations in each stratum.

3.2.3 Nonquantitative variables

A passenger can choose an airport simply because it was recommended by a travel agent or a friend or a relative. This recommendation, however, could be based upon variables already listed (e.g. access time to airport, frequency of flight, etc.). The passenger can also decide to fly from a particular airport because he is familiar with it from previous trips. These variables can enter the model as dummy variables. The use of such measures, however, is not useful for planning applications.

3.3 Selection of variables

The previous section has shown that a large range of variables can enter into the airport choice model. Of all these factors, however, many are not readily available or are incapable of being quantified or they are highly inter-correlated. Before selecting the variables which will be included in the model, it is necessary to discuss some modelling aspects which have a direct impact on the a priori selection of variables.

First, a number of different hypotheses can be made regarding the structure of passengers' decision making. The passengers' whole choice process will involve a number of decisions, for example:

- Whether to fly or not
- On which day of the week to fly
- Which airport to fly from
- How to get to the airport

With the data available, it is not possible to construct models to examine the first two decisions above. The choice of departure airport and choice of surface access mode, however, are decisions which can be examined using the CAA survey data. It is also not possible to know a priori the structure of passengers' choices in these two decisions. A passenger might choose first an airport and then an access mode or first an access mode, then an airport or he might make a simultaneous choice of an airport and an access mode. Because none of the three Central England airports, which are considered in this study, had direct rail connections and because the automobile was by far the predominant mode of access (9 out of 10 passengers used the car), the two structures are nearly indistinguishable and can be reduced to the situation where the passenger is only choosing the airport from which to fly. However, it should be acknowledged that different classes of passengers might have different choice structures. The most appropriate structure for the business passengers may not be applicable to the inclusive tours passengers. Indeed, if the business and leisure passengers structure could be a single choice of airport, the inclusive tours passengers could, for example, choose first a holiday destination or a travel company and then an airport or they could have a more complicated choice structure. The data available do not permit the examination of the particularities of the inclusive tours passengers' hierarchy of decision making and it will be assumed that the structure is a choice of airport for all the passengers irrespective of their trip purpose. Consequently, the same combination of variables will be included in the model for each category of passengers which will be stratified by trip purpose.

Secondly, the purpose of model building has a direct influence on the strategies of model building and thus on the selection of variables. There is little point in modellers producing models with increasing and varied levels of sophistication, if they still fail to provide the inputs to the analyses necessary to the policy planner. The balance between model complexity and efficiency should always be kept and a good model should be responsive to policy questions (Hansen and Rogers 1979; Domencich and McFadden 1975). The airport system attributes are not only crucial to the passengers who base their choices on them, they also constitute very important parameters in the hands of the policy makers since they are under their control and also provide a basis for evaluating alternative policies. The aim of this work is to build a pragmatic model which could be easily applied in practice by airport and airline managements. Indeed, these managements have under their direct control such variables as frequency of flights and air fare and although they have little control over the access to the airport (clearer road signs, easier parking,etc.), they can strive for regional transport improvement to ease it. Therefore, if the policy sensitive variables can explain with an acceptable degree of accuracy the passengers' choice of airport, the analysis will be confined to these variables in order to avoid limitations to the model's practical applications. However, it must be admitted that this strategy of model building, where the emphasis is only on the airport characteristics, could be rather odd as far as the behavioural dimension of the model is concerned. The effect of the socioeconomic variables can also be reflected by their use as the basis for stratifying the passengers but again the gain in model accuracy has to be balanced against the requirements of substantially larger data sets.

Thus, three main criteria should guide the selection of the variables to be included in the model:

- These variables should reflect important determinants of airport choice
- These variables should, as much as possible, be policy sensitive
- The data relating to these variables are available.

The determinants of airport choice can be best identified by asking the passengers the reason for choosing a particular airport. This question has been asked during the 1978 CAA survey but, unfortunately, at Luton airport only.

From a sample of 6,450 international passengers at Luton, 3,004 passengers had their holiday available only from Luton airport. For the remaining 3,446 passengers, the reason for choosing Luton airport was given as:

- 1,287 passengers (37.56%) : my home is easily accessible to the airport
- 281 passengers (8.15%) : holiday was cheaper at Luton
- 259 passengers (7.52%) : convenient flight times
- 976 passengers (28.32%) : recommended by a travel agent

- 64 passengers (1.86%) : recommended by friend or relative
- 61 passengers (1.77%) : familiar with airport from previous trip
- 15 passengers (0.44%) : like the airport and amenities
- 503 passengers (14.60%): other and no answer.

It is clear that for those passengers who had a choice of airport, accessibility, recommendation by a travel agent, cost and convenience of flight times were the main reasons for choosing Luton airport. Thus, access to the airport, cost and frequency of flights variables will be selected to enter the airport choice model. The determinant "recommendation by a travel agent" which is important for the inclusive tours passengers will not be included because of lack of data (the question about the reason of choosing an airport had not been asked during the 1975 CAA survey). Future research should, however, investigate this variable. It should be, however, acknowledged that Luton airport may not be typical of all airports, unless one is concerned only with the inclusive tours passengers.

3.3.1 Convenience of airport location

The problem of airport access, in many parts of Europe, remains the principal bottleneck in the total air trip from origin to destination. It is not unusual for a centre-to-centre trip from London to Paris to take 4½ hours of which the actual air time for the block from Heathrow to Orly airport constitutes less than an hour (Ashford and McGinity,1975). For most users, airport access is about equivalent to the duration of air travel. Furthermore, the great environmental technical pressures for a planner to choose an airport site far from the city increase passenger difficulty in reaching air services and this may depress the level of air traffic. In any event, it has been seen that transferring airline services to a less accessible facility can significantly decrease traffic especially for air trips over short distances. This is illustrated by an example from de Neufville (1976) for short haul traffic at Detroit airport in Figure 3.1.

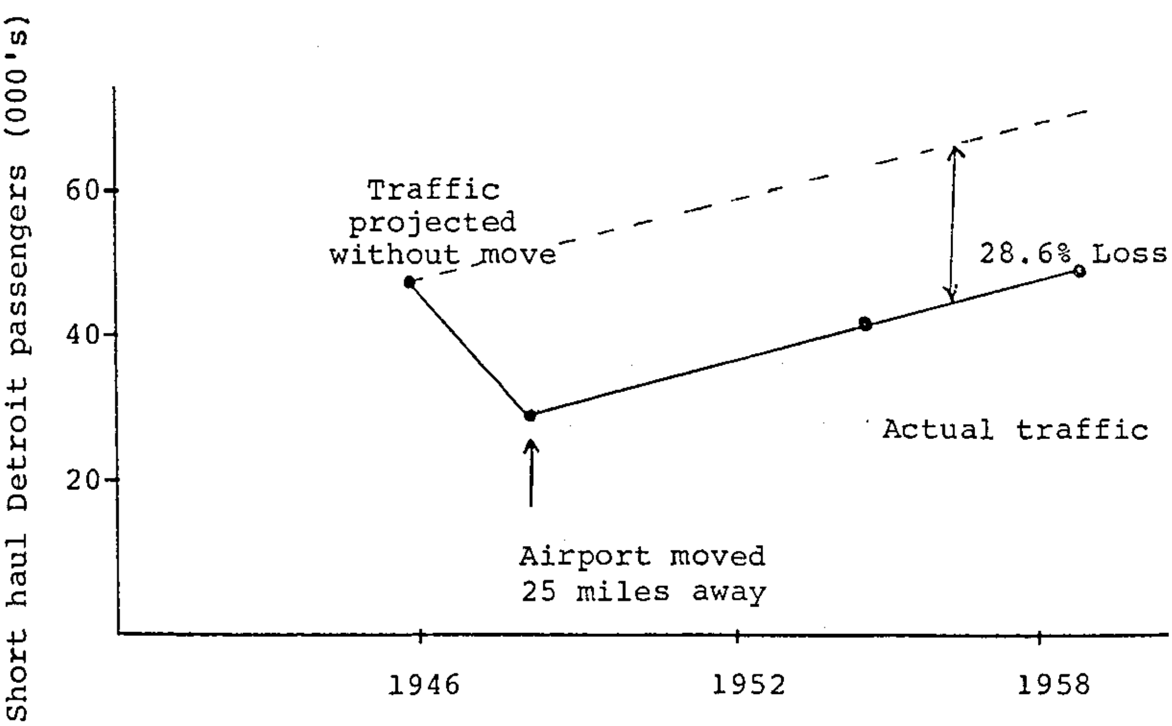

Figure 3.1 Sensitivity of Air travel over short distances to changes in accessibility.

What happened elsewhere, at another time and in a different environment should not be assumed to apply to a new situation. But similar drops in traffic have been associated with the opening of other remote airports (Genest, 1970), thus emphasizing the importance of airport location on the level of traffic at this airport.

This determinant of airport choice will be included in the airport choice model in the form of ground access time from the passenger's origin to the airport. The use of an access time variable instead of a distant variable will take into account the quality of the road network.

In general, airport choice behaviour may be affected by available modes of ground access. Table 3.1 shows the primary mode of transport used by passengers at Central England airports in 1975.

	Percentage of passengers		
Airport	Car (private, hire, taxi)	Bus (public, charter, airport)	Other
Birmingham	93	5	1
East Midlands	84	15	1
Manchester	88	11	1

Table 3.1 Airport by mode of transport (Source: CAA, 1976).

None of the three airports had direct rail connections. The automobile was the predominant mode (93%, 84% and 88% of terminating passengers at Birmingham, East Midlands and Manchester airports respectively used the car to access the airport). Thus there is no need to consider a complicated access time measure which would have taken into account all the access modes to the airport. Such a measure would not materially improve the model.

3.3.2 Frequency of air service

Frequency of air service is often a crucial factor for a person contemplating which airport to use. The airport with more flights to a particular destination will almost inevitably offer more convenient departures and thus decreasing the passenger's schedule delay (the time between the desired departure time and the best available flight). Also, more flights will increase more abstract elements of convenience, such as the availability of backup departures, if the planned one is missed. Likewise, a person travelling to a city with several airports often prefers to use the one with the greater service because it offers more possibilities for transferring to connecting flights.

The relationship between frequency of service and its attractiveness is generally represented by S-shaped curves of the type appearing in Figure 3.2 (Source: de Neufville, 1976). Figure 3.2 shows that the airport with more flights will serve more than a proportional percentage of passengers. Indeed, if the market/supply ratio was constant, the observed curve would be the straight line. This tendency of the passengers to "go with the winner", i.e. to travel disproportionately from the airport that offers most services is well known in the air transport industry (Gelerman and de Neufville,1973). Specifically, Figure 3.2 shows when Gatwick offers about 30 per cent of the flights to another city, it only obtains 20 per cent or less of the market.

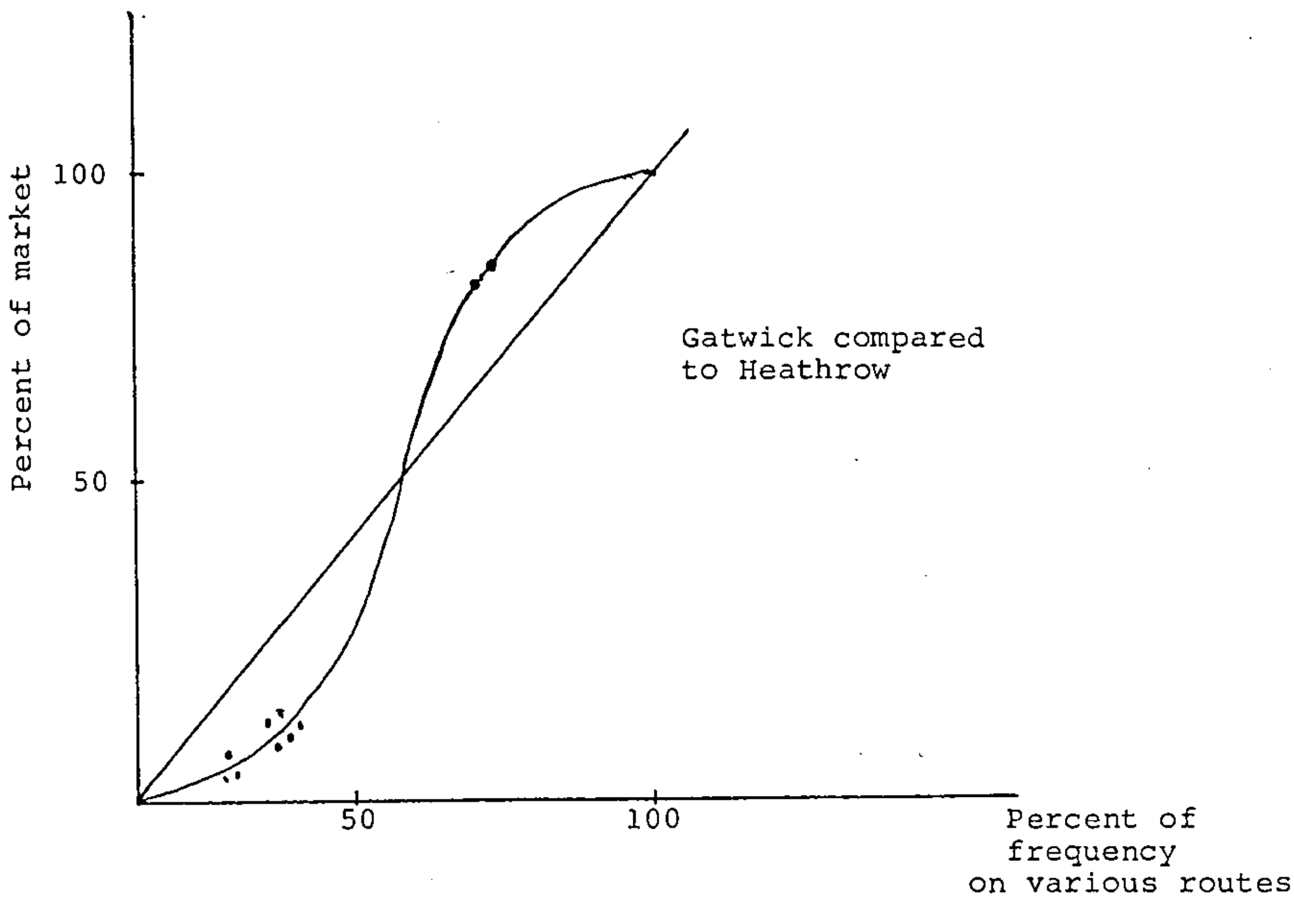

Figure 3.2 Attractiveness of frequency of service.

In a survey conducted by Caves et al (1983), in the East Midlands area, businessmen have been reported saying they will use a direct international flight from East Midlands airport instead of London if the frequency of flights offered at East Midlands airport is at least daily (Monday to Friday).

Many passengers may, of course, attach little importance to frequency of flight. For example, holiday travellers leaving on a charter flight may only be concerned about a single specific departure. This behaviour does not, however, contradict the general rule which is that frequency of air service is a major factor in determining the attractiveness and use of an airport.

This determinant of airport choice will be included in the airport choice model in the form of number of direct flights per day offered from an airport to the traveller's given destination. Since the routes to be examined are among the densest routes that most passengers travel on non-stop flights and since non-direct flights provide a different airport to airport travel time, only direct flights have been considered.

It will be advantageous to discuss some aspects of the influence of expressing the frequency of flights in a manner more directly related to time, particularly for the business passengers whose preference to arrive in time for meetings etc., is well known. It is clear that a flight at 5.a.m. has not the same attraction as a flight at 8.a.m. and a specification of the frequency variable in terms of departure time rather than in terms of number of flights per day could be more precise. However, the number of scheduled daily flights from the Central England airports to the selected destinations in 1975 was low (usually 1 or 2 flights per day) and the passenger was looking more for the availability of a flight than for the convenient departure time. Furthermore, air carriers' competitive scheduling attempts to space flights in prime periods so that if the number of flights from two competing airports to a given destination is similar, these flights will certainly be departing in the same slot of time. Despite these arguments, it should be acknowledged that the purpose is in fact to measure the waiting time of the passenger which can be broken down into two components:

- the frequency delay which is the difference between a traveller's desired departure time and the closest scheduled flight time

- the stochastic delay which results from the possibility that the selected flight may be sold out.

The sum of these delay times is probably the most relevant variable in the airport choice model. It is a function of flight frequency and average load factor. The degree of difficulty in calibrating that variable (e.g. as the passenger's desired departure time is not known, a simulation probably being necessary) suggests that a simpler form could be acceptable as a first approximation. In any case a topic for further research could compare the different specifications of the frequency variable and the consequent gain/loss in the model's accuracy.

3.3.3 Air fare

The passenger's choice of airport is also clearly determined by the out of pocket cost for the total trip. Travellers are more likely to choose the departure airport which leads to the smallest total cost. This total cost includes the air fare and the access cost. The access cost varies with the distance between the passenger's origin and the airport and then is collinear with travel time to the airport which was covered in paragraph 3.3.1. Thus, the cost variable will be represented by the air fare.

People will often undertake longer and more costly journeys in pursuit of the low fare. This is illustrated by the example in North America in 1983 where for a variety of reasons, including airline deregulation in the United States, fares on Canadian airlines were generally higher than those on US carriers. Canadian travellers crossed in great numbers the natural border to take advantage of the low fares at Burlington airport (Vermont) located 56 km from the U.S. Canadian border and were the major reason for the airport's growth rate of more than 50% in 1983 (Airport Forum,1983).

Other evidence that air fare is usually considered is also reported by Caves et al (1983) for the businessmen in the East Midlands. These potential passengers said that they will use a direct international flight from East Midlands airport if the air fare is not "significantly" higher than the air fare from London.

The air fare variable will be introduced in the airport choice model. The normal economy class (Y) fare in force in summer 1975 has been systematically recorded. Despite all the criticism to which it is open, this value is a significant factor to the extent that it applies to a "basic" clientele, i.e. those who pay the full fare and justify the existence of scheduled services, and to the extent that it is used as a reference for the calculation of reduced fares. The air fare assumption about the inclusive tours passengers is presented in Chapter 4.

3.3.4 Other variables

Beside access time, frequency of flight and air fare, other variables might influence, to a lesser degree, travellers' choice of airport.

The influences on choices are likely to be different for different lengths of haul; domestic, short haul international, long haul international. From a theoretical point of view, it is reasonable to postulate that short-haul passengers would be more sensitive particularly with respect to access time and more discriminating in their choice of airport than long-haul passengers. The air network serving Birmingham, East Midlands and Manchester airport in 1975 was essentially made up of domestic and European routes (no long haul traffic). Thus, as the passengers will be stratified in domestic and international in this study, this length of haul variable is not important and will not be included in the model.

Changes between modes of access transport could also have some influence in the choice of airport. However, in this study, the percentage of passengers with only one mode of transport was 94.4, 92.8 and 91.1 at Birmingham, East Midlands and Manchester airports respectively. This variable will then not be considered being of prime importance and will not be included in the model. Other variables such as airline nationality or passenger's experience in flying, could also play a role in the passenger's choice of airport but are not such vital criteria as those already selected. They will not be included in the model to avoid limitation of the model's potential for practical applications and would be a refinement which could be attempted later if the modelling concept proved successful.

The passenger's age is also important in regard to decision making. Thus, only passengers who were over 16 have been retained.

3.4 Conclusion

In this chapter, we defined the set of variables which might influence travellers' choice of airport as:

- surface access time
- frequency of flight
- air fare

Although there is no limit to the number of explanatory variables that could be included in the model, the efforts in this analysis were confined to these three variables which intuitively seem to be the most dominant factors in the choice decision. It was felt that anything more complicated would severely limit the model's potential for practical applications, and in any case, would be a refinement which could be attempted later if the modelling concept proved successful.

CHAPTER 4

DATA REQUIREMENTS AND PREPARATION

4.1 Introduction

In Chapter 1, it has been stated that the airport choice model should be a disaggregate behavioural model whose calibration uses data collected at the most disaggregate level (i.e. the individual level). For each passenger, the data needed are:

- His surface origin
- His flight destination
- His age
- The day of the week his flight takes place on
- His trip purpose
- His selected airport
- The non-chosen airports
- The travel time from his surface origin to all the competing airports
- The number of flights from the competing airports to the selected destination for that particular day of the week
- The air fare from the competing airports to the selected destination.

It is then essential to design and carry out a survey to collect these data. Because of the nature and characteristics of this survey, a mail survey will be by far less efficient than the direct interviewing method. The costs involved and practical considerations did not allow the initiation of such survey for the purpose of this research only. However, the sort of data needed were available from past CAA surveys (C.A.A., 1976, 1980) which were carried out for the Civil Aviation Authority by National Opinion Polls Ltd. The information collected during these two origin/ destination surveys make up the basic data for this research.

The other primary source of information was the ABC World Airways (1975, 1979) for the collection of air fare and frequency of flight data. The distances between the passengers' origins and the competing airports have been drawn from the Big Road Atlas of Great Britain (1981).

The preparation of the data consisted of:

- identifying the passengers from the study area with a choice of airport
- extracting from the survey data, the observations relating to those passengers, and
- defining for each passenger, the attributes of both the choice made and the choices rejected.

4.2 The surveys

In 1975, 40,000 passengers were interviewed between July 21st and November 30th in Scotland and Central England at the four B.A.A. Scottish airports and at Manchester, Birmingham and East Midlands airports. These airports accounted for 19% of the 42 million terminal passengers at United Kingdom airports in 1975.

In 1978, 91,086 passengers were interviewed between July 17th and December 22nd at Heathrow, Gatwick and Luton airports. In 1978, these airports handled 70% of the 52 million terminal passengers at United Kingdom airports.

At all airports, inward and outward passengers were interviewed. Systematic random samples of passengers were taken at each of the airports. The survey period was chosen to cover the peak holiday period, the shoulder and the trough in an attempt to eliminate seasonal bias.

These surveys have been designed to obtain information of a kind that could not be collected on a routine basis from the available air transport industry data. To cause minimum disruption to passenger flow, it was necessary to impose a constraint on questionnaire length and content. The question wording was chosen so that no ambiguity or double meaning could occur. The main questions included in the questionnaire were about:

- the flight : scheduled, charter, private flight.
- the route: domestic, international; arrival, departure.
- the age and sex of the passenger.
- the passenger's nationality: UK, foreign.
- the home in UK/the area visited in UK.
- the surface origin/destination. (The point at which the journey is considered to start or end).
- the method of surface transport: private car, taxi, public bus, etc.

- the journey time to the airport
- the length of trip
- the airport of current journey
- the airline of current journey
- the passenger socio-economic group
- the passenger's income (for business passengers only)
- the household income (for leisure passengers only).

The 1975 and 1978 questionnaires are in appendix A and B to this work.

The surveys data are recorded on magnetic tapes. Each tape contains a number of records corresponding to the number of passengers interviewed. Each record which corresponds to a single questionnaire contains 80 characters which are the coded answers to the questions asked. Indeed, a coding system has been used for, among others, areas (e.g. Loughborough = 310, Sheffield = 240), airports (e.g. Brussels = 2200 , Manchester = 0008), airlines (e.g. British Midlands Airways = 0830, Alitalia = 3200), main business, age and income. The area of residence of UK residents was recorded using a zonal system: regions, areas and zones have been defined in terms of the boundaries of, respectively, economic planning regions, counties and district council areas.

The magnetic tapes have been made available, for this research, by the Department of Transport Technology at Loughborough University with the permission of the CAA which provided all necessary coding information for reading and deciphering the data.

4.3 Study Area

A preliminary decision of some importance is to define what airports and what study area are to be considered. The future of Manchester, Birmingham and East Midlands airports which have been under discussion following the Central England Airports Study and the 1985 governmental decision to develop Stansted as Third London Airport make it desirable that more information should become available on the nature of air travel in this region. Consequently, Central England has been selected as the study area.

The counties included in the study area are:

- Metropolitan counties : West Midlands, Greater Manchester, South Yorkshire and West Yorkshire.
- Non-metropolitan counties : Derbyshire, Leicestershire, Nottinghamshire, Lancashire and Staffordshire.

Ten cities in this area have been selected: Birmingham, Stoke-on-Trent, Manchester, Sheffield, Leeds, Derby, Leicester, Nottingham, Blackpool and Coventry. Being the most populated centres in the study area (the population of each city ranges from 149,900 in Blackpool to 1,074,500 in Birmingham (mid-1975 figures)), these cities are more likely to generate more traffic than the less populated cities.

Air transport services are available from a number of airports in or around the study area. The airports considered in this study are Manchester, Birmingham, East Midlands, Luton (for Inclusive Tour flights only) and London Heathrow. They have been selected because:

- They are the main airports in the area, or,
- Their attractiveness is important because of the extensive services they provide (i.e. Luton for inclusive tours or Heathrow).
- The data relevant to the origin/destination of passengers are available (CAA 1976, CAA 1980).

Gatwick airport has been omitted considering that if the scheduled passengers from the study area choose to fly from a London airport, they will certainly select Heathrow.

Then, the following destinations were selected:

- Domestic : Belfast, Jersey, Glasgow, Aberdeen.
- International : Dublin,Amsterdam, Frankfurt, Brussels.
- Inclusive Tours : Palma, Alicante, Ibiza.

In this selection, the following criteria were considered:

- Destination cities must be served by direct flights by at least two of the study area airports.
- The routes serving them from the study area airports have a high density traffic.
- Destination cities should have only one major air carrier airport, otherwise the choice of airport in the study area might be influenced by the destination airport alternatives (consequently, Paris has not been retained).

4.3.1 Stratification

Stratification is often desirable when an heterogeneous population is investigated. In developing travel models, it is common to calibrate separate models for groups that differ in some respect that is believed a priori to affect the process being modelled. For the airport choice model, the passengers have been stratified on the basis of:

- Domestic
- International business (business/official, armed services, airline)
- International leisure (visiting friends/relatives, migration, studies, holiday)
- International inclusive tours (package holiday),

reasoning that these four categories of passengers are making different travel decisions.

4.4 Data Collection

Data relevant to the study include individual air passenger trip records, frequencies of flights, air fares and ground access time to the airport. They consist, for each selected passenger, of the definition of the attributes of both the choice made and the choices rejected. As an example, for a selected passenger originating from Nottingham and flying to Amsterdam on a business trip from Manchester airport on Thursday, the following actions were taken:

a) For the choice made:

- Computation of the travel time from Nottingham to Manchester airport
- Extraction from the pertinent ABC World Airways of the number of flights leaving Manchester airport to Amsterdam on Thursday as well as the economy air fare.

b) For the choices rejected:

- Computation of the travel times from Nottingham to Birmingham airport, to East Midlands airport and to Heathrow airport.
- Extraction from the pertinent ABC World Airways of the number of flights on Thursday to Amsterdam from Birmingham airport, East Midlands airport and Heathrow airport as well as the respective economy air fares.

4.4.1 Air passenger trip record

The magnetic tapes recording the surveys information have been read on the prime computer at Loughborough University. Each survey record which corresponds to a single questionnaire (see appendix A and B), contains typical air passenger survey information including those required in this study: passenger local origin, destination, airport used, access time, flight number, trip purpose and day of interview. The information relating to the passengers from the selected origins in the study area to the selected destinations using the selected airports have been extracted from the total data file. A number of records provide incomplete information so that only the fully interviewed passengers have been considered (around 75% of total passengers approached). Thus the number of records that can actually be used (passengers fully interviewed, from a selected origin in the study area to a selected destination) is diminished substantially to 2577 with the domestic data file containing 683 records, the international business file containing 214 records, the international leisure file containing 331 records and the inclusive tours file containing 1349 records.

The data for Manchester, Birmingham and East Midlands airports were drawn from the 1975 survey while the data for Heathrow (Luton for inclusive tours) were drawn from the 1978 survey. Therefore the 1978 data have to be adjusted to achieve consistency with 1975 figures. It has been assumed that the growth of air travel between 1975 and 1978 has been the same throughout a planning region (i.e. East Midlands, North West). Therefore the percentage change in the number of passengers from a selected city using Heathrow (Luton for inclusive tours) between 1975 and 1978 is the same as the percentage change in the number of passengers using Heathrow (Luton for inclusive tours) and originating from the corresponding planning region (the one where this city is) for the same period of time. As an example, if the number of passengers using Heathrow and originating from the West Midlands planning region has increased by 10% between 1975 and 1978, the assumption is that the number of passengers originating from Coventry (or Birmingham) and using Heathrow has also increased by 10% during that period of time. The number of passengers using Heathrow (Luton) in 1975 and in 1978 are available for each planning region. Thus, the percentage change has been computed for each planning region and is taking values between 9% and 14%. Therefore, the 1978 figures have been adjusted during the data compilation by omitting the tenth passenger in each group of ten passengers.

The final samples are given in Table 4.1. These samples are large enough to be representative of the population under study and there is no need to select more passengers from other origins in the study area or to other destinations.

SAMPLE / AIRPORT	INTERNATIONAL BUSINESS	INTERNATIONAL LEISURE	INCLUSIVE TOURS	DOMESTIC	TOTAL
Manchester Airport	59	65	322	68	514
Birmingham Airport	54	197	453	240	944
East Midlands Airport	33	46	470	362	911
Heathrow Airport*	68	23	104	13	208
TOTAL	214	331	1349	683	2577

* Luton airport for inclusive tours

Table 4.1 Study samples (in passengers)

These samples broken down by origin/destination are displayed in Tables 4.2, 4.3, 4.4 and 4.5 for the business, leisure, inclusive tours and domestic passengers respectively. They show that the business, inclusive tours and domestic passengers are evenly distributed by origin and by destination while for leisure there is a predominance of Birmingham as origin and Dublin as destination. Tables 4.6, 4.7, 4.8 and 4.9 present the use of airport broken down by origin for the business, leisure, inclusive tours and domestic passengers respectively. They show that Manchester airport is used mainly, as far as the study area is concerned, by passengers whose origin is north of a line Stoke - Derby. Birmingham is used by passengers originating from locations south of Nottingham. East Midlands airport is mainly used by passengers from the East Midlands and Yorkshire/Humberside regions. Heathrow attracts passengers from all the study area but in a much smaller degree for non-business passengers than for business passengers which suggests that the non-business passengers are less prepared to travel an extra distance to benefit from the better frequency of flights offered at Heathrow. The tendency to use the nearest airport is more obvious for the domestic passengers than for the international passengers.

4.4.2 Flight frequency

This variable will be measured by the number of direct flights per day to a particular destination. This figure was

Destination / Origin	Dublin	Amsterdam	Frankfurt	Brussels	Total
Blackpool	0	0	5	0	5
Manchester	8	10	4	0	22
Birmingham	14	12	13	3	42
Coventry	6	5	9	1	21
Leicester	10	7	9	3	29
Nottingham	10	7	14	1	32
Derby	3	3	9	1	16
Stoke	5	1	0	3	9
Sheffield	5	10	8	2	15
Leeds	2	3	6	2	13
Total	63	58	77	16	214

Table 4.2 Origin/Destination - business passengers

Destination / Origin	Dublin	Amsterdam	Frankfurt	Brussels	Total
Blackpool	0	0	1	1	2
Manchester	44	5	2	2	53
Birmingham	138	10	4	0	152
Coventry	26	0	3	3	32
Leicester	19	4	4	3	30
Nottingham	20	5	0	1	26
Derby	14	2	0	0	16
Stoke	2	2	2	1	7
Sheffield	4	1	1	1	7
Leeds	4	1	1	0	6
Total	271	30	18	12	314

Table 4.3 Origin/destination - leisure passengers

Destination / Origin	Palma	Alicante	Ibiza	Total
Blackpool	33	23	3	59
Manchester	62	44	16	119
Birmingham	168	88	60	316
Coventry	47	27	14	88
Leicester	55	15	21	91
Nottingham	85	49	41	175
Derby	44	22	17	83
Stoke	39	17	21	77
Sheffield	95	56	25	176
Leeds	103	46	16	165
Total	734	384	231	1349

Table 4.4 Origin/destination - inclusive tours passengers

Destination / Origin	Belfast	Jersey	Glasgow	Aberdeen	Total
Blackpool	2	1	0	2	5
Manchester	25	10	3	0	38
Birmingham	44	102	12	5	163
Coventry	18	25	6	4	53
Leicester	29	25	15	5	74
Nottingham	20	82	15	3	120
Derby	20	42	31	6	99
Stoke	10	32	3	3	48
Sheffield	5	54	6	0	65
Leeds	4	14	0	0	18
Total	177	387	91	28	683

Table 4.5 Origin/Destination - Domestic passengers

Origin \ Destination	Manch-ester	Birm-ingham	East Midlands	Heathrow	Total
Blackpool	4	0	0	1	5
Manchester	22	0	0	0	22
Birmingham	0	36	0	6	42
Coventry	0	9	0	12	21
Leicester	0	9	6	14	29
Nottingham	0	0	11	21	32
Derby	0	0	13	3	16
Stoke	8	0	0	1	9
Sheffield	14	0	3	8	25
Leeds	11	0	0	2	13
Total	59	54	33	68	214

Table 4.6 Use of airport by origin - business passengers

Destination / Origin	Manch-ester	Birm-ingham	East Midlands	Heathrow	Total
Blackpool	1	0	0	1	2
Manchester	51	0	0	2	53
Birmingham	0	149	0	3	152
Coventry	0	28	0	4	32
Leicester	0	19	5	6	30
Nottingham	0	0	23	3	26
Derby	0	0	16	0	16
Stoke	4	1	0	2	7
Sheffield	5	0	2	0	7
Leeds	4	0	0	2	6
Total	65	197	46	23	331

Table 4.7 Use of airport by origin - leisure passengers

Destination / Origin	Manchester	Birmingham	East Midlands	Luton	Total
Blackpool	59	0	0	0	59
Manchester	119	0	0	0	119
Birmingham	0	306	0	10	316
Coventry	0	65	3	20	88
Leicester	0	29	40	22	91
Nottingham	0	3	162	10	175
Derby	0	0	83	0	83
Stoke	27	50	0	0	77
Sheffield	30	0	116	30	176
Leeds	87	0	66	12	165
Total	322	453	470	104	1349

Table 4.8 Use of airport by origin - inclusive tours passengers

Destination / Origin	Manch-ester	Birm-ingham	East Midlands	Heathrow	Total
Blackpool	5	0	0	0	5
Manchester	38	0	0	0	38
Birmingham	0	161	0	2	163
Coventry	0	51	0	2	53
Leicester	0	6	65	3	74
Nottingham	0	2	116	2	120
Derby	0	0	98	1	99
Stoke	13	20	15	0	48
Sheffield	6	0	58	1	65
Leeds	6	0	10	2	18
Total	68	240	362	13	683

Table 4.9 Use of airport by origin - domestic passengers

drawn from the 1975 ABC World Airways Guide for the domestic and international flights. For the inclusive tours, this information was collected from the magnetic tapes by reading for each day the airline numbers of the flights leaving from each airport to each selected destination.

Tables 4.10, 4.11 and 4.12 give the number of daily direct flights from the airports considered to the destinations considered for respectively the international, the domestic and inclusive tours passengers. They reveal that, London excepted, the number of flights offered daily is between zero and three (only Manchester airport exceeds this, offering on two weekdays four flights to Dublin) with a poor level of service at Birmingham and East Midlands airports for the scheduled flights during the weekend. Therefore, the frequency variable which is more precisely specified in terms of departure time, might be, for that particular case, expressed in the more simple form of number of flights per day.

If only one airport serves a particular destination, the passenger has no choice but to fly from that airport. Therefore, the selection of destinations should reflect the passenger's possibility of choice of airport. The air network serving Birmingham, Manchester and East Midlands airports in 1975 was essentially made up of domestic and European routes (no long haul traffic). The routes selected are all domestic or European and therefore are a representative subset of the total routes offered from these airports. The selected destinations are also those where the competition between airports is the highest. Indeed, if the criterion for selecting a particular destination is that it should be served by at least four flights per week from at least two of the airports considered, only Copenhagen, Dusseldorf and Zurich for the international flights could have been added. Their routes, however, carry much less traffic than the selected ones. For the domestic flights, only Edinburgh could have been added to the selected destinations. Further, only the selected destinations answer the criterion of being served by at least four flights per week from at least three airports. However, development in competition in air transport since the late seventies has resulted in a larger spread of destinations served and a more widespread difference between airlines' fares and frequencies of flights, and this will probably have implications for any model to be constructed from more recent data.

4.4.3 Air fare

Unfortunately data are not available on the range of promotional and discount fares in 1975. Thus, the economy (Y) fare has been retained as a measure of the cost variable. This fare has been drawn from the 1975 ABC World Airways Guide for the domestic and international flights. For the inclusive

to / from	Dublin				Amsterdam				Frankfurt				Brussels			
	MAN	BMG	EMA	LHR	MAN	BMG	EMA	LHR	MAN	BMG	EMA	LHR	MAN	BMG	EMA	LHR
Mon.	3	2	1	10	3	1	1	15	1	1	1	10	2	1	1	11
Tues.	3	2	1	9	3	1	1	16	1	1	1	11	2	1	1	10
Wed.	3	2	1	9	3	1	1	14	1	1	1	13	2	1	1	11
Thrs.	3	2	1	10	3	1	1	15	1	1	1	12	2	1	1	10
Fri.	4	2	1	12	3	1	1	14	1	1	1	12	2	1	1	10
Sat.	3	2	0	8	2	0	0	9	1	0	0	11	0	0	0	8
Sun.	4	2	0	9	2	0	0	13	1	0	0	11	1	0	0	8
Air fare (in £'s)	19.3	20.8	20.8	23.8	35.1	33.8	33.8	29.4	54.4	47.7	47.7	41.6	39.2	36.0	36.0	32.0

Table 4.10 Number of scheduled daily direct flights to selected international destinations in 1975 with respective air fare.

to / from	Belfast				Jersey				Glasgow				Aberdeen			
	MAN	BMG	EMA	LHR	MAN	BMG	EMA	LHR	MAN	BMG	EMA	LHR	MAN	BMG	EMA	LHR
Mon.	3	2	2	5	1	1	1	5	2	2	2	12	1	2	0	5
Tues.	3	2	2	5	1	1	1	5	2	2	2	12	1	2	0	5
Wed.	3	2	2	5	1	1	1	5	2	2	2	12	1	2	0	5
Thrs.	3	2	2	5	1	1	1	5	2	2	2	12	1	2	0	5
Fri.	3	2	2	5	1	1	1	5	2	2	2	12	1	2	0	5
Sat.	1	2	0	5	1	1	1	5	1	0	0	7	0	0	0	4
Sun.	1	2	0	4	1	1	1	5	1	0	0	6	1	0	0	3
Air fare (in £'s)	13.5	16.8	19.0	19.0	19.7	18.2	16.3	13.8	15.7	19.6	18.8	19.0	23.8	26.0	-	25.2

Table 4.11 Number of scheduled daily direct flights to selected domestic destinations in 1975 with respective air fare.

to / from	Palma				Alicante				Ibiza			
	MAN	BMG	EMA	LUT	MAN	BMG	EMA	LUT	MAN	BMG	EMA	LUT
Mon.	1	2	2	4	2	1	0	0	2	0	2	1
Tues.	2	1	0	5	0	2	2	5	0	0	0	1
Wed.	1	2	2	5	0	2	1	3	0	1	0	2
Thrs.	1	3	2	4	1	2	1	0	1	2	0	0
Fri.	3	1	2	6	2	2	0	5	1	1	0	1
Sat.	3	3	2	10	2	2	1	6	1	3	2	2
Sun.	3	2	2	3	2	2	2	1	1	2	1	6
Air fare (in £'s)	29.7	28.1	28.1	25.7	33.4	31.8	31.8	29.4	30.6	29.0	29.0	26.7

<u>Table 4.12</u> Number of daily charter flights to selected international destinations in 1975 with respective air fare.

tours, the fare covers not only the air fare but also the holiday. Thus the air fare has been assumed to be 40% of the global price during July/August 1975 after private discussions with local travel agencies who have also supplied the package holiday prices charged in July/August 1975 from the airports considered to the selected destinations.

Tables 4.10 and 4.11 give the economy (Y) fare from the airports considered to the selected destinations for the international and domestic flights respectively. Table 4.12 gives the assumed air fare for the inclusive tours flights. Tables 4.13 and 4.14 show the differences between the cheapest and the dearest air fare offered at the airports considered for each destination, along with the ratio of these differences to the dearest fare. In only one international case (Frankfurt) is this ratio greater than 20% while in the domestic case, only one destination (Aberdeen) has a ratio lower than 20%. This suggests that domestic passengers might be more sensitive to fare differences than international passengers.

Destination	Dearest fare	Cheapest fare	Δ	Δ/dearest fare
Belfast	19.0	13.5	5.5	29%
Jersey	19.7	13.8	5.9	30%
Glasgow	19.6	15.7	3.9	20%
Aberdeen	26.0	23.8	2.2	9%

Table 4.13 Fare ratios for domestic flights (1975)

Destination	Dearest fare	Cheapest fare	Δ	Δ/dearest fare
Dublin	23.8	19.3	4.5	19%
Amsterdam	35.1	29.4	5.7	16%
Frankfurt	54.4	41.6	12.8	23%
Brussels	39.2	32.0	7.2	18%

Table 4.14 Fare ratios for international flights (1975)

4.4.4 Access time

Three distinct values of access time can be identified, namely perceived, reported and measured values. The perceived value is the value actually perceived by the passenger. The reported value is the value the passenger reports when questioned about his access trip, while the measured value is obtained by some computation. In a behavioural sense, it is clear that perceived values are those upon which individual decisions are made and, thus, are the most appropriate for modelling choices. They are, however, probably impossible to determine.

The reported value may represent an individual's best attempt to translate his perceptions into standard measurement. During the survey, the passengers have been asked how long did their journey from the starting point to the airport take. Thus, the reported access time from a passenger's origin to the airport used is available from the magnetic tape. However, there is no information for the non-chosen airports. On this basis, a measured value of access time can be justified.

The distance between each of the 10 cities and each of the five airports has been computed from the big road atlas of Great Britain. The car average speed has been considered equal to 60 mph for a motorway and 35 mph for a primary road (rural roads plus urban roads). The motorway network and AA recommended routes were used. Having these distances and the car speed, the car travel time from each city to each of the five airports has been computed. Since no significant differences have been found between the computed car travel time and the average of the reported travel times read from the tapes, the computed car travel time has been retained for use in the model. Even if it is a simple measure, it takes into account the quality of the road network and the fact that other modes of ground access play a negligible role; the car being the most predominant mode (93%, 84% and 88% of terminating passengers at respectively Birmingham airport, East Midlands airport and Manchester airport in 1975 used the car to access the airport). There is no loss in accuracy and this presents the advantage of simplicity in future use. Furthermore, it eliminates the possible discrepancy between the true access time and the access time reported by the passenger. Table 4.15 displays the computed travel time between each of the 10 cities and each of the five airports.

From \ To	Manch- ester airport	Birming- ham airport	East Midlands airport	Heath- row airport	Luton airport
Blackpool	68	140	170	276	220
Manchester	12	90	102	221	165
Birmingham	80	12	60	152	96
Coventry	104	19	45	135	79
Leicester	130	39	30	149	93
Nottingham	118	68	24	173	117
Derby	87	68	12	179	123
Stoke	44	50	90	191	135
Sheffield	72	104	60	213	157
Leeds	70	140	80	235	179

Table 4.15 Travel times to the airports (in minutes).

4.5 Sources of errors in model estimation

A good model cannot perform satisfactorily with poor data and it is useful to discuss the major sources of error in model estimation which are specification error, measurement error and sampling error (Hensher and Johnson, 1981). Specification error results from the simplifying process associated with the construction of models to represent choice. It is one source of the violation of the independence from irrelevant alternative property (I.I.A.) in the multinomial logit model. In Chapter 3, it was stated that the aim of this work is to build a pragmatic model which could be easily applied in practice by airport and airline managements and there could be a risk that the model could be mispecified due to the omission of one or more relevant variables. However, in Chapter 6, the validity of the IIA property for the problem under study is tested and this, along with the good performance of the model, suggests the absence of specification error. A second source of error is measurement error which is primarily associated with errors made in the measurement of variables. Most of the data in this study was provided from an external source and of course there is a possibility of errors in measurement, particularly during the survey of some passenger answering questions in a way to give "a more flattering image of himself" (e.g. indicating an income group above the correct one). Another possible source of measurement error is the consequence of the necessity to adjust the 1978 data to achieve consistency with the 1975 data. Further, even if great care is taken in computing the access time variable, a measurement error could be associated with the assumptions used. It can be reasonably assumed that there is no measurement error in the frequency of flights and air fare variables except for the inclusive tours passengers where an assumption about the air fare level has to be made. Thirdly, there are many sources of error in sampling most commonly a deliberate selection of a "representative" sample or an "average" sample, a selection of a random sample in which the random selection process is not strictly adhered to and the substitution of additional members of population when difficulties are encountered in sampling the original sample observation. At each airport in the 1975 survey, systematic random samples were taken. Each passenger who crossed a predetermined counting line in the right direction was counted and the passenger corresponding to the sampling interval (15 passengers) was interviewed. This sampling technique was used so as to minimise sample bias.

4.6 Conclusion

In this chapter, it was described how the data relevant to this study were collected and prepared and how the explanatory variables will be measured. Descriptive tabulations to reveal features of the data were also provided. The possible errors in model construction were also discussed. The samples considered are large enough to be a representative subset of the population.

CHAPTER 5

METHODOLOGY

This chapter attempts to specify the features and general form desirable for constructing an airport choice model. In the first section of this chapter, the potential advantages of disaggregate models are justified. Subsequently, a choice theory based upon the hypothesis of random utility maximization is presented. The alternative models which may be used are examined. The structure of passenger's decision-making is defined. Finally statistical measures of model validity which can assist in the assessment of empirically estimatable individual choice models are described.

5.1 Disaggregate versus aggregate models

A model is called aggregate when the demand function has been specified in terms of the average of the relevant attributes. When a model is based on samples of individual observations, it is called a disaggregate model. The disaggregate approach attempts to use variables resulting from a hypothesis regarding the behavioural choice process of the decision-maker. The dependant variable in a disaggregate model is an individual observation, whilst the explanatory variables are represented by data collected at the level of the decision maker.

Although the concern is with aggregates of people, their behaviour can probably best be understood by considering the behaviour of individual travellers. Furthermore, the disaggregate models offer several advantages over aggregate models, focusing on the travel decision making unit. Thus, all the problems created by the aggregation of data, such as variability in characteristics aggregated in the same observation, the substantial decrease of the number of observations available for analysis, are avoided.

Once a model has been calibrated on individual observations, the computation of aggregate demand can be accomplished by direct aggregation over values of the explanatory variables. The correct approach is to use the values of the independent variables relevant to each individual to predict the individual probabilities of choosing a particular alternative, and then to sum these probabilities. The predictions obtained by the individual disaggregate approach (often referred to as the enumeration method) will be more accurate than those obtained by the use of the average sample values of the independent variables (often referred to as the naive aggregation method) because the average of a nonlinear function is not equal to the function evaluated at the average of the explanatory variables.

5.2 Choice Theory:

5.2.1 Random utility maximization:

Many important decisions which an individual must take in his life involve selection from a limited and constrained set of discrete alternatives, e.g. choice of a house, occupation, the mode of travel on a work trip and so on. In these circumstances, conventional "marginalist" micro-economic consumer theory with its assumption that individual demand is the result of utility maximization by a representative consumer whose decision variable is continuous (i.e. who is selecting fractional quantities from a continuously divisible choice continuum) is patently unrealistic.

Over the past ten years, in the work of McFadden, Hensher, Manski, Ben-Akiva, Williams, Daly, Daganzo, Lerman, Gaudry, Horowitz, and many others, significant progress has been made in modifying conventional consumer theory and it has been shown that a logically consistent discrete choice theory can be developed based upon the hypothesis of random utility maximization.

There are two available interpretations of random utility maximization. These have been termed the interpersonal and intrapersonal interpretations respectively.

The first is characteristic of most of the discrete choice modelling literature in economics, transportation science and geography. In this interpretation, the distribution of demand in the population is conceived to be the result of individual preference maximization but preferences are viewed as being influenced, in part, by variables which are unobserved by the modeller. Because certain choice relevant attributes are unobserved and because the valuation of observed attributes may vary from individual to individual, a random element enters an individual's utility function and utility functions are assumed to vary over the population of decision-makers.

The second interpretation is that found in the literature of psychology. This assumes that each individual draws a utility function from a random distribution each time a decision must be made. That is to say, the individual is a classical utility maximizer given his state of mind but his state of mind varies randomly from one choice situation to the next.

The two interpretations are formally indistinguishable in their effects on the observed distribution of demand, and the random element in each implies that discrete choice problems must be handled using probabilistic choice models.

Adopting the first interpretation* of random utility maximization, the conventional derivation of a probabilistic discrete choice model proceeds as follows:

a) First it is assumed that each decision maker k is faced with a set A of G available choice alternatives

$$A = (A_1, A_2 \quad \ldots . A_G)$$

*A use of the intrapersonal interpretation of random utility maximization can be found in Luce and Suppes (1965)

b) Second, it is assumed that a utility value U_{gk} is associated with the choice of alternative g by each individual k and that each individual k wishes to select an alternative which yields maximum utility.

c) Third, the utility is assumed as $U_{gk} = U(Z_g, S_k)$ where Z_g is a vector of attributes of the choice alternative g and S_k is a vector of socioeconomic characteristics of each individual k.

d) Finally, it is assumed that the modeller knows the structure of the U function up to a finite parameter vector, has observed specific values of the subset of the many possible attributes Z_g and socioeconomic characteristics S_k and knows up to a finite parameter vector the distribution of unobserved characteristics across the population.

In most practical applications, the random utility is assumed to have a linear in parameter additive form. One reason is because it can easily be estimated while other more general forms present enormous difficulties (Sheffi, Hall and Daganzo, 1980). Note that the characteristics may enter in straight linear form as logarithms or as various powers. The random utility can be written as:

$$U_{gk} = a.X_{gk} + \varepsilon_{gk} \qquad (5.1)$$

The first term of the right hand side of (5.1) is referred as the representative component of utility and, subsequently it will often prove convenient to denote it as V_{gk}. The second term is the random component. It represents the deviation of individual utility from the representative or "group average" utility component, and it accounts for the unobserved attribute of the choice alternative g, the unobserved socioeconomic characteristics of individual k and also the idosyncratic tastes of individual k (i.e. the deviation of the tastes of k from the group average).

In selecting the alternative that yields the highest utility, the individual k will select alternative g among a set of alternatives G if

$$U_{gk} > U_{rk} \qquad r = 1,\ldots.G$$

because only a part of the utility, the V_{gk}, can be measured, the event $U_{gk} > U_{rk} \; \forall \; r = 1, \ldots G$ will occur with some probability. That is a probability must be assigned to any individual decision, as shown in equation (5.2)

$$P_{gk} = \text{Prob } (U_{gk} > U_{rk} \quad r = 1,\ldots.G) \qquad (5.2)$$

where P_{gk} denotes the probability that individual k faced with a choice set of G alternatives will select alternative g.

Prob = denotes "probability that"

and using equation (5.1), this may be written as

$$P_{gk} = \text{Prob } (\varepsilon_{rk} - \varepsilon_{gk} < V_{gk} - V_{rk} \quad r = 1,\ldots.G)$$

If we assume that the random elements in utility (i.e. ε_k's) are independent across alternatives and that they are identically distributed, then a suitable distribution is the double exponential or Weibull distribution. The Weibull distribution in terms of ε_k's is defined as:

$$\text{Prob} \quad (\varepsilon_k \leqslant \varepsilon) = (\exp - \exp(-\varepsilon)) = e^{-e^{-\varepsilon}} \tag{5.3}$$

Under these assumptions, the choice probabilities have the form:

$$P_{gk} = \frac{e^{a.X_{gk}}}{\sum_{r=1}^{G} e^{a.X_{rk}}} = \frac{e^{V_{gk}}}{\sum_{r=1}^{G} e^{V_{rk}}} \tag{5.4}$$

Hensher and Johnson (1981) give the details of the derivation of eq.(5.4) from eq.(5.2) and eq.(5.3).

The resulting model is the well-known multinomial logit model (M.N.L.). Perhaps the most general of the alternative discrete choice models is that produced when it is assumed that the vector of random components has a multivariate normal distribution with zero mean and an arbitrary variance-covariance matrix. It relaxes the assumption of ε being independent and identically distributed. This model is known as the multinomial probit model (M.N.P.).

These two alternative models will be examined in more detail in the next sections. The econometric derivation of the random utility model is well documented (for example McFadden (1974), Hensher and Johnson (1981)).

5.2.2 <u>Multinomial logit model:</u>

As seen in the previous section, under the assumption that the random elements in utility are independent across alternatives and that they are identically distributed, the resulting model is the multinomial logit model

$$P_{gk} = \frac{e^{V_{gk}}}{\sum_{r=1}^{G} e^{V_{rk}}} \qquad (5.5)$$

where

P_{gk} = probability that alternative g will be chosen by individual k

$V_{rk} = a_0 + a_1 . X_1 + a_2 . X_2 + ... + a_n . X_n$ is

the representative component of the utility assumed to be a linear combination of the explanatory variables. An important application of the above formulation is that the ratio P_{ik}/P_{gk} of choosing alternative i over alternative g is independent of the presence or absence of third alternatives, satisfying the equation

$$\text{Ln } (P_{ik}/P_{gk}) = V_{ik} - V_{gk} \qquad (5.6)$$

This property is termed "Independence of/from irrelevant alternatives" (I.I.A.). The I.I.A. condition is both the principal strength and the principal weakness of this model.

It is a strength because this restriction allows the introduction of a new alternative without re-estimation of the model, once a numerical functional form of V has been established. This is done by simply adding the new term to the denominator of eq.(5.5) for each alternative and recalculating the probability of choosing each alternative (including the probability of choosing the new alternative). This procedure is possible because the addition of an alternative cannot change the relative ratios with which the previous alternatives are selected.

It is a weakness because it requires that the alternatives be perceived as completely distinct and independent. In accordance with the I.I.A. property, the share obtained by a new alternative from each of the other alternatives is directly proportional to their original share; clearly this property is valid if the new alternative competes equally with each existing alternative. An example of a choice setting in which the I.I.A. property is inappropriate is the classic blue bus/red bus (Hensher and Johnson,1981) where two transport services differ only in colour and therefore are not distinct alternatives since they are unlikely to have distinct ε's. One of the modes is clearly irrelevant and should not be introduced as a mode in the multinomial logit model.

There are some statistical approaches to the estimation of the parameters a_0, a_1 ... a_n of the multinomial logit model but in many cases, the procedure known as maximum likelihood estimation (M.L.E.) is preferable to any of the other available methods (e.g. generalised least squares). Suppose we have a random sample of R individuals and for each we observe the choice actually made. If we order our observations so that the first n_1 individuals are observed to have chosen alternative 1, the next n_2 to have chosen alternative 2, etc., the likelihood function of our sample may be written as:

$$L = \prod_{k=1}^{n_1} P_{1k} \cdot \prod_{k=n_1+1}^{n_1+n_2} P_{2k} \cdot \ \cdots \ \prod_{k=R-n_g+1}^{R} P_{gk}$$

where Π is an operator implying multiplication in the same way as Σ refers to addition and G denotes the number of alternatives.

It is possible to simplify this expression for L slightly by defining a dummy variable f_{gk} such as $f_{gk}=1$ if alternative g is chosen and $f_{gk}=0$ otherwise. The likelihood function can now be written as:

$$L = \prod_{g=1}^{G} \prod_{k=1}^{R} (P_{gk})^{f_{gk}}$$

and the log likelihood function L^* may be written

$$L^* = \sum_{g=1}^{G} \sum_{k=1}^{R} f_{gk} \ln P_{gk} \qquad (5.7)$$

If we replace P_{gk} in eq.(5.7) by the expression (5.5) the result is an equation which is a function of the unknown parameters contained in the expression V_{gk} since all other quantities are known. L^* is then maximised with respect to the a's in the usual manner, the resulting estimates being the MLEs for the model's parameters (Hensher and Johnson, (1981) describe the entire process of deriving the maximum likelihood estimates of the parameters).

These maximum likelihood estimates are reached when the first partial derivatives of eq.(5.7) $\delta L^* / \delta a_i$ for i = 1, ...n approach zero, and the second partial derivatives $\delta^2 L^* / \delta^2 a_i$ are negative.

5.2.3 Multinomial probit model:

Probit analysis can also be used to estimate the relationship between a dichotomous dependent variable and a set of explanatory variables. As seen in section 5.2.1 the multinomial probit model relaxes the assumption of independent and identically distributed (I.I.D.) random terms in the utility function by substituting the multivariate normal distribution for the Weibull as the distribution of the ε. Unfortunately the choice function of an MNP model cannot be easily written in closed form, except for the case of two alternatives and thus must be evaluated numerically. The difficulty lies in the

N choice situation where the MNP model necessitates the evaluation of a multiple integral of dimension N-1 which cannot be reduced to an analytical expression.

As with the MNL, the maximum likelihood method is actually the most efficient way of calibration. Daganzo (1979) discuss in detail the process and the problems associated with finding the maximum likelihood estimates of the MNP from disaggregate data.

5.2.4 Other choice models:

There has been a search for what Wrigley (1982) termed 'half way house' models which lie between the generality and complexity of the MNP model and the restrictiveness but tractability of the MNL model.

One of these models is the nested (or structured or hierarchical) logit model. In contrast to the MNL, the nested logit model permits correlation between the random components of the choice alternatives, is not constrained by the I.I.A. properties of cross-substitution and thus overcomes the similarity of choice alternative problems. As a result, applications of nested logit models are now multiplying rapidly (Sobel 1980; McFadden 1981; Hensher and Manefield 1981). Finally it should be noted that an extension of the nested logit known as the cross-correlated logit model has recently been proposed by Williams (1981) and Williams and Ortuzar (1982).

Another 'half way house' which has attracted a considerable amount of attention is the Dogit model proposed by Gaudry and Dagenais (1979). A potentially useful aspect of the dogit model is that it allows some pairs of alternatives to exhibit the independence from irrelevant alternative property while other pairs may not exhibit the I.I.A. as basic logit does and hence the name dogit. However, empirical applications of the dogit model are, as yet, few in number (Gaudry and Willis 1979; Gaudry 1980) and McFadden (1981) has questioned whether the dogit model has sufficient flexibility to accommodate reasonable patterns of cross-substitution. A review of 'half way house' models is given by Wrigley (1982).

5.2.5 Discriminant analysis

One of the earliest non-linear techniques to be considered for building choice models is discriminant analysis. The technique is based upon the assumption that there exists in a population two or more distinct subgroups that can be distinguished by means of a discriminating function. The method is not described in this chapter but descriptions may be found in several references (Benabi, 1983; de Donnea,1974). Empirical tests of discriminant analysis (Watson,1974; Stopher and Lavender,1972) seem to confirm that the technique may be incorrect. In comparison with estimates from probit and logit analysis, relative coefficient values are markedly different and uniform goodness-of-fit statistics have been found to be significantly inferior for discriminant analysis than for the other two techniques. Therefore, it does not appear to be appropriate to use discriminant analysis as a technique for building behavioural choice models.

5.2.6 Specification of the attributes in the utility expression

There are two main types of explanatory variables, generic variables and alternative-specific variables. The distinction relates to the way a variable enters the data, not the way it is estimated. Generic variables vary in level across choice alternatives whereas alternative-specific variables have an identifiable correspondence between choice alternatives. In a choice between the three following alternatives: fully detached house, town house and flat, an alternative specific variable is, for example, one which gives the appropriate level of establishment size for the "flat" alternative and which takes a zero value for all other alternative establishment types. A generic variable, for example, could be one which gives the relevant level of establishment size for all the alternatives in the set. In this study, the variables are generic such that they relate to attributes common to all alternatives. This means that each variable will have the same parameter in each of the utility functions associated with different airports.

5.3 Decision making process:

To understand the nature of the individual choices that will be modelled, we must first identify the range of choices open to the passengers. Secondly, we shall have to make an assumption on the structure of their decision-making process.

5.3.1 Binary versus multinomial models:

Most of the published disaggregate behavioural models explain binary choices. As de Donnea (1971) pointed out this may be partly explained by the fact that computer programs for calibrating multinomial models were developed later than programs for binary choice. But it may also be partly explained by the fact that some of the first disaggregate models were developed for modal choice in U.S. cities where the only choice open to people is between car and some form of public transport, especially when walking is excluded because of the distance travelled.

Most, if not all, passengers in our problem have a genuine choice of at least two airports. This implies that we must either build multinomial models or try to decompose the choice into a sequence of binary choices (de Donnea,1971; Watson,1974). Is a multinomial model covering all the possible airports a realistic representation of the passenger's choice process? The use of a multinomial model is based on the implicit behavioural assumption that people simultaneously consider all the alternatives open to them and weigh their relative merits in a total decision. If this assumption may be questioned in modal choice when the alternatives (e.g. car, walking, bus, bike, etc...) open to travellers show differences on a wide range of characteristics, it is clear that in our case the alternatives (airports) will offer quite similar characteristics in many respects and that they will only differ significantly on a few characteristics.

On the basis of these considerations and that computer programs required for calibrating multinomial models are now available, we have decided to use multinomial models to explain the choice of airport.

5.3.2 Structure of passenger's decision-making: Hensher's p.97

Two alternative structures can be defined, simultaneous and recursive. They are based on different hypothesis about the underlying decision-making process. The recursive structure represents a specific conditional decision structure (e.g. a passenger might decide first an airport and then an access mode or first an access mode, then an airport). Generally, there are no a priori reasons to justify a selection among sequences assumption. The simultaneous structure is very general and does not require any specific assumption. However, it is very complex because of the large number of alternatives that a passenger faces in making his decision (e.g. 9 options in the case of a choice between Heathrow airport and Gatwick airport in the C.A.A. study, 1984). Where

Ben-Akiva (1974) recommended the simultaneous model structure, Liou and Talvitie (1974) in 'disaggregate access mode and station choice models for rail trips' indicate that the traveller's decision-making process is behaviourally separate, the sequence being station choice followed by access mode choice. It was acknowledged in Chapter 3, that different categories of passengers might have different choice structures. The data available however, do not permit the investigation of the particularities of each category of passenger hierarchy of decision making and it has been assumed that the passenger decision making structure is a choice of airport for all the passengers irrespective of their trip purpose.

5.4 Model evaluation:

There are a number of statistical measures of model validity which can assist in the assessment of empirically estimatable individual choice models. The object of this section is to describe these measures.

Hensher's p.48

5.4.1 Significance of coefficients:

Using the maximum likelihood procedure, we can calculate asymptotic standard errors for the coefficient estimate(a's) in the model and use these to test the significance of the a's using an asymptotic t-test in much the same way as a regression coefficient is tested using a t-test.

$$t = \frac{a_i}{\Delta a_i}$$

where Δa_i is the standard deviation of the parameter estimate a_i.

Δa_i is estimated as the square root of the inverse of the second partial derivative of the log likelihood with respect to the parameter a_i, i.e. $\Delta a_i = \left| \delta^2 L^* / \delta^2 a_i \right|^{-\frac{1}{2}}$.

5.4.2 Goodness of fit measures:

Hensher's p. 49 McFadden

A goodness of fit measure is a summary statistic indicating the accuracy with which a model approximates the observed data. When using maximum likelihood techniques, two statistical tests are now widely used : the likelihood ratio test and the likelihood ratio index. A number of studies has also proposed test of prediction success which involve a comparison of the summed probabilities from the model (e.g. Hensher and Bullock, 1979; McFadden, 1979).

5.4.2.1 Likelihood ratio test:

This test proposes a null hypothesis on the basis of equal-shares or market shares that the probability Pi of an individual choosing alternative i is independent of the explanatory variables in the estimated model. For an equal-share hypothesis, it is assumed that the true values of the model parameter are all zero. That is, in a four alternatives set, the same probability (i.e. 0.25) is assumed for the choice of either alternative for any observation. The likelihood ratio test can be determined as:

$$LRT = -2\ (\overset{*}{L}_o - \overset{*}{L}) \qquad (5.8)$$

where $\overset{*}{L} = \overset{*}{L}\ (\hat{a}_o, \hat{a}_1 \ldots\ldots \hat{a}_n)$ is the loglikelihood for the fitted model

$\overset{*}{L}_o = \overset{*}{L}\ (0,0,\ldots.0)$ is the loglikelihood for the null hypothesis.

Wils (1962) shows that the likelihood ratio is approximately distributed like chi-square with the number of degrees of freedom equal to the number of the model parameters. A value of $-2\ (\overset{*}{L}_o - \overset{*}{L})$ greater than the corresponding tabulated χ^2 at a given confidence level indicates that the hypothesis of independence between the model probability and the explanatory variables can be rejected at the given confidence level.

Another test considers the null hypothesis of market-shares. In this case, it is assumed that the true values of the model parameter are all zero except for a constant. The value of the constant is taken to replicate the market-shares. The log-likelihood for this case is:

$$\overset{*}{L}_o = L\ (a_o,\ 0,\ldots.0)$$

The likelihood ratio test can again be determined as

$$LRT = -2\ (\overset{*}{L}_o - \overset{*}{L})$$

5.4.2.2 Likelihood ratio index:

The likelihood ratio index, formulated as:

$$\rho^2 = 1 - \overset{*}{L}/\overset{*}{L}_o \qquad (5.9)$$

can be used, much as R^2 is in regression, to measure the goodness of fit of the model. As $\overset{*}{L}$ is larger than $\overset{*}{L}_o$ (smaller in absolute value but they are both negative numbers), $\overset{*}{L}/\overset{*}{L}_o$ will be between 0 and 1 and subsequently ρ^2 takes a value between 0 and 1. It should be noted, however, that values of ρ^2 between 0.2 and 0.4 are considered good fits so that the analyst should not be looking for values in excess of 0.9 as is often the case when using R^2 in ordinary regression (McFadden,1979). Domenchich and McFadden (1975) stated that in terms of consistency and statistical properties, the ρ^2 index appears to provide a practical and theoretically sound index of goodness of fit.

5.4.2.3 Prediction success:

A third method of assessing the fit of an estimated model is to examine the proportion of successful predictions, by alternative and overall. McFadden (1979) has defined a prediction success table (table 5.1) with the entry N_{ij} in row i and column j giving the number of individuals who are observed to choose i and predicted to choose j. The formula for N_{ij} is

$$N_{ij} = \sum_{n=1}^{N} F_{in} P_{jn}$$

where $F_{in} = 1$ if i is chosen and $F_{in} = 0$ otherwise.

Column sums give predicted shares for the sample; row sums give observed shares. The proportion of alternatives successfully predicted, $N_{ii}/N_{.i}$ indicates that fraction of individuals expected to choose an alternative who do in fact choose that alternative. An overall proportion successfully predicted $(N_{11}+\ldots+N_{jj})/N_{..}$ can also be calculated. (A dot subscript indicates summation over the corresponding index, e.g. $N_{i.} = \sum_j N_{ij}$).

Because the proportion successfully predicted for an alternative varies with the aggregate share of that alternative, an appropriate goodness of fit measure is the prediction success index which may be written as

$$\sigma_i = \frac{N_{ii}}{N_{.i}} - \frac{N_{.i}}{N_{..}} \tag{5.10}$$

where $N_{ii}/N_{.i}$ is the proportion of individuals expected to choose an alternative who actually choose that alternative. $N_{.i}/N_{..}$ is the proportion which would be successfully predicted if the choice probabilities for each sampled individual were assumed to equal the predicted aggregate share.

This prediction success index will usually be non-negative with a maximum value of $1 - N_{.i}/N_{..}$ An overall prediction success index is

$$\sigma = \sum_{i=1}^{j} \left(\frac{N_{ii}}{N_{..}} - \left(\frac{N_{.i}}{N_{..}}\right)^2 \right)$$

Again this index will usually be non-negative with a maximum value of

$$1 - \sum_{i=1}^{J} \left(\frac{N_{.i}}{N_{..}} \right)^2$$

It can be normalised to have a maximum value of one.

		Predicted Choice 1	2	........J	Observed Count	Observed Share
Observed Choice	1	N_{11}	N_{12}	N_{1j}	$N_{1}.$	$N_{1}./N..$
	2	N_{21}	N_{22}	N_{2j}	$N_{2}.$	$N_{2}./N..$
	.					
	J	N_{J1}	N_{j2}	N_{jj}	$N_{j}.$	$N_{j}./N..$
Predicted count		$N._{1}$	$N._{2}$	$N._{j}$	$N..$	1
Predicted share		$\frac{N._{1}}{N..}$	$\frac{N._{2}}{N..}$	$\frac{N._{j}}{N..}$	1	
Proportion successfully predicted		$\frac{N_{11}}{N._{1}}$	$\frac{N_{22}}{N._{2}}$	$\frac{N_{jj}}{N._{j}}$	$\frac{N_{11} + \ldots + N_{jj}}{N..}$	
Success index		$\frac{N_{11}}{N._{1}} - \frac{N._{1}}{N..}$	$\frac{N_{22}}{N._{2}} - \frac{N._{2}}{N..}$	$\frac{N_{jj}}{N._{j}} - \frac{N._{j}}{N..}$	$\sum_{i=1}^{j} \left(\frac{N_{ii}}{N..} - \left(\frac{N._{i}}{N..}\right)^{2}\right)$	

Table 5.1 A prediction success table

5.5 Conclusion

A number of conclusions can be drawn from this chapter:

- A disaggregate approach is preferable with the traveller being the decision-making unit.

- A logically consistent discrete choice theory can be developed based upon the hypothesis of random utility maximization which implies that discrete choice problems must be handled using probabilistic choice models. The concept of random utility produced the choice model, i.e. logit, probit model. These alternative models have been described.

- The use of a multinomial model (by contrast to a binary model) is justified.

- It has been assumed that the traveller examines only the choice of airport. ? (simultaneous or recursive)

- The goodness of fit measures are available for the assessment of the model and have been presented : t-test for the parameters and likelihood ratio test, likelihood ratio index and prediction success table for the model.

CHAPTER 6

AIRPORT CHOICE MODEL

6.1 Introduction:

In Chapter 5, it was established why the airport choice model should be a disaggregate behavioural multinomial model; either logit or probit. The aim of this chapter is to outline their potential advantages and shortcomings in order to select, the technique which is more appropriate.

6.2 Model selection:

6.2.1 Logit versus probit models:

It would seem that the MNP model is to be preferred to the MNL in that it can allow for taste variation across traveller and as Vanlierop and Nijkamp (1984) pointed out it provides more possibilities of bringing theory closer to reality. On the other hand, it is very difficult to calibrate even if recent developments in algorithms (Lerman and Manski, 1977; Hausman and Wise,1978; Daganzo, 1979) have made it computationally feasible.

The MNL model is empirically tractable and has a satisfactory theoretical justification in terms of the underlying behaviour of individual decision-makers. Domencich and McFadden (1975) suggest that the MNL model should be limited to multiple-choice situations where the alternatives can plausibly be assumed to be distinct and independent in the eyes of the decision maker.

As noted in Chapter 1, the aim of this work is also to assess the responsiveness of airport choice to changes in its determinants. Therefore it is necessary to have an adequate model which has highly acceptable elasticities properties. The advantage of the logit model, in this sense, is that its elasticities are easily interpretable: for a one per cent change in any determinant X there will be a P_g percentage change in the probability of using airport g. The elasticities in the probit model are harder to interpret as they involve the use of standard deviation: each unit change in a determinant X corresponds to a Pg standard deviation unit change in the probability of using airport g.

If we can ensure that the IIA assumption is not violated, the MNL model can be used as the airport choice model and thus we avoid the additional computational burden imposed by the MNP model. We are going to test the validity of the IIA property in our case, using a test based on conditional choice and developed by McFadden, Train and Tye (1977).

6.2.2 Diagnostic test for the IIA property:

6.2.2.1 Test based on conditional choice:

If two dependent alternatives are included in the calibration sample, a set of model coefficients will be generated other than those generated by a model in which one of the dependent alternatives is eliminated, i.e. violation of the IIA property will cause the maximum likelihood parameter estimates to be biased. If the IIA property is valid, however, the coefficients estimated from the full choice set will coincide with the coefficients for a smaller choice set. A test of the validity of the IIA property is whether or not the coefficients estimated from a reduced choice set are statistically different from those estimated from the full choice set. In applying this test, the estimation is performed on the subsample of individuals who choose an alternative in the subset of alternatives to be tested for dependence. The coefficients of representative utility are estimated on the subsample and the log likelihood at convergence is calculated. The log likelihood is also calculated on the subsample with the coefficients restricted to the values estimated from the full choice set.

By using the likelihood ratio test statistic $\chi^2 = 2$ (log likelihood for subsample - log likelihood with coefficients restricted) which is an asymptotically distributed chi-square with the number of degrees of freedom equal to the number of parameters, the hypothesis that the coefficients estimated on the subsample are the same as those estimated from the full choice set, is tested.

6.2.2.2 Results

The test based on conditional choice has been carried out for the four categories of passengers (domestic, international business, international leisure, inclusive tours). First, the coefficients from the full choice set (4 airports) have been estimated for each category of passengers. Then, the coefficients were estimated on the subsample of passengers who chose an airport in a set of 3 airports. This subsample has been obtained by excluding from the total sample the passengers who chose the airport to be tested for dependence. (The computer program used for these estimations is described in Chapter 7). The log-likelihood at convergence and the log-likelihood with the coefficients restricted to the values estimated from the full choice set have been computed for each case. The likelihood ratio statistic was then calculated for each case.

The results of the test are given in Tables 6.1, 6.2, 6.3 and 6.4., for the business, leisure, inclusive tours and domestic samples respectively, which show that the critical value of χ^2 at 95 percent for the corresponding degrees of freedom is higher than the value of the test statistic in all cases.

We can therefore, accept the hypothesis that the coefficients estimated on the subsamples are the same than those of the total sample (i.e. total set) for each one of the four categories of passengers, and thus we can conclude non-violation of the IIA assumption.

	Alternatives included in subset			
	All except Manchester	All except Birmingham	All except East Midlands	All except Heathrow
Log-likelihood at convergence for sub-sample choosing an alternative within subset	-25.07	-27.47	-9.61	-4.92
Log-likelihood with coefficients restricted to values estimated from the full choice set	-25.62	-29.60	-12.58	-7.60
Test	1.10	4.26	5.94	5.36
Critical $\chi^2_{(0.05,2)}$	5.99	5.99	5.99	5.99

Table 6.1 : Test based on conditional choice for international business sample

	Alternatives included in subset			
	All except Manchester	All except Birmingham	All except East Midlands	All except Heathrow
Log-likelihood at convergence for subsample choosing an alternative within subset	-31.24	-5.74	-17.26	-29.23
Log-likelihood with coefficients restricted to values estimated from the full choice set	-31.58	-7.54	-17.87	-30.24
Test	0.68	3.60	1.22	1.82
Critical $\chi^2_{(0.05,3)}$	7.83	7.83	7.83	7.83

Table 6.2 : Test based on conditional choice for international leisure

	Alternatives included in subset			
	All except Manchester	All except Birmingham	All except East Midlands	All except LUTON
Log-likelihood at convergence for subsample choosing an alternative within subset	-129.51	-205.13	-153.82	-217.90
Log-likelihood with coefficients restricted to values estimated from the full choice set	-131.81	-208.06	-156.29	-220.20
Test	4.60	5.86	4.94	4.60
Critical $\chi^2_{(0.05,2)}$	5.99	5.99	5.99	5.99

Table 6.3 : Test based on conditional choice for inclusive tours

	Alternatives included in subset			
	All except Manchester	All except Birmingham	All except East Midlands	All except Heathrow
Log-likelihood at convergence for subsample choosing an alternative within subset	-30.80	-14.53	-20.61	-50.17
Log-likelihood with coefficients restricted to values estimated from the full choice set	-33.32	-17.20	-23.04	-52.67
Test	5.04	5.34	5.58	5.0
Critical χ^2 (0.05,3)	7.83	7.83	7.83	7.83

Table 6.4 : Test based on conditional choice for domestic sample

6.3 Conclusion:

Because we ensured that the IIA assumption is not violated and for its tractability advantages, it emerges that the MNL used is, on balance the most appropriate tool for the airport choice.

CHAPTER 7

MODEL EVALUATION AND ANALYSIS

7.1 Introduction

This chapter presents and analyses the results of the model calibration. Separate models were calibrated for business, leisure, inclusive tours and domestic air passengers for the reason stated in chapter 4. The original computer program was written by Ben-Akiva (1973) of the Massachussets Institute of Technology. This program, based on a maximum likelihood technique, is not bounded as the number of observations, parameters to be estimated or alternatives. It provides the probability of selecting an alternative, the model's parameters and their corresponding t-tests and the likelihood ratio statistics for the equal share hypothesis. This computer program is in appendix D to this work.

The model is also assessed on its ability to predict the actual choice by the presentation of the prediction success table. An evaluation of the responsiveness of choice to changes in the significant explanatory variables is also determined using an elasticity measure*.

7.2 Calibration results

The utility function of the model can be written as:

$$U = a_1 . TT + a_2 . FREQ + a_3 . FARE \qquad \text{where:}$$

* Additional programs have been written to calculate the direct and cross elasticities and the prediction success table.

TT is the travel time to the airport

FREQ is the number of flights per day

FARE is the air fare

a_1, a_2, a_3 are the coefficients to be estimated.

The parameter estimates are displayed in table 7.1.a with the corresponding t-values, the likelihood ratio test and the likelihood ratio index. (These goodness of fit measures have been presented in Chapter 5).

	Business	Leisure	Inclusive Tours	Domestic
a_1	-0.15581 (-6.25)*	-0.1 3788 (-6.47)*	-0.17231 (-13.54)*	-0.23254 (6.71)*
a_2	2.0034 (6.15)*	1.07 (5.87)*	2.0587 (13.58)*	2.6957 (6.61)*
a_3	0.34998 (3.32)	-1.2035 (-4.23)**	0.096607 (3.25)	-0.74645 (-5.22)**
likelihood ratio test	535.92	847.02	3134.27	1746.60
likelihood ratio index	0.90	0.92	0.84	0.92
$\chi^2(0.01,3)$	11.34	11.34	11.34	11.34

* significant at 99 percent level (t values shown in parentheses)

** significant at 95 percent level

Table 7.1.a Calibration results

The likelihood ratio test values are much larger than the tabulated χ^2 at 99 percent confidence level and so implies an excellent fit. The likelihood ratio index values of 0.90, 0.92 and 0.84 reinforce this verdict. The travel time and frequency parameters are significant at 99 percent level for all the samples.

The fare parameter is found to be significant at 95 percent level for the international leisure and domestic samples. It has the 'wrong' sign for the international business and inclusive samples (this parameter is expected to be negative).

Thus, the model was run again leaving out this variable for the business and inclusive tours samples. The re-estimation of the model would reveal whether or not this variable was important or not from the viewpoint of overall fit. The results of the second calibration are given in Table 7.1.b.

The likelihood ratio test values and the likelihood ratio index values show that the overall fit is still excellent.

The variable parameters are similar in sign and in magnitude little different from those of Table 7.1.a. They are significant at 95 or 99 percent confidence level. This result suggests that the fare variable can be dropped from the analysis for the international business and inclusive tours passengers. Thereafter, model will mean model II.

	Business	Leisure	Inclusive Tours	Domestic
a_1	-0.13605 (-6.93)**	-0.13788 (-6.47)*	-0.17787 (-11.23)*	-0.23254 (-6.71)*
a_2	1.6607 (6.79)**	1.07 (5.87)*	2.069 (10.69)*	2.6957 (6.61)*
a_3	- -	-1.2035 (-4.23)**	- -	-0.74645 (-5.22)**
likelihood ratio test	521.03	847.02	3123.67	1746.60
likelihood ratio index	0.88	0.92	0.84	0.92
$\chi^2(0.01,df)$	9.21	11.34	9.21	11.34
df	2	3	2	3

*significant at 99 percent confidence level

** significant at 95 percent confidence level

(t values shown in parentheses)

Table 7.1.b. Calibration results. Model II.

Another goodness of fit measure is the prediction success tables which were presented in Chapter 5 and are given in Tables 7.2, 7.3, 7.4 and 7.5 for the international business, leisure, inclusive tours and domestic samples respectively. The overall success index after normalisation (so as to have a maximum value of 1) is also presented in each case. They indicate that the model predicts correctly 94.85%, 94.56%, 90.21% and 95.75% of the passengers' choices for the business, leisure, inclusive tours and domestic samples respectively with a normalised overall success index of 0.93, 0.91, 0.86 and 0.93 respectively. These high values confirm again the excellent fit of the model.

	Predicted choice					
	Manchester	Birmingham	E.Midlands	L.H.R.	observed share	observed share
Manchester	57	0	0	2	59	27.57
Birmingham	0	54	0	0	54	25.23
E.Midlands	0	0	31	2	33	15.42
L.H.R.	1	0	6	61	68	31.78
Predicted count	58	54	37	65	214	
Predicted share	27.10%	25.23%	17.29%	30.27%	1	
Proportion successfully predicted	98.3%	100%	83.78%	93.84%	94.85%	

overall proportion successfully predicted: 94.85%

normalized overall success index: 0.93

Table 7.2 Prediction success table - business sample

	Predicted choice					
	Manches-ter	Birming-ham	E.Mid-lands	L.H.R	observed count	observed share
Manchester	65	0	0	0	65	19.63%
Birmingham	0	183	10	4	197	59.52%
E.Midlands	0	0	46	0	46	13.90%
L.H.R.	1	2	1	19	23	6.95%
Predicted count	66	185	57	23	331	
Predicted share	19.94%	55.89%	17.22%	6.95%	1	
Proportion success-fully predicted	98.48%	98.92%	80.70%	82.61%	94.56%	

overall proportion successfully predicted: 94.56%

normalized overall success index : 0.91

Table 7.3 Prediction success table - leisure sample

	Predicted choice					
	Manches-ter	Birming-ham	E.Mid-lands	Luton	observed count	observed share
Manchester	277	10	35	0	322	23.87%
Birmingham	19	430	4	0	453	33.58%
E.Midlands	13	19	433	5	470	34.84%
Luton	3	2	22	77	104	7.71%
Predicted count	312	461	494	82	1349	
Predicted share	23.13%	34.17%	36.62%	6.08%	1	
Proportion success-fully predicted	88.78%	93.28%	87.65%	93.9%	90.21%	

overall proportion successfully predicted: 90.21%

normalized overall success index: 0.86

Table 7.4 Prediction success table - Inclusive tours sample

	Predicted choice					
	Manchester	Birmingham	E.Midlands	L.H.R.	observed count	observed share
Manchester	65	0	3	0	68	9.96%
Birmingham	17	214	5	4	240	35.14%
E.Midlands	0	0	362	0	362	53.00%
L.H.R.	0	0	0	13	13	1.90%
Predicted count	82	214	370	17	683	
Predicted share	12.00%	31.33%	54.17%	2.49%	1	
Proportion successfully predicted	79.27%	100%	97.84%	76.47%	95.75%	

overall proportion successfully predicted: 95.57%
normalized overall success index: 0.93

Table 7.5 Prediction success table - Domestic sample

The predicted share of each airport is close to the observed share for all the samples. Table 7.6 and Figure 7.1 summarize this result.

	Business		Leisure		Inclusive Tours		Domestic	
	Obs.%	Pred.%	Obs.%	Pred.%	Obs.%	Pred.%	Obs.%	Pred.%
Manchester	27.57	27.10	19.63	19.94	23.87	23.13	9.96	12.00
Birmingham	25.23	25.23	59.52	55.89	33.58	34.17	35.14	31.33
E.Midlands	15.42	17.29	13.90	17.22	34.84	36.62	53.00	54.17
LHR*	31.78	30.27	6.95	6.95	7.71	6.08	1.90	2.49

* LUTON for Inclusive Tours

Table 7.6 Observed and Predicted shares

Because of the strategy of model building considered, only policy sensitive variables have been selected which meant that the variables representing passengers' characteristics have been deliberately excluded. The predictive performance of the models show that the remaining proportion that could be successfully predicted by the inclusion of the omitted variables has a maximum of 5% for the business, leisure and domestic models and 10% for the inclusive tours model. This suggests that the strategy of pragmatic model building used avoids limitations to the model's practical applications with no great loss in accuracy.

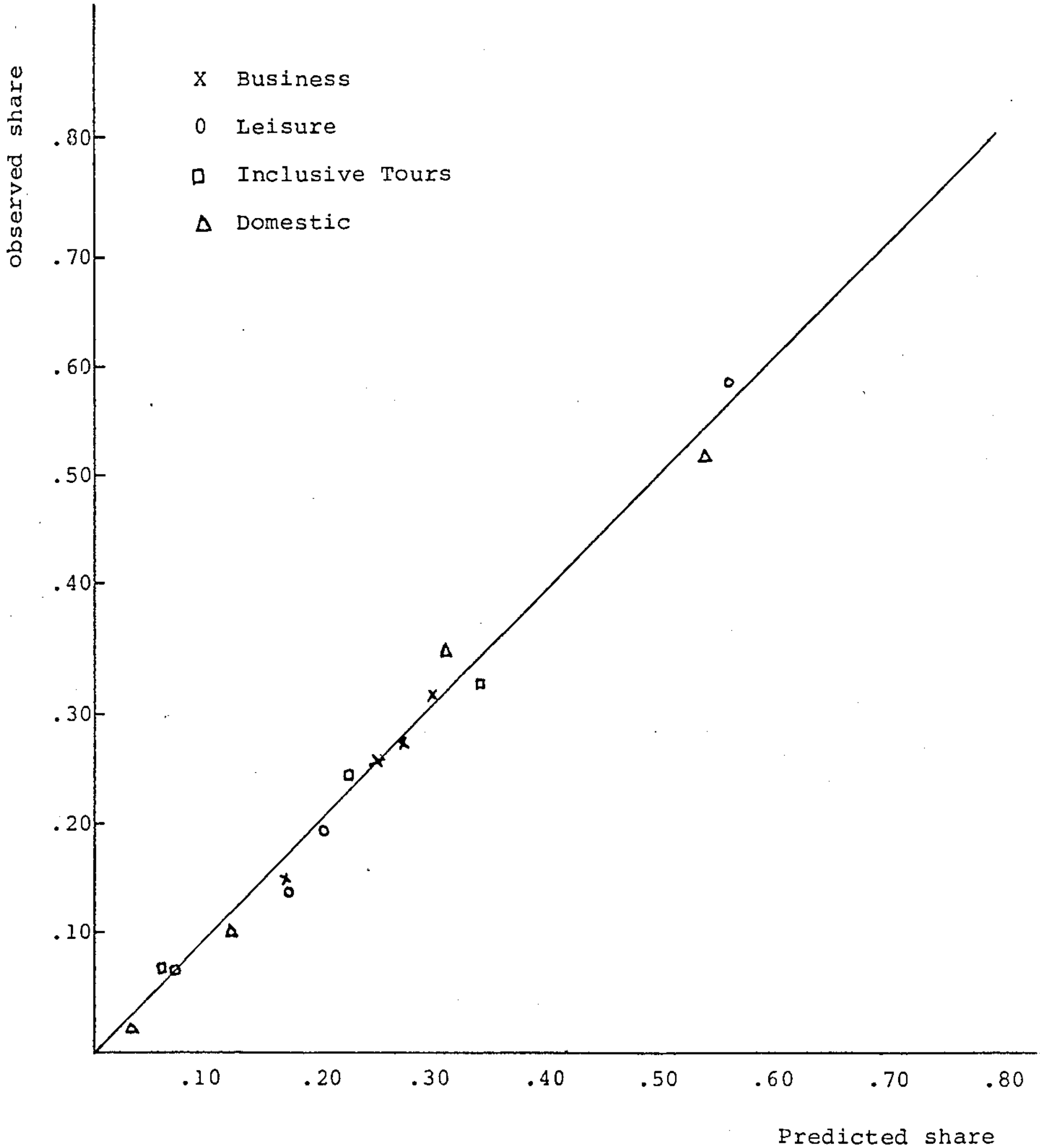

Figure 7.1 Observed vs.Predicted shares

7.3 Responsiveness of airport choice to changes in the explanatory variables

An import piece of information which should be provided by an airport choice model is a measure of the responsiveness of airport choice to changes in explanatory variables. A classical measure of responsiveness is the elasticity measure.

Direct and cross elasticities can be obtained from the model. Direct elasticity is the percentage change in the probability of choosing a particular airport in the choice set with respect to a given percent change in an explanatory variable which describes the utility of that airport.

A cross-elasticity is the percentage change in the probability of choosing a particular airport in the choice set with respect to a given percent change in an explanatory variable which describes the utility of a competing airport.

The elasticity of Y = f(X) with respect to the variable X is $(\frac{\Delta Y}{Y})/(\frac{\Delta X}{X})$ which reduces to $(\frac{\delta Y}{\delta X}) \cdot (\frac{X}{Y})$ as ΔX approaches zero. Therefore, direct point elasticities in the MNL model can be written:

$$E^{P_{gk}}_{X_{gk\ell}} = \frac{\delta P_{gk}}{\delta X_{gk\ell}} \cdot \frac{X_{gk\ell}}{P_{gk}}$$

where P_{gk} = probability of choosing airport g by individual k

$X_{gk\ell}$ = ℓ^{th} explanatory variable describing airport g for individual k

$E^{P_{gk}}_{X_{gk\ell}}$ = elasticity of the probability of choosing airport g with respect to a change in the ℓ^{th} explanatory variable which describes airport g for individual k.

Considering the equation for P_{gk} given by eq.(5.4), and recalling the definition of the utility function V_{gk} and using the rule

$$\frac{\delta e^{az}}{\delta z} = a.e^{az}, \text{ we have:}$$

$$E^{P_{gk}}_{X_{gk\ell}} = a_{g\ell} \cdot X_{gk\ell} \cdot (1 - P_{gk}) \qquad (7.1)$$

Thus, the logit model assumes that direct elasticity of choice for a particular airport with respect to a particular explanatory variable is a function of the level of that variable and the share (probability) this airport could still gain.

For the cross elasticity, we follow a similar path by evaluating $\frac{\delta P_{gk}}{\delta X_{jk\ell}}$ and then evaluate the cross-elasticity as:

$$E^{P_{gk}}_{X_{jk\ell}} = \frac{\delta P_{gk}}{\delta X_{jk\ell}} \cdot \frac{X_{jk\ell}}{P_{gk}} = -a_{j\ell} \cdot X_{jk\ell} \cdot P_{jk} \qquad (7.2)$$

We can notice that this cross elasticity only depends on variables associated with alternative j and is independent of g. Therefore, cross elasticities with respect to a variable associated with alternative j are the same for all $i \neq j$ (e.g. Birmingham, East Midlands and Heathrow will have the same cross elasticity with respect to a variable associated with Manchester airport). This constrained result arises because of the I.I.D. assumption of the logit model but can be considered reasonable after the test carried out in Chapter 6.

A simple way in which eq.(7.1) and eq.(7.2) may be combined to yield a single point elasticity formula for the MNL is given by:

$$E^{P_{gk}}_{X_{jk\ell}} = a_{j\ell} \cdot X_{jk\ell} \cdot (\delta_{gj} - P_{jk}) \tag{7.3}$$

where

$\delta_{gj} = 1$ if $g = j$ (direct point elasticity)

$\delta_{gj} = 0$ if $g \neq j$ (cross point elasticity).

Eq.(7.3) gives elasticities for each individual. To find aggregate elasticities to determine the market demand elasticities which are needed for policy analysis, we cannot evaluate eg.(7.3) at the sample average $\dot{X}_{j\ell}$ and $\hat{P}_j$ since the MNL is non-linear and the estimated logit function need not pass through the point defined by these sample averages. Hensher and Johnson (1981) consider that a preferable approach is to evaluate eq.(7.3) for each individual k and then aggregate, weighting each individual elasticity by the individual's estimated probability of choice. This technique is known as the sample enumeration method, the formula for which is:

$$E^{\bar{P}_g}_{X_{jk\ell}} = \left(\sum_{k=1}^{R} \hat{P}_{gk} \cdot E^{P_{gk}}_{X_{jk\ell}} \right) / \left(\sum_{k=1}^{R} \hat{P}_{gk} \right) \tag{7.4}$$

where

$\hat{P}_{gk}$ = estimated choice probability

$\bar{P}_g$ = aggregate probability of choice of alternative with respect to a variable $X_{jk\ell}$.

The values of aggregate direct and cross elasticities with respect to the variables found significant at the 95 percent level are estimated and presented in Tables 7.7, 7.8, 7.9 and 7.10 for business, leisure, inclusive tours and domestic samples respectively.

It should be noted that in the cross elasticities table, each column represents the cross elasticity with respect to a 1 percent change in a variable of the utility function of the airport not mentioned (e.g. first column is the cross elasticity of Birmingham, East Midlands and Heathrow with respect to a change in a variable associated with Manchester airport).

	Direct elasticities				Cross elasticities			
	M	B	EM	LHR	B.EM LHR	M.EM. LHB	M.B. LHR	M.B. EM.
TT	-0.77	-0.33	-0.85	-1.94	4.98	2.21	2.76	21.05
FREQ	0.31	0.26	0.24	1.79	-4.33	-3.49	-2.12	-26.87

Table 7.7 Direct and cross-elasticities - business sample

The interpretation of the estimated elasticities is a straightforward task. For example, for the direct elasticities, Table 7.7 shows that 1 percent increase in travel time will, all else remaining constant, cause a 0.77%, 0.33%, 0.85% and 1.94% decrease in the probability of choosing Manchester, Birmingham, East Midlands and Heathrow respectively. In the same way, 1% increase in the frequency of flight will cause a 0.31%, 0.26%, 0.24% and 1.79% increase in the probability of choosing Manchester, Birmingham, East Midlands and Heathrow respectively.

For the cross elasticities, a 1% increase in travel time to Manchester, everything else remaining constant, will result in 4.98% increase in the probability of choosing, Birmingham, East Midlands and Heathrow. Other estimated elasticities may be interpreted in a similar manner throughout this section. In addition, the higher the absolute value of the elasticity for any variable, the more sensitive is the choice of airport to the value of that variable. By convention, an absolute value of the elasticity greater than unity is said to be elastic, equal to unity it is known as unit elastic and less than unity it is called inelastic.

Furthermore, when cross elasticities are lower (higher) than direct elasticities, it suggests that airport choice decisions are more (less) responsive to changes in the variables characterizing the airport in question than to changes in the variables of the other alternative airports.

Great care, however, must be exercised when comparing cross and direct elasticities. Their relative sizes are intuitively doubtful in all cases, as the cross elasticities are always greater than the direct elasticities. Appendix E suggests that this comes from the aggregation approach in the sample enumeration method where each individual elasticity is weighted by the individual's estimated probability of choice. Consequently, when a model performs very well, as the airport choice model considered in this work does, the aggregate choice elasticities will inevitably be higher than the aggregate direct elasticities. As such, a mechanistic use of the direct and cross elasticities should be avoided. These comments apply to the interpretation of Tables 7.7, 7.8, 7.9 and 7.10. Table 7.7 shows that:

- Travel time has the most influence on airport choice.
- Frequency of flight has a lower direct elasticity but is still found to have some influence especially when we consider that this variable measure is not continuous (e.g. at 5 flights per day, additional flights can be added only in 20% increments). Furthermore, there are no small changes in this variable (e.g. increase from 4 to 5 flights per day is a 25% increase).
- The two variable elasticities are much higher for Heathrow than for the other 3 airports. This could be explained by the fact that in the logit model the elasticities are a

function of the value of the associated variables and these values are large for Heathrow: travel time value is large because of the geographical location of Heathrow outside the study area and frequency value is large because of the large number of daily flights offered at Heathrow to the chosen destinations.

- The cross elasticities suggest that each of the airports considered is less responsive to changes in its own variables than to changes in the variables of the other alternative airports.

- The high values of the cross-elasticities when there is a change in a variable characterizing Heathrow (last column of Table 7.7) show that there is most scope for drawing back business traffic, originating from Central England, from Heathrow to the three regional airports by considering restrictions on the growth of Heathrow.

	Direct elasticities				Cross elasticities			
	M	B	EM	LHR	B.EM. LHR	M.EM. LHR	M.B. LHR	M.B. E.M.
TT	-0.22	-0.19	-0.67	-4.38	3.23	1.90	2.56	18.74
FREQ	0.07	0.08	0.13	2.89	-3.21	-2.03	-0.95	-14.23
FARE	-0.97	-1.26	-4.25	-6.74	25.21	25.5	23.11	31.7

Table 7.8 Direct and cross elasticities - Leisure sample

Table 7.8 indicates that the fare variable is the attribute having the most influence on airport choice for the international leisure passenger when frequency of flight is the less dominant factor. Manchester, Birmingham and East Midlands are more sensitive to changes in the attributes of the other alternative airports than to changes in their own variables. Heathrow, on the other hand, is more sensitive to changes in its travel time and frequency variables and less sensitive to changes in its fare variable than to changes in the corresponding variable of the other 3 airports. Again, there is most scope for drawing back leisure passengers originating from Central England, from Heathrow to the three regional airports by considering restrictions on the growth of Heathrow.

	Direct elasticities				Cross elasticities			
	M	B	EM	LUT	B.EM. LUT	M.EM. LUT.	M.B. LUT.	M.B. EM
TT	-1.73	-0.52	-1.38	-7.20	6.97	2.79	5.29	11.83
FREQ	0.58	0.49	0.48	6.69	-4.02	-3.98	-3.24	-13.57

Table 7.9 Direct and cross elasticities - Inclusive tours sample

Table 7.9 shows that travel time is again the variable having the most influence with the frequency variable still found to be important (keeping in mind the comments about Table 7.7).

A comparison of direct and cross elasticities shows that Luton is more responsive to changes in its attributes than to changes in the attributes of the 3 other airports.

On the other hand, Manchester, Birmingham and East Midlands are more sensitive to changes in the attributes of the other alternative airports than to changes in their own attributes.

In Table 7.10, fare is found to be the variable having the most influence (except for Heathrow where it is travel time) followed by travel time. Manchester, Birmingham and East Midlands are more sensitive to changes in the attributes of the other alternative airports than to changes in their own variables. Heathrow is more responsive to changes in its travel time and frequency variable and less sensitive to changes in its fare variable than to changes in the corresponding variable of the other 3 airports. Again, restrictions on the growth of Heathrow seem to be the best way of drawing back domestic passengers originating from Central England to the three regional airports.

	Direct elasticities					Cross elasticities			
	M	B	EM	LHR		B.EM. LHR	E.EM. LHR	M.B. LHR	M.B. EM.
TT	-1.77	-0.62	-0.46	-9.13		7.07	3.23	6.83	28.67
FREQ	0.42	0.28	0.21	8.72		-4.41	-3.51	-3.39	-27.38
FARE	-1.97	-0.87	-0.62	-3.68		10.63	12.48	12.21	8.28

Table 7.10 Direct and cross elasticities - Domestic sample

7.4 Results analysis

One of the primary objectives of calibrating the model was to measure the relative importance of the frequency of flights, travel time to the airport and air fare as determinants of airport choice and consequently to suggest some implications regarding the role of these attributes as airport policy tools.

7.4.1 Relative importance of variables

In the calibration of the model, frequency of flights and travel time have been found significant for the four samples. The fare was not significant for the business passengers, a result one could expect especially when we consider the regulatory environment as far as air fare is concerned (e.g. the difference between the cheapest and the dearest air fare offered to Dublin from the airports considered was £1.5). For the inclusive tours, it seems that the hidden benefits such as free holiday for children or the hotel class etc. which were not taken into account, counterbalanced the fare which has been assumed to be 40% of the global cost (air fare plus holiday).

Table 7.11 indicates that the absolute value of the ratio of coefficient of frequency to the coefficient of travel time is greater for the business passenger than for the non-business passenger.

	Ratio of Freq.coefficient to TT coefficient	Ratio of fare coefficient to TT coefficient
Business	12.21	-
Leisure	7.76	8.72
Inclusive Tours	11.63	-
Domestic	11.59	3.21

Table 7.11 Comparison of calibration results by purpose

The non-business passenger appears more concerned with accessibility or less concerned with flight frequency, or both, than business passenger when choosing between airports.

In the same way, the leisure passenger appears more concerned with fare or less concerned with flight frequency and accessibility, or both, than the other 3 categories of passenger when choosing between airports. These results confirm what one might expect intuitively.

The high ratio for the inclusive tours passengers (11.63) who are only concerned about a single specific departure may seem surprising but could be explained by the fact that these passengers are also less sensitive to an extra travel time to catch their flight from a distant airport.

The relative importance of travel time, frequency of flights and air fare as determinants of airport choice can be measured using the elasticities figures in Tables 7.7 to 7.10.

Tables 7.7 and 7.9 show that for the business and inclusive tours passengers, the absolute value of the direct elasticity of travel time is in all cases higher than the corresponding value of the direct elasticity of frequency of flights. From this result we can conclude that the travel time variable is the dominant factor. However, it should be noted that the assumption made about the inclusive tours air fare levels could explain the non-significance of this variable and thus weakens this particular finding as far as the inclusive tours passengers are concerned. Future research should investigate this variable more thoroughly.

Similarly Tables 7.8 and 7.10 indicate that the fare variable is the dominant factor for the leisure and domestic passengers. This importance of the fare variable for the domestic passenger could be found rather odd but could be explained by the competition of other modes (rail, bus). A low fare adds to the attractiveness of an airport for passengers who could otherwise travel by surface transport. These con-conclusions are summarized in Table 7.12.

	Business - Inclusive Tours	Domestic - leisure
1st dominant	TT	FARE
2nd dominant	FREQ	TT
3rd dominant	-	FREQ

Table 7.12 Relative importance of variables

Examples can be constructed to illustrate the magnitude of this relative importance of variables. For a business passenger with a travel time of 60 minutes, a 20% increase in flight frequency (e.g. from 4 to 5 flights per day) is equivalent to a decrease in travel time of 5 minutes, 9 minutes, 4 minutes and 7 minutes for Manchester, Birmingham, East Midlands and Heathrow respectively. For a domestic passenger, the same increase in flight frequency is equivalent to a decrease in travel time of 3 minutes, 5 minutes, 6 minutes and 11 minutes for Manchester, Birmingham, East Midlands and Heathrow. Examples can be constructed in a similar way for the other samples. While these results do appear believable qualitatively, their magnitude needs some comments. We have already indicated that the frequency variable is not a truly continuous variable. There is also a tendency to be no small changes in this variable (e.g. from 2 to 3 flights per day, there is a 50% increase) so that the differences between travel time and frequency elasticities which appear large are not exaggerated. Secondly, air carriers competitive scheduling attempts to space flights in prime periods,(airlines often adjust their frequencies at a given period of time in response to the output sold during the corresponding period of the previous year), so that the difference between 4 flights at one airport and 5 flights at another airport may not be perceived as a very significant difference by a passenger.

In the same way, for a domestic passenger with a travel time of 60 minutes, a 10% increase in fare is equivalent to 6 minutes, 8 minutes, 8 minutes and 4 minutes increase in travel time in terms of the net effect on the probability of choosing Manchester, Birmingham, East Midlands and Heathrow respectively. A 20% increase in flight frequency is equivalent to a decrease in fare for a passenger going to Jersey of 84p, 117p, 127p and

654p for Manchester, Birmingham, East Midlands and Heathrow respectively. Endless illustrations can be constructed in the same way. They are in the same order of magnitude as the examples depicted previously.

7.4.2 Policy implications

The model results suggest some significant implications regarding the role of travel time, frequency of flight and air fare.

Unless very substantial fare differentials are applied, air fare will have a limited success as a policy tool to encourage business passengers to choose an airport offering less flights or higher accessibility time. On the other hand, air fare can be very effective in shifting the leisure and domestic passengers' choice of airport.

Differences in frequency of flights can be an effective policy tool for shifting the business and inclusive tours passengers' choice of airport. For the leisure passengers, these differences would have to be so substantial that it could result in one airport being left with virtually no service to the affected destinations.

Improvements in airport access can be an efficient policy tool in shifting the passengers' choice of airport regardless of their trip purpose. However, as an entire region has been considered as a study area, this improvement should be oriented to the entire area and not only to a single city and this could prove to be a difficult and costly task.

From the elasticities tables and the airports market shares, the changes in traffic level resulting from changes in variables can be derived. For example, a 10% decrease in travel time will result in a 7.7% increase in the probability of choosing Manchester airport for the business passengers. By multiplying this increase in the probability of choice by the business market share of Manchester airport, the percentage change in the business market share is obtained as 2.09%. Tables 7.13, 7.14 and 7.15 give the percentage change in traffic level at each airport for each category of passengers resulting from a 10% decrease in travel time, a 10% increase in flight frequency and a 10% decrease in fare respectively. In each case, the two other variables are assumed to be unchanged.

	MAN	BMG	EMA	LHR
Business	2.09	0.83	1.46	5.81
Leisure	4.38	1.06	1.15	2.00
Inclusive tours	4.00	1.77	5.05	4.37
Domestic	2.12	1.94	2.49	2.27

Table 7.13 % change in traffic levels for 10% decrease in TT

	MAN	BMG	EMA	LHR
Business	0.84	0.66	0.42	5.42
Leisure	0.14	0.44	0.23	2.00
Inclusive tours	1.34	1.67	1.76	4.06
Domestic	0.51	0.88	1.13	0.92

Table 7.14 % change in level of traffic for 10% increase in flight frequency

	MAN	BMG	EMA	LHR
Leisure	1.9	7.0	7.4	4.7
Domestic	3.5	2.7	3.3	1.0

Table 7.15 % change in level of traffic for 10% decrease in fare

Confining the analysis to the three regional airports, Tables 7.13, 7.14 and 7.15 show that the same change in a given variable results in different changes in traffic level between airports. The higher the change in traffic level, the greater is the possibility of attracting more passengers by applying the consequent policy. Thus, it can be concluded for the regional airports, as far as the passengers from the study area are concerned, that:

- An access improvement policy would give the best results in the attraction of more passengers if it is applied at East Midlands airport for the domestic and inclusive tours passengers and at Manchester airport for business and leisure passengers.

- A frequency of flights policy would give the best results if it is applied at Manchester airport for the business passengers and at East Midlands airport for the domestic and inclusive tours passengers.

- A fare policy would give the best results in the attraction of more passengers if it is applied at Manchester airport for the domestic passengers and at East Midlands airport for the leisure passengers.

Therefore, Manchester airport has the potential to develop into a "hub" airport. This finding supports the 1985 UK government White Paper on airport policies which states that:

> *The government is fully committed to maintaining and further developing Manchester airport as a gateway ... for long haul services, and as a domestic and European hub.*

East Midlands airport could also attract more passengers, particularly those on domestic, inclusive tours and leisure trips, if the required policies are applied. Birmingham airport seems to suffer more from the proximity of the London airports and must make large changes in the variables considered to impact the market significantly.

These model results give the airport managements, the possibility of selecting from all the measures open to them, that policy option which results in a higher increase in their market shares. As an example, if the measures considered are the changes in variables already cited above, the consequent options that should be selected are:

- At Manchester airport: the airport management should strive for regional transport improvements to ease access to the airport as the decrease in travel time is the best option to attract more business, leisure and inclusive tours passengers. The decrease in air fare is the best option to attract more domestic passengers.

- At Birmingham airport: the decrease in travel time will attract more business and inclusive tours passengers while the decrease in fare will attract more leisure and domestic passengers.

- At East Midlands airport: select the decrease in travel time for attracting more business and inclusive tours passengers. Select the decrease in air fare for attracting more leisure and domestic passengers.

These options are the best ones of the measures considered. Different hypotheses about the relative changes in variables (e.g. 5% decrease in fare, 20% increase in frequency, 5% decrease in travel time) will result in different courses of recommended action.

7.5 Conclusions

In the light of the results and analysis presented in the previous sections, a number of conclusions can be made:

- The models are good in terms of their explanatory ability and in predicting the actual choice of airport. There is, certainly, a major improvement on similar work (see Chapter 2) when the models were calibrated with aggregate data.

- Variables expected a priori to be important, are important explanatory variables in fact.

- The a priori stratification of passengers into four categories reasoning that they make trade offs differently, was justified by the models results.

- The models confirmed that non-business passengers are more concerned with accessibility or less concerned with flight frequency than business passengers.

- The models indicate that the three explanatory variables do not have an equal importance as determinant of airport choice. The accessibility variable is more important than the frequency of flights variable for all the passengers. The fare variable is the most important determinant for the leisure and domestic passengers. As a result, improvement in airport access, flight frequency differentials (for the business traffic) and air fare differentials (for the leisure and domestic traffic) can be effective tools in shifting the passengers' choice of airport.

- As far as the passengers originating from the study areas are concerned, Manchester and East Midlands airports have a greater potential than Birmingham airport which seems to suffer more from the proximity of the London airports.

- The model results can be used in practice by airport managements in the choice, from all the options open to them, of the policy which will result in a predictable increase in their airport market share.

CHAPTER 8

TRANSFERABILITY OF THE AIRPORT CHOICE MODEL

8.1 Introduction

Individual choice models have been credited with great accuracy in estimating and predicting travel choice. It has been argued (Hensher and Johnson,1981) that the parameters of such models, developed at the level of the decision-making unit, should remain stable in predicting travel behaviour not only for one area in different time periods but also for different groups of people in different areas. The basis of this belief is that behavioural decision processes are aspatial, that is, regularities in behaviour of individuals would allow a single model estimated in one place at one time to be used in applications in other places at other times, thereby making the estimation of new models unnecessary. Transferring existing models to new situations and areas provides a low-cost alternative to the development of new models which is relatively expensive in data requirements and analyst's time.

This chapter tests the hypothesis of transferability of the models estimated on Central England in 1975 (Chapter 7) to the London area in 1978. First, past studies of transferability of model coefficients are presented. Secondly, the data sets are described. Thirdly, the set of transfer models are estimated and tested for parameter transferability. Finally, the models are tested for predictive performance.

8.2 Past studies of transferability of model coefficients

Studies dealing with transferability have sought to answer the question: Can disaggregate travel choice models be transferred from one area to another without modification of the coefficients? This section reviews some of the past studies which have attempted to test the transferability of model parameters.

Watson and Westin (1975) studied the transferability of binary logit mode choice models among different subareas within the Edinburgh-Glasgow area of Scotland. Their models contained level of services variables and a mode specific constant but no socio-economic variable. The data were grouped into six categories according to whether the trip origins and destinations were in the central city, the suburbs or the area peripheral to the urban area. Each of the six models was then used to predict the mode splits of the other five samples. In the three categories that contained at least one trip end in the central city, they found that within the group, the models predicted well. The remaining three categories performed poorly within themselves. As between the non-central groups, the coefficients were significantly different. Watson and Westin concluded that the predictive ability of the model of the central city was fairly favourable to transferability, but that the results for the other group indicated a need to refine the models for locational differences.

Atherton and Ben-Akira (1976) explained travel behaviour in Los Angeles and New Bedford, Massachusetts, from a work trip mode choice model estimated on Washington, D.C. data. Their model predicted the probability of choosing to drive alone, share a ride or use public transit. The independent variables included network-derived travel times and costs, income, car availability, a dummy variable indicating whether the trip maker was household head, and mode specific constants. The test was to use the variables for the original (Washington) model to estimate new models with Los Angeles and New Bedford data, and then to compare the coefficients of the new models with those of the old model. The comparison consisted of statistical tests of the null hypothesis that the coefficients of the new models are equal to those of the old. For both models, the coefficients of the car availability variables were the only ones significantly different from their Washington counterparts. The authors concluded that evidence was encouraging but as no model is ever completely specified and hence perfectly transferable, they suggested that up-dating procedures were required.

Ortuzar and Fernandez (1985) estimated the same model structure using two different data sets. The first one gathered in1975 for suburban corridor in Leeds, England, and the second one gathered in 1981 for an urban corridor in Santiago, the capital of Chile. In both cases alternatives range from car alone to rail (underground in Santiago) with several combinations like park-and-ride and feeder bus services. The model examined was the multinomial logit model. The attributes considered were travel times and costs, and socio-economic variables of each individual's household. The authors concluded that the results obtained do not grant transferability of MNL models between different cultural settings.

8.3 Data collection

For the transferability test, two distinct data sets were required. The first one was the data set for 1975 Central England described in Chapter 4. The second data set concerned the London area in 1978. These data were collected from the 1978 CAA origins/destinations survey and the 1978 ABC World Airways. In the same way as the data set for Central England, the observations relating to the passengers from the study area (London area) with a genuine choice of airport have been extracted from the magnetic tapes recording the 1978 CAA Survey. The attributes of both the choice made and the choice rejected have been defined.

The airports considered were Heathrow, Gatwick and Luton (for inclusive tours only). The passengers' origins were randomly selected in this area and were: Kensington and Chelsea, Barnet, Ealing, Sutton, Croydon, Bromley, Bexley and Canterbury.

The passengers have been again stratified on the basis of:

- domestic
- International business
- International leisure
- International inclusive tours

The destinations were as follows:

Domestic:	Belfast, Jersey, Glasgow, Aberdeen, Manchester and Dusseldorf.
International:	Dublin, Amsterdam, Frankfurt, Zurich and Dusseldorf.
Inclusive tours:	Palma, Ibiza, Alicante, Tenerife and Malta.

In an attempt to achieve consistency between the two data sets, the same criteria were used for the selection of the above destinations. It can be noted that all the destinations selected in the Central England data set are also considered in the London area data set. This aim of consistency resulted in the rejection of one of the most important destinations, namely New York, which has characteristics (e.g. Concorde service, Laker "fare wars", choice at the other end between Kennedy and Newark airports) which would make it misleading to analyse data in the same manner as other destinations. Furthermore, only short and medium haul destinations were considered in the Central England data set and consequently the same criteria was applied to the London area data set. Travel time to the airport has been computed in the way described in Chapter 4. The frequency of flights data were drawn from the 1978 ABC World Airways Guide. As the air fares from Heathrow and Gatwick were the same for the destinations considered, this variable will not be included in the London area models as it is not a determinant of airport choice. The consequences of the non-inclusion of this variable on the model transferability will be discussed in section 8.4. Table 8.1 gives the samples considered for estimating the London area models.

	Business	Leisure	Inclusive	Domestic	Total
Heathrow*	299	131	38	93	561
Gatwick	30	37	236	82	385
Total	329	168	274	175	946

* Luton for inclusive tours

Table 8.1 1978 samples - London area

8.4 Estimating and testing the transfer model

A test of whether a model is transferable is to compare the parameter estimates obtained for the different areas (R.A.Galbraith and D.A.Hensher,1982). The transferability test statistically tests the differences between the two sets of coefficients where the null hypothesis is that the two sets of coefficients are equal. The test is given by:

$$\text{calculated } \chi^2 = 2\ (L^* - L^*_{CR})$$

where L* is the log-likelihood of the London area coefficients on the London area sample;

L^*_{CR} is the log-likelihood of the London area sample when the coefficients are restricted to the value of the Central England model.

The calculated χ^2 value is compared against the critical χ^2 value (at the 0.05 level). If the calculated $\chi^2 >$ critical χ^2, then the hypothesis that the two sets of coefficients are equal is rejected.

Thus,we should first estimate the model on the London area data in the same way as in Chapter 7.

8.4.1 Model estimation

The model has been estimated using the same computer program than the calibration of the Central England model and presented in Chapter 7.

The parameter estimates are displayed in Table 8.2 with the corresponding t-values, the likelihood ratio test and the likelihood ratio index.

	Business	Leisure	Inclusive tours	Domestic
TT	-0.11928	-0.18111	-0.12154	-0.51258
	(-5.00)*	(-4.37)*	(-5.63)*	(-4.38)*
FREQ	0.92781	1.0106	1.0196	2.8905
	(5.94)*	(4.36)*	(5.81)*	(4.45)*
Likelihood ratio test	371.8	167.93	294.71	197.82
Likelihood ratio index	0.81	0.72	0.78	0.81
$\chi^2(0.01,2)$	9.21	9.21	9.21	9.21

*significant at 95 percent confidence level

Table 8.2 Calibration results (1978)

Here again, the likelihood ratio test values which are much larger than the tabulated χ^2 and likelihood ratio index values show that the fit is good. The travel time and frequency parameters are signwise and significant at 95 percent confidence level for all the samples.

8.4.2 Transferability test

Table 8.3 shows the results of the transferability test described in the previous section. It indicates that the critical value of χ^2 at 95 percent confidence level for two degrees of freedom is lower than the value of the test in all cases.

We can therefore reject the hypothesis that the coefficients estimated on Central England samples are the same than those estimated on London Area samples for all the four categories of passengers and thus we conclude that the model estimated on 1975 Central England data is not transferable to the London Area in 1978.

	Business	Leisure	Inclusive Tours	Domestic
Log-likelihood at convergence	-42.14	-32.46	-42.57	-22.39
Log-likelihood at convergence with coefficients restricted	-84.13	-50.29	-55.65	-238.56
Test	83.98	35.66	26.16	432.34
$\chi^2(0.05,2)$	5.99	5.99	5.99	5.99

Table 8.3 Transferability test

Likely sources of problems in transferability are differences in sampling procedures, differences in estimation techniques and true behavioural differences.

The same sampling technique was used during the two surveys (1975, 1978). The questions relating to similar variables were worded consistently. It has already been noted that in an attempt to achieve consistency between the two data sets, the same criteria for data collection and preparation were used.

One significant difference between the two sets of data is the distribution of the frequency variable where the values are predictably much higher for the London model than for the Central England model.

As for the range of explanatory variables, a difference between the two models is the non-inclusion of the fare variable in the London model for the leisure and domestic passengers. This means that a model estimated on one group (Central England) was applied to another group (London area) which has a representative utility different from the first group. This suggests that the error associated with the fare variable was likely to be an important factor in the non-transferability of the two models.

The difference between individual coefficients can be evaluated by the statistic for the absolute difference between the transfer and unrestricted coefficients (last column of Table 8.4).

A t-statistic of less than 1.96 indicates that the null hypothesis that a pair of coefficients are equal cannot be rejected at the 5% level. The t-statistic used is the difference of the two coefficients divided by the square root of the sum of the variances of the two coefficients.

Table 8.4 shows that at the 5% level:

- For the business model, the travel time coefficients are not statistically different while the frequency of flights coefficients are statistically different.
- For the leisure model, the two pairs of coefficients are not statistically different.
- For the inclusive tours model, the two pairs of coefficients are statistically different.
- For the domestic model, the travel time coefficients are statistically different while the frequency of flights coefficients are not statistically different even if their difference (-0.195) is large because the standard error for this variable is large.

		Transfer model		Unrestricted model		test	
		a	t	a	t	difference	t
Business	TT	-0.13605	-6.93	-0.11928	-5.00	-0.0168	-0.54
	FREQ	1.6607	6.79	0.92781	5.94	0.733	2.53
Leisure	TT	0.13788	-6.47	-0.18111	-4.37	0.043	0.92
	FREQ	1.07	5.87	1.0106	4.36	0.06	0.20
Inclusive Tours	TT	-0.17787	-11.23	-0.12154	-5.63	-0.056	-2.09
	FREQ	2.069	10.69	1.0196	5.81	0.997	3.81
Domestic	TT	-0.23254	-6.71	-0.51258	-4.38	0.28	2.29
	FREQ	2.6957	6.61	2.8905	4.45	-0.195	-0.25

<u>Table 8.4</u> Comparison of pairs of coefficients

Although transferability is frequently linked to validation of the disaggregate modelling approach, model validity is a necessary but not sufficient condition for transferability. Theoretically, the argument for transferability may be realistic. However, the potential decrease as we move from theory to practical estimation, and as additional assumptions and/or imperfect data reduce the range and validity of the model. Whatever improvements are implemented, no model will be perfectly specified and therefore perfectly transferable, hence the motivation for the application of updating procedures for the model coefficients.

8.5 Procedures for updating models to improve transferability

Up until this point, transferability has been discussed in terms of using information from one particular geographic area to predict in other situations. Updating procedures attempt to use information available from the area to which the model is to be transferred in order to improve the model's predictive ability. The five main procedures investigated to date are:

(1) The 'do-nothing' case which is a default option if data is not available from the application area.

(2) Adjustment of the alternative-specific constant terms, using aggregate data of behaviour, on the grounds that there is no theoretical basis for transferring terms which account for all the dimensions not explicitly explained by the model.

(3) Enrichment by re-estimation of the coefficients with a small disaggregate sample. It is assumed that at least a small sample of observations on individual behaviour representative of the study area will be available for use in updating the original model.

(4) Re-estimation of the constant terms and estimation of a scalar to weight all other coefficients so that the ratios between them are unchanged.

(5) Bayesian updating using the original coefficients (procedure 1) and the coefficients resulting from the small disaggregate sample (procedure 3).

Updating procedures are discussed by Ben-Akiva and Atherton (1976) and Hensher and Johnson (1981). It was concluded that the Bayesian updating procedure is the most effective procedure.

8.5.1 Bayesian updating

This procedure can be used for both transferability over space and transferability over time. The Bayesian procedure combines the information contained in the original sample and the new sample, by computing the updated coefficient on any one variable as the weighted average of the coefficient of that variable as estimated in the original model, and the coefficient as estimated with the sample from the new area. The weights used are the inverses of the variances of the two coefficient estimates. Thus:

$$a_{upd} = \left[\left(\frac{a_{ce}}{s^2_{ce}}\right) + \left(\frac{a_{\ell a}}{s^2_{\ell a}}\right)\right] / \left[\left(\frac{1}{s^2_{ce}}\right) + \left(\frac{1}{s^2_{\ell a}}\right)\right]$$

where a_{ce} and $a_{\ell a}$ are the coefficients estimated on Central England sample and London area sample respectively. s_{ce} and $s_{\ell a}$ are standard deviations of the Central England and London area coefficients respectively. A_{upd} is the updated coefficient.

Table 8.5 gives the updated coefficients obtained from applying the Bayesian update procedure.

Variable	Business	Leisure	Inclusive Tours	Domestic
TT	-0.129	-0.161	-0.158	-0.255
FREQ	1.110	1.046	1.489	2.750

Table 8.5 Updated coefficients

8.5.2 Transferability test with updated coefficients

The transferability test presented in section 8.4 is again carried out. This time, it tests the differences between the updated coefficients presented in Table 8.5 and the coefficients estimated on the London area samples where the null hypothesis is that the two sets of coefficients are equal. Table 8.6 shows the results of this transferability test.

	Business	Leisure	Inclusive Tours	Domestic
Log-likelihood at convergence	-42.14	-32.46	-42.57	-22.39
Log-likelihood at convergence with coefficients restricted to the values of the updated coefficients	-44.64	-36.76	-45.41	-207.38
Test	5.00	8.60	5.68	369.98
$\chi^2(0.05,2)$	5.99	5.99	5.99	5.99

Table 8.6 Transferability test with updated coefficients.

Since the critical value of χ^2 at 95 percent confidence level is higher than the values of the test for the business and inclusive tours (5.00 and 5.68 respectively),we can therefore accept the hypothesis that the coefficients are the same and thus, we conclude that the business model and the inclusive tours model are transferable to the London area in 1978.

For the leisure passengers, the test value (8.60) is only marginally greater than the critical value (5.99) at 95 percent confidence level. In fact, at the 99 percent confidence level with a critical value of χ^2 equal to 9.21, the hypothesis that the models are not significantly different could not have been rejected.

For the domestic model, the test value is still very large and the null hypothesis of statistical transferability is therefore rejected.

The non-transferability of the leisure and domestic models could be explained for a large part by the part of the fare variable. Indeed, as the air fares from Heathrow and Gatwick were the same for the destinations considered, this variable has not been included in the London area models, considering that it was not reflecting a determinant of airport choice for this particular case. However, in Chapter 7, it was established that the air fare variable was the most important factor for the leisure and domestic passengers originating from Central England 1975.

8.6 Predictive performance of the transfer models

So far, we have shown that, if the transferability of the Central England models is to be judged on statistical criteria, two out of four models are transferable. However, in practice a level of transferability is required which is sufficient for specific planning needs. It would thus seem relevant to gauge the predictive ability of the transfer models in relation to the estimates which could be obtained from unrestricted London area models.

The prediction success table is normally a goodness of fit measure, used on the sample which provided estimation of the model. Here it can be used to test prediction because the transfer model coefficients are used to predict the airport choice of a different population.

Tables 8.7, 8.8, 8.9, and 8.10 present the prediction success table for the unrestricted model and the transferred model for the business, leisure, inclusive tours and domestic respectively.

They show that the business and inclusive tours transfer models which were found to be transferable perform very well respectively (93.3% and 93.06% successfully predicted) and predict correctly the airport's share. The leisure transfer model which was found statistically not to be transferable performs as well (91.66%) as the unrestricted model. The domestic transfer model is the least satisfactory (78.28%) and

	Unrestricted model				Transferred model			
	H	G	count	share	H	G	count	share
Heathrow	289	10	299	90.88%	294	5	299	90.88%
Gatwick	12	18	30	9.12%	17	13	30	9.12%
Predicted count	301	28	329		311	18	329	
Predicted share	91.49%	8.51%			94.53%	5.47%		
Proportion successfully predicted	96.0%	64.3%			94.53%	77.22%		
Overall proportion successfully predicted:				93.3%				93.3%

Table 8.7 Prediction success table - business

	Unrestricted model				Transferred model			
	H	G	count	share	H	G	count	share
Heathrow	122	9	131	77.98%	128	3	131	77.98%
Gatwick	5	32	37	21.02%	11	26	37	21.02%
Predicted count	127	41	168		139	29	168	
Predicted share	75.6%	24.4%			82.74%	17.26%		
Proportion successfully predicted	96.06%	78.05%			92.09%	89.66%		
Overall proportion successfully predicted:				91.67%				91.66%

Table 8.8 Prediction success table - leisure

	Unrestricted model				Transferred model			
	L	G	count	share	L	G	count	share
Luton	35	3	38	13.87%	35	3	38	13.87%
Gatwick	15	221	236	86.13%	16	220	236	86.13%
Predicted count	50	224	274		51	223	274	
Predicted share	18.25%	81.75%			18.61%	81.39%		
Proportion successfully predicted	70%	98.66%			68.62%	98.65%		
Overall proportion successfully predicted				93.43%				93.06%

<u>Table 8.9</u> Prediction success table - Inclusive tours

	Unrestricted model				Transferred model			
	H	G	count	share	H	G	count	share
Heathrow	89	4	93	53.14%	93	0	93	53.14%
Gatwick	10	72	82	46.86%	38	44	82	46.86%
Predicted count	99	76	175		132	44	175	
Predicted share	56.57%	43.43%			74.85%	25.15%		
Proportion successfully predicted	89.9%	94.73%			70.99%	100%		
Overall proportion successfully predicted:				92%				78.28%

<u>Table 8.10</u> Prediction success table - Domestic

largely underpredicts Gatwick's share. Table 8.11 gives the predicted shares by the London area model and the transfer model.

	Business		Leisure		Inclusive Tours		Domestic	
	H	G	H	G	L	G	H	G
London area model	91.49%	8.51%	75.6%	24.4%	18.25%	81.75%	56.57%	43.43%
Trans-ferred model	94.53%	5.47%	82.74%	17.26%	18.61%	81.39%	74.85%	25.15%

Table 8.11 Predicted shares - unrestricted and transfer models

8.7 Conclusion

This chapter has sought to evaluate the hypothesis of transferability of the 1975 Central England model on 1978 London Area.

On statistical criteria, the original models were found not to be transferable to the London area in 1978. However, after a Bayesian updating procedure was applied, the business and inclusive tours models were transferable. The leisure model, statistically not transferable, has a predictive ability which can give a level of transferability which could be sufficient. The domestic model was found not transferable.

These results demonstrate that the airport choice models require refinement before they can be consistently transferred in space and in time. This refinement, however, is a smaller and cheaper task than the development of new models.

CHAPTER 9

GENERAL CONCLUSIONS AND RECOMMENDATIONS

This chapter presents a summary of the approach adopted to airport choice modelling. It also summarises the main findings of this work and draws planning implications from the major points raised in the study. Finally, it suggests some directions for further research.

9.1 Summary of the adopted approach

The main purpose of this study was to research into passengers' choice of airport so that competing airports can be evaluated on a more reliable basis. This implied making contributions to

- the specification of the determinants which influence passengers' choice
- the specification of the passengers' sensitivites to changes in the values of these determinants.

The review of the literature (Chapter Two) has suggested that non-behavioural models perform poorly even when including the a priori most important factors affecting the airport choice. In response to these shortcomings, it was decided to construct a disaggregate behavioural airport choice model. Disaggregate implies that the model explains airport choice at the level of the individual decision making unit. Viewing the decision process at this level allows the analyst to appreciate better the real choices the decision-maker faces and the factors which determine these decisions. Thus, the variables that airlines and transport planners are able to control, can be incorporated and thus models responsive to policy oriented questions can be estimated. Moreover, all the problems created by the aggregation of data such as variability in characteristics aggregated in the same observation, the substantial decrease of the number of observations available for analysis, are avoided. The description "behavioural" implies that the model attempts to describe the causal relationship between socio-economic and transport system characteristics on the one hand and travel decisions on the other. This entails that the model's parameters reflect the motivations of the decision-makers in general, and thus a model calibrated for one area could be transferred to another allowing re-use without expensive re-estimation analysis, given the decision-making units do not differ in character. The various reasons that make disaggregate behavioural choice models attractive are given in Chapters 2 and 5.

The next step was to select the determinants of airport choice and specify the form of these variables to be used in the model. This task is dealt with in Chapter 3. It was felt that the most important variables which might influence passengers' choice of airport are surface access time, frequency of flight and air fare. Although there is no limit to the number of explanatory variables that could be included in the model, the efforts in this analysis were confined to these three variables which seem to be the most dominant factors in the choice decision.

The validity and reliability of the results from any research project are directly dependent on the soundness of the data collected. In this research, the basic data required were collected during two origin/destination surveys previously carried out for the Civil Aviation Authority. The frequency of flight and air fare data were drawn from the relevant ABC World Airways. The preparation of the data consisted of:

- identifying the passengers from the study area with a genuine choice of airport;
- extracting from the survey data, the observations relating to those passengers;
- defining for each passenger the attributes of both the choice made and the choices rejected.

Chapter 4 describes the data collection and preparation process. The models calibrated in this study are probabilistic in that they yield the probability that each passenger will use a given airport.

The random utility theory founded on the principle of utility maximization, provides a useful basis for deriving a plausible mathematical formulation for a probabilistic choice model. Two curve fitting procedures namely logit analysis and probit analysis that are suited to calibrate disaggregate probabilistic choice models have been outlined in Chapter 5. Their potential advantages and shortcomings have been presented in Chapter 6. After carrying out a test based on conditional choice, the multinomial logit model has been selected as the most appropriate tool for modelling the passengers' choice of airport. After calibrating the model and analyzing the results (Chapter 7), the hypothesis of transferability of this model to the London Area in 1978 has been tested (Chapter 8).

9.2 Main findings

The results show that the multinomial logit model used for the airport choice is good in terms of its explanatory ability and successful in predicting the choices actually made. The predicted share of each airport is also close to the observed share for the four categories of passengers.

The results justified the a priori stratification of passengers into four categories and confirmed some intuitive expectations such as that non-business passengers are more concerned with accessibility or less concerned with flight frequency than business passengers when choosing between airports. Elasticity analysis has also been conducted in this study. It shows that the choice is not equally responsive to changes in the determinants and thus access time, flight frequency and air fare cannot be viewed as equal determinants of airport choice. The accessibility variable is more important than the frequency of flights variable for all the passengers. The fare variable, found significant only for the leisure and domestic passengers, is the most important determinant for these two categories of passengers.

By influencing these determinants, it appears that there exists room for the transport planner to shift traffic from one airport to another to have an economically and/or environmentally efficient airport system. Air fare differentials can be very effective in inducing the leisure and domestic passengers' choice from one airport to another. Differences in flight frequency can be an effective tool for shifting the business passengers' choice of airport. Improvement in airport access can be an efficient policy tool of shifting the passengers' choice of airport regardless of their trip purpose. As this access improvement should be orientated to the entire study area, this could prove to be a difficult and costly task. As far as the three regional airports are concerned, Manchester airport has the potential attraction to develop into a "hub" airport. East Midlands airport could also attract more non-business passengers if the required policies are applied. Birmingham airport seems to be in the shadow of the London airports and must make large changes in the variables (frequency, fare, travel time) to impact the market significantly. However, the major investments in Midland motorways over the last few years means that with the analysis done on current data, Birmingham would probably be in a much better situation.

The airport choice model developed could be used in practice for marketing purposes and for planning purposes.

The elasticities of choice with respect to the significant variables have been estimated for all the airports. In Chapter 7, it was shown how the changes in traffic level, as a consequence of changes in variables, can be derived from the direct elasticities and the airport market shares. Therefore, airport managements can know in advance the effect on traffic of all the options open to them with regard to influencing the determinants of airport choice and therefore can select and implement the most appropriate policy. For example, if the options are either to:

(a) improve access by decreasing travel time by 10%

(b) increase frequency of flight by 10%

(c) decrease air fare by 10%

The management at each of the three airports should select the first option for increasing their business and inclusive tours market shares and the third option for increasing their domestic share. For the leisure passengers, the fare decrease should be selected at Birmingham and East Midlands airports while the decrease in travel time should be the chosen option at Manchester airport. It should be noted that although the airport management has little control over travel time to the airport, they could strive for regional transport improvements to ease the access to the airport.

Because the model predicts each airport share, it could be used in forecasting the redistribution of passengers among airports if, for example, a new airport is added to the system. It could also help in the determination of the optimum location of such a new airport, particularly as the access variable was found to be very important in the choice making. Indeed, as the IIA property was found valid for the specification of the representative utility of the airport choice model (test in Chapter 6), this model need not be re-estimated if a new airport is considered. The new probability of choice of a given airport could be relcalculated by simply adding the exponential of the representative utility of the new airport in the denominator of the equation of the logit model and thus the new predicted airport shares could be derived. If the new airport location has not yet been selected, the prediction of airport shares could be carried out as many times as the number of possible sites and the location which has the greatest market share, everything else being equal, could be selected. In general, inferences about air fares and frequency need to be made in the case of a new airport.

On statistical criteria, the original models were found not to be transferable to the London area in 1978.

However, after a Bayesian updating procedure was applied, the business and inclusive tours models were transferable. The leisure model was not statistically transferable but had a good predictive capability while the domestic model remained not transferable. The reason for the non-transferability of the domestic model is that the representative utility of these passengers is different between the London area and Central England. These results demonstrate that the airport choice models can be, under circumstances, transferable.

Any model developed for use in the field of air transport will contain limitations as to its use and accuracy. This is true for the airport choice model although most of the limitations are recognizable. Some of these could be eliminated or reduced by further research and these areas will be discussed in the next section. The major limitations are:

- The Central England airports survey took place in 1975, while the London area airports survey took place in 1978. An assumption had to be made to adjust the 1978 data to achieve consistency with 1975. This limitation could be eliminated by covering in future surveys a larger number of airports or by carrying out similar surveys in the same period of time.

- The difficulty of determining the part of the air fare in the total holiday cost for the inclusive tours passengers. Again an assumption on the air fare level had to be made.

In conclusion, disaggregate behavioural models of airport choice provide an important new tool for the airport planners and managers. Although, conceptually, they are more difficult than the models currently in use, their advantages in accuracy and reduced data requirements argue strongly for their adoption.

9.3 Recommendations for future research

The next step of research in this field should be in the area of data refinement and availability. It is quite possible that basic improvements of the data could lead to more consistent models with greater explanatory and predictive power. To this end, the following efforts need particular attention:

- inclusion of other variables
- refinement of the way the variables used are measured.

In order to identify other determinants of airport choice, a question about the passenger's reason for choosing a particular airport should be included in future airport surveys. The variable "recommendation by a travel agent" should be investigated for the inclusive tours passengers. Other variables, such as airline nationality and airport parking costs, could also be investigated in future airport choice models. The adopted approach to model building resulted in the exclusion of travellers' characteristics variables. Future research could include such variables as income and sex as explanatory variables or as a basis for stratification.

The pragmatic approach to the problem resulted also in the variables, particularly frequency, being specified in a simple form. The access variable could be refined by considering peak and off peak speeds. In circumstances where there is no predominant mode between the various available modes of access to the airport, an accessibility measure provided by an access modal split model could be developed. The frequency variable could be defined more precisely in terms of departure time. It was shown in Chapter 3, that the relevant variable in the airport choice model is the waiting time of the passenger which is the sum of the frequency delay and the stochastic delay. A topic for future research could compare the different specifications of the frequency variable and the consequent gain/loss in model's accuracy. It would also be interesting to collect information on the range of promotional and other discount air fares particularly as the recent development in competition in air transport has resulted in more widespread differences between airlines fares and to investigate more thoroughly this variable for the inclusive tours passengers.

Another area of research is to investigate different hypotheses about the structure of the passenger decision-making process, for example, a single simultaneous choice of airport and access mode or a nested choice structure. This investigation should be broken down by trip purpose as different classes of passengers might have different choice structures.

Appendix E has shown that when a multinomial logit model performs very well, aggregate cross elasticities are inevitably higher than aggregate direct elasticities as a result of the aggregation approach in the sample enumeration method. Future research should investigate other ways of aggregating individual elasticities.

An exercise could also be carried out with the introduction of Stansted airport as a choice alternative by considering various hypotheses about air fares and flight frequencies levels at this airport, after the recent decision to develop Stansted as London's third airport.

APPENDIX A

1975 QUESTIONNAIRE
(Arrival)

QUESTION	Variable	Column numbers	Codes/Comments
	NOP	1-5	NOP Market Research Ltd.
	AIRPORT	6	1 Aberdeen 2 Edinburgh 3 Glasgow 4 Prestwick 5 Manchester 6 Birmingham 7 East Midlands
	Contact Number	7-10	Office use only
	Flight	11	1 scheduled 2 charter 3 shuttle 4 helicopter 5 don't know
	Route	12	1 Domestic departure 2 Domestic arrival 3 International departure 4 International arrival
	Day of interview	13	1 Monday 2 Tuesday 3 Wednesday 4 Thursday 5 Friday 6 Saturday 7 Sunday

QUESTION	Variable	Column numbers	Codes/Comments
	Shift	14	1 A.M. 2 P.M. 3 Night
	Sex	15	1 Male 2 Female 3 don't know (no contact only
	Sub period	16	
	Sample weight	18-19	
	Outcome	20	1 complete interview 2 partly completed 3 refusal 4 ineligible 5 no interviewer 6 no time 7 no English 8 other
In which country are you living at present?	Country	21	1 United Kingdom 2 3 4 5 6 7 other 8 9 0 X Y
Where is your home in the U.K?	Home in UK	22-24	First number refers to planning region, second number to main area and third number to district

QUESTION	Variable	Column numbers	Codes/Comments
Have you been living in...for the last 12 months?		25	1 Yes 2 No
Which is the last country you have lived in for 12 months or more?		26	1 United Kingdom 2 3 4 5 6 7 other 8 9 0 X Y
What sort of transport do you expect to use when leaving the airport?		27	1 Air 2 Other
Where in the UK is your destination after leaving this airport?		28-30	
What are the types of transport you expect to use on your journey from this airport?	Method of surface transport	31-34	
Which station will you go to in order to catch your train?		35	Manchester only
On your present trip, have you used this airport before?		36	1 Yes 2 No
If yes, when was it?		37	1 Today 2 Yesterday 3 3-7 days ago 4 8-14 days ago 5 Over 2 weeks ago 6 Don't know

QUESTION	Variables	Column numbers	Codes/Comments
While you were away, was the car you will be leaving in, parked at the airport?		38	1 Yes 2 No
How many, if any, family, friends or colleagues will be leaving this airport today with you in the same vehicle?		39-40	
How many,if any, of these were flying with you?		41-42	
Which airport will you be flying to?		43-46	all leaving airport by air only
Which flight will this be?		47-50	all leaving airport by air only
At which airport did you join the flight you just arrived on?	Airport	51-54	
Which flight were you travelling on?	Airline	57-58	
Did you start your air journey at... or did you fly there simply to catch this plane?		59	1 started journey 2 caught plane
At which airport did you start your air journey?		60-63	

QUESTION	Variables	Column numbers	Codes/Comments
How long before you first used your air ticket did you book it?		64	1 less than a week 2 1 week - under 1 month 3 1 month - under 2months 4 2 months - under 3 months 5 3 months - under 6 months 6 6 months or over 7 don't know/other
What is the chief purpose of your present trip?	Trip purpose	65	1 Business/official 2 Armed services/on duty 3 Airlines(on duty) 4 Holiday-inclusive tour/package holiday 5 Holiday 6 Visiting friends/ relatives 7 Migration 8 Studies 9 Other
What is the main business of your firm or organisation?	Main business	66-67	Business/official only
Is your journey connected with the oil industry?		68	Scotland only
Can you indicate from this card which income group applies to you before tax and other deductions?	Personal income	69	

QUESTION	Variables	Column numbers	Codes/Comments
How many times, if ever, have you travelled on holiday by air before?		70	holiday passengers only 1 Never 2 1-4 times 3 5 or more times 4 don't know
How many children, if any, are in your family, under 16 years old?		71	
How many children, if any, are there under 6 years old?		72	
Which of these age groups do you come into?	Age	73	
Can you indicate the total income of all those in your family and living in your home before tax and deductions are made?	House-hold income	74	

APPENDIX B

1978 QUESTIONNAIRE

(Departure)

QUESTION	Variable	Column numbers	Codes/Comments
	NOP	1-5	NOP Market Research Ltd
	Contact number	6-10	
	Sample weights	11-14	
	Airport/ Terminal	15	1 Gatwick domestic 2 Gatwick international 3 Heathrow domestic 4 Heathrow Terminal 1 international 5 Heathrow Terminal 2 international 6 Heathrow Terminal 3 international 7 Luton international 8 Luton domestic
	Day of interview	16	1 Monday 2 Tuesday 3 Wednesday 4 Thursday 5 Friday 6 Saturday 7 Sunday
	Shift	17	1 Early 2 A.M. 3 P.M. 4 Night

QUESTION	Variable	Column numbers	Codes/Comments
	Flight	18	1 scheduled 2 scheduled (diversion) 3 charter 4 charter (diversion) 5 private flights 6 helicopter Y don't know
	Route	19	1 domestic departures 2 domestic arrivals 3 international departures 4 international arrivals
	Sex	20	1 male 2 female Y don't know
	Outcome of interview	21	1 complete interview 2 partial interview 3 refusal 4 ineligible 5 no time 6 no English 7 other
In which country have you lived for most of the last 12 months?	National-ity	22-23	UK = 01 foreign = 02-99
Where is your home in the UK? In which town did you spend most of your time in the UK on this trip?(if foreign)	Home in UK /area visited in UK	24-26	First number refers to planning region, second number to main areas and third number to district.

QUESTION	Variable	Column numbers	Codes/Comments
Have you arrived at this airport by air within the last 24 hours?	Type of passenger	27-28	
Was this just to change planes or did you have some other reason for coming here?	Interline/ Terminating		
Have you spent any time away from the airport?	Interline passengers - leaving the airport	29	1 spent time away from airport 2 no time spent away from airport
Where did you begin your journey in the UK to catch this plane?	Surface origin/ destination	30-32	
What method of transport did you use to arrive at this airport?	Method of surface transport	33	1 private car 2 self-drive hire car 3 taxi/minicab 4 airline coach 5 British rail coach 6 hotel coach 7 charter coach 8 London transport bus 9 other public bus 0 underground (British rail) X other Y don't know
How long did your journey from ... to the airport take?	Journey time to the airport	34-37	departing passengers only

QUESTION	Variable	Column numbers	Codes/Comments
Will you be returning to this airport on your present trip or have you completed your visit?	Starting or ending trip	38	1 starting 2 ending 3 neither
When will that be?	Length of trip	39	1 today 2 tomorrow 3 2-7 days from now 4 8-14 days from now 5 over 2 weeks Y don't know
How many friends, relatives or colleagues who are now flying with you also accompanied you to the airport?	Number of people flying with passenger	40-41	departing passengers at Heathrow and Gatwick only (Reason for choosing Luton at Luton airport)
How many people came to the airport just to see you off?	Number of people seeing passenger off	42-43	departing passengers at Heathrow and Gatwick only
Which airport did you fly from?	Interline passengers - airport at other part of journey	44-47	interliners only
Which airline did you travel with?	Interline passengers - airline on other part of journey	48-51	interliners only

QUESTION	Variable	Column numbers	Codes/Comments
Which airport are you travelling to on the flight you are joining now?	Airport of current journey	52-55	
Which airline are you travelling with?	Airline of current journey	56-59	
Are you completing your air journey at...or are you flying there solely to catch another plane?	Passengers travelling on to further airports	60	1 completing journey 2 staying over 3 flying on
At which airport will you complete your air journey?	Final airport of current journey	61-64	
Have you ever travelled by air on business?	Flown before on business	65	1 Yes 2 No
In the last 12 months how many times have you travelled by air on business from this airport?	Number of business trips from airport of interview in last 12 months	66-67	
Have you ever travelled by air on leisure?	Flown before on leisure	68	1 Yes 2 No
In the last 12 months how many times have you travelled by air on business from this airport?	Number of leisure trips from airport of interview in last 12 months	69-70	

QUESTION	Variable	Column numbers	Codes/Comments
What is the purpose of your present trip?	Trip purpose	71	1 business 2 Armed services/on duty 3 Airlines (on duty) 4 holiday inclusive tour 5 holiday 6 visiting friends/ relatives 7 migration 8 studies 9 other
What is the occupation of your head of household?	Socio-economic group	72	
What is the main business of firm or organisation?	Main business	73-74	
Can you indicate from this card which income group applies to you before tax and other deductions?	Personal income	75	business passengers only
How many children under 16 are living in your home?	Number of children under 16	76	leisure passengers only
How many of these are under 6 years old?	Number of children under 6	77	leisure passengers only
Which of these age groups do you come into?	Age	78	leisure passengers only
Can you indicate from this card the total income of all your family living in your home before tax and other deductions?	Household	79	leisure passengers only

APPENDIX C

NOTATION

The following symbols are used:

g	=	alternative
k	=	individual
A	=	set of alternatives
U_{gk}	=	utility associated with the choice of alternative g by individual k
V_{gk}	=	representative component of U_{gk}
ε_{gk}	=	random component of U_{gk}
P_{gk}	=	probability that individual k will select alternative g
L	=	likelihood function
L^*	=	log likelihood function
f_{gk}	=	dummy variable
LRT	=	likelihood ratio test
ρ^2	=	likelihood ratio index
TT	=	travel time to airport
FREQ	=	number of flights per day
FARE	=	air fare
$X_{gk\ell}$	=	ℓ^{th} explanatory variable describing airport g for individual k
$E^{P_{gk}}_{X_{gk\ell}}$	=	elasticity of the probability of choosing airport g by individual k with respect to a change in the ℓ^{th} explanatory variable which describes airport g for individual k.

APPENDIX D

AIRPORT. CHOICE. MODEL

```
c        MAXIMUM LIKELIHOOD ESTIMATION OF THE
c        MULTINOMIAL LOGIT MODEL
      open(unit=45,mode="in",form="formatted",file="business")
      open(unit=46,mode="out",file="results")
      common/prm/tol,sqtol,k,n,nn,nd,iset,itend,eps,iverge,ifirst,
                                                          iter,kk
      common/beta/b
      common/dbeta/db
      common/nbeta/bneg
      common/ndbeta/dbneg
      common/moment/xx
      common/exwhy/xy
      common/xe/exb,xexb
      common/xxe/xxexb
      common/cont/ic
      common/moshes/sxx
      common/moshet/txx,sxxl
      dimension b(3),db(3),bneg(3),dbneg(3),data(9)
      real*8 xx(6),xy(3),exb,xexb(3),xxexb(6),rexb,r,xb
      real*8 sxx(3)
      dimension txx(3),sxxl(3)
c     input data set number
      iset=45
c     convergence parameter - option 1
      tol=.005
c     convergence parameter - option 2
      sqtol=.05
c     loss of significance tolerance for matrix inversion
      eps=.0001
c     convergence option -----
c                         1  1-delb(i)/b(i)<=tol,all i
c                         2  2-sig(delb(i)/b(i))**2<=sqtol
c                         3  both options 1 and 2 are required
      iverge=1
```

```
c       1--sets b(i)=0,all i on first iteration
c       0--supply own initial values for b(i),all i
        ifirst=1
c       number of data records (observations) to be read
        nn=10000
c       number of parameters
        k=3
        n=9
c       kk=k*(k+1)/2
        kk=6
c       max.size of a logical record
c       nd=k*(nt-1)
c       nt=number of alternatives
        nd=9
c       max. number of iterations
        itend=20
        if(ifirst.eq.1) go to 1002
        read(5,1000)(b(i),i=1,k)
1000    format(8f10.2)
        do 1001 i=1,k
1001    b(i)=-b(i)
1002    continue
        call begin
1       call itrat
        call comp
        call calc($1)
        stop
        end
        subroutine begin
        common/prm/tol,sqtol,k,n,nn,nd,iset,itend,eps,iverge,ifirst,
                                                            iter,kk
        common/beta/b
        common/dbeta/db
        common/moment/xx
        common/exwhy/xy
        common/xe/exb,xexb
        common/xxe/xxexb
        common/cont/ic
```

```
      real*8 xx(1),exb,xexb(1),xxexb(1),rexb,rrexb,r,xb
      dimension b(1),db(1)
      write(46,110) nn,nd,iset,k,iverge,itend,eps,ifirst
      iter=0
      ic=0
      return
110   format('logit analysis',i5,'data records,each of length',i5,
     &'words will be read from data set',i5,'the model contains',13
     &'explanatory variables','convergenceoption',i3,'has been
                                                               chosen
     &and a maximum of ',i3,'iterations will be performed','eps
                                                               tolera
     &nce has been specified as',f7.4,'and the initialization of b i:
     &handled by option',i3)
      end
      subroutine itrat
      common/prm/tol,sqtol,k,n,nn,nd,iset,itend,eps,iverge,ifirst,
                                                              iter,kk
      common/beta/b
      common/dbeta/db
      common/moment/xx
      common/exwhy/xy
      common/xe/exb,xexb
      common/xxe/xxexb
      real*8 xx(6),xy(3),exb,xexb(6),rexb,rrexb,r,xb
      dimension b(3),db(3)
      exb=1.d0
      do 10 i=1,k
      xy(i)=0.d0
10    xexb(i)=0.d0
      do 15 i=1,kk
      xx(i)=0.d0
15    xxexb(i)=0.d0
      iter=iter+1
      if(iter.ne.1.or.ifirst.ne.1) ifirst=0
      return
      end
      subroutine comp
```

```
      common/prm/tol,sqtol,k,n,nn,nd,iset,itend,eps,iverge,ifirst,
                                                            iter,kk
      common/beta/b
      common/dbeta/db
      common/moment/xx
      common/exwhy/xy
      common/xe/exb,xexb
      common/xxe/xxexb
      common/cont/ic
      real*8 xx(6),xy(3),exb,xexb(3),xxexb(6),rexb,rrexb,r,xb,dexp
      real*8 rdata,rxexb
      real*8 likfn
      real*8 pca,poa
      real*8 zlik,tlrl
      dimension b(3),db(3),data(9)
      icase=0
      likfn=0.d0
      do 2 ii=1,nn
      read(45,*,end=12)(data(ik),ik=1,n)
          l=0
      if(ifirst,eq.1) go to 5
      do 22 jj=1,n,k
      jl=jj-1
      xb=0.d0
      do 10 i=1,k
10    xb=xb+data(jl+i)+b(i)
      if(xb.lt.-170.d0) go to 22
      if(xb.gt.170.d0) go to 13
      go to 11
13    jjj=jj
      go to 50
11    r=dexp(xb)
      exb=exb+r
      do 25i=1,k
      rdata=data(jl+i)*r
      xexb(i)=xexb(i)*rdata
      do 25 j=1,i
      l=l+1
```

```
25    xxexb(l)=xxexb(l)+rdata*data(jl+j)
22    l=0
      go to 32
5     do 42 jj=1,n,k
      icase=icase+1
      jl=jj-1
      exb=exb+1
      do 35 i=1,k
      xexb(i)=xexb(i)+data(jl+i)
      do 35 j=1,i
      l=l+1
35    xxexb(l)=xxexb(l)+data(jl+i)*data(jl+j)
42    l=0
32    rexb=1.d0/exb
      do 45 i=1,k
      rxexb=xexb(i)*rexb
      xy(i)=xy(i)+rxexb
      do 45 j=1,i
      l=l+1
45    xx(l)=xx(l)+(xxexb(l)-rxexb+xexb(j)*rexb
      go to 24
50    do 51 i=1,k
51    xy(i)=xy(i)+data(jjj+i-1)
24    likfn=likfn-dlog(exb)
      if(ic.eq.0) go to 200
      pca=1.d0/exb
      poa=1.d0-pca
      write(46,1000) ii,pca,poa
200   continue
      exb=1.d0
      do 34 i=1,k
34    xexb(i)=0.d0
      do 44 i=1,kk
44    xxexb(i)=0.d0
2     continue
12    if(ifirst.eq.1) write(46,140) icase
      if(iter.eq.1) zlik=likfn
```

```
      if(ic.eq.0) go to 300
      tlrl=-2.d0*(xlik-likfn)
      idfl=k
      write(46,1002) idfl,tlrl
300   continue
      write(46,141) likfn
      return
140   format(1x,i6,'cases were read')
141   format(//'likelihood=',f10.5)
1000  format(1h,5x,i6,16x,f7.4,14x,f7.4)
1002  format('degrees of freedom -2,times log likelihood ratio'/
     &16x,i3,20x,f14.4////)
      end
      subroutine calc(*)
      common/prm/tol,sqtol,k,n,nn,nd,iset,itend,eps,iverge,ifirst,
                                                              iter,kk
      common/beta/b
      common/dbeta/db
      common/nbeta/bneg
      common/ndbeta/dbneg
      common/moment/xx
      common/exwhy/xy
      common/xe/exb,xexb
      common/xxe/xxexb
      common/cont/ic
      common/moshes/sxx
      common/moshet/txx,sxxl
      real*8 xx(6),xy(3),exb,xexb(3),xxexb(6),rexb,rrexb,r,xb
      real*8 sxx(3)
      dimension txx(3),sxxl(3)
      dimension b(3),bneg(3),db(3),dbneg(3)
      if(ifirst.ne.1) go to 70
      do 50 i=1,k
50    b(i)=0
70    call dsinv(xx,k,eps,ier)
      if(ier.ne.0) go to 98
12    continue
      if(ic.eq.1) go to 20
      do 72 i=1,k
      db(i)=0.
```

```
      do 72 j=1,k
      call loc(i,j,ir,k,z,ms)
      db(i)=db(i)-xx(ir)*xy(j)
72    continue
      do 75 i=1,k
75    b(i)=b(i)+db(i)
      do 76 i=1,k
      bneg(i)=-b(i)
76    dbneg(i)=-db(i)
20    continue
      write(46,100) iter,(bneg(i),i=1,k)
      l=0
      do 45 i=1,k
      l=l+i
      sxx(i)=dsqrt(xx(l))
      sxxl(i)=sxx(i)
      txx(i)=bneg(i)*sxxl(i)
45    continue
      write(46,110)(sxx(i),i=1,k)
      write(46,120)(txx(i),i=1,k)
      write(46,200)(dbneg(i),i=1,k)
      write(46,307)(xy(mm),mm=1,k)
      write(46,308)(xx(mm),mm=1,kk)
      if(ic.eq.1) return
      call conv($99)
      if(iter.ge.itend) go to 97
      rewind(unit=45)
      return 1
99    write(46.101) iverge
      ic=1
      write(46.130)
      rewind(unit=45)
      return 1
98    if(ier.eq.-1)go to 96
      write(46.102) ier
      go to 12
96    write(46,103)
      return
```

```
97     write(46,104)
       return
100    format(//////////1x,'iteration number ',i4/1x,'new values of
      &b'/(1x,8(e11.5,5x)))                                   b'/
200    format(1x,'changes in b from preceeding iteration'/
      &(1x,8(e11.5,5x)))
101    format(/////1x,'convergence according to convergence option',
      &13,'has been completed')
102    format(/////1x,'warning.loss of significance.ier= ',i3,'.
      &execution continues.')
103    format(/////1x,'execution terminated.moment matrix cannot be
      &inverted or n or k has been misspecified')
104    format(/////1x,'execution terminated.maximum number of
      &iterations has been performed and the desired level of
      &convergence has not'/1x,'been achieved')
307    format(1x,'xy'/(1x,8(d11.5,5x)))
308    format(1x,'xx inverse'/(1x,8(d11.5,5x)))
110    format(1x,'sxx=dsqrt(xx inverse)'/(1x,8(d11.5,5x)))
120    format(1x,'txx=b/sxx'/(1x,8(e11.5,5x)))
130    format(1h1,'choice probabilities'//1h,'observation
      &selected alternative  other alternatives'/)
       end
       subroutine conv(*)
       common/prm/tol,sqtol,k,n,nn,nd,iset,itend,eps,iverge,ifirst,
                                                            iter,kk
       common/beta/b
       common/dbeta/db
       common/moment/xx
       common/exwhy/xy
       common/xe/exb,xexb
       common/xxe/xxexb
       real*8 xx(6),xy(3),exb,xexb(3),xxexb(6),rexb,rrexb,r,xb
       dimension b(3),db(3)
       do 10 i=1,k
       a=db(i)/b(i)
10     if(abs(a).gt.tol) return
       return 1
       end
       subroutine dsinv(xx,k,eps.ier)
       dimension xx(6)
```

```
      double precision xx,din,work
      call dmfsd(xx,k,eps,ier)
      if(ier) 9,1,1
1     ipiv=k*(k+1)/2
      ind=ipiv
      do 6 i=1,k
      din=1.d0/xx(ipiv)
      xx(ipiv)=din
      min=k
      kend=i-1
      lanf=k-kend
      if(kend) 5,5,2
2     j=ind
      do 4 ll=1,kend
      work=0.d0
      min=min-1
      lhor=ipiv
      lver=j
      do 3 l=lanf,min
      lver=lver+1
      lhor=lhor+1
3     work=work+xx(lver)*xx(lhor)
      xx(j)=-work*din
4     j=j-min
5     ipiv=ipiv-min
6     ind=ind-1
      do 8 i=1,k
      ipiv=ipiv+i
      j=ipiv
      do 8 ll=i,k
      work=0.d0
      lhor=j
      do 7 l=ll,k
      lver=lhor+ll-i
      work=work+xx(lhor)*xx(lver)
7     lhor=lhor+1
      xx(j)=work
```

```
8       j=j+11
9       return
        end
        subroutine dmfsd(xx,k,eps,ier)
        dimension xx(6)
        double precision dpiv,dsum,xx
        if(k-1) 12,1,1
1       ier=0
        kpiv=0
        do 11 ll=1,k
        kpiv=kpiv+ll
        ind=kpiv
        lend=ll-1
        tol=abs(eps*sngl(xx(kpiv)))
        do 11 i=ll,k
        dsum=0.d0
        if(lend) 2,4,2
2       do 3 l=1,lend
        lanf=kpiv-l
        lind=ind-l
3       dsum'dsum+xx(lanf)*xx(lind)
4       dsum=xx(ind)-dsum
        if(i-ll) 10,5,10
5       if(sngl(dsum)-tol) 6,6,9
6       if(dsum) 12,12,7
7       if(ier) 8,8,9
8       ier=ll-1
9       dpiv=dsqrt(dsum)
        xx(kpiv)=dpiv
        dpiv=1.d0/dpiv
        goto 11
10      xx(ind)=dsum*dpiv
11      ind=ind+i
        return
12      ier=-1
        return
        end
        subroutine loc(i,j,ir,k,z,ms)
```

```
      z=k
      mx=1
      ix=i
      jx=j
      if(mx-1) 10,20,30
10    irx=k*(jx-1)+ix
      goto 36
20    if(ix-jx) 22,24,24
22    irx=ix+(jx*jx-jx)/2
      go to 36
24    irx=jx+(ix*ix-ix)/2
      go to 36
30    irx=0
      if(ix-jx) 36,32,36
32    irx=ix
36    ir=irx
      return
      end
```

APPENDIX E

AGGREGATE DIRECT AND CROSS ELASTICITIES IN THE M.N.L.

The formula for the aggregate elasticity using the sample enumeration method is

$$E_{X_{jk\ell}}^{\bar{P}_g} = \left(\sum_{k=\ell}^{R} \hat{P}_{gk} \cdot E_{X_{jk\ell}}^{P_{gk}} \right) \Big/ \sum_{k=1}^{R} \hat{P}_{gk}$$

where $\hat{P}_{gk}$ is an estimated choice probability

$\bar{P}_g$ refers to aggregate probability of choice of alternative g

$E_{X_{jk\ell}}^{P_{gk}}$ is a joint elasticity the formula for which is:

$$E_{X_{jk\ell}}^{P_{gk}} = a_{j\ell} \cdot X_{jk\ell} \cdot (\delta_{gj} - P_{jk})$$

where δ_{gj} = 1 if g = j (direct point elasticity)

δ_{gj} = 0 if g ≠ j (cross point elasticity)

g = alternative

k = individual

$X_{jk\ell}$ = ℓ^{th} explanatory variable describing alternative j for individual k

$a_{j\ell}$ = parameter of the ℓ^{th} variable.

The relative size of the aggregate direct and cross elasticities obtained in this work needs some justifications and comments as to why the aggregate cross elasticities are systematically higher. The explanation found refers to the performance of the model. The better the model fits the data, the higher the aggregate cross elasticities compared to the aggregate direct elasticities.

Indeed, for an individual k choosing alternative g, the direct point elasticity with respect to a change in the ℓ^{th} variable will approach zero because P_{gk} will approach unity if the model performs very well. The contribution of this individual's direct elasticity in the aggregate direct elasticity will be near zero. The cross point elasticity will also be near zero (if P_{gk} near 1, then P_{jk} near 0) and thus the contribution of this individual's cross elasticity in the aggregate cross elasticity will also be near zero.

The same direct point elasticity when individual k has not chosen alternative g will be near $a_{g\ell} . X_{gk\ell}$ and the contribution of this individual's direct elasticity in the aggregate direct elasticity will be very small ($\hat{P}_{gk}$ being near 0 because the model performs very well). If this individual has chosen alternative j, the cross point elasticity for alternative g will be near $- a_{j\ell} . X_{jk\ell}$ (P_{jk} near 1) and the contribution of this individual's cross elasticity in the aggregate cross elasticity will also be near $- a_{j\ell} . X_{jk\ell}$. Similarly, if the individual has chosen an alternative other than g and j, the contribution of this individual's direct and cross elasticity (with respect to a variable associated with alternative j) to respectively the aggregate direct and cross elasticity will be near zero.

Thus, when the summation over all individuals is performed, inevitably, if the model performs very well, the aggregate cross elasticities will be higher than the aggregate direct elasticities.

In the same way,it can be shown that if the model does not perform very well, the relative size of aggregate direct and cross elasticities does not follow the above rule.

These conclusions are directly linked to the aggregation approach in the sample enumeration method where each individual elasticity is weighted by this individual's estimated probability of choice and future research should investigate other ways of elasticities aggregation.

R E F E R E N C E S

- ABC World Airways Guide (1975)

- ABC World Airways Guide (1978)

- ADLER, T.S. and BEN-AKIVA, M.E. (1976) - 'Joint choice model for frequency, destination and travel mode for shopping trip' - Transportation Research Record No.569, pp.136-150.

- AIRPORT FORUM (1983) - 'Canadian passengers filling airport in Vermont' - No.5, October 1983, p.30.

- ALBRIGHT, R., LERMAN, S. and MANSKI, C.(1977) - 'The multinomial probit estimation package : features and operational tests' - Cambridge Systematics, Inc.,Boston.

- ALTHERTON, T.J. and BEN-AKIVA, M.E. (1976) - 'Transferability and updating of disaggregate travel demand Models' - Transportation Research Record, 610, pp.12-18.

- ASHFORD, N.J. and McGINITY, N. (1975) - 'Access to airports using high speed ground modes' - High Speed Ground Transportation Journal vol.9, No.1.

- AUGUSTINUS, J.G. (1973) - 'An air passenger airport distribution model for the New-York - New Jersey Area' - G.P. Howard (ed), Airport Economic Planning, MIT Press, Cambridge, Mass.

- BENABI, B. (1983) - 'Disaggregate behavioural freight modal choice models : A case study of manufacturing industry' - Unpublished PhD thesis - Loughborough University of Technology.

- BEN-AKIVA, M.E. (1973) - Program for maximum likelihood estimation of the multinomial logit model - MIT, Cambridge, Mass.

- BEN-AKIVA,M.E. (1974) - 'Structure of passenger travel demand models' - Transportation Research Record No.526, pp.26-42.

- BEN-AKIVA, M.E. and ALTHERTON, T.J. (1977) - 'Choice model prediction of carpool demand : methods and results' - Transportation Research Record No.637, pp.13-17.

- BIG ROAD ATLAS of Great Britain (1981).

- CAVES, R.E. and KENWORTHY, P. (1983) - 'Competition between direct and indirect services for international trips of East Midlands businessmen - Unpublished paper - Loughborough University of Technology, Department of Transport Technology.

- CIVIL AVIATION AUTHORITY (1976) - Passengers at major airports in Scotland and Central England : origin and destination survey, July - November 1975' - CAP 394 - CAA, Cheltenham.

- CIVIL AVIATION AUTHORITY (1980) - 'Passengers at the London Area airports in 1978' - CAP 430 - CAA, Cheltenham.

- CIVIL AVIATION AUTHORITY (1984) - 'Research into passengers' choice of airport' - August 1984; unpublished document.

- DAGANZO, C; BOUTHELIER, F. and SHEFFI, Y. (1977) - 'Multinomial probit and qualitative choice : a computationally efficient algorithm' - Transportation Science Vol.11.

- DAGANZO, C. (1979) - 'Multinomial Probit : the Theory and its Application to Demand Forecasting' - Academic Press, New York.

- DE DONNEA, F.X. (1971) - 'The Determinants of Transport Mode Choice in Dutch Cities' - Rotterdam University Press.

- DEPARTMENT of Trade (1976) - Airport strategy for Great Britain : a consultation document, part II : the Regional Airports' - HMSO.

- DOMENCICH, T.A. and McFADDEN, D. (1975) - Urban travel demand : a behavioural analysis - North Holland Publishing Co., Amsterdam.

- FAA (1979) - 'FAA aviation forecasts' - June 1979.

- GALBRAITH, R.A. and HENSHER, D.A. (1982) - 'Intra-metropolitan transferability of mode choice models' - Journal of Transport Economics and Policy, pp. 7-29.

√ - GAUDRY, M.J.I. and DAGENAIS, M.G. (1979) - 'The dogit model' - Transportation Research, 13B. pp.105-111.

- GAUDRY, M.J.I. and WILLS, M.J. (1979) - 'Testing the dogit model with aggregate time series and cross sectional data' - Transportation Research, 13B. pp.155-166.

√ - GAUDRY, M.J.I. (1980) - 'Dogit and logit models of travel mode choice in Montreal' - Canadian Journal of Economics, 13, pp.268-279.

√ - GELERMAN, W. and de NEUFVILLE, R. (1973) - 'Planning for satellite airports' - Transportation Engineering Journal, August 1973.

- GENEST, B.A. (1970) - 'An analysis of accessibility effects on terminal location and configuration' - PhD thesis, Department of Civil Engineering, M.I.T.

√ - HANSEN, S. and ROGERS, K.E. (1979) - 'The role of dis-aggregate modelling tools in the policy area' - in HENSHER, D.A. and STOPHER, P.R. (eds.), Behavioural Travel Modelling, Croom Helm, London, pp.395-409.

- HAUSMAN, J.A. and WISE, D.A. (1978) - 'A conditional probit model for qualitative choice : discrete decisions recognizing interdependence and heterogeneous preferences' - Econometrica 46(2).

- HENSHER, D.A. and BULLOCK, R. (1979) - 'Price elasticity of commuter mode choice : effect of 20 percent rail fare reduction' - Transportation Research 13A. pp.193-202.

- HENSHER, D.A. and JOHNSON, L.W. (1981) - Applied Discrete, Choice Modelling' - John Wiley & Sons,Inc. New York.

- HENSHER, D.A. and MANEFIELD, T. (1981) - A structured logit model of automobile acquisition and type choice : some preliminary evidence - School of Economic and Financial Studies, Macquarie University, Australia.

- HOROWITZ,J. (1981) - 'Identification and diagnosis of specification errors in the multinomial logit model' - Transportation Research 15B, pp.345-360.

- LAVE, C. (1968) - 'Modal choice in urban transportation : a behavioural approach' - unpublished Ph.D. thesis - Department of Economics - Stanford University, California.

- LERMAN, S. (1975) - 'A behavioural model of urban mobility decision' - Ph.D.dissertation, M.I.T.,Massachusetts.

- LIOU, P.S. and TALVITIE, A.P. (1974) - 'Disaggregate access mode and station choice models for rail trips' - Transportation Research Record No.526. pp.42-52.

- LISCO, T. (1967) - 'The value of commuter's travel time : a study in urban transportation' - unpublished Ph.D. thesis, Department of Economics, University of Chicago.

- LUCE, R.D. and SUPPES,P. (1965) - 'Preference, utility and subjective probability' - in LUCE, R.D.; BUSH, R. and GALANTER, E.(eds). Handbook of Mathematical Psychology, Vol.3. New York, pp.249-410.

v - MANSKI, C.F. and McFADDEN,D. (1980) - Structural Analysis of Discrete Data : with Econometric Applications' - MIT Press, Cambridge, Mass.

- MANSKI, C.F. and WISE, D.A. (1983) - College Choice in America - Harvard University Press, Cambridge, Mass.

- McFADDEN, D. and REID (1974) - 'Aggregate travel demand forecasting from disaggregate behavioural models' - Transportation Research Record No.534, pp.24-37.

- McFADDEN, D. (1974) - 'Conditional logit analysis of qualitative choice behaviour' - in P. ZAREMBKA (ed): Frontiers in Econometrics.

- McFADDEN, D., TRAIN, K. and TYE, W.B. (1977) - 'An application of diagnostic tests for the I.I.A. property of the multinomial logit model' - Transportation Research Record No.637, pp.39-46.

- McFADDEN, D. (1979) - 'Quantitative methods for analysing travel behaviour of individuals : some recent developments' in HENSHER, D.A. and STOPHER, P.R. (eds.) : Behavioural Travel Modelling, Croom Helm, London, pp.279-318.

- McFADDEN, D. (1981) - 'Econometric models of probabilistic choice' - in MANSKI, C.F. and McFADDEN, D., op cit, pp.198-272.

- de NEUFVILLE, R. (1976) - Airport Systems Planning - MIT Press.

- de NEUFVILLE, R., and KING, C.R. (1979) - 'Access, Fares, Frequency : Effects on airport traffic' - Transportation Engineering Journal of ASCE, Vol.105, March 1979, pp.109-125.

- ORTUZAR, J.D. and FERNANDEZ, E. (1985) - 'On the stability of discrete choice models in different environment' - Transportation Planning and Technology, vol.10, pp.209-218.

- RASSAM, P; ELLIS, R. and BENNET, J. (1971) - 'The N-dimensional logit model : development and application' - Highway Research Record No.369, pp.135-147.

- ROUDIER, J. (1977) - 'Evaluation of demand - freight transport' - Contribution to the 7th international symposium on theory and practice in transport economics, London, 26-29 Sept.1977.

√ - RUBIN, D. and FAGAN, L.N. (1976) - 'Forecasting air passengers in a multiairport region' - Transportation Research Record No.588, pp.1-5.

- SHEFFI, Y.; HALL, R. and DAGANZO, C. (1980)- 'On the estimation of the multinomial probit model' - International conference on research and applications of disaggregate travel demand models, University of Leeds; 14-16 July.

- SOBEL, K.L. (1980) - 'Travel demand forecasting with the nested logit model' - paper presented at the Transportation Research Board 59th annual meeting, January 1980.

- SPECTOR, L.C. and MAZZEO, M. (1980) - 'Probit analysis and economic education' - Journal of Economic Education, Vol.11, Spring 1980.

- STOPHER, P.R. and LAVANDER, J.O. (1972) - 'Disaggregate behavioural demand models : empirical tests of three hypothesis' - Transportation Research Forum Proceedings, XIII(1).

- STOPHER, P.R. and MEYBURG, A.H. (1976) - Behavioural Travel Demand Models - Lexington Books, Massachusetts.

- VAN LIEROP, W. and NIJKAMP, P. (1984) - 'Perspectives of disaggregate choice models on the housing market' - in PITFIELD, D.E. (ed.), Discrete Choice Models in Regional Science.

- WATSON, P.L. (1974) - 'Choice of estimation procedure for modes of binary choice : some statistical and empirical evidence' - Regional and Urban Economics, 4, (2).

- WATSON, P.L. and WESTIN, R.B. (1975) - 'Transferability of disaggregate mode choice models' - Regional Science and Urban Economics, vol.5.

- WHITE, K.J. (1975) - 'Consumer choice and use of bank credit cards : a model and cross-section results' - Journal of Consumer Research Vol.2, June 1975.

✓ - WILKS, S.S. (1962) - Mathematical Statistics - John WILEY & Sons, Inc., New York.

✓ - WILLIAMS, H.W.C.L. (1981) - 'Random utility theory and probabilistic choice models' - in WILSON, A.G. et al. Optimization in Locational and Transport Analysis, Chichester, pp.46-84.

- WILLIAMS, H.W.C.L. and ORTUZAR, J.D. (1982) - 'Behavioural theories of dispersion and the mis-specification of travel demand models' - Transportation Research 16B. pp.167-219.

✓ - WRIGLEY, N. (1982) - 'Quantitative methods - developments in discrete choice modelling' - Progress in Human Geography, Vol. 6. pp.547-561.