

**MODELING THE TRAVEL DEMAND BEHAVIOUR FOR WORK TRIPS
IN SEBERANG PERAI SELATAN**

by

BADERUL AMIN BIN ABDUL HAMID

**Thesis submitted in fulfillment of the
requirements for the degree of
Master of Science**

March 2008

ACKNOWLEDGEMENTS

I would like to express my deep appreciation to my supervisor Associate Professor Dr Ahmad Farhan Mohd Sadullah, for his valuable advice and guidance throughout my master program and dissertation process. Sincere appreciation and gratitude are also extended to Associate Professor Dr Wan Hashim Wan Ibrahim for his interest and encouragement. In addition, I would like to thank En Kamaruddin Che Lah, Director of Engineering, Majlis Perbandaran Seberang Perai and all his staff, for their cooperation and kindness. I am very grateful to the many individuals who listened to my ideas and directed my studies, Finally, I am greatly indebted to my parents, my wife Noraishah Hashim , and my children (Muhammad Saiful, Nur Dini, Muaz and Faiz) for their love over years of learning.

TABLE OF CONTENTS

| | PAGE | |
|-----------------------|---|----|
| Acknowledgement | ii | |
| Table of Contents | iii | |
| List of Tables | vii | |
| List of Figures | xi | |
| List of Abbreviations | xii | |
| Abstrak | xiv | |
| Abstract | xv | |
| | | |
| CHAPTER 1 | INTRODUCTION | |
| | | |
| 1.1 | Introduction | 1 |
| 1.2 | Problem Statement | 4 |
| 1.3 | Objectives of the Research | 5 |
| 1.4 | Scope of Research | 6 |
| 1.5 | Approach of the Study | 6 |
| 1.6 | Thesis Organization | 7 |
| | | |
| CHAPTER 2 | LITERATURE REVIEW | |
| | | |
| 2.1 | Introduction | 9 |
| 2.2 | Mode choice models: Comparison Between Disaggregate | 10 |

| | | |
|------------------|--|----|
| | and Aggregate Models | |
| 2.3 | Disaggregate Mode Choice Models | 14 |
| | 2.3.1 Discrete Choice Model: General Assumptions | 15 |
| | 2.3.2 Random Utility Theory | 16 |
| | 2.3.3 The Logit Model | 18 |
| | 2.3.4 Disadvantages of Logit Model | 21 |
| 2.4 | Revealed Preference and Stated Preference Techniques | 22 |
| 2.5 | Summary | 27 |
| | | |
| CHAPTER 3 | METHODOLOGY | |
| | | |
| 3.1 | Introduction | 28 |
| 3.2 | Study Area Selection | 30 |
| 3.3 | Identifying Attributes for Analysis | 31 |
| 3.4 | Selection of Attributes Levels | 33 |
| 3.5 | Questionnaire Design | 33 |
| 3.6 | Survey Administration | 35 |
| 3.7 | Model Calibration and Data Analysis | 37 |
| 3.8 | Likelihood Ratio Test | 41 |
| | | |
| CHAPTER 4 | SERVICE ATTRIBUTE SELECTION | |
| | | |
| 4.1 | Introduction | 44 |
| 4.2 | Study Area | 45 |
| 4.3 | The Initial Survey Design | 47 |

| | | |
|-----|---|----|
| 4.4 | The Descriptive Analysis of the Initial Survey Data | 49 |
| 4.5 | Selecting the Travel Attributes | 57 |
| 4.6 | Conclusions | 60 |

CHAPTER 5 RESULTS AND ANALYSIS OF THE STATED PREFERENCE SURVEY

| | | |
|-------|---------------------------------|----|
| 5.1 | Introduction | 61 |
| 5.2 | The Levels of Travel Attributes | 62 |
| 5.2.1 | Level of Fare Attribute | 62 |
| 5.2.2 | Level of Service Headway | 63 |
| 5.2.3 | Level of Travel Time Attribute | 63 |
| 5.2.4 | Level of Comfort Attribute | 64 |
| 5.3 | Fractional Factorial Design | 65 |
| 5.4 | Questionnaire Design | 71 |
| 5.4.1 | Design of Choice Sets | 73 |
| 5.5 | Survey Administration | 77 |
| 5.5.1 | Selection of Respondents | 79 |
| 5.5.2 | Survey Method | 79 |
| 5.5.3 | Pilot Test | 80 |
| 5.6 | Stated Preference Data | 80 |
| 5.7 | Conclusions | 88 |

CHAPTER 6 MODEL ESTIMATION RESULTS

| | | |
|-----|--|-----|
| 6.1 | Introduction | 89 |
| 6.2 | Model Estimation Strategy | 90 |
| 6.3 | Estimation of MNL Model Specification for Simpang Ampat Area | 91 |
| 6.4 | Estimation of MNL Model Specification for Sungai Bakap Area | 109 |
| 6.5 | Estimation of MNL Model Specification for Nibong Tebal Area | 125 |
| 6.6 | Discussions of the Final MNL Model Specifications | 141 |

CHAPTER 7 CONCLUSIONS AND FUTURE RESEARCH

| | | |
|------------|--|-----|
| 7.1 | Conclusions | 149 |
| 7.2 | Future Research | 153 |
| References | List Of References | 155 |
| Appendix 1 | Initial Survey Form | 161 |
| Appendix 2 | Final Survey Form A | 164 |
| Appendix 3 | Final Survey Form B | 169 |
| Appendix 4 | Final Survey Form C | 174 |
| Appendix 5 | Model Estimation Results for Simpang Ampat | 179 |
| Appendix 6 | Model Estimation Results for Sungai Bakap | 185 |
| Appendix 7 | Model Estimation Results for Nibong Tebal | 195 |

LIST OF TABLES

| | | Page |
|------------|---|-------------|
| Table 4.1 | Age Distribution of Commuting Mode Users | 50 |
| Table 4.2 | Income Distribution of Commuting Mode Users | 51 |
| Table 4.3 | Work Place Distribution of Commuting Mode Users | 53 |
| Table 4.4 | Reasons For Using Commuting Mode | 55 |
| Table 4.5 | Public Transport Attribute Ratings | 57 |
| Table 5.1 | Travel Time of Local Bus | 64 |
| Table 5.2 | Attributes And Levels | 65 |
| Table 5.3 | Treatment Combinations | 69 |
| Table 5.4 | Treatment Combinations Having the Same Pair of Values | 70 |
| Table 5.5 | Treatment Combinations – (1,0) and (2,1) | 70 |
| Table 5.6 | Levels of Travel Attributes -Distance Less Than 10 km | 74 |
| Table 5.7 | Levels of Travel Attributes - Distance Between 10 km and 20 km | 74 |
| Table 5.8 | Levels of Travel Attributes - Distance Between 20 km and 30 km | 75 |
| Table 5.9 | Levels of Travel Attributes - Distance Between 30 km and 40 km | 75 |
| Table 5.10 | Levels of Travel Attributes - Distance Between 40 km and 50 km | 75 |
| Table 5.11 | Treatment Combinations – (1,0) and (2,1) | 76 |
| Table 5.12 | Example of Paired Choice Sets | 77 |
| Table 5.13 | Respondent’s Characteristics | 81 |

| | | |
|------------|--|-----|
| Table 5.14 | Percentage of the Chosen Mode By Scenario Set | 84 |
| Table 5.15 | Percentage Of Mode Change From Motorcycle To Bus | 86 |
| Table 5.16 | Percentage Of Mode Change From Car To Bus | 87 |
| Table 6.1 | Model SA1 for Simpang Ampat Area | 93 |
| Table 6.2 | Estimation Results for Model SA1 for Simpang Ampat Area | 95 |
| Table 6.3 | Model SA2 for Simpang Ampat Area | 97 |
| Table 6.4 | Estimation Results for Models SA1 and SA2 for Simpang Ampat Area | 98 |
| Table 6.5 | Model SA3 for Simpang Ampat Area | 100 |
| Table 6.6 | Estimation Results for Models SA1, SA2 and SA3 for Simpang Ampat Area | 101 |
| Table 6.7 | Model SA4 for Simpang Ampat Area | 102 |
| Table 6.8 | Model SA5 for Simpang Ampat Area | 102 |
| Table 6.9 | Estimation Results for Models SA1, SA2, SA3 and SA4 for Simpang Ampat Area | 104 |
| Table 6.10 | Estimation Results for Models SA4 and SA5 for Simpang Ampat Area | 106 |
| Table 6.11 | Model SA6 for Simpang Ampat Area | 107 |
| Table 6.12 | Estimation Results for Models SA4, SA5 and SA6 for Simpang Ampat Area | 108 |
| Table 6.13 | Estimation Results for Models SB1 for Sungai Bakap Area | 110 |
| Table 6.14 | Estimation Results for Models SB1 and SB2 for Sungai Bakap Area | 112 |

| | | |
|------------|---|-----|
| Table 6.15 | Estimation Results for Models SB1 and SB3 for Sungai Bakap Area | 114 |
| Table 6.16 | Estimation Results for Models SB1, SB3, SB4 and SB5 for Sungai Bakap Area | 116 |
| Table 6.17 | Estimation Results for Models SB3 and SB6 for Sungai Bakap Area | 118 |
| Table 6.18 | Estimation Results for Models SB3, SB6, SB7 and SB8 for Sungai Bakap Area | 121 |
| Table 6.19 | Estimation Results for Models SB3, SB6, SB8 and SB9 for Sungai Bakap Area | 123 |
| Table 6.20 | Estimation Results for Models SB8, SB9 and SB10 for Sungai Bakap Area | 124 |
| Table 6.21 | Estimation Results for Model NT1 for Nibong Tebal Area | 126 |
| Table 6.22 | Estimation Results for Models NT1, NT2 and NT3 for Nibong Tebal Area | 128 |
| Table 6.23 | Estimation Results for Models NT1 and NT4 for Nibong Tebal Area | 130 |
| Table 6.24 | Estimation Results for Models NT1, NT4, NT5 and NT6 for Nibong Tebal Area | 132 |
| Table 6.25 | Estimation Results for Models NT4 and NT7 for Nibong Tebal Area | 134 |
| Table 6.26 | Estimation Results for Models NT7, NT8 and NT9 for Nibong Tebal Area | 136 |

| | | |
|------------|--|-----|
| Table 6.27 | Estimation Results for Models NT7, NT8, NT9 and NT10 for Nibong Tebal Area | 138 |
| Table 6.28 | Estimation Results for Models NT10 and NT11 for Nibong Tebal Area | 140 |
| Table 6.29 | Final Model Specification | 142 |

LIST OF FIGURES

| | | Page |
|------------|-------------------------------------|-------------|
| Figure 3.1 | Flow Chart of Research Methodology | 29 |
| Figure 4.1 | Study Area | 46 |
| Figure 5.1 | Five Segments of Commuting Distance | 73 |

LIST OF ABBREVIATIONS

| | |
|--------------|--|
| ASC | Alternative specific constants |
| CB | travel cost of bus (RM) |
| CC | travel cost of car (RM) |
| CM | travel cost of motorcycle (RM) |
| COMFB | dummy variable equal to 1 for bus with air-condition and 0 otherwise |
| CV | Critical value |
| DOF | Degrees of Freedom |
| exp | exponential |
| HW | the service headway of the bus (minutes) |
| HWDT | service headway of bus over distance (minutes/km) |
| IIA | Independence of irrelevant alternatives |
| INC | monthly income of respondent (RM) |
| IVTB | in-vehicle bus travel time (minutes) |
| $LL (\beta)$ | the log-likelihood of beta |
| LPKP | Commercial Vehicle Licensing Board |
| LRT | Likelihood ratio test |
| MNL | Multinomial logit |
| N | the sample size |
| NOC | number of cars |
| NOM | number of motorcycles |
| OVTB | out-of-vehicle bus travel time (minutes) |

| | |
|---------------|---|
| $P_n(\beta)$ | the probability of an outcome occurring for an individual n is beta |
| RP | Revealed Preference |
| SP | Stated Preference |
| SPSS | Statistical Package for the Social Sciences |
| SST | Statistical Software Tools |
| TB | total travel time of bus (minutes) |
| TC | travel time of car (minutes) |
| TM | travel time of motorcycle (minutes) |
| U | utility equation |
| V | the deterministic part of the utility equation |
| ε | The random part of the utility equation |
| β | parameters to be estimated using maximum likelihood method |
| Σ | summation |

**PEMODELAN KELAKUAN PERMINTAAN PERJALANAN KE TEMPAT KERJA
DI SEBERANG PERAI SELATAN**

ABSTRAK

Objektif utama kajian ini adalah untuk membangunkan model kelakuan permintaan perjalanan berasaskan memaksimumkan utiliti, iaitu dengan menggunakan kaedah model “multinomial logit” untuk permintaan pengangkutan ke tempat kerja di Seberang Perai Selatan. Proses membangunkan pemodelan ini merangkumi tiga fasa penting: (1) menjalankan bancian awal untuk menentukan atribut yang paling penting. (2) menjalankan bancian menggunakan kaedah gabungan “revealed preference” dan “stated preference”, dan (3) menguji dan menganggarkan spesifikasi fungsi utiliti untuk model “multinomial logit” bagi perjalanan ke tempat kerja. Keputusan dari kajian awal mendapati kos perjalanan, masa diambil untuk setiap perjalanan, kekerapan perkhidmatan dan keselesaan adalah atribut yang paling penting. Atribut ini seterusnya digunakan untuk merekabentuk dan melaksanakan bancian kombinasi “revealed preference” dan “stated preference”. Di dalam kajian seterusnya, data yang diperolehi ini telah digunakan untuk membangunkan model “multinomial logit” untuk permintaan perjalanan ke tempat kerja yang merangkumi beberapa koridor dalam kawasan kajian. Untuk mendapatkan spesifikasi model yang terbaik untuk setiap koridor, beberapa spesifikasi alternatif fungsi utiliti dicuba dan diuji. Model “multinomial logit” terakhir menunjukkan kesepadanan yang memuaskan. Model-model tersebut mempamerkan bahawa atribut-atribut polisi dan atribut-atribut bukan polisi adalah penting untuk menerangkan mod pilihan pengguna bagi perjalanan ke tempat kerja.

**MODELING THE TRAVEL DEMAND BEHAVIOUR FOR WORK TRIPS IN
SEBERANG PERAI SELATAN**

ABSTRACT

The ultimate objective of this research is to develop disaggregate models based on a behavioural principal called utility maximization. The travel demand behaviour will be modeled using the multinomial logit model for work trips in Seberang Perai Selatan. The development process of the multinomial logit model for work trips are composed of three main phases of work: (1) to conduct an initial survey in order to determine the most relevant attributes to the travelers, (2) to conduct a combined revealed preference and stated preference survey and (3) to test and estimate the utility function specifications for the multinomial logit (MNL) models for work trips. The results of initial survey indicate travel cost, travel time, service headway and comfort were the most important attributes. These attributes were used to develop a stated preference survey of mode choice selection for work trips. In the next stage of research, the data acquisitions were used to build multinomial logit models of mode choice selection for work trips for different corridors of the study area. In order to determine the “best” model specification for each corridor, a series of alternative specifications of the utility functions of the MNL models were explored and tested. The final multinomial logit models showed a satisfactory goodness of fit. The models reveal that the policy attributes and non-policy attributes are important in explaining travelers’ mode choice for work trips.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Public transport is a very important part of any society in the world. It is an organised passenger service for the general public who do not own cars or who perhaps prefer to avoid driving or using other private transport means.

With the increasing public dissatisfaction about public transport especially the bus services, transport planners need to come out with appropriate strategic planning to improve public transport. In order to come out with effective action plans, a prior analysis on the mode choice needs to be made. The mode choice is a general term for the different kinds of transport alternatives that are often used to transport people. The mode choice analysis allows transport planners to forecast what mode of transport will be used. It will also indicate the level of market shares of the transport modes.

The importance of mode choice in transportation policy analysis and decision making has led to a variety of methods for predicting the effects of policy measures on travelers' mode choices.

The usual forecasting methods for mode choice are based on cross-classification data of origins and destinations from each subzone within study area. This prediction methods are well-known as aggregate mode split modelings which have been used widely for transportation demand analysis in Malaysia. It has been shown that the aggregate

modelings have the limitation of forecasting and estimating of mode choice with aggregated zonal data. Moreover these models are subject to serious biases and errors owing to their reliance on aggregate travel data rather than records of individual choice (Horowitz et al, 1986).

Disaggregate models which appeared in 1980's offered substantial advantages over the aggregate alternative (Shen, 2005). These models offer a powerful tool for helping the transport planners understand the complexities of travel behaviour.

These modern approaches to mode choice modeling are based on a behavioural principal called utility maximization. Utility maximization is used to develop a mathematical function which predicts an individual's choice based on the utility equations. These prediction methods are also well-known as disaggregate mode choice modelings or discrete choice modelings. The disaggregate mode choice modeling is a mathematical representation of traveler preferences that provides estimate of the utility or value that travelers place on the different features or benefits of a mode.

The models can be used to estimate the total number of people who may change their behaviour in response to an action. They can also be used to predict how changes we make in the transportation system will affect the individual traveler's choice. As a result, the change in both non-motorized and motorized trips can be estimated.

This is a hard task since the analyst is not able to observe the determinants of traveller choice. He is only able to attribute choice probabilities to individuals. The modeling reveals the relationship between the probability of choosing an alternative and the attributes or benefits that characterize that alternative.

In this thesis, the attempt is to analyse and identify the attributes and parameters of each alternative of mode choice for Malaysia conditions. The attributes that have been identified are likely to explain the mode choice of a traveler. By changing the numerical values of these attributes we can predict the outcomes on the mode choice of the traveler.

Travelers can be classified into two groups : choice riders and captive riders. Choice riders have two travel modes to select: public transport or private transport (car and motorcycle) whereas captive riders have only one travel mode option to select i.e. public transport.

The aim of this study is to derive a travel demand model which is disaggregated at the individual level. This model is considered to be relatively new in Malaysia where aggregate modeling is often being used. Seberang Perai Selatan is selected as a case study for this modeling purpose.

In Seberang Perai Selatan, public transport should play an important role in the society because of the relatively low standard of living of the population in the area.

The public transport services in this area are poorly served and some of the rural state roads are not served by the system. This typical situation is due to the low level of patronage and the operators are not interested to provide the services due to expected poor ridership and high costs of operation. As a result it has created difficulty to the public and forced the public to depend on private modes of transports such as cars and motorcycles.

1.2 Problem Statement

Seberang Perai Selatan is currently facing rapid economic growth and urbanization. With this, the demands for both private and public transport have been increasing. The state government of Penang has planned to expand bus services for this district to encourage bus transport usage. However, the anticipated demand for bus transport for these areas has never been estimated quantitatively since these areas are presently not served by the bus transport. It is a known fact that the economics of bus service does not permit frequent services if there is insufficient demand. Without understanding the travelers' needs and preferences and to just rely on the considerations of the supply side, the planning of bus services in these areas will be problematic. So far, the travelers' needs and preferences of the case study corridors have not been well understood.

In view of the current need to understand and evaluate the expansion strategies that will meet the needs of the travelers, it is necessary to estimate the travel demand for bus service relatively to the existing travel mode choice. It is a challenging task as there is not much information available regarding the attributes and determinants of travelers' choice behaviour.

These issues are the motivation to study the travel demand for these affected areas. However, this study focuses on home-to-work trip. It enables us to understand travel demand behaviour to work trips and constraints that travelers face.

The study attempts to estimate the demand for bus transport for work trips if the bus transport is introduced in selected areas of Seberang Perai Selatan. The study is also a test of the ability of the hypothetical bus mode of transportation to attract work trip makers

from existing private modes of transportation since work trips contribute to congestion problems at peak hours. In addition, the study was made to identify the aspects of public and private transportation which are truly relevant determinants of the choice of transport mode for work trips.

A detailed study need to be undertaken to answer the following questions

- a. What is the demand for public transport in this area ?
- b. How will policy measures affect the demand for public transport ?

1.3 Objectives Of The Research

The aim of this research is to answer some of the questions raised above. In order to reach to concluding answers, the following objectives of the research need to be achieved.

- a. To provide a quantitative explanation of the choices of travel mode for work trip in Seberang Perai Selatan.
- b. To develop a model of travel mode choice for work trip in Seberang Perai Selatan.
- c. To understand the public attitude towards public transportation.

1.4 Scope Of Research

The scope of this research will be limited to the areas of Seberang Perai Selatan which presently do not have bus services and they are as follows:-

- a. From Simpang Ampat to Junjung.
- b. From Sg. Bakap to Sg. Duri.
- c. From Nibong Tebal to Sungai Acheh, Sungai Chenaam and Sungai Bakau.

The study was done based on revealed and stated preference survey data collected on a sample of individual travellers from these areas.

1.5 Approach of The Study

This study comprises of three main phases of work. The first phase relates to the process of selection of the service travel attributes of the hypothetical bus mode. It involved designing initial survey form, administering the survey, and the analysis of the survey data. The process is important to determine the attributes which were most relevant to the travelers in the study area. The resulting attributes would be included in the stated preference survey.

The second phase of the work involved the process of conducting a stated preference survey from start to finish, including selection of level attributes, the design and implementation of the survey instrument, the collection and descriptive analysis of the data.

The stated preference survey was used to collect disaggregate data based on travelers responses towards change in mode choice selection for work trips. The data acquisition would be used in the final phase to build multinomial logit models of mode choice selection for work trips.

Finally, the third phase involved the testings and the estimations of the utility function specifications for the multinomial logit (MNL) models for work trips in Seberang Perai Selatan. The MNL models were developed for different corridors of the study area. A series of alternative specifications of the utility functions were explored and tested in order to determine the “best” model specification for each corridor of the study area.

1.6 Thesis Organization

This thesis is organized into 7 chapters. The thesis first presents the introduction on the background, problem definition, objective and approach of the study. Chapter 2 reviews briefly the literature related to discrete choice behavior in different field of research including the current literature on transportation planning models, and aggregate and disaggregate mode choice models.

Chapter 3 describes the methodology and approach for the analysis and evaluation of the results. It also describes the explanation of the theoretical foundation of the proposed mode choice methodology.

Chapter 4 describes the first phase survey, which focused on the selection of the relevant attributes. The selection of the relevant attributes included designing the initial survey

form, administering the survey, and the descriptive analysis of the survey data. The findings from this chapter are used to develop the stated questionnaire in Chapter 5.

The following Chapter 5 presents the second phase survey and deals with the survey method employed and data collection for the development of multinomial logit model for work trips. It describes the stated questionnaire used in behavioural analysis of commuting mode choice for work trips and method of analysis as well as the results of the stated preference survey along the transport corridors of the study area.

Chapter 6 describes the development of multinomial logit model for work trips. It begins with the estimation procedure, the development, specification, estimation and interpretation of discrete choice model of multinomial logit type. It describes the estimation results of travel demand behaviour model for work trips.

Finally, Chapter 7 concludes this study with main findings from these behavioural experiments and how the objectives of the research have been addressed. This chapter also provides some thoughts for future research in this area.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Transportation modeling plays an important role in supporting transportation planning. One of the major roles of transportation modeling is to forecast travel demand based on changes in the transportation system. There are many different types of models that have been developed to simulate actual travel patterns of people and existing demand conditions. The models are used to predict changes in travel and utilization of the transportation system in response to changes in land-use, demographics and socio-economic conditions.

One of the important aspects of transportation modeling is to predict the travel choice behavior. The travel choice behaviour is also referred to as traveler mode choice, which is the most frequently modelled travel decision. It involves a specific aspect of human behavior dedicated to choice decisions. With a model, as simplified representation of a part of reality provides a better understanding and interpreting of these complex systems.

Traditionally aggregate models are used in dealing with the travel choice behavior of individual travelers. However, the aggregate models have the limitation of forecasting and estimating of travel choice with aggregated zonal data.

Disaggregate behavioral demand models which became increasingly popular during the 1980's offer substantial advantages over the aggregate counterparts (see Ben-Akiva and

Lerman, 1985 and Shen, 2005). Many models and methodologies have been developed since then.

Disaggregate behavioral demand models are based on the observed choices behavior of individual travelers. These models consider that the demand is the result of several decisions of each individual traveler. A discrete choice analysis is the methodology used to analyze and predict the travel decisions. The discrete choice model is a mathematical function which estimates the probability of selecting individual travel choice based on the utility maximization principle or relative attractiveness of competing alternatives. The logit function is one common mathematical form used in discrete choice modeling.

Revealed preference and stated preference survey data which contains data sets of individual decisions, characteristics of the individual and alternative choices for the trip is used to develop the discrete choice models.

2.2 Mode Choice Models: Comparison between disaggregate and aggregate models

An overview of publications related to disaggregate and aggregate models was conducted. Mode choice models have received a lot of attention by the transportation research community and have been studied extensively. It has been shown that the aggregate modelings have the limitation of forecasting and estimating of mode choice with aggregated zonal data. Aggregate models also have limitations in treating policy questions. Moreover these models are subject to serious biases and errors owing to their reliance on aggregate travel data rather than records of individual choice (Horowitz *et al*, 1986). Some of the main factors which contributed to the above result include the fact that the

disaggregate approach rectifies several deficiencies associated with the aggregate approach, and that it is more behaviourally sound. These factors are explained in more detail below.

- Prediction ability

Empirical evidence indicates that disaggregate mode choice models have the ability to simulate the phenomenon under study and to produce reliable forecasts for future planning purposes.

Research by Watson (1973, 1974) showed that the disaggregate mode choice model provides a better statistical explanation of mode choice behavior than its aggregate counterpart. He found that prediction errors associated with the aggregate approach are over 12 times larger than those associated with the disaggregate approach. His conclusions were further supported by Parody (1977) whose results showed that disaggregate models can be used successfully to predict future travel behaviour.

- The behaviour aspect

As pointed out by Richards (1982), the behavioural aspect of a model is a required property which reflects the analysts' understanding of the mechanisms influencing travel behavior and which depends on the model's ability to capture casual relationships. Daly (1982) indicated in his study that the disaggregate statistical analysis is the most useful technique in the quantitative analysis of travel behaviour. In this technique, theory of behaviour is assumed *a priori* and then empirically tested and refined. Disaggregate models are based on the observable behaviour of individual decision-makers. Therefore, statistically, they are efficient in extracting the maximum

possible information through the averaging procedure. Moreover disaggregate models are much more efficient to simulate actual travel patterns of people through causal factors and interrelationships of travelers and system attributes.

- Model Calibration

Berkowitz *et al* (1985) demonstrated that disaggregate parameter estimates in model calibration process are more reliable than the aggregate estimates derived from the same data. Parameter estimation bias can lead to serious planning deficiencies and limitations when misleading information or erroneous parameter signs are obtained. A subsequent impact of the latter is possible omission of important variables from the analysis.

- The transferability

Transferability of mode choice models between regions and over time plays an important part in travel demand analysis. Based on the assumption that people with similar characteristics and under similar circumstances behave similarly, disaggregate models have potential for transferability. In a study by Koppelman and Rose (1983), four models of different specifications were transferred between geographical sectors. Updating the disaggregate models with aggregate data consistently gave transferability improvements close to those achieved when these models were updated with disaggregate data. On the other hand, aggregate models are highly dependent on the zoning structure. Therefore, they are not easily transferable, geographically, and may not even apply to the same area if changes in the zoning system occur. It must be borne in mind, however, that this transferability property of disaggregate models can only be

attained if they are properly selected and calibrated, and that, unfortunately, proper validation of the choice probability function is a very difficult task requiring enormous data-gathering efforts. (Daganzo, 1979)

- Data requirements

Disaggregate data contain more information than aggregate data and are therefore more valuable for calibration purposes. Moreover disaggregate models can be developed by using data less than one tenth of the data required by aggregate models (Horowitz *et al*, 1986). Some transportation studies have indicated that much of the dispersion in some attributes, such as income and age, is lost when one used aggregate data (McCarthy, 1969). Since the choice probability function of a disaggregate demand model is perfectly suited for calibration with high-quality disaggregate data, it is possible to use just a few hundred data points in order to obtain accurate estimates of parameters. The efficiency with the information in the data is handled with disaggregate demand models is perhaps their most desirable feature, since it greatly simplifies the data-collection process. Berkowitz *et al*. (1985) demonstrated that a disaggregate model performs as well as an aggregate model (comparable goodness of fit results) when employing as little as 13.3% of the data used to develop the latter.

Contrary to this quantitative efficiency, only high quality data are adequate for disaggregate analysis. The lack of such data to date has been the major obstacle to employ disaggregate models. However, the improvements in disaggregate model transferability as noted in Koppelman and Rose (1983), and other ongoing research on improving the method of modeling human choice behavior, help to reduce the problem of insufficient data.

Miller (1985) outlines the important features of an ideal database for disaggregate transportation analysis. Three key attributes are presented: first, comprehensiveness in covering all travel modes, purposes, duration, and characteristics of the trip and trip-maker; second, consistency in the data collection methodology and timing and in the definition of terms and variables throughout all phases of the travel demand analysis; and third, spatial disaggregation by designing survey forms which provide information about trips and trip-makers at the individual level and by constantly updating these surveys. Data availability does not make aggregate models any better than the disaggregate ones. It only makes it easier to continue employing them in practice. Therefore, there is an urgent need to rectify this data problem and further improve model transferability in order to benefit from the disaggregate approach and achieve the best possible results by integrating it in the complete demand forecasting process.

2.3 Disaggregate Mode Choice Models

Disaggregate mode choice modeling is based on behavioural theory which is the result of choices made by individuals. It represents the choice behaviours of an individual. It is to predict a decision made by an individual (choice of mode, choice of route etc.) as a function of any number of variables, including factors that describe a motorcycle or pedestrian facility improvement or any policy change.

2.3.1 Discrete Choice Model: General Assumptions

The framework for a discrete choice model can be presented by a set of general assumptions as follows (Ben-Akiva and Bierlaire,1999):

- decision-maker

In discrete choice models, decision-maker is assumed to be an individual. The individual decision-making entity will depend on the particular application. For example a group of people, a household or an organization can be considered as the decision-maker. Because of its disaggregate nature, a discrete choice model has to include the characteristics or attributes of the decision-maker such as the socio economic variables such as age, gender, education and income.

- Alternatives

This assumption determines what is the possible options for a decision-maker. In this context, option is referred to as a set of alternatives available for decision-maker to choose. The set of considered alternatives is referred to as the choice set.

- Attributes

This assumption identifies the attributes of each potential alternative that the decision-maker is taking into account to make his decision.

- Decision rule

The decision rule is the process used by the decision-maker to evaluate the attributes of alternatives in the choice set and determine a choice.

Most models used for travel demand forecasting are based on utility maximization theory, which assumes that the decision-maker's preference for alternative is captured by a value, called utility, and decision-maker selects the alternative in the choice set with the highest utility. The concept of utility associated with the alternatives plays an important role in the context of discrete choice model.

2.3.2 Random Utility Theory

According to the utility maximization theory, there is a mathematical function U , called a utility function, whose numerical value depends on attributes of the available options and the individual. The utility function has the property that its value for one option exceeds the value for another if and only if the individual prefers the first option to the second. Thus, the ranking of the available options according to the individual's preferences and the ranking according to the values of the utility function are the one with the highest utility-function value (Horowitz *et al*, 1986 and McFadden, D.,2000).

However this theory yields a simple model of decision rule that makes deterministic prediction of travel choices but does not treat the variations in travel behavior. In other words this theory does not take into account the uncertainty in the predicted choices. The complexity of human behavior suggests that the decision

rule should include a probabilistic dimension in order to tackle of the issue of uncertainty.

Random utility theory which is the most common theoretical basis of discrete choice model is used to capture the uncertainty (Ben-Akiva and Bierlaire ,1999). The resulting models are called “random utility models” or “probabilistic choice models” because they describe preferences and choice of individual in terms of probabilities. Instead of predicting that an individual will choose a particular mode with certainty, these models give probabilities that each available modes will be chosen.

The utility, U is modeled as a random variable in order to reflect the uncertainty. That is

$$U = V + \varepsilon \quad (2.1)$$

where V is the deterministic part of the utility and ε is the random part of the utility, capturing the uncertainty.

The deterministic part of the utility is commonly assumed to be linear in the parameters. The deterministic part can be written that individual n is associating with alternative i as

$$V_{in} = V_{in} (X_{in}) \quad (2.2)$$

where X_{in} is a vector containing all attributes, both of individual n and alternative i . The function defined in (2.2) is commonly assumed to be linear in the parameters and is denoted as follows

$$V_{in} = \sum \beta_k X_{ink} \quad k = 1, 2, \dots, K \quad (2.3)$$

where parameters β are to be estimated and K is the total number of parameters.

Parameters β are generally estimated using maximum likelihood method. This method is widely used for statistical estimation since its concept is easy to understand and its procedure is straightforward. It can be stated that a maximum likelihood estimator is the value of parameters for which the observed sample is most likely to have occurred. Horowitz *et al* (1986) stated that the maximum likelihood method yields estimates of the coefficients and predictions of choice probabilities that have the greatest possible accuracy.

2.3.3 The Logit Model

Many different probabilistic choice models can be derived by making different assumptions about the distribution of the random part of the utility which is also referred to as the stochastic part of utility. For example, bivariate normal distribution yields the binary probit model, which has its multivariate generalization in Multinomial Probit discrete choice mode; a Gumbel distribution gives rise to the conditional or Multinomial Logit model; and a Generalized Extreme Value distribution gives rise to models such as the Nested Multinomial

Logit and the Ordered Generalized Extreme Value (see Adamowicz *et al*, 1998 and Bhat, 2003).

From a variety of probabilistic choice models, the logit-based models are by far the most popular among researchers and have been widely used for travel demand analysis (Ben-Akiva and Bierlaire,1999). Some of the most used logit models are the Multinomial Logit model, the Nested Logit Model, the Cross-Nested Logit Model and the Generalized Extreme Value Model.

In this thesis, the Multinomial Logit Model (MNL) has been used to model the travel demand. MNL is the most commonly applied choice model due to its computational advantages and convenient properties, the availability of software to estimate such model, and the large number of areas which have models that could be transferred to other areas. Its appeal basically evolves from its relative amenability to computation due to its independence of irrelevant alternatives (IIA) property. In spite of its computational advantage, the IIA property is also known to be a weakness of the logit model. The IIA property is a limitation for some practical applications.

As mentioned earlier the logit models are derived from the assumption that the random part of utility functions are independent and identically Gumbel distributed. This assumption gives a great advantage in calculating the choice probability functions. It makes the calculation very simple and easy.

The logit models were first introduced in the context of binary choice models, where the logistic distribution is used to derive the probability which means that

the relative probability of choice between two alternatives depends only on their measured attractiveness and is independent of all other available alternatives (Horowitz *et al*, 1986).

Binary or binomial logit model is about the choices between only two alternatives. In such a model, the probability that alternative 1 is chosen when the choice set consists of alternatives 1 and 2 is given by the following formula:

$$P(1) = \frac{\exp(V_1)}{\exp(V_1) + \exp(V_2)} \quad (2.4)$$

$$P(2) = \frac{\exp(V_2)}{\exp(V_1) + \exp(V_2)} \quad (2.5)$$

Equation 2.4 is simplified by dividing the numerator and denominator with $\exp(V_1)$ to obtain Equation 2.6 as follows

$$P(1) = \frac{1}{1 + \exp \left[-(V_1 - V_2) \right]} \quad (2.6)$$

It shows that $P(1)$ depends only on the difference between V_1 and V_2 . $P(1)$ is affected equally by increases in the value of V_1 and decreases in the value of V_2 .

The binomial logit model can easily be extended to accommodate choices among more than two alternatives. In the extended model, the probability that alternative i is chosen is

$$\Pr(i) = \frac{\exp(V_i)}{\sum_k \exp(V_k)}, k = 1, \dots, K \quad (2.7)$$

where i is a chosen alternative

This generalization to more than two alternatives is referred to as multinomial logit model (Bierlaire, 1997).

2.3.4 Disadvantages Of Logit Model

Consider the logit model and look at the relative odds of choosing one alternative over another, say i over j :

$$\frac{P(i) = \exp(V_i)}{P(j) = \exp(V_j)} \quad (2.8)$$

This indicates that the relative odds between any two alternatives are only a function of the attributes of these two and are independent of any other alternatives that may be available. This property of choice models is referred to as the *independence of irrelevant alternatives* and is considered a weakness of models that have it, such as the logit model. For example, if one is dealing with urban mode choices, then this property might imply that the relative odds of taking an automobile over taking a bus is independent of whether there is train service to the same destination. This is not likely since the presence of a train as a third alternative is likely to affect the probability of choosing the bus more so than that of choosing the automobile and hence likely to change their relative odds (Horowitz *et al*, 1986).

2.4 Revealed Preference (RP) and Stated Preference (SP) Techniques

Many researchers have attempted to model travel demand and travelers' behavior using revealed preference and/or stated preference survey data. These two techniques are used as complementary tools to elicit the preferences of the decision-makers (see Adamowicz *et al* 1998, Bhat 2003, Phanikumar *et al* 2004 and Mansyur and Kabit 2004). The revealed preference and stated preference techniques conveniently provide data for the development of disaggregate travel forecasting models.

Revealed preference technique, which is based on the actual choice of the individual, enables researchers to employ true observed behaviour into their studies. However, it seems difficult to observe or measure concepts such as alternatives of mode, which are not available at the moment of decision.

With theoretical advances in econometrics, a stated preference technique, which uses the decision-makers' statements about their preferences in a set of transport options, is accepted for the purpose of transportation behavior studies. The transport options are typically descriptions of transport situations constructed by the researcher (Shen, 2005). In other words, the stated preference technique requires purpose-designed survey questions for collection of data. This technique becomes an attractive option in transportation modeling since it presents the decision-maker's choice and behavioural pattern under different hypothetical scenarios. Many researchers use this technique to understand the unpredictable behavior of decision-makers under conditions that are new or hypothetical.

Phani Kumar *et al* (2004) have found that the stated preference data are effective for developing a utility model comprising of different attributes of travel, even in a nonurban

scenario. Multinomial logit model has been used in this study to develop the utility equations. The research has identified that the values of in vehicle travel time, service headway and comfort level are significant and should be considered for formulating improvement proposals for rural public transportation systems.

A limited amount of work has been undertaken in Malaysia to study and understand travel demand behavior. Most recent study in Malaysia is the Kuching Public Transportation System. In this study, Mansyur and Kabit (2004) used stated preference technique to evaluate the travel demand and travelers' behavior. Multinomial logit model was used to examine the relationship between independent variables (travel time, waiting time, fares and comfort) and dependent variables (choice of respondent whether to use old bus or choose new bus services). They found that for trip purpose, fares and comfort were the primary factors that reflected the choice of the decision-makers.

The stated preference technique has become a convenient tool for researchers to analyze data especially for the non-existing scenarios. Ortuzar and Willumsen (2001) have summarized the main features of stated preference technique as follows:

- i. It is based on the elicitation of respondents' statements of how they would respond to different hypothetical (travel) alternatives;
- ii. Each option is represented as a 'package' of different attributes like travel time, price, headways, reliability and so on;
- iii. The researcher constructs these hypothetical alternatives so that individual effect of each attribute can be estimated; this is achieved using experimental design techniques

that ensure the variations in the attributes in each package are statistically independent from one another;

- iv. The researcher has to make sure that interviewees are given hypothetical alternatives they can understand, appear plausible and realistic, and relate to their current level of experience;
- v. The respondents state their preferences towards each option by either ranking them in order of attractiveness, rating them on a scale indicating strength of preference or simply choosing the most preferred option from a pair or group of them;
- vi. The responses given by individuals are analyzed to provide quantitative measures of the relative importance of each attribute.

In study related to container port competition, Gonzalez and Gualda (2000) have summarized a series of steps to be followed:

- i Statement of the mathematical form of the "Utility Function" of each of the alternatives to be investigated;
- ii Statement of the variables of interest (attributes) and of the variations of its values (levels);
- iii Statement of the structure of the experiment (if a "choice", "ranking" or "rating" experiment);