Intelligent simulation system for studying impacts of multimodal traveller information on commuters' travel behaviour

Abdul Ahad Memon

2009

Abdul A. M. (2009). Intelligent simulation system for studying impacts of multimodal traveller information on commuters' travel behaviour. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/46719>

<https://doi.org/10.32657/10356/46719>

Nanyang Technological University

Downloaded on 26 Apr 2023 12:25:16 SGT

**AN INTELLIGENT SIMULATION SYSTEM FOR
STUDYING IMPACTS OF MULTIMODAL TRAVELLER
INFORMATION ON COMMUTERS' TRAVEL
BEHAVIOUR**

ABDUL AHAD MEMON

School of Civil and Environmental Engineering

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Doctor of Philosophy

2009

ACKNOWLEDGEMENTS

To begin with, I would like to express my deep gratitude to my supervisors, Dr. Lam Soi Hoi and Dr. Wong Yiik Diew, for all the encouragement, patience, wisdom, guidance and support throughout my PhD study and development of this thesis. For all of the help and support they gave me beyond my PhD study at Nanyang Technological University (NTU), I am also truly grateful.

I also want to express my gratitude to other Academic Staff for their continuous constructive and useful comments and suggestions on my study; to the technical executives in Transportation Laboratory, for their continuous support to build up a good environment for our studies; and to the fellow research students, the discussions with them were very interesting and helpful.

The financial support of the NTU and technical support from Singapore Land Transport Authority (LTA) are deeply appreciated.

I would like to express my heartfelt gratitude to my family, especially my parents who have given me endless love, support and confidence not only throughout my graduate work, but during my entire lifetime. I would also like to thank my wife and my Kids for their patience and understanding during the entire course of study. Finally, I would like to kneel and thank Allah (S.B.T.) for His Kindness and Mercy on me and my family for the span of life here and hereafter.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS..... ii

TABLE OF CONTENTS..... iii

LIST OF TABLES..... xi

LIST OF FIGURES xiv

LIST OF ABBREVIATIONS..... xviii

LIST OF SYMBOLS xxi

SUMMARY xxiii

CHAPTER 1 26

 INTRODUCTION 26

 1.1 Background..... 26

 1.2 Problem Statement..... 28

 1.3 Study Approach and Scope of Work 32

 1.4 Objectives of the Study 35

 1.5 Structure of the Thesis..... 37

CHAPTER 2 39

 LITERATURE REVIEW 39

 2.1 Introduction 39

 2.2 Traveller Information Systems 39

 2.2.1 Unimodal Traveller Information 41

 2.2.2 Multimodal Traveller Information 43

 2.2.3 Integrated Multimodal Traveller Information 44

 2.3 Impact of Pre-Trip Information on Travel Behaviour 45

 2.4 Travel Behaviour 47

2.4.1	Travel Behavioural Models.....	48
2.4.2	Travel Behaviour Survey	53
2.5	Travel Behaviour Measurements in the Information Rich Environment...	56
2.6	Traffic/Transportation Simulation Tools	58
2.6.1	INTEGRATION	59
2.6.2	DYNASMART	59
2.6.3	AIMSUN2.....	60
2.6.4	MITSIMLab.....	61
2.6.5	VISSIM.....	62
2.6.6	PARAMICS	63
2.7	Agent Based Simulation Model	66
2.8	Conclusion.....	70
CHAPTER 3		72
METHODOLOGY		72
3.1	Introduction	72
3.2	Travel Behaviour Survey	73
3.2.1	Study Area	75
3.2.2	Data Overview	76
3.2.3	Survey Instrument.....	79
3.2.4	Sampling Procedure	80
3.3	Travel Behaviour Modelling	81
3.3.1	Mode Choice Logit Model.....	81
3.3.2	Ordered Probit Model	85
3.3.3	Bounded Rationality Model.....	87
3.3.4	Statistical Characteristics of Mode Choice Model.....	94

3.3.5	Sensitivity Analysis.....	97
3.4	Rule-based Expert System.....	101
3.4.1	Architecture and Working Mechanism	103
3.4.2	General Decision Framework	104
3.5	Rule Flow Structure.....	107
3.5.1	Execution of Rule Flow (Macro Level)	109
3.5.2	Execution of Rule Flow (Micro Level).....	110
3.6	The Mode Choice Model.....	111
3.6.1	Pure Rule-Based Model	111
3.6.2	Discrete Choice Model.....	113
3.6.3	Probabilistic Model	114
3.7	Intelligent Transportation Network Simulation Model	116
3.7.1	Working Mechanism of INSIM	116
3.7.2	INSIM Commuter	119
3.7.3	API Functionalities.....	120
3.7.4	Data Transfer Interface.....	122
3.7.5	Transportation Network Simulation Model	123
3.7.6	Calibration of the INSIM Network	124
3.8	Traveller Information Systems	127
3.8.1	Experimental Factors	128
3.8.2	Measures of Effectiveness.....	130
3.8.3	Experimental Setup	131
3.9	Summary.....	133
CHAPTER 4		134
TRAVEL BEHAVIOUR SURVEY		134
4.1	Background.....	134

4.2	Travel Behaviour Framework	135
4.3	Travel Behaviour Survey Of Commuters.....	136
4.3.1	Survey Technique	136
4.3.2	Instrument Design.....	136
4.3.3	Pilot Survey.....	137
4.3.4	Survey Administration	137
4.3.5	Data Requirements.....	139
4.3.6	RP Section.....	140
4.3.7	SP Section	141
4.3.8	The Survey	149
4.4	Empirical Findings of Travel Behaviour Survey	151
4.4.1	Socio-Economic and Travel Attributes.....	152
4.4.2	Traveller Information Sources	155
4.4.3	Respondent's Desire to Access Traveller Information	155
4.4.4	Influence of Traveller Information on Respondents' Travel Plan	158
4.4.5	Pre-Trip and En-Route Traveller Information	160
4.4.6	Respondents' Mode Switching Behaviour.....	162
4.5	Summary	170
CHAPTER 5		172
TRAVEL BEHAVIOUR MODELLING AND ANALYSIS.....		172
5.1	Introduction	172
5.2	Impact of Traveller Information on Commuters' Travel Behaviour....	173
5.2.1	Commuters' Desire to Access Traveller Information.....	173
5.2.2	Commuters' Desire to Change Travel Plans.....	176
5.2.3	Impact of Multimodal Traveller Information on Commuters' Mode Choice Behaviour	178

5.3	Commuters' Mode Switching Propensity	181
5.3.1	Marginal Effect of the Estimated Probit Variables	184
5.3.2	Predicted Probabilities based on the Explanatory Variables.....	185
5.3.3	Marginal Effect on the Probability of Commuter's Mode Switching Decision	189
5.4	Impact of Integrated Traveller Information on Commuters' Mode Choice Behaviour	192
5.4.1	Marginal Effect of the Estimated Logit Variables	196
5.4.2	Predicted Probabilities based on the Explanatory Variables.....	196
5.4.3	Marginal Effect on the Probability of Commuters' Mode Switching Propensity	197
5.5	Discussion.....	199
CHAPTER 6		201
DEVELOPMENT OF RULE-BASED MODE CHOICE MODEL – INSIM EXPERT SYSTEM.....		201
6.1	Introduction	201
6.2	The Architecture and Working Mechanism	202
6.2.1	The INSIM Expert (IE)	204
6.2.2	The INSIM Commuter (IC)	205
6.2.3	Microscopic Traffic Simulator	208
6.3	Characteristics of the IES	209
6.3.1	Reliability	209
6.3.2	Adaptive Learning.....	213
6.4	Reliability Analysis	215
6.4.1	Evaluation Criterion	215
6.4.2	Experimental Design.....	216
6.4.3	Results and Discussion.....	216

6.5	Summary	220
CHAPTER 7		223
INTELLIGENT NETWORK SIMULATION MODEL		223
7.1	Introduction	223
7.2	Study Network.....	224
7.2.1	Data Collection and Network Coding.....	224
7.3	Calibration of the INSIM Network	228
7.3.1	Study Site	229
7.3.2	Basic Input Data.....	229
7.3.3	Required Data for Calibration.....	232
7.3.4	OD Matrix	235
7.3.5	Preliminary Testing.....	237
7.3.6	Trip Release Profile	240
7.3.7	Calibration of Route Choice Model	242
7.3.8	Secondary Testing.....	242
7.3.9	Calibration of Driving Behaviour Model.....	243
7.3.10	Calibration of Public Mode of Transport	246
7.4	Discussion	248
CHAPTER 8		250
SIMULATION EXPERIMENTS USING THE INTELLIGENT NETWORK SIMULATION MODEL		250
8.1	Introduction	250
8.2	Experimental Design	251
8.2.1	Rule-Based Mode Choice Model.....	251
8.2.2	Transportation Network Model.....	252
8.2.3	Commuter Generation.....	252

8.3	Experimental Setup and Results	253
8.3.1	First Experiment – Base Case	253
8.3.2	Second Experiment –Traveller Information Dissemination Schemes	256
8.3.3	Third Experiment – Information Update Time	260
8.3.4	Fourth Experiment – Access to Traveller Information	263
8.3.5	Fifth Experiment – Compliance with Traveller Information	265
8.3.6	Sixth Experiment – Feeder Bus Frequency	267
8.3.7	Seventh Experiment – Introduction of LRT service	269
8.3.8	Eighth Experiment – Incident Management.....	270
8.4	Discussion.....	272
CHAPTER 9		275
CONCLUSION		275
9.1	Contributions	276
9.2	Significant Findings.....	279
9.2.1	Sophistication of the Mode Choice Model.....	279
9.2.2	Traveller Information Dissemination Schemes.....	280
9.2.3	Information Update Time.....	280
9.2.4	Access to Traveller Information.....	281
9.2.5	Compliance with Traveller Information.....	281
9.2.6	Feeder Bus Frequency	282
9.2.7	Introduction of LRT service	282
9.2.8	Incident Management.....	282
9.3	The Global Approach	283
9.4	Limitation and Future Extensions	283
9.4.1	Relaxation of Assumptions	283

9.4.2 Mode Choice Models.....284

9.4.3 Policies Introducing Financial Disincentives.....285

9.4.4 The Attractiveness of Free or Subsidised Public Transport.....285

REFERENCES286

APPENDIX A.....305

APPENDIX B.....320

APPENDIX C.....327

LIST OF TABLES

Table 3.1	General Expressions for Error Term Elements.....	89
Table 3.2	Variance-Covariance Matrix.....	89
Table 3.3	Expressions for Error Term Structure Elements.....	90
Table 3.4	Expressions for Error in Matrix Form	90
Table 4.1	Summary Report on Personal Attributes of the Respondents.....	152
Table 4.2	Influence of Delay on Commuters' Desire to Access Traveller Information	157
Table 4.3	Influence of Delay on Commuters' Desire to Change their Usual Travel Plan.....	160
Table 4.4	Influence of Socio-Economic Characteristics of Car Users on Mode Choice Behaviour	163
Table 5.1	Ordered Probit Model Estimating the Likelihood that the Respondent would access Traveller Information under the Influence of Delay..	175
Table 5.2	Ordered Probit Model Estimating the Likelihood of Commuters' Desire to Change their Usual Travel Plans.....	177
Table 5.3	Binary Logit Model Estimating the Respondent's Mode Choice Behaviour, given Multimodal Traveller Information	179
Table 5.4	Ordered Probit Model Estimating the Willingness of the Commuter to Switch His/Her Mode of Travel, given Integrated Multimodal Traveller Information.....	183

Table 5.5	Marginal Effects on the Probabilities of Commuters' Mode Switching Decision	190
Table 5.6	Mode Choice Logit Model Estimating the Respondent's Mode Choice Behaviour, given Integrated Multimodal Traveller Information.....	193
Table 5.7	Impact on Modal Split by Increasing Travel Time and Travel Cost of Private Mode of Transport.....	197
Table 5.8	Marginal Effects of the Variables on the Probability of Commuters' Mode Switching Propensity.....	198
Table 7.1	Zone-Based Trip Distribution (%).....	237
Table 7.2	Output Data Comparison with the Un-calibrated Model.....	238
Table 7.3	Calibrated Percentages of Different Types of Vehicles	239
Table 7.4	Output Data Comparison Based on the Calibrated Model	240
Table 7.5	Traffic Count Calibration Results for Release Profile RP4.....	244
Table 8.1	Modal Split and Measures of Effectiveness for Base Case Scenario	255
Table 8.2	Effects of Different Schemes of Traveller Information on Modal Split and Network Performance	257
Table 8.3	Impact of Percentage of Commuters Having Access to Traveller Information on Modal Split and Network Performance	264
Table 8.4	Effects of Commuters Level of Compliance with the Traveller Information on Modal Split and Network Performance	266
Table 8.5	Impact of Feeder Bus Frequency on the Modal Split and Overall Network Performance.....	268

Table 8.6 Comparative Statistics with and without the Provision of LRT Service
..... 270

Table 8.7 Comparative Statistics with and without the Accident 271

LIST OF FIGURES

Figure 3.1	Details of the Working Procedure of Research	74
Figure 3.2	Probability Density Function of “ y^* ” and its Relationship to “ y ”	86
Figure 3.3	Working Mechanism of the Rule-Based Intelligent Agent (Source: Russell and Norvig, 1995)	102
Figure 3.4	Architecture of the INSIM Expert System	103
Figure 3.5	Macro Level Rule Flow Structure Imitating the Expertise of a Traveller Information Provider.....	108
Figure 3.6	Generic Rule Pattern for PRB Model	112
Figure 3.7	Generic Rule Pattern for DCM Model	113
Figure 3.8	Venn diagram showing the Partitioning of Different Variables.....	114
Figure 3.9	Partitioning of the Sample Space Showing the Union of Probabilistic Distribution of Different Mutually Exclusive Variables	115
Figure 3.10	Architectural Design of Intelligent Network Simulation Model ..	117
Figure 3.11	Complete Functional Cycle of INSIM.....	118
Figure 4.1	Generic Template for Scenario (SP1).....	143
Figure 4.2	Generic Template for Scenario (SP2).....	144
Figure 4.3	Generic Template for Scenario (SP3).....	146
Figure 4.4	Generic Template for Scenario (SP4).....	147
Figure 4.5	Generic Template for Scenario (SP5).....	149

Figure 4.6 Tree Diagram Illustrating the Presentation Style of RP and SP Slides .
..... 151

Figure 4.7 Commuters’ Access to Different Traveller Information Sources 155

Figure 4.8 Distribution of Respondents’ Desire to Access Traveller Information
in Congested Environment..... 156

Figure 4.9 Distribution of Respondents’ Desire to Change their Usual Travel
Plan under the Influence of Traveller Information 159

Figure 4.10 Respondents’ Attitude towards Pre-Trip and En-Route Traveller
Information..... 161

Figure 4.11 Mode Choice Behaviour of Male Respondents Based on their Age,
Education and Income 165

Figure 4.12 Mode Choice Behaviour of Female Respondents Based on their
Age, Education and Income 166

Figure 4.13 Impact of Delay on Commuters’ Mode Switching Propensity 167

Figure 4.14 Impact of Estimated Time Saving on Commuters’ Mode Switching
Propensity..... 168

Figure 4.15 Impact of Estimated Cost Saving on Commuters’ Mode Switching
Propensity..... 169

Figure 4.16 Impact of Waiting Time (min) and Seat Availability (1 if seat is
available, else 0) on Mode Switching Propensity 170

Figure 5.1 Different Regions Illustrating the Commuter’s Desire to Access
Travel Information corresponding to the Delay 176

Figure 5.2 Change in Modal Share due to Increase in Private Mode Travel Cost..
..... 181

Figure 5.3 Predicted Probabilities at Different Levels of Access Time 185

Figure 5.4	Predicted Probabilities at Different Levels of Waiting Time	186
Figure 5.5	Predicted Probabilities at Different Levels of Travel Time Difference	187
Figure 5.6	Predicted Probabilities at Different Levels of Cost Difference	188
Figure 5.7	Marginal Effects of the Socio-Economic Variables	191
Figure 5.8	Marginal Effects of the Transport Facility Variables	192
Figure 6.1	Architecture of the INSIM Expert System	203
Figure 6.2	Rule Flow Illustration	211
Figure 6.3	Mode Choice Decision Rule Flow for Missing Data	213
Figure 6.4	Automated Knowledge Acquisition Module	214
Figure 6.5	Impact of Range on PE Considering a Rule-Set of 30 Rules	217
Figure 6.6	Impact of Range on PE Considering a Rule-Set of 50 Rules	218
Figure 6.7	Impact of the Number of Rules on the Overall PE	219
Figure 7.1	Overall View of the Simulated Transportation Network	225
Figure 7.2	Coded Phasing Pattern Schemes	227
Figure 7.3	Detailed Map and the Overview of the Simulated Network	230
Figure 7.4	Detailed Simulated Network showing Zones, Links and Nodes	231
Figure 7.5	Different Trips Release Profiles	241
Figure 7.6	Calibrated Demand Profile	245
Figure 7.7	Observed and Simulated Point to Point Travel Time	245
Figure 7.8	Observed and Simulated Travel Time by Bus Route SBS 242	247

Figure 7.9 Observed and Simulated Travel Time by Bus Route SBS 199 247

Figure 8.1 Simulated Transportation Network Showing Details about the
Expressways, Road Network, MRT Lines, and the Proposed LRT
Service 252

Figure 8.2 Average Vehicle Travel Time by Private Mode (min) 254

Figure 8.3 Impact of Information Update Time Interval on Average Vehicle
Travel Time by Private Mode 262

Figure 8.4 Marginal Effect of the Percentage of Commuters Having Access to
Traveller Information..... 265

Figure 8.5 Marginal Effect of the Percentage of Commuters Complying with the
Traveller Information..... 267

LIST OF ABBREVIATIONS

AD	Acceptable Delay
AI	Artificial Intelligence
AMTIS	Advanced Multimodal Traveller Information Systems
API	Application Programming Interface
APTS	Advanced Public Transportation Systems
ATIS	Advanced Traveller Information Services
BLI	Boonlay Interchange
CBD	Central Business District
COM	Probabilistic Model
DCM	Discrete Choice Model
DLL	Dynamic Linking Library
DOS	Department of Statistics, Singapore
DTI	Data Transfer Interface
DVU	Driver Vehicle Unit
EMAS	Electronic Monitoring and Advisory System
EMAS	Expressway Monitoring and Advisory System
ERP	Electronic Road Pricing
GUI	Graphical User Interface
HAR	Highway Advisory Radio
HOVs	Heavy Occupancy Vehicles
IA	Intelligent Agent
IA's	Intelligent Agent's

IC	INSIM Commuter
IE	INSIM Expert
IE	Intelligent Expert
IES	INSIM Expert System
IIA	Independence from Irrelevant Alternatives
IID	Independent and Identically Distributed
IMTI	Integrated Multimodal Traveller Information
IMTIS	Integrated Multimodal Traveller Information Systems
INSIM	Intelligent Network Simulation Model
ISTEA	The Intermodal Surface Transportation Efficiency Act
ITS	Intelligent Transportation Systems
KB	Knowledge-Base
LRT	Light Rapid Transit
MDB	Microsoft Access Database
MNL	Multinomial Logit
MOE	Measures of Effectiveness
MRT	Mass Rapid Transit
MTI	Multimodal Traveller Information
MTI	Multimodal Traveller Information
NTU	Nanyang Technological University
NV	Number of Vehicles
OD	Origin and Destination
PE	Prediction Error
PRB	Pure Rule-Based
RN	Random Number

RP	Revealed Preferences
RSSD	Root Sum Square Difference
SMRT	Singapore Mass Rapid Transit
SP	Stated Preference
Sp1	Hypothetical Scenario
SP1	Scenario 1
SP2	Scenario 2
SP3	Scenario 3
SP4	Scenario 4
SP5	Scenario 5
TCD	Travel Cost Difference
TIS	Traveller Information System
TMS	Traffic Management Simulator
TTD	Travel Time Difference
VMS	Variable Message Signs

LIST OF SYMBOLS

(Y)	Dependent Variable
(X)	Independent Variables
C_n	Choice Set
U_{in}	Utility of Individual n with Alternative i
V_{in}	Deterministic Utility
ε_{in}	Random Error Term
z_{in}	Vector of Attributes
S_n	Vector of Commuters Characteristics
β	Vector of Parameters
α	Scale Parameter
$\Phi()$	Standard Normal Cumulative Distribution Function
ϕ_{im}^k	Mode Switching Decision Indicator Variable
θ_{im}^k	Vector of Parameters
$L(0)$	Log Likelihood of the Restricted Model
$L(\beta')$	Log Likelihood of the Unrestricted Model
ρ^2	Likelihood Ratio
$\bar{\rho}^2$	Adjusted Likelihood Ratio
K	Degrees of Freedom

χ^2	Chi Square Distribution
n	Required Number of Runs
μ	Mean Travel Time
s	Standard Deviation
ε	Desired Margin of Error (percent of “ μ ”)
t	(Student’s t-statistic) Confidence Coefficient
$(1-\alpha)$	Desired Confidence Level
e	Tolerance Error
p	Population Proportion

SUMMARY

Travel mode choice behaviour, as referring to the manner travellers select their travel mode(s), is inherently influenced by travellers' (perceived) knowledge about travel conditions of available choices; this knowledge about travel conditions is collectively termed as traveller information. Timely provisions of strategic traveller information can contribute significantly in improving transport network efficiency while affording better transport services to commuters.

The advents of Intelligent Transportation Systems, and particularly Integrated Multimodal Traveller Information Systems (IMTIS), are bringing much potential towards improving public transport services. However, despite rapid advancements in mode choice behaviour, much remains to be known about the impacts from IMTIS application, which motivates the present research to develop an Intelligent Network Simulation Model (INSIM) and to apply the INSIM system to study the impacts of providing Integrated Multimodal Traveller Information (IMTI). In particular, several IMTI schemes/strategies were applied via INSIM to study how regular car users may be influenced to switch to public transport.

Three types of mode choice models, namely Pure Rule-Based (PRB) Model, Discrete Choice Model (DCM), and Probabilistic (COM) Model, were developed that characterise commuters' mode choices. The PRB Model uses 'IF and THEN' rules, the DCM is based on the estimated mode choice logit model, and the COM model employs the Bayesian approach. A travel behaviour survey was conducted that provided data to calibrate the models, and the PRB was found to be the best model.

The calibrated mode choice models provided the decision rules in INSIM's mode choice module, the INSIM Expert (IE), which simulates commuters' mode choice decisions under the influence of real-time IMTI. A Transportation Network Model for the central and western areas of Singapore was simulated in PARAMICS and integrated with IE by means of an Application Programming Interface (API) to

form the INSIM. Upon calibration, INSIM was able to realistically present complicated scenarios in which real-time IMTI is provided to commuters and network performance measures are recorded.

The functionality of INSIM was demonstrated through a set of simulated experiments under congested travel conditions. The experiments showed that private-to-public mode switch propensity bears a strong and direct relation with the amount of disseminated IMTI as well as information update time. Increase in the amount of provided traveller information can increase mode switching propensity in commuters, whereas the increase in information update time during peak periods can result in less reliable information. The other aspect that was analysed from this set of experiments was the capability of INSIM Expert to imitate the skills of the traveller information provider.

Also analysed during the experiments were influential factors that included degree of accessibility and compliance to IMTI, introduction of public transport and transit facilities and congestion-related events such as accidents. It was observed that a higher level of compliance with Traveller Information improves the overall network performance. Furthermore, access to traveller information as compared to the level of compliance has more significant influence on mode switching and the overall network performance.

The influence of higher feeder bus frequency on the modal split and overall network performance is a slight increase in the transit share. However, improving the transit service encourages many private mode users to use transit mode. The comparative statistics with and without the provision of LRT service shows that with the provision of LRT service the overall network performance is improved.

It was observed during the simulated environment that imitated accidents that disseminating traveller information regarding the occurrence of an accident and corresponding travel environment increased the mode switching propensity in the commuters. It was also observed that the mode switching was occurring not due to the increase in average travel time by private mode or due to the congestion factor,

but it was due to the provided information regarding the accident and corresponding delay on the specific road.

This research shows that the impacts of IMTI in an urban transportation network can be studied using a simulation system such as INSIM. The research findings contribute to advancements on several substantive issues that have not been systematically investigated to date. The limitations in modelling and simulation techniques are also highlighted which provide scope for future extension to this study.

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Population growth and economic development in many countries have generated heavy demand for passenger (and freight) transportation. Expansion of transportation supply is the common solution adopted by these countries – but it is usually costly and not timely enough to meet the demand. Even if it works, the expanded facilities may again induce additional travel demand, which would soon require expansion of infrastructure, resulting in a vicious circle.

Another strategy is to manage the transport demand. It is aimed at reducing or redistributing travel demand spatially or temporally in order to achieve a more balanced use of the transport supplies. To be more specific, this approach is normally carried out by developing an efficient public transportation system, and implementing traffic management strategies for the private transport.

Having an efficient public transportation system alone would not be effective if there is continuous increase in car ownership and usage, which results in undesirable levels of congestion on the road networks. To mitigate the congestion, different traffic management strategies can be implemented, at macro and micro levels. The traffic management strategies may resolve the congestion problem, but only under certain specific conditions, as they can only be successfully implemented in an environment where travel demand is within the capacity of the provided transportation infrastructure. Once the travel demand becomes more than the capacity, the transportation system would again face congestion problems.

Intelligent Transportation Systems (ITS) form part of modern day traffic management tools to relieve traffic congestion (Machado and Figueiredo, 2007). Common ITS (Chorus et al., 2006) are:

- (a) Advanced Traveller Information Systems (ATIS).

- (b) Advanced Public Transportation Systems (APTS).
- (c) Advanced Traffic Management Systems (ATMS).

By providing real-time traveller information to users, ITS offer promising means to manage traffic and improve operations within multimodal transportation networks (Lunt et al., 2004). ITS aims to provide enhanced transit level of service through provision of multimodal traveller information. The key component of this approach is the multimodal transportation network (i.e. integration of private and existing or planned public modes of transport) facilitated by Advanced Multimodal Traveller Information Systems (AMTIS). The functions of AMTIS is to provide latest multimodal traveller information to commuters, and most importantly reduce to a more acceptable level the uncertainties in services such as arrival times and estimated travel times associated with public modes of transport (Lam and Memon, 2003). Providing real-time information to users on network congestion, availability and status of transit modes, opportunities for easy transfer and parking availability could contribute to an efficient integration among the existing modes. The availability of AMTIS allows transit operators to closely monitor their operations and improve the reliability of their services (Grotenhuis et al., 2007).

The ability of AMTIS to accomplish its mission critically depends on how successfully the traveller information can influence the commuters' behaviour in the desired direction, and how effectively it supports transportation management decisions. In this study, the commuter trips are taken into consideration for analysis purpose, because these trips are made on a regular basis from a worker's home to a worksite with a regularly scheduled arrival time during certain peak hours on weekdays. Similarly, the school trips are also considered in this study as these trips are also part of the morning peak traffic. Furthermore, it is highly expected that in certain regions the school trips do get integrated with the commuter trips within certain spatio-temporal distribution.

The provision of real-time traveller information (both pre-trip and en-route) can influence commuters' travel behaviour, e.g. route, mode, and departure time choices, which directly affects the network flow conditions. Thus, understanding the commuters' travel behaviour dynamics in response to traveller information

facilitates the determination of how travel demand may distribute over the entire network (Rietvala, 2003; Hoogendoorn-Lanser, 2005).

1.2 PROBLEM STATEMENT

The success and effectiveness of transportation demand and supply management schemes depend on, among others, the understanding of the interactions between the following four major factors in influencing transportation system performance:

- (a) Commuters' travel behaviour in response to information;
- (b) Commuters' knowledge and experience of network conditions;
- (c) The spatial-temporal distribution of commuters' travel demand in the network; and
- (d) Real-time traveller information.

The AMTIS can be effective to influence commuters' travel decisions by providing commuters with real-time and reliable multimodal traveller information on the available modes and routes of travel. The contents, formats, and forms of dissemination of such information can affect the effectiveness to influence in causing changes to the ways people normally travel. These changes, which may be mode switching and/or route switching, could collectively bring improvement to the network performance – subject to the conditions that the commuters' behavioural dynamics is captured and the information is provided accordingly.

(a) Modelling Impacts of AMTIS on Commuters' Mode Decisions

A number of studies have attempted to explore the potential of traveller information provisions on influencing the ways commuters normally travel (Hoogendoorn-Lanser, 2005). Polak and Jones (1993) observed in their study that the regular car users had significant tendency to obtain multi-modal pre-trip traveller information to organise their trips during the peak hours. In another study, Hickman and Wilson (1995) found that real-time transit information yields a slight improvement in passenger service measure such as the origin to destination travel times. Abdel-Aty (2001) reported that even commuters who do not use transit would likely consider using transit if improved information were provided. Jou et al. (2005) found that the

effectiveness of traveller information depends on the types and format of how it is disseminated. Research has also been conducted on the impacts of ATIS on travel behaviour, mainly focusing on the study of commuters' route choice behaviour. The important factors that influence route choice include system performance attributes such as trip time and congestion (Khattak et al., 1992); experience factors such as scheduled delay and familiarity (Mahmassani and Stephan, 1988; Adler et al., 1992; Mannering et al., 1994; Abdel-Aty et al., 1995; Buliung et al., 2007) and the nature, extent, and quality of ATIS information (Bonsall and Parry, 1991; Abdel-Aty et al., 1994; Adler, 2001). ATIS information also indirectly influences route choice through users' expectations of system performance and their perception of feedback on actual performance measures on alternative routes (Srinivasan and Mahmassani, 1999; Chen et al., 1999; Bradley, 2006).

However, the above referred studies did not focus on the impacts of traveller information on commuters' travel choice behaviour, as reported in Mahmassani (1996), Mahmassani and Stephan (1988), Iida et al. (1992), Polydoropoulou et al. (1996), and Abdel-Aty, (2001). More precisely, the interaction between mode choices, among others, was not satisfactorily captured in most studies, especially when the dynamic nature of traffic flows on the road network is considered. Due to such limitations, the impacts of real-time traveller information on commuters' modal choices have not been modelled in a way to enable estimation of the mode switching propensity of commuters – the practical side of modelling of travel behaviour in multimodal transportation systems. Thus, the effects of real-time multimodal traveller information on commuters' mode choice behaviour have not been well established.

(b) Modelling Multimodal Travel Decisions

Furthermore, in most studies, the impacts of traveller information were observed by providing information to specific mode users, e.g. the impacts of ATIS were analysed by providing traveller information regarding the private modes to private mode users only. In such a framework, where the information is provided for specific modes only, the possibility to study the commuters' mode choice behaviour is very limited (Chorus et al., 2005). Considering the existing multimodal nature of

transportation systems, e.g. in Singapore, where the availability of different modes and the provision of integrated traveller information allows the commuters to plan their trips by integrating different modes or to choose between public and private modes of travel, it becomes necessary to study the commuters' mode choice phenomenon (Luk and Yang, 2003; Bize, 2004). As such, this research sets out to explore a more efficient way to estimate the effects of multimodal traveller information, given by an information system like AMTIS, on commuters' mode choice behaviour and how it quantitatively affects the transport network (Hoogendoorn-Lanser, 2005).

(c) *Studying Impacts of Integrated Multimodal Traveller Information*

AMTIS provides real-time multimodal traveller information, which enables faster and frequent evaluation and decisions about alternative modes of transport. Currently, almost all existing AMTIS presents multimodal information to the users, but without providing the comparative analysis between all the available modes of transport and recommended mode(s) of transport. This approach leaves the users to integrate the information in order to gain the understanding of the state of the transportation systems and make decisions about their trips (Grotenhuis, 2005; Grotenhuis et al., 2007). The lack of understanding of the cognitive processes underlying the complex nature of commuters' decision-making, especially in the presence of real-time traveller information, also provides the scope for substantial research progress. This line of investigation emanates the need to model the commuters' mode choice behaviour in the presence of traveller information as a complex decision process involving many independent decision-makers interacting with each other in complex and non-linear ways in the transportation system. The cognitive and decision processes are necessary to associate decision-makers' characteristics, attributes of the traffic environment, real-time traveller information, and constraints in the decision environment with behavioural outcomes of interest. It is the aim of this research to understand the cognitive processes underlying user behaviour that could enable the identification and representation of stable aspects in behavioural processes that can improve the prediction power of the relevant choice models. Such robust and parsimonious models are desirable because of their

increased computational efficiency, enhanced forecasting accuracy, and improved interpretability. Inquiry into cognitive processes also has important implications for the evaluation of policy measures aimed at managing demand, and assessment of traffic control strategies.

(d) Evaluating Impacts of AMTIS and User Responses

This research is also motivated by the need to develop an integrated tool to evaluate impacts particularly due to implementation of intelligent transportation systems. The model is aimed to overcome many of the known limitations of static tools used in current practice. Limitations of past and existing approaches are related to the types of alternative measures to be represented and evaluated, and the policy questions that planning agencies are increasingly asked to address. The new model approach captures the interaction between users' mode choices and the interactions with a dynamically changing environment under different information provision strategies and network control schemes. It adopts a rule-based expert system, which is used to simulate commuters' decisions in which the modes to use are based on a range of choice criteria.

The model assumes a stochastically diverse set of commuters in terms of underlying preferences, as well as their accessibility and response to the supplied information. To overcome the modelling limitations in handling the modal distribution and dynamism in mode choice behaviour, an Intelligent Network Simulation Model (INSIM) is proposed. INSIM is developed as a rule-based model with application programming interface to integrate with existing traffic simulation software packages e.g. PARAMICS, VISSIM, City-Traffic, etc. The intelligent system approach facilitates the discovery and implementation of rules and their dynamics to govern the simulation of a multimodal transportation network. The rules are used to imitate the commuters' decision process and their interaction with the network dynamics at a micro level, and traveller information systems operator's process at a macro level, at the same time. The rules to be implemented are determined by a hierarchical process that best satisfies the agent's utility. The use of a rule-based also allows the development of a dynamic knowledge-base that can evolve along with the continuing applications of the intelligent agent.

(e) *The Research Question*

Based on the above discussion, the research question can be summarised as follows:

“How to develop an intelligent simulation system that can provide a platform to imitate the impacts of integrated multimodal traveller information on commuter mode switching behaviour and allows one to analyse the influence of mode switching dynamics within the existing travel environment?”

1.3 STUDY APPROACH AND SCOPE OF WORK

In this research, an INSIM is developed, which is aimed at analysing the impacts of different traveller information schemes. The core component of INSIM is the INSIM Expert, which is capable of simulating the commuters' mode choice decisions under the influence of real-time multimodal traveller information. To evaluate the performance of the developed model a multimodal transportation network for a certain specified area of Singapore is simulated, and the network performance is measured after implementing different traveller information schemes.

The pre-trip traveller information can influence the commuters' travel choice behaviour and bring changes in the network mobility aspects resulting in a better spatial-temporal travel demand distribution. Based on this assumption, the INSIM is designed in a way such that it is capable of capturing and analysing the overall multimodal network conditions and evaluating different traveller information strategies. Different pre-trip traveller information supply strategies can also be analysed subject to the condition that they minimise the total travel time in the network. Thus, the information is supplied at different strategic levels and its impact on the travel time can be evaluated. The information supply strategy that gives the maximum reduction in travel time can be implemented and the network performance measures can be analysed. Therefore, the impacts of different multimodal traveller information and travel demand management strategies can be evaluated in a more realistic and efficient manner.

This research first develops the foundation to understand and gain knowledge about the commuters travel behaviour in an information-rich multimodal transportation system. In this research the information-rich environment is considered as the environment where the commuters have full access to IMTIS before and during their travel period. This knowledge is necessary for the design and development of the rule-based mode choice model. To gain this knowledge a travel behaviour survey is conducted. The survey target group is the car users/drivers, as the objective is to analyse the mode switching behaviour of commuters. The respondents are presented with different travel scenarios and their travel behaviour is gathered. To analyse the impact of traveller information, the respondents' behaviour towards access of traveller information and their level of compliance to the provided information is analysed and modelled.

Inferential testing is performed based on the modelled variable coefficients and associated statistics. The relative importance of various explanatory variables on the choice process is analysed by examining the sign, magnitude and statistical significance of the estimated parameters. The suitability of alternative specifications is ascertained using various statistical tests. The ordered probit models are estimated to study the socio-economic and commute variables that can influence the commuters travel behaviour in response to the provided traveller information. The logit models are estimated to provide the insights about the commuters' mode choice preferences corresponding to different forms and levels of traveller information. This part of the study gives the basis to support the idea that the commuters are willing to access and comply with the provided traveller information. It also provides the knowledge to design the rules to develop the rule-base, which is an integral part of the knowledge-base that acts as the logical or rational component of the commuters' mode choice model (i.e. INSIM Expert).

The knowledge-base is designed and developed on the expertise gained from the travel behaviour survey. A comprehensive rule-base is developed, which consists of a variety of rules. Each rule imitates a different respondent's travel behaviour. The imitation of commuter's behaviour is based on the cause and effect pattern such as, if this condition is true then this action will be taken else alternate action will be

taken, and so on. These IF and THEN conditions are related to the commuters' socio-economic characteristics and travel environment, and their mode choice decisions that are captured during the travel behaviour survey. To simulate the mode choice decision of a commuter, the INSIM Expert applies a heuristic search and finds the nearest matching commuter(s) within the rule-base. The weighted decision of matching commuter(s) is then predicted as the decision of the provided commuter.

The INSIM Expert resides in the shell of INSIM Expert System which is integrated with the transportation network simulation model. The simulator used in this research is PARAMICS. It provides a powerful Application Programming Interface (API) to control the simulation environment and to integrate it with other platforms. The API is designed and developed to collect real-time traveller information on public mode i.e. buses and trains, as well as private modes of transport. The well known Dijkstra's algorithm is implemented to solve the problem of finding the shortest path with respect to travel time from any origin to any destination for public mode of transport. The travel times that are used in the algorithm incorporate the waiting time for the buses and trains in that specific time stamp.

Performance measures such as travel time or travel speed are considered as real-time traveller information, and are gathered by the API. After the loading of trips onto the network, the changes in performance measures at every node and link are captured by the API at specified time stamps. The developed API is capable of communicating with the INSIM Expert System. It sends all the required traveller information to the INSIM Expert System to disseminate within the commuters that are to be loaded onto the network in the commencing time stamp.

In the simulated environment the INSIM Expert System continuously analyses the system performance and provides real-time traveller information to the commuters that are to enter the network in the commencing time stamp. Once the commuters receive the traveller information, the INSIM expert simulates their mode choice decisions. The commuters with their update modes of travel are collected by the API and loaded onto the transportation network simulation model.

The role of supply conditions, AMTIS, and commuter behaviour raise several substantive issues. The use of INSIM circumvents the limitations that preclude the evaluation of different traveller information supply strategies. A variety of traveller information schemes are implemented with different blends of commuters' travel behaviour and levels of travel demand. The network performance measures are constantly observed while exploring different aspects of real-time multimodal traveller information strategies. To address these questions, a series of experiments are designed and conducted to examine the influence of AMTIS and commuter behaviour on overall network performance. In the first set of experiments, interest is centred on the rationality of the INSIM Expert, and the effects of varying traveller information schemes. The second set of experiments focuses on the update timings and the reliability of information. The third set of experiments addresses the commuters' access and level of compliance. Lastly, the impact of changing the supply side (i.e. the transport infrastructure) is analysed by improving the public transport facilities.

The results of this research can provide important foundations that help the operations of a multimodal transportation system to regulate the travel demand based on the provision of multimodal real-time traveller information. Through the use of carefully calibrated travel behavioural models, more robust network state prediction capabilities can be achieved, with applications in the design of effective control strategies and traffic planning, as well as the evaluation of congestion relief techniques and development of travel demand management strategies for long term applications.

1.4 OBJECTIVES OF THE STUDY

The research encompasses several enabling applications in the disciplines of commuter travel behaviour modelling, transport operational planning, dissemination of multimodal traveller information, and artificial intelligence. The objectives outlined for this research are as follows:

1. To model commuters' mode and route switching propensity based on pre-trip real-time multimodal traveller information.

2. To design and develop a microscopic traffic simulation model, which is aimed at:
 - a. representing a multimodal transportation network;
 - b. simulating the provision of real-time multimodal traveller information; and
 - c. being integrated with a rule-based intelligent agent.
3. To develop a rule-based expert system, which
 - a. integrates with a dynamic traffic simulation model;
 - b. imitates the behaviour of a traveller information systems operator, and disseminates traveller information; and
 - c. imitates the commuters' mode choice behaviour through the rule-based intelligent agent.
4. To study the network performance measures under the influences of:
 - a. various types and forms of integrated multimodal traveller information supply strategies; and
 - b. different travel demand distributions and travel behaviour patterns to achieve optimal distribution of travel demand in a multimodal transportation network.

The first objective is aimed at understanding better the travel characteristics of commuters in Singapore and their behaviour and preferences towards mode and/or route switching decision. In this regard a travel behaviour survey is to be conducted. Models are established to capture and analyse the impacts of multimodal traveller information on commuters' travel behaviour.

In the second objective, a microscopic traffic simulation approach is adopted, and a multimodal transportation network is simulated. The simulated network provides the details about the traffic and travel environment. The traffic details are then utilised to evaluate the network level of performance, whereas the details about the travel environment are used as traveller information, which is provided to a rule-based intelligent agent representing the commuters via an Application Programming Interface.

The third objective deals with the design and development of the intelligent agent, which is designed as a rational agent to both mimic commuters' travel behaviour and a system operator to control the travel information. The intelligent agent provides a sophisticated yet flexible way to integrate with the microscopic traffic simulator, from where the agent collects the real-time traffic information on different modes of transport. The agent performs two basic functions; at the macro level it disseminates the traveller information to the commuters, and at the micro level it imitates the commuters' mode choice behaviour.

The first and third objective results in the development of a rule-based mode choice model that can simulate the commuters' mode choice behaviour in response to the provided real-time traveller information. The combination of second objective with the first and third objectives provides a platform to analyse the impact of different multimodal traveller information schemes. The fourth objective addresses the problem of evaluating different kinds of traveller information strategies under different travel conditions and blends of travel behaviours. It analyses the impacts on, and improvement in, the overall performance of the network, with respect to the spatial and temporal distribution of travel demand. Such study provides new perspectives to influence the dynamics of the traffic conditions and to control travel demand in the network. Furthermore, it can also analyse the reliability of the information and commuters' level of compliance to such information.

1.5 STRUCTURE OF THE THESIS

This thesis is structured as follows. First, the introduction chapter provides an overview of the problem definition and motivation, as well as the research objectives of this study. A literature review of related work is presented in the second chapter. The review pertains to four related areas: the role of real-time multimodal traveller information; the commuters' travel behaviour; models of commuter decisions, and the underlying cognitive and decision processes; and the expert systems and their applications. The third chapter describes the methodology used in this study to analyse and simulate the commuters' travel behaviour under real-time traveller information, and to evaluate the impact of different multimodal traveller information schemes. Chapter 4 presents the details about the travel

behaviour survey. It elaborates the design and development of the survey tool along with the conduct of survey. It also presents the empirical findings of the travel behaviour survey. The formulation of different models, which estimate the commuter's response to the provided traveller information, is discussed in Chapter 5. The logit model, ordered probit model, and bounded rationality model are developed. Chapter 6 gives the description of the design and development of the INSIM Expert System. The design and functionality of rule-base, knowledge-base, and the search process of matching rules along with the learning capabilities of the INSIM expert are detailed in this chapter. The network simulation model is presented in Chapter 7. The framework and structure of the model are presented along with the description of different components of the model. These details about the calibration and validation of the model are also presented. Chapter 8 presents the design and results of a variety of simulation experiments for multimodal transportation networks. The principal objectives of these experiments are to: (i) show the significance of including the mode choice dimension in the dynamic traffic assignment framework, and (ii) quantify the network level of service under alternative information provision strategies, different level of commuters' access and compliance, transit operation plans, and improved transport infrastructure. Finally, Chapter 9 discusses the overall conclusions from the research. This chapter also identifies future research needs and indicates possible directions for further research.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

This chapter presents the review of travel behaviour dynamics in an information-rich environment. It discusses the issues related to existing transportation network modelling approaches with the provision of real-time travel information. The main purpose of this review is to provide a comprehensive status of the current research and operational activities, with respect to the commuters' mode choice decision-making process and the modelling approaches within the framework of the objectives that are stated in Chapter 1. The review shall provide the author with the important aspects that are to be considered in his research study. It will also highlight the scope and limitations related to the research issues of interest. In view of the objectives of this study, the literature review will focus on four aspects i.e. Intelligent Transportation Systems (ITS), impact of travel information on commuters' travel choice behaviour dynamics, transportation network simulation modelling, and application of Intelligent Agents (IA) in network simulation modelling.

2.2 TRAVELLER INFORMATION SYSTEMS

Traffic congestion is a root cause of resource consumption and environment degradation. To mitigate the adverse impacts of congestion, innovative technologies are being designed, developed and implemented by transportation researchers, planners and operators under the Intelligent Transportation Systems (ITS) umbrella. ITS have emerged in the transportation arena with the aim to leverage on information technology to better manage transportation system, which holds promises to alleviate congestion. The evolving ITS technologies essentially gather real-time traveller information for dissemination to different components of the transportation system, which in turn can influence the travel behaviour. ITS can also improve the level of service in the transportation system by providing opportunities for more efficient navigation through the network, thereby resulting in more

balanced distribution of travellers across the available routes and modes of transport (Adler, 2001; Machado and Figueiredo, 2007). A high level of response to ITS can significantly reduce travel time, delay, fuel consumption, and environment pollution.

Commuter's decision-making as influenced by ITS travel information takes place at two principal instances i.e. at the beginning of the trip (trip origin) or during the trip (en-route). The travel information that influences the travel decisions at trip origin (home or workplace) is referred as "pre-trip traveller information" and it can be provided through PDA, mobile phone, television, radio, telephone enquiry, computer on-line services, etc. Pre-trip information can reduce the degree of unexpectedness when travellers encounter travel time variation on their routes. The users of transportation systems could make trips more efficiently by changing most attributes of the journey, including departure time, route, mode, destination, or even whether to make the trip at all based on the pre-trip traveller information. However, the success of transportation system depends greatly on the ability to provide useful information round the clock and is accessible to all travellers. The Intermodal Surface Transportation Efficiency Act (ISTEA) bill passed by Congress in 1990 further emphasised the fact that effective dissemination of information regarding transportation services is essential in promoting a balanced use of different transportation modes, which would, in turn, alleviate traffic congestion (Kikuchi et al., 1994).

The era of development of traveller information systems has been grouped into two generations. The first generation covered the emergence of computer technologies and traffic surveillance and control systems during the late 1960s and early 1970s. These systems served to improve flow at localised points in a network, such as a heavily congested freeway-to-freeway interchange, or to make travellers aware of non-recurring congestion, such as special events or incidents. Variable Message Signs (VMS) and Highway Advisory Radio (HAR) are examples of 1st generation systems. The 2nd generation systems are the burgeoning advanced traveller information systems. They encompass a wide range of new technologies, which are being designed to provide travellers with dynamic route guidance (Hidas, 2002),

real-time traffic information (Ferman et al., 2005) and advanced traveller information services (Adler et al., 1998; Chorus et al., 2006; Dia and Panwai, 2007).

The provision of advanced traveller information can be distinguished in three different categories, depending on the nature and scope of its content and the type of users (Kenyon and Lyons, 2003). The first category belongs to the Unimodal Traveller Information (UTI), which covers information within a particular service domain, relating to a single mode of travel. The second category is the Multimodal Traveller Information (MTI) which is concerned with information on more than one mode of travel within a single source. The third category is Integrated Multimodal Traveller Information (IMTI), which covers information concerning different mode choice options in response to a particular journey specified by the user, within a particular information service.

2.2.1 Unimodal Traveller Information

The UTI ranges from low technology paper timetables and road atlases, to dynamic itinerary planning facilities or real time service information. The UTI system may integrate information about a number of operators, but it is limited to providing information about a single mode. The provision of such information can facilitate in making better decisions regarding the travellers' mode under consideration, but is less likely to provide any stimuli that may bring/develop any mode switching attitude in the travellers (Potgraven and Van Leusden, 1995; Stradling, 2002).

(a) Advanced Traveller Information Systems (ATIS)

The ATIS are the best example of UTI-type system. ATIS provide drivers with en-route travel information regarding traffic conditions, incidents, construction activities, weather conditions, hazardous road conditions, and recommended safe speeds. The information allows the drivers to select the best route, or to change routes during their trips (Zhang and Verhoef, 2006). Travellers with longer travel times operating at higher level of uncertainty (about the network) have shown greater propensity to make pre-trip decisions and to switch their routes and departure times under the influence of ATIS (Khattak et al., 1999; Abdel-Aty and

Abdalla, 2006). For example travellers can choose to avoid long queues from comparing the difference between actual and desired arrival times and opt to switch to travel conditions with a shorter travel time (Ben-Akiva, 1985; Abdel-Aty and Abdalla, 2005).

Modern ATIS provide travellers with instructions on how to reach to their destinations and even identify a suggested route to reach the specified destination (Dongjoo et al., 2007). They acquire information to help the motorists to improve way-finding performance in order to reduce travel time (Adler, 1998). Route guidance for commuters could be determined readily from simple roadway network models or in a more advanced manner by using real-time information describing the current traffic conditions, incidents, road closures, etc. Dynamic route guidance allows a broader spread of routing options across the network. It is clear that the overall system performance under the influence of ATIS depends on the level of traveller's response to ATIS information. A higher degree of response can be achieved by providing prescriptive information rather than descriptive information (Jou et al., 2005). A hierarchy of achievable response rates exists where the highest degree of response is obtained for reliable, predicted information and least response for random information (Chen et al., 1999; Paz and Peeta, 2008).

(b) Advanced Transit Information Systems

Advanced transit information systems are another example of UTI type system. These systems provide travellers with information about available transit services before and during their trips. Travellers can access this information at home, work, transportation centres, wayside stops, and while on board vehicles through a variety of media such as telephones, monitors, cable television, variable message signs, kiosks, and personal computers, PDA, and mobile phones. Some systems with links to automatic vehicle location provide real-time information about available transit service, such as arrival times, departure times, and delays (Rodier et al., 1998; Grotenhuis, 2007).

In addition, on-board vehicle sensors automatically monitor data such as passenger loading, fare collection, drive-line operating conditions, etc., and disseminate to

travellers and operators for real-time management response. The pre-trip transit information provides travellers with accurate and timely information about transit services. It can increase travellers' awareness and reduce some of the uncertainties related to the use of transit. For some trips, the provision of such information can make travel by transit more appealing than travelling by car. A few studies have examined the effect of transit information systems on travellers' choice of mode. Abdel-Aty (1995) used computer-aided telephone interviews in the Sacramento and San Jose areas of California to identify the transit service information that was desired by non-transit users. Customised stated preference choice sets were used to identify the likelihood of a commuter's choice to use transit. The study found that 38% of the respondents who did not use transit would likely consider using transit if improved information was provided.

2.2.2 Multimodal Traveller Information

Due to the limited scope of UTI a more advanced traveller information system was introduced which comprises a series of UTI services covering several modes that are integrated and provided to travellers in a single medium. As compared to UTI, which only reduces the uncertainties of the trip by the chosen mode, the MTI has been observed to be more effective as the information covers more than one mode (Lappin, 2000; Egeler, 2001; Kenyon and Lyons, 2003; Bize, 2004).

(a) Advanced Multimodal Traveller Information Systems (AMTIS)

The AMTIS are systems which disseminate MTI that can affect the travel demand patterns (Shank and Roberts, 1996). An AMTIS informs the potential travellers of current overall conditions with respect to available modes and routes of the transportation system, and helps them to best assess their travel options before or during their trips. Metropolitan authorities typically look to regional multimodal traveller information systems to enable the travelling public to make more informed mode and route choices (Adler and Blue, 1998). The developments concerning the integration of such information across modes and routes are well documented. Casey et al., (1998) summarises that in the USA, many regional information systems have become multimodal and offer travellers an opportunity to make fully

informed mode choice decisions about travelling in those locality (FTA, 2003). In Europe, efforts have been made to develop a Europe-wide citizens' network which will promote the passenger transport services and information (Caulfield and O'Mahony, 2004). The Netherlands Ministry of Transport is in the process of developing a real-time multimodal information system. It will enable the travellers to make better choices regarding their mode, route and departure time, both before and during their journeys (Van Toorenburg, 1999; Bize, 2004; Grotenhuis, 2005).

Though MTI services do provide a centralised access system to information about multiple modes, there is typically no interaction linkage on the information among various modes. The user has to actively seek information about each mode via separate enquiries and such mode specific query does not promote the consideration of choosing alternative modes. The results from some pre-implementation studies into user attitudes towards MTI had suggested that users may respond to such information services with rational consideration of modal choices but a significant modal shift in response to MTI has yet to be observed (Vaughn et al., 1999). Thus, it can be seen that MTI services are used in the same way as UTI services in that the mode is first selected before the associated information is consulted (Kenyon and Lyons, 2003).

2.2.3 Integrated Multimodal Traveller Information

To address the transport problems of tomorrow in an efficient manner, the key actors in the transport sector will have to transform existing information services into a more integrated and holistic one. In this regard, the IMTI can highlight more effectively the benefits to travellers, non-travellers, information suppliers, transport operators and local and national governments (Kenyon and Lyons, 2003). As a system where a number of modal options for a journey are available, the IMTI presents the user with comparative and detailed information on the desired and/or available modes in a single enquiry. The IMTI can be visualised as a repository for current, comprehensive, and accurate roadway and transit performance data. It directly receives data from a variety of public and private sector sources, combines and packages that data, and provides the resulting information to travellers and other customers via a variety of distribution channels.

The provision of IMTI can provide the travellers with the opportunity to choose a particular route, mode, departure time, or even decide whether or not to make the trip, or to bypass congestion and delays within a single enquiry. Using the IMTI services, the travellers may switch their travel modes, routes and/or delay departure times in the event of congestion as the mode, route and departure time choices are important components of a traveller's decision regarding trip-making (Bhat, 1998; Grotenhuis, 2007). IMTI may be more effective and more practical given the reduced effort required in locating and comparing information about alternative modes (Wright and Egan, 2000). The IMTI services are also aimed at promoting travel by public transport or by non-motorised alternatives (Lyons et al., 2001).

Since many of the IMTI systems/technologies are aimed at relieving flow conditions during congested travelling periods of the day, the behaviour of travellers must be treated as a central element in the formulation and implementation of these relief measures. On the other hand, pre-trip traffic information systems are important because they potentially provide travellers with the greatest flexibility. Therefore, a better understanding of how pre-trip information will influence travellers' choice behaviour is required. At the same time, the rapid growth of traveller information-related systems/technologies extended the scope and scale in analysing operational alternatives and policy issues under consideration. The emphasis on long-term planning of surface street networks has shifted to short-run planning issues that focus on the management of integrated multimodal transportation systems. A major aspect related to such issues is to forecast the changes in travel demand induced by the pre-trip IMTI. The decision on whether to support the developments of various IMTI-related systems/technologies requires a very careful analysis of their impacts on travel behaviour (McFadden et al., 1977; Polak and Jones, 1993; Abdel-Aty, 2001; Grotenhuis, 2005).

2.3 IMPACT OF PRE-TRIP INFORMATION ON TRAVEL BEHAVIOUR

There have been extensive investigations into departure time and route choices in interactive laboratory-like experiments (Mahmassani and Liu, 1997; Lunt et al., 2004). Mahmassani and Liu (1999) focused on the day-to-day dynamics of travellers' choice behaviour in response to the supplied information. Vaughn et al.

(1999) developed an interactive experiment to investigate drivers' learning and pre-trip route choice behaviour. The results indicated that drivers can rapidly identify the accuracy level of provided information and adjust their behaviour accordingly. Although these studies provided valuable insights into the complex human decision behaviour, they are primarily based on theoretical concepts or simulated experiments, and the transferability of these insights to travellers' choice behaviour has not been adequately established.

Polak and Jones (1991) investigated travellers' requirements for different types of travel information and methods of enquiry. They examined the relation between the processes of information acquisition to changes in travel behaviour. The study was carried out using a Stated Preference (SP) approach, built on the use of a microcomputer-based simulation of an in-home pre-trip information system offering information on travel times from home to city centre, by bus and by car. The study focused mainly on investigating which mode the respondents enquired about and the first-ranked alternative (the chosen alternative). The results indicated that even amongst regular car users in the study area, there is a requirement for multi-modal pre-trip travel information. Up-to-date information on travel conditions and services helps travellers make more "informed" decisions about the when, where and how to travel. In turn these informed decisions help to optimise the use of available transport infrastructure (Lunt et al., 2004).

Kitamura et al. (1995) conducted in-laboratory interviews with 50 subjects who used a PC-based transit pre-trip information system prototype. The subjects' ratings of the system indicated that age is an important variable that defines market segments for such information systems. Stephanedes et al., (1992) developed empirical models that can provide a better understanding of the relationship between pre-trip departure time and route choice decisions, and the duration of the announced delay. In a survey of Seattle commuters, pre-trip traffic information influenced route choice for only 11% of the respondents and departure time for 13% (Beaton and Sadana, 1994). Survey conducted in the New York metropolitan area indicated that if people are able to obtain more timely and reliable traffic information, they are more likely to use it (Harris and Konheim, 1995).

Khattak et al. (1995) proposed a combined revealed and SP model to explore how travellers' pre-trip decisions are affected by ATIS in the context of unexpected congestion. Khattak et al. (1996) surveyed technology suppliers of transit information systems about the features, functions, and performance testing and deployment in transit agencies, and their potential impacts on travellers and transit operators. The survey dealt with Advanced Public Transportation Systems (APTS) in general and information systems specifically. The survey results suggested a trend toward transfer of data in real time through electronic media and increased automation. Abdel-Aty et al. (1994) found that commuters might value and use pre-trip information more than en-route, because it gives them the situation on their routes in advance, which enables them to change their travel plan. Abdel-Aty et al. (1997) further investigated and found that receiving pre-trip traffic information was a significant factor in the model. Jha et al. (1998) developed a Bayesian updating model to capture the mechanism by which travellers update their travel time perceptions from one day to next in light of information provided by ATIS and their previous experience.

2.4 TRAVEL BEHAVIOUR

The potential efficiency and success of various policy measures for eliminating or reducing traffic problems in metropolitan areas largely depend on how people will respond to them. Lack of public acceptance is an important issue that has been highlighted in recent years (Emmerink, 1995; Garling, 2004; Lam and Trinh, 2006). Whether and how travel actually changes is, however, an equally important issue that is far from settled. As has been highlighted by leading travel behaviour researchers (Axhausen, 1992; Vilhelmson, 1999), it has become increasingly evident that travel results from choices that people make are both interdependent and dependent on desires or obligations to participate in activities. It may therefore be a mistake to focus any measure solely on target behaviour such as car use. In an alternative approach (Kitamura, 1998; Garling, 1998), travel choice is viewed as an adaptation to changes where people try out different choice options over time.

2.4.1 Travel Behavioural Models

Existing travel behaviour models focus on the ways in which one can simplify and abstract important relationships underlying the provision and use of transportation facilities. The accuracy of these models and methodologies depends to a large extent on their forecasting capabilities and their compatibility towards the evolving transportation system. A substantial amount of research has been devoted in recent years, towards developing promising new models and methodologies to study travel choice behaviour. (Adler and Ben-Akiva, 1979; Pas, 1983; Bhat and Koppelman, 1994; Recker, 1995; Mahmassani and Liu, 1997; Paz and Peeta, 2008). The different methodologies adopted for modelling travellers' behaviour, two major categories can be identified, namely: application of mathematical programming concepts, and application of discrete choice models.

The mathematical programming approach is relatively limited in modelling the decision process underlying travellers' behaviour, principally because the decision process is an extremely complex problem to formulate and solve which necessitates prohibitively costly computational requirements. Another reason is that it is very difficult to formulate a mathematical optimisation problem that can adequately capture the human decision-making process (Jou et al., 2008). These limitations in the mathematical modelling approach has led to the application of discrete choice models that are widely adopted by travel behaviour researchers.

The discrete choice models have been in existence since the early 1960's. They first appeared as the result of academic research in the field of transportation economics (Beesley, 1965). These models were used to evaluate the relative importance of certain transportation variables in trip-making decisions. Since then, a substantial amount of research has focused on the following issues (Adler and Blue, 1998):

Developing the theoretical aspects of individual choice behaviour;

- (a) Simplifying the computational requirements of model building;
- (b) Identifying new and more powerful explanatory variables;

- (c) Resolving the issues that limit the application of individual choice models to other travel demand decisions; and
- (d) Demonstrating the capabilities of these models in solving practical planning problems.

The discrete choice modelling procedure estimates the probability of an individual making a particular choice based on modal attributes and the individual's socio-economic characteristics (Domencich and McFadden, 1975; Ben-Akiva and Lerman, 1985). These models provide a technique which replicates the decision theory of microeconomics related to the choices among the discrete sets of alternatives. The discrete choice models, like other models of microeconomics, assume that an individual's preferences among the possible alternatives can be described by a utility function (Anderson et al., 1992). They account for the effect of uncertainty of human behaviour by using a random error component. Manheim (1979) listed the factors which have contribution to this randomness as: service attributes (comfort or other non-quantifiable attributes), the level of information or knowledge about the available choices, and randomness in the travellers' behaviour.

The random utility theory is the most common theoretical base for generating discrete choice models, and it has been introduced because it is impossible to define exactly the utility of an alternative. The random utility theory assumes that all individuals belong to a homogenous population, and have detailed information. They act rationally and choose the alternative offering the highest utility. The individual's choice set is predetermined and each choice is attached to a net utility for every individual (Williams, 1977). In theory the individuals facing identical attributes would evaluate the alternatives identically, behave identically and choose the same alternative. In practice, this is not true so only the observable attributes are included in the attractiveness function. The individuals perceive some utility for each alternative, which differs from the measured attractiveness, and their choices are based on this utility.

Several types of discrete choice models exist. The main difference among them is essentially based on the variation in the assumed probability distribution of their error terms. If the error term is assumed to have a multivariate normal distribution,

this leads to a family of probit models. This is the most general type of model for discrete choice and it allows both dependent error terms and error terms with different variances. However, probit models have computational drawbacks since the choice probabilities can only be approximately evaluated. Daganzo (1979) has described the theory and application of probit model. Jou (2001) established the model framework to investigate the effects of pre-trip information on commuter's decision including departure time and route. Two scenarios, with and without pre-trip information, were analysed. Instead of performing estimations for these two scenarios (models) separately and then comparing the difference and similarity between these two models, a joint latent variable incorporating with and without pre-trip information was introduced and derived for simplifying the estimation procedures. Because of the assumption of normal distributed error term in the latent variable, probit framework was applied to take advantage of the allowance of more flexible model specification through parameters in the variance covariance matrix. The results indicated that the model does possess a good explanatory power.

On the other hand, if the error term is assumed to have Gumbel or Weibull distribution, the result is the family of logit models. The simplest logit type model is the Multinomial Logit (MNL) model, where all the alternatives are assumed to be independent and to have equal variance. Luce (1959) stated the axiom of Independence from Irrelevant Alternatives (IIA), which made it possible to extend the model to multiple choice situations. For the IIA property to hold, it is necessary that all variances of the error terms are equal and no correlation between the alternatives is allowed. This is the basic assumption of MNL model (Ben-Akiva and Lerman, 1985). Henn (2000) states that the route choice under the influence of pre-trip information can be modelled by adopting MNL model. The travellers' route evaluation process specifically in dynamic environment can be based on variables that can barely be strictly ordered. Thus, different routes cannot be ordered classically, but the probability that the variable under consideration might be the smallest can be calculated. Such a probability then can be seen as a score given to each path and represents its level of optimality – the higher the score, the better the path. Munizaga et al. (2000) states that the MNL model is remarkably robust and can be used reliably to evaluate the effects of even substantial policy changes in the

presence of heteroskedasticity. In this case, its performance in response analysis is comparable to the more computationally demanding multinomial probit model.

If correlation between alternatives is assumed in MNL model, the more advanced Nested Logit (NL) model can be used. This model is regarded as “state-of-the-art” (Lundgren, 1989). In nested logit models or structured logit models, the correlation, which is the dependence among the error terms of the available modes in the utility function, can be handled. The choice situation is described as a sequential multi-level choice. In a two-level situation, the initial choice is assumed to be among a number of sets of alternatives, and the subsequent choice is among the alternatives within the chosen set (McFadden, 1984; Riddington et al., 2000). In a general choice situation, several types or levels of choices can be defined. The combined trip distribution and modal split problem can be formulated as a two-step choice problem. Either the destination choice is step one and the mode choice is step two, or the mode choice is considered first and the distribution choice second. A nested logit model can describe this choice situation in detail (Lundgren, 1989).

When the alternatives are having some specific sequence or do follow some order then the ordered probit model can be used to model characteristics that are not explicitly observed in the population. It is specifically useful to model an ordered set of outcomes where distances between choices are not equal. Assumptions are made on the underlying continuous unobserved variable and thus different boundaries are thereafter determined to recognise discrete nature of the observed dependent choice variables (Bhattacharjee et al., 1997). A study carried out by Abdel-Aty (2001) adopted the ordered probit modelling technique. The objectives of the study were to investigate whether advanced transit information would increase the acceptance of transit, and to determine the types and levels of information that are desired by commuters. The survey included a customised procedure that presents realistic choice sets, including the respondent’s preferred information items and realistic travel times. The results indicated promising potential of advanced transit information in increasing the acceptance of transit as a commute mode.

The latest development in discrete choice models is the logit kernel model (Srinivasan et al., 2005). It is a discrete choice model in which the disturbances (of the utilities) consist of both a probit-like portion and an additive independent, identically distributed Gumbel portion (i.e., a multinomial logit disturbance). Multinomial logit (MNL) has its well-known blessing of tractability and its equally well-known curse of a rigid error structure leading to the IIA property. The nested logit model relaxes the rigidity of the MNL error structure and has the advantage of retaining a probability function in closed form. Nonetheless, nested logit is still limited and cannot capture many forms of unobserved heterogeneity including, for example, random parameters. The logit kernel model with its probit-like disturbances completely opens up the specification of the disturbances so that almost any desirable error structure can be represented in the model. As with probit, however, this flexibility comes at a cost, namely that the probability functions consist of multi-dimensional integrals that do not have closed form solutions. Standard practice is to estimate such models by replacing the choice probabilities with easy to compute and unbiased simulators. The beauty of the additive Independent and Identically Distributed (IID) Gumbel term is that it leads to a particularly convenient and attractive probability simulator, which is simply the average of a set of logit probabilities. The logit kernel probability simulator has all of the desirable properties of a simulator including being convenient, unbiased, and smooth (Walker, 2001). Srinivasan and Mahmassani (2003) proposed a logit kernel framework to model dynamics in route-switching behaviour in the presence of ATIS. In the light of limitations of cross-sectional travel demand models and the shift towards activity-based approaches, the proposed Dynamic Kernel Logit framework was found to be particularly well-suited for calibrating dynamic travel behaviour models with a large number of periods. Hensher and Greene (2003) also adopted the mixed logit model to analyse the stated choice experiment for long distance travel between six different cities in New Zealand. Mixed logit and multinomial logit models were also estimated and compared on California households' revealed and stated preferences for alternative fuel vehicles (Brownstone et. al., 2000).

2.4.2 Travel Behaviour Survey

The formulation of discrete choice models is based on data collected from travel behaviour surveys. In these surveys, the samples are drawn from the pool of commuters from whom the socio-demographic and commute-related characteristics are gathered. The conduct of such surveys is not an informal procedure. Rather it is a series of logical and interconnected steps, which progresses towards the final end product of the survey.

At the disaggregate level, the travel behaviour characteristics under real-time information are captured in a dynamic environment. To accomplish such an objective, two major methods are generally employed to gather behavioural data, namely Revealed Preference (RP) and Stated Preference (SP) methods. RP and SP data are the outcomes of surveys that elicit responses regarding decisions made about one or more preferences. RP data are concerned with actual choices and/or behaviours in real transport environments, while SP data are concerned with travel choices and/or behaviours in hypothetical transport environments (Srinivasan, 2000). Both of these methods entail a wide array of possible ways of asking commuters about their preferences in terms of mode choices, manner of using options, usage frequencies and so on (Hensher et al., 2008).

(a) Revealed Preference (RP) Method

RP information basically covers the data collected in existing conditions imparting information about the current situations. By definition, the RP data describe only those alternatives that exist, which implies that the modelled results cover only existing attribute levels and attendant correlation between attributes. This limits the freedom of allocating specific attributes to the alternative(s) under consideration and there may not be sufficient variability to permit identification of certain levels of personal attributes of income or age, etc. RP data do, however, possess high reliability and face validity. Though RP data are well suited for short-term forecasting, they are quite inflexible and are often not suitable for forecasting to a market other than the historical one (Louviere et al., 2000).

Ortuzar (2002) has highlighted that RP data are often based on recall of past experience, which can be biased or unrealistic with the passage of time (between survey and actual events). Furthermore, the data collected using RP methods on new schemes may not provide sufficient variability for constructing good model for evaluation and forecasting purpose. The respondent's observed behaviour might well be dominated by a few factors which make it rather difficult to detect the relative importance of other variables. The RP method is expensive and inflexible to the extent that RP data are not considered as being well suited for modelling responses to new schemes.

(b) Stated Preference (SP) Method

In cognisance of the RP's limitations, the stated preference (SP) method is used predominantly in the modelling of discrete choice behaviour given the suitability of SP data when evaluating new choice alternatives. The SP method does appear to be able to better simulate environmental changes, such as changes to urban structure, transport systems and policies (Brownstone, 2002). The SP approach is based on systematically generating situations by varying alternatives, attribute levels, whereby the respondents state their preference of alternatives in each choice set. Typically, SP data are managed by ranking, rating or first-choice methods with the first-choice method being considered the most appropriate approach in undertaking choice behaviour analysis. However, it is noted that respondents might have similar or even the same preference of different alternatives when stating their preference of a given choice set under conditional trade-off of attributes (Tzeng and Chiu, 1997).

The development of a better understanding of the dynamic aspects of travel behaviour is a major challenge in behavioural research. SP method offers a number of potential advantages in this context, including the capability to study behaviours that occur infrequently or are otherwise difficult to observe, and the ability to control for covarying factors and to distinguish between genuine changes as opposed to random variations in behaviour. However, the major issue that confronts the potential use of SP method in the context of behavioural dynamics is the implicit time scale of SP responses. One way to deal with this issue is to attempt to

identify such factors that might impede adaptation in behaviour within the SP exercise itself, and make some estimate of the distribution of these factors within the population. While one of the great advantages of the SP method is its ability to deal with new and possibly quite radically different travel options, the practical experience is that SP works best in cases of marginal behavioural change, where the new behaviour closely mirrors the aspect of existing behaviour (Polak and Jones, 1997; Hensher et al., 2008). In past studies, the combination of RP and SP data has mostly focused on a single outcome such as choice of mode, choice of destination (Brownstone and Train, 1999; Hensher, 1999; Hensher et al., 2008), though some of these choices co-occur and give rise to second, third or higher order choices.

(c) *Advantages and Disadvantages of RP and SP Methods*

The RP Survey data concerns choice between the existing scenarios or prevailing situations. It has been demonstrated that when RP data are not measured with a high level of precision, model structures and functional forms which would be appropriate with a fully disaggregate (i.e. properly measured) data set may not be selected leading to unknown bias in forecasting (Daly and Ortúzar 1990).

SP data, designed to overcome most RP problems, allow researchers to have good quality information (since the design is under the analyst's control). However, use of SP data may mask a potentially large problem: it is easy to achieve what appears to be good modelling results with almost any SP survey, but if the technique is not used appropriately (for example using a non-customised design in a general context instead of focusing on specific behaviour), serious problems may remain undetected until forecasts are eventually compared with actual outcomes (Ortúzar and Willumsen 2001).

The SP Survey data consider the choices related to futuristic situations or existing scenarios with futuristic situations. A SP survey can be tailored to understand very specific aspects of behaviour. The design and execution of the survey needs to consider the problem that is under consideration, and a choice situation relating to the problem needs to be selected. The attributes describing the alternatives need to

be identified, and the survey form needs to be designed and the technology needs to be selected.

The SP data are analysed by applying discrete choice models, which assign values to the different attributes that were included in the survey. Positive values indicate components that make the configuration more attractive, while negative values indicate components that make respondents less likely to choose an alternative. The mixed RP/SP data set using the artificial tree structure method can be used to estimate Logit models that allow for correlation among RP alternatives.

2.5 TRAVEL BEHAVIOUR MEASUREMENTS IN THE INFORMATION RICH ENVIRONMENT

In the course of modelling the behavioural changes under the influence of real-time travel information, the time dimension in travellers' behaviour is critically important, as existing information services can influence future decisions. Several measurement issues arise in this context of observing travellers' decision dynamics in the presence of real-time travel information and they have been summarised (Mahmassani and Herman, 1990; Srinivasan and Mahmassani, 2003; Srinivasan and Guo, 2004) as follows:

- (a) Commuters' decision dynamics must be measured simultaneously with associated service levels in congested systems.
- (b) The observational framework should allow for the consequences of commuters' decisions to depend on the result of their own decisions and those of a large number of other decision-makers.
- (c) The process must be observed in a realistic setting with complex interaction occurring between the users and the traffic system.

Despite the obvious need for assessing users' acceptance and the potential impacts of these systems in terms of improving traffic conditions for individual drivers and the overall transportation system, there has been a lack of models to evaluate their full impacts (Ben-Akiva et al., 1997). The range of approaches being applied to observe and investigate includes analytical models, field surveys, operational tests,

laboratory experiments, and simulation models. The idealised analytical models are not appropriate for observing commuters' behaviour dynamics as they lack the ability of representing decision-making processes occurring in a complex dynamic environment. On the other hand, treating travel behaviour within the framework of discrete choice models may not be that effective because, with the spatial and temporal dimensions, the choice set becomes astronomically large if one wishes to gain adequate levels of precision in forecasts (Kitamura et al., 2000). The traveller information systems have also been evaluated through operational tests and travel surveys (Bonsall and Parry, 1991). Given the fact that the real-time travel information systems are still progressing, it may perhaps be a premature attempt to conduct operational tests and surveys, particularly considering the lack of theoretical constructs to guide acquisition and analysis of such observations (Mahmassani and Herman, 1990). Furthermore, these tests and experiments are very expensive to conduct and do not allow for effective evaluation of different alternatives.

Realising the limitations of other approaches with respect to the dynamics of travellers' behaviour under the influence of real-time information, several researchers have invested considerable effort in laboratory experiments conducted using simulators (Adler et al., 1993; Chen and Mahmassani, 1993; Koutsopoulos et al., 1994; and Miller, 1996). As compared to operational tests, the computer simulation models allow for testing alternative system designs before conducting operational tests. These laboratory experiments enjoy the advantages of experimental control, limited resource requirements, and manageable number of participants (Srinivasan, 2000). The simulation can replicate the behaviour of complex systems or processes, and is therefore suited for the representation of travel behaviour, which is a complex behaviour. The factors that make travel behaviour complex include: the multitude of contributing factors and decision rules involved; constraints that govern the behaviour; inter-personal interactions; multiple planning horizons; and the complexity of activity-travel decision making (Pas, 1990). Thus, simulation is an effective approach to study such a complex phenomenon which facilitates its practical, yet realistic, representation.

2.6 TRAFFIC/TRANSPORTATION SIMULATION TOOLS

The advancement of computer technology and provision of sophisticated modelling options have made simulation an effective tool. The traffic/transportation simulation models have become increasingly popular for analysing a variety of dynamic problems, which are not amenable to study by other means (May, 1990). Many simulation models are capable of simulating individual vehicle's movement in real time on large networks in personal computers, with graphical animations (Yang et al., 2000; Quadstone, 2008; Gettman, 2008). A computer-based simulation program provides the analyst with a platform for describing the network and traffic characteristics using suitable interfaces (Miller et al., 2004). The numerical results from simulation provide the analyst with detailed quantitative description of what is likely to happen (Chandrasekar et al., 2002). Basically, simulation is used as an evaluation tool. It is treated as a real world setting where algorithms, models, and control strategies can be easily tested.

If the simulation model is used to explore design issues such as the relative effectiveness of different information channels, the effect of reducing time lag between the receipt of information in a control centre and its onward transmission to travellers, or the advantage to be gained by broadcasting forecasts of traffic conditions, it becomes necessary to represent system dynamics in some details. This implies the use of simulation model that can represent the dynamics of travellers' behaviour as well as the traveller information components. In such models, the travellers' knowledge-base could be modified to an appropriate extent and at an appropriate instant depending on the provision of information. A number of simulation models have been designed specifically to explore the dynamic effects of information provision on network performance. Interesting practical examples include INTEGRATION (Rakha and Van Aerde 1996), THOREAU (Wang and Glassco, 1995), DYNASMART (Jayakrishnan et al., 2001), AIMSUN2 (Barceló and Garda, 2007), and PARAMICS (PTV AG, 2001).

2.6.1 INTEGRATION

INTEGRATION (Rakha and Van Aerde 1996) model was conceived during the mid 1980's as an integrated simulation and traffic assignment model. What makes the model unique is that the model's approach utilises the same traffic flow logic to represent both freeway and signalised links, and that both the simulation and the traffic assignment components are also microscopic, integrated and dynamic. In order to achieve this mix of attributes, traffic flow is represented as a series of individual vehicles that each follows macroscopic traffic flow and assignment relationships. The combined use of individual vehicles and macroscopic flow theory has resulted in the model being considered mesoscopic by some. During the 1990's the INTEGRATION model evolved considerably from the original mesoscopic roots. This evolution took place through the addition, enhancement and refinement of various new features during the application of the model in the classroom as well as in the field. Some of these improvements have enhanced the fundamental traffic flow model, such as the addition of car-following logic, lane-changing logic, and more dynamic traffic assignment routines. However, others have extended the model's application domain, such as the inclusion of features for modelling toll plazas, vehicle emissions, weaving sections, and Heavy Occupancy Vehicles (HOVs). In addition, some features, such as the real-time graphics animation and the extensive vehicle probe statistics, have been added to simply make the model easier to understand, validate and calibrate.

2.6.2 DYNASMART

DYNASMART (Jayakrishnan et al., 2001) is a macroscopic model for ATMIS that incorporates efficient modelling of path dynamics for large urban networks. This model was developed at the University of Texas at Austin and at the University of California at Irvine. DYNASMART was specifically developed for studying the effectiveness of alternative information-supplying strategies, as well as alternative information/control system configurations. DYNASMART was originally used as a simulation tool to find dynamic assignment solutions and was extended to multi-user class real-time assignment. The model is based on simulating individual vehicle movements according to a macroscopic flow model, with the driver path

selection behaviour under information being explicitly modelled. The path-processing component is designed for efficient application of the framework to large and realistic networks. DYNASMART does model individual vehicles, though based on macro flow models. Due to the idealised network links, the number of nodes in the network model may not be significantly higher than the decision nodes in the actual network. Using macroscopic traffic speed-flow equations, DYNASMART models link travel times as well as network level traffic details such as path travel times effectively. The path dynamics is modelled based on the route or routes that drivers have in their minds, as well as the routes provided by ATMIS guidance. The flexibility for modelling various driver response mechanisms and information supply strategies comes from the ability to find and store multiple paths efficiently, using networks of reasonable sizes.

2.6.3 AIMSUN2

AIMSUN2 (Barceló and Garda, 2007) is a microscopic traffic simulator developed to analyse ITS components. It can deal with different traffic networks such as urban networks, freeways, highways, ring roads, arterial and any combination of them. The traffic loading patterns in AIMSUN2 are based on two different types of simulation environments. In one simulation environment, the traffic inputs are traffic flows and turning proportions and vehicles are distributed across stochastically across the network. In the other simulation environment, O-D matrices and route selection models are used to assign vehicles to specific routes from the start of their journey to their destination. The vehicle behaviour models (car-following, lane-change, gap-acceptance, etc.) are functions of several parameters that allow modelling of different types of vehicles: cars, buses, trucks, etc. They can be classified into groups, and reserved lanes for given groups can also be taken into account. Due to the detailed modelling of each vehicle in the network, AIMSUN2 can simulate any kind of measurable traffic detectors. AIMSUN2 has a powerful Graphical User Interface (GUI) which allows the user to access any information in the model and define traffic incidents before or during the simulation run. A list of incidents may be stored for use in subsequent simulation runs. To cope with the requirements of simulating ATIS, specific Dynamic Linking Library

(DLL) has been developed. This library gives AIMSUN2 the ability to communicate with almost any of the following external applications:

- (a) Adaptive Traffic Control, Traffic Management Systems and Incident Management Systems;
- (b) Vehicle Guidance, Fuel Consumption and Emissions;
- (c) Public Vehicle Scheduling and Control Systems.

2.6.4 MITSIMLab

MITSIMLab (Yanget al., 2000) is a microscopic traffic simulation laboratory developed for the design and evaluation of advanced traffic management systems and advanced traveller information systems. It consists of a Microscopic Traffic Simulator (MITSIM) and a Traffic Management Simulator (TMS). MITSIM models traffic flows in the network at the vehicle level, including driver behaviour. It uses a microscopic simulation approach, in which movements of individual vehicles and operations of traffic control and surveillance devices are represented in detail. This representation is necessary for evaluating dynamic traffic management systems at the operational level, since it allows for capturing the stochastic nature of traffic flow, drivers' response to real-time traffic information, and operations of surveillance sensors. TMS mimics the logic behind the traffic control and traveller information systems under evaluation. TMS has a generic structure that allows testing of a wide range of control and guidance strategies. It supports a rolling horizon implementation of control and route guidance and is capable of simulating ATMS/ATIS systems with advanced capabilities including traffic prediction. The traffic control and route guidance generated in TMS, according to the strategies to be evaluated, are fed into MITSIM and that affects the behaviour of individual drivers and hence, traffic flow characteristics. The changes in traffic flows are in turn measured by the surveillance system which provides TMS the traffic information that is utilised to generate control and routing strategies. The interaction between the traffic flows in the network and the control and route guidance is a critical element for modelling dynamic traffic management systems.

MITSIMLab provides a laboratory environment for the coupling of traffic management with traffic flows and is designed to represent a wide range of traffic management systems, model drivers' response to real-time traffic information and controls, and calculate measures of effectiveness that are necessary for the evaluation of traffic management systems and road network designs. MITSIMLab has an integrated graphical user interface (GUI) for visualising the simulation process. The GUI features animation of the vehicle movements, graphical display of traffic data and state of control devices. It is an essential tool for verification of input data and presentation of simulation output. MITSIMLab is implemented in C++ using the object-oriented programming model and can operate in a distributed environment (Yang, 2000).

2.6.5 VISSIM

VISSIM (PTV AG, 2001) is a microscopic, time step and behaviour-based simulation model developed to model urban traffic and public transit operations. The program can analyse traffic and transit operations under constraints such as lane configuration, traffic composition, traffic signals, transit stops, etc., thus making it a useful tool for the evaluation of various alternatives based on transportation engineering and planning measures of effectiveness.

VISSIM can be applied as a useful tool in a variety of transportation problem settings. The simulation package consists internally of two different programs, exchanging detector calls and signal status through an interface. The simulation generates an online animation of traffic operations and offline output files that gather statistical data such as travel times and queue lengths.

The traffic simulator in VISSIM is a microscopic traffic flow simulation model including car following and lane changing logics. The signal state generator is signal control software polling detector information from the traffic simulator on a discrete time step basis (as small as one tenth of a second). It then determines the signal status for the following time stamp and returns this information to the traffic simulator. VISSIM simulates traffic flow by moving "driver-vehicle-units" through a network. Every driver with specific behaviour characteristics is assigned to a

specific vehicle. Consequently, the driver behaviour corresponds to the technical capabilities of his/her vehicle. Essential to the accuracy of a traffic simulation model is the quality of the actual modelling of vehicles for example the methodology of moving vehicles through the network. In contrast to less complex models using constant speeds and deterministic car-following logic, VISSIM uses the psychophysical driver behaviour model.

The basic concept of this model is that the driver of a faster moving vehicle starts to decelerate to become a slower moving vehicle as he/she reaches his individual perception threshold. Since he/she cannot exactly determine the speed of that vehicle, the speed will fall below that vehicle's speed until he/she accelerates again after reaching another perception threshold. This results in an iterative process of acceleration and deceleration. Stochastic distributions of speed and spacing thresholds replicate individual driver behaviour characteristics. VISSIM's traffic simulator not only allows drivers on multiple lane roadways to yield for two preceding vehicles, but also two neighbouring vehicles on the adjacent travel lanes (Verkehr, 2000).

2.6.6 PARAMICS

PARAMICS (Quadstone Ltd., 2008) is an advanced suite of software tools for microscopic traffic simulation developed by Quadstone Limited. The complete suite of PARAMICS software comprises five software modules:

- (a) **Modeller:** It performs the three fundamental operations of model build, traffic simulation (with 3-D visualisation) and statistical output accessible through a powerful and intuitive graphical user interface.
- (b) **Processor:** It is a simulation configuration tool that allows the user to set up network simulations to be run in batch mode.
- (c) **Analyser:** It is an analysis tool for displaying the outputs obtained from PARAMICS traffic simulation. The primary aim of Analyser is to display and report on statistical data produced by running the simulation through Modeller and/or Processor.

- (d) Programmer: It is an Application Programming Interface for traffic modelling research.
- (e) Monitor: It is the interface to pollution emissions model and calculates the levels of traffic emission pollution on a road network.

The traffic simulation model in PARAMICS follows a deterministic, fixed time step approach. At each time step, individual vehicles are moved at the prevailing local speed on the same link or transferred to another link, and modelled in fine detail for the duration of their entire trip. The modelling provides accurate traffic flow, transit time and congestion information. PARAMICS allows vehicle routing according to routing tables but does not allow storage of sufficient path trees and storage of individual vehicle's routes, which are essential requirements for the simulation of driver response towards information supply and the resulting route choice. The primary difficulty with microscopic simulations is the inability to handle path dynamics in large networks. The problem arises from the detailed network descriptions used in such microscopic simulation models. PARAMICS scalability permits vehicle simulation of very large networks with additional processors, but if detailed driver response modelling and path processing are to be incorporated, the model can only be used to simulate small to medium-sized urban areas. This is because many network algorithms show nonlinear increase in storage and computational requirements as network sizes increase (Jayakrishnan, 2001).

PARAMICS interfaces to standard data formats, and such data can be obtained from induction loops and optical sensor installed in the field. It excels in modelling highly congested networks and ITS infrastructures including a variety of traffic management, information and control strategies. One of the important features of PARAMICS model is that it can be customised. Access to the core model is available through a functional interface or Application Programming Interface (API). At the fundamental level of the PARAMICS simulator, a number of functions are used to control the simulation environment and with the provision of API these functions can be controlled.

An API module exchanges dynamic data with the core PARAMICS model and other API modules through the Dynamic Link Library (DLL) mechanism. These

modules operate on data structures that describe individual vehicles and the environment around a particular vehicle. The API facility widely supports the modellers and researchers to fine tune the driving behaviour of simulated drivers, vehicle models and parameters to reproduce specific observed behaviour. With the aid of API, the core models can be changed to other models for designing specific simulation environment.

PARAMICS limitations include a lack of equilibrium traffic assignment, and limited options in modelling traveller information systems. It means that the model updates the routing instructions at each intersection instead of being path based, which may result in myopic travel paths with extensive twists and turns. Other weaknesses include not being able to explicitly model a number of control options such as bus signal pre-emption from mixed lanes, and limited user options in modelling incidents and work zones (Boxill and Yu, 2000).

While all the above mentioned simulation models have been successfully applied in particular studies, a common shortcoming is the relatively limited range of their applications. Some of them are designed for particular applications and are useful only for specific purposes, while others do not support the provision of integrated traveller information systems and transit operations. No model has the integrated component and functionality required for evaluation of dynamic control and route guidance strategies in a multimodal environment on surface street networks. The lack of an integrated simulation environment with realistic user behaviour for real-time traffic management studies has become a bottleneck in APTS and ATIS research and development. The need for the development of a more sophisticated methodology and simulation tools has been pointed out by researchers, including Santiago and Kanaan (1993), and Grotenhuis, (2005).

The best developed application in the area of transportation network simulation models, in which a number of operational (and often commercially supplied) software packages exist, is the microsimulation. This programming technique has been applied with increasing frequency over the past decade or more in the field of transportation system analysis. In microsimulation the dynamic behaviour of

individual agents is explicitly simulated over both time and space to generate aggregate system behaviour.

The operations of individual road and/or transit vehicles are modelled second-by-second. Examples of such micro simulators include PARAMICS (Quadstone, 2008), INTEGRATION (Rakha and Van Aerde 1996), DYNASMART (Mahmassani, 2002), and VISSIM (PTV, 2008). Models of urban travel behaviour in which the temporal, spatial and modal distributions of trips in an urban area are predicted, are increasingly being developed and applied in a microsimulation framework (Bonsall and Parry, 1991; Kikuchi et al., 2002; Miller et al., 2004; Bradley et al., 2006). Extensive reviews of microsimulation applications in transportation are provided by Miller (2006) and Miller and Salvini (2001), which also place transportation-related microsimulation modelling within the larger context of the microsimulation modelling literature as a whole.

2.7 AGENT BASED SIMULATION MODEL

A long tradition exists in the field of transportation to bring the various transportation-related components into an integrated modelling framework. An ideal model is sought that is comprehensive, logically consistent and theoretically sound. The provision of current theories, modelling methods and computational capabilities (Miller et al., 2004) provides an open platform for extensive discussion of both the technical and policy rationales for developing such integrated transportation system models.

Research focusing into development of integrated transportation models has shown that many past models either had failed to become fully operational or be sufficiently useful as policy analysis tools. Integrated models are inherently computationally intensive, data hungry, require extreme demands on theoretical understanding, require complicated methodological capabilities for capturing theoretical understanding within operational computer code, and extremely difficult in estimating and validating of the models (Miller et al., 2004). Nevertheless, with rapid progress in the past few decades, computational capabilities have improved far beyond expectations.

Similarly, great advancements have been made in the modelling methodology, empirical and theoretical understanding of spatial and temporal processes, and data resources to support modelling activities. Research and development with respect to integrated models has thus proceeded to the point that operational models are in use in a number of locations worldwide (Dia, 2002).

Currently operational integrated models, including several commercially available software packages, are not sufficiently capable whereby all the existing ITS components can be evaluated. Nevertheless some of these software packages do provide researchers with very strong API modules which can communicate with the core model or other programs. Such capability provides the platform to further develop available models in the desired manner so that they can be utilised for specific objectives.

As in many modelling applications, a more disaggregated approach to modelling socio-economic processes such as travel behaviour etc. is generally desirable in order to reduce model aggregation bias, enhance its behavioural fidelity, etc. (Goulias and Kitamura, 1992). Similarly, it is increasingly recognised that the dynamic evolution of ITS systems must be explicitly captured if future ITS system states are to be properly estimated. Putting these two observations together leads inevitably to the adoption of a sophisticated microsimulation approach to model such ITS systems (Miller et al., 2004; Grotenhuis, 2005).

To bring the simulation closer to reality, production-rule system approach can be adopted. The premise underlying production rule systems is that choice behaviour in a certain context can be better described by a series of “if-then” rules. If the condition in the “if” part is true, then the action of the “then” part follows (Clark and Smith, 1993). This approach utilises an external element (agent) that is integrated with the system to rationalise the “if-then” process within the system. Such an approach is known as agent-based simulation with the potential to model complicated cognitive processes (Russell and Norvig, 2003).

Agent-based models are ones in which each individual actor within the system of interest is modelled as an autonomous agent, possesses identity, attributes and the

capability to behave i.e. to make decisions and to act within the system. It is basically a combination of software components which can perceive through sensors, think with some knowledge and take actions via effectors (Wooldridge and Jennings, 1995).

Agent-based modelling has been recognised as an extremely powerful design paradigm across virtually the gamut of socio-economic modelling, including travel-related behaviour (Pollack and Ringuette, 1990, Dia, 2002). It provides an extremely efficient, effective and natural way of both conceptualising and implementing complex, dynamic, disaggregate models of human decision-making. As the software industry-standard approaches to designing and programming complex software systems with the programming languages (Java, C++, etc.), agent-based models have exploded as a practical, operational possibility over the last decade or more with the emergence of object-oriented design principles (Taylor, 1990).

Research of agent-based models has typically focused on how the technology can improve the efficiency with which a user makes a decision, and improve the effectiveness of decision-making (Weiss, 2001). Over the last two decades research has evolved to include several additional concepts and views and have extended the scope from personal or small group use to the corporate level. With the use of knowledge-based decision support systems and artificial intelligence, the agent can provide smarter support for the decision-makers (Timmermans et al., 2002).

In this framework, intelligence comprises the search for problems, design involves the development of alternatives, and choice consists of analysing the alternatives and choosing one for implementation. An intelligent agent deals with the problem subjected to its knowledge-based rule and gives the judgement on behalf of the decision-maker (Pearson and Shim, 1995). It is necessary to mention that the complexity of the behavioural rules and the total number of agents do strongly affect the computational requirements of agent-based systems (Dia, 2000). And it has also been found from other studies that agent-based systems are suitable for implementation in real-time traffic management systems as compared to traditional approaches (Park et al., 2007).

The verification of the Intelligent Agent (IA) involves the assessment of its adherence to the specifications. It is normally required to demonstrate the consistency and completeness of the knowledge (rules) and the correctness of the inference process (logical verification). Furthermore, the validation of IA is concerned with assessing the quality of the decision, thus focusing on its effectiveness. Relevant aspects of the IA validation are also related to how it addresses end-user needs, in terms of user acceptance and usefulness in the field. Validation involves testing IA to ascertain whether it achieves acceptable performance levels.

In recent years, the validation of IA has been recognised as the corner stone of the evaluation process (Fillipo, 2001). Attempts to develop true agent-based microsimulation models of transportation network and travel behaviour interactions have been relatively rare to date. Either because such models were developed prior to the agent/object-based approach was widely understood and operationally practical (Mackett, 1985; Miller, 2004; Mackett, 1990) or they had evolved over time out of more aggregate, non-agent-based approaches.

The purpose of the present research is to describe and to develop a fully agent-based, integrated micro-simulation model that can evaluate the impacts of real-time multimodal traveller information on travellers' behaviour. It is hypothesised that agent-based micro-simulation, in fact, represents the best approach currently available to model large, complex, dynamic, and open-ended transportation systems which can also simulate the travellers' choice behaviour with respect to their socio-economic and travel characteristics (Park et al., 2007). In this regard it is believed that the agent-based microsimulation models may prove to be the most computationally efficient and practical approach to model highly complex systems. Nevertheless, classic issues in developing and applying integrated models (which include data requirements, computational feasibility, model parameter estimation, model validation, robustness of model results, etc.) clearly exist and must be addressed in detail before the hypothesis of the agent-based, micro-simulation approach can be accepted (Miller, 2004).

2.8 CONCLUSION

This chapter provides a comprehensive review of current research and operational activities with respect to commuters' travel choice decision-making process and the modelling approaches within the framework of the objectives that are stated in Chapter 1. The travel choice decisions are influenced by the provision of travel information taking place at two principal instances, one of which is at the beginning of the trip. It is the pre-trip information that greatly influences the travel decisions at the beginning of the trip which can reduce the degree of uncertainty when travellers encounter travel choice variation.

Furthermore, the success of transportation system is also greatly dependent on the ability to provide useful travel information that is available round the clock and is accessible to all travellers. The provision of traveller information can be differentiated into three categories namely, Unimodal Traveller Information (UTI), Multimodal Traveller Information (MTI) and Integrated Multimodal Traveller Information (IMTI). The IMTI provides information concerning different mode choice options in response to a particular journey specified by the user. The impact of these traveller information systems has been analysed by different researchers and it has been found that commuters might value and use pre-trip information more than en-route information as it gives them the situation on their routes in advance, thereby allowing them to change their travel plan in advance.

However, from the transportation system's perspective, the potential efficiency and success of various policy measures for eliminating or reducing traffic problems in metropolitan areas largely depend on the level of public acceptance of the traveller information. A variety of travel behaviour models based on revealed and stated preference surveys have been developed to study the nature and dynamics of traveller's choice behaviour. Despite the obvious need for assessing user acceptance and the potential impacts of these systems in terms of improving traffic conditions for individual drivers and the overall transportation system, there is still a lack of models to evaluate their full impacts. Realising the limitations of other approaches with respect to the dynamics of travellers' behaviour under the influence of real-

time information, several researches have invested considerable effort in laboratory experiments being conducted with the support of the simulation models.

Currently, the simulation models are not sufficiently capable such that all the existing ITS components can be evaluated. Furthermore, in many modelling applications a more disaggregated approach to model socio-economic processes is generally desirable to reduce the model aggregation. These two observations together have led inevitably to the adoption of a sophisticated micro-simulation approach to model the ITS systems in an intelligent environment where the simulation is closer to reality based on the production-rule system approach.

The premise underlying production-rule system is that choice behaviour in a certain context can be better described by a series of “if-then” rules. If the condition in the “if” part is true, then the action of the “then” part follows. This approach utilises an external element (agent) that is integrated with the system to rationalise the “if-then” process within the system. Such an approach is known as agent-based simulation and has been shown to have the potential to model complicated cognitive processes. This is also the approach taken in the present research.

CHAPTER 3

METHODOLOGY

3.1 INTRODUCTION

The prime objective of this research is to recommend a *modus operandi* for the design and development of a transportation network model that is capable of evaluating the impacts of integrated real-time traveller information on commuters' behaviour in different scenarios. Adopting the simulation-based approach, an intelligent transportation network simulation model (INSIM) is developed. The INSIM consists of two main components: INSIM Expert System and transportation network simulation model. The integration of INSIM Expert System with the transportation network simulation model can enhance the simulation capabilities of network simulation model. Furthermore, it can provide the opportunities to simulate the commuters' travel behaviour and the behaviour of any traveller information provider in a multimodal environment. The impacts of integrated traveller information on the commuters' travel behaviour that changes the overall travel environment within the multimodal transportation system has not been fully integrated and analysed through the analytical modelling techniques. The use of INSIM to simulate traffic system performance and to evaluate different integrated traveller information schemes can overcome the limitations that are related to the analytical formulation of the prime objective.

To analyse the working mechanism of INSIM a set of simulation experiments have been designed. These experiments have a two-fold impact; one is to validate the working mechanism of INSIM, and the other is to recommend traveller information supply strategies that can improve the transportation system performance. The effectiveness of different information supply strategies can be evaluated in a congested environment, during which the travelling time as well as the travelling cost are the highest; if travel information is made available to the travellers, they may decide to switch mode and/or route in order to minimise relevant travel time and delay. As the provision of relevant traveller information can affect the

commuters' travel behaviour, it provides the network controller with the capability to use information as a tool in achieving better spatial and temporal travel demand distribution in the network.

A working procedure to achieve the prime objective of this research is presented in Figure 3.1. There are five inter-linked sections in the working procedure that are as follows:

- (a) Travel behaviour survey;
- (b) Travel behaviour modelling;
- (c) Rule-based expert system;
- (d) Intelligent transportation network simulation model; and
- (e) Evaluation of an integrated traveller information system;

These five sections are described in this chapter as follows. In Sections 3.2, the details about a travel behaviour survey are discussed. This survey was conducted to gather existing and proposed travelling conditions and the corresponding travel behaviour of commuters. Section 3.3 discusses the travel behaviour modelling techniques and the analysis approach to evaluate the impact of traveller information. The methodology to design and develop the rule-based expert system for INSIM is presented in Section 3.4. The ways to integrate different components of INSIM and the microscopic traffic simulation model via Application Programming Interface (API) are detailed in Section 3.5. Lastly, the strategic approach to evaluate different integrated traveller information systems and their impact on network performance is described in Section 3.6.

3.2 Travel Behaviour Survey

The understanding of commuters' travel behaviour is a key requirement to achieve the prime objective of this research. In this regard, a travel behaviour survey was conducted to gather the commuters' travel choice decisions with respect to public versus private modes in congested but information-rich multimodal transportation environment, specifically considering the influence of integrated traveller information.

Figure 3.1 Details of the Working Procedure of Research

The traveller information was assumed to be integrated for which comparison was provided on all available modes along with a recommended mode and/or route of travel. The survey focused on the perception and reaction of the commuters under the following conditions: at different levels of congestion, frequency of facing certain congestion, types of information provided, and commuters' mode or route switching propensity.

3.2.1 Study Area

To find insights and details about commuters' travel behaviour in an information-rich environment, it is necessary to select a study area where a multimodal transportation system exists, and the commuter population has ready access to the Traveller Information System (TIS). Based on these requirements, Singapore is a suitable place where a multimodal transportation system is in operation for a long time, and the commuters have access to TIS. The transportation system in Singapore supports commuter mobility by public and private modes of transport, with seamless connectivity and a high level of integration (Lam and Toan, 2006). The prevailing TIS provides information on different modes of transport, along with certain personal requirements like shortest path or cheapest path.

According to the statistics published by Department of Statistics, Singapore (DOS) the total population was 4.839 million residing in a land area of 710.2 square kilometres (as of April 2009). The population of Singapore-registered motor vehicles was 851,336 units of which there were 451,745 cars and 24,446 taxis. The total road length was 3,297 km, which includes 153 km of expressways. Public transport was served by well-integrated bus, Light Rail Transit (LRT), and Mass Rapid Transit (MRT) services throughout Singapore. There were 22 bus interchanges and 15 bus terminals. The overall bus network consisted of 4,560 bus stops located throughout Singapore. The average bus fleet were 3,255 buses and the number of bus routes in operation was 325 services. The total length of MRT system was 109.4 km and the LRT system was 28.8 km. The average daily ridership for MRT was 1,564,000 passenger-trips, and for LRT it was 81,000 passenger-trips. These statistics apply to the period of 2007 to 2008.

The prevailing traveller information systems in Singapore serve to provide traffic updates for private and public modes of transport. Information for private mode users is provided via internet, radio, hand/mobile phones, and Electronic Monitoring and Advisory System (EMAS). Information can be obtained regarding the travel time, travel distance, incidents, road works, etc. Information for public mode users is provided by travel guides, travel pamphlets, internet, and kiosks at bus interchanges, MRT/LRT stations, and MRT trains.

3.2.2 Data Overview

Knowledge about commuters' perception and reaction to the travel information as provided can be analysed in relation to commuters' socio-economic, travel and behavioural characteristics. In this regard, a survey was conducted based on a combination of Revealed Preference (RP) and Stated Preference (SP) methods. The combined RP-SP approach allows the development of SP scenarios that can be easily conceptualised by the respondents given that the scenarios are based on actual travel information. In particular, Hensher and Swait (2000) stated that the strength of RP and SP data sources could be exploited and weaknesses ameliorated by pooling both data sources. This process is called enrichment and provides more robust parameter estimates and increases confidence and accuracy in predictions (Ben-Akiva et al., 1994; Hensher and Bradley, 1993).

The RP component is aimed at gathering commuters' socio-economic and travel characteristics, their knowledge/experience about actual scenarios reflecting different levels of congestion, and available traveller information options (Ben-Akiva and Lerman, 1985). The RP component in this survey was divided into four parts in which information about commuters' socio-economic characteristics, usual trip preferences, availability and usage of public and private modes was gathered.

The commuters' socio-economic characteristics were surveyed in order to observe the impact of these characteristics on commuters' travel choice preferences (Ben-Akiva and Lerman, 1979). The socio-economic characteristics often taken into consideration are: age, gender, education, occupation, personal income, household income, household size, postcode, marital status and type of vehicle owned (Dia,

2002). For this survey, the socio-economic attributes were covered in the first part of the RP component. Age was divided into four groups of 18 to 35, 36 to 45, 46 to 55, and above 55 years. Gender was a binary variable, corresponding to two groups i.e. males and females. Education was divided into four groups of A-level and below, bachelors, masters, and PhD degrees. The occupational categories considered were: student, academics, administrative, technical, professional, housewife, and retired. Personal monthly income was divided into five groups of S\$1500 (US\$ 990) or less, S\$1501 (US\$ 991) to S\$3000 (US\$ 4530), S\$3001 (US\$ 4531) to S\$6000 (US\$ 9060), S\$6001 (US\$ 9061) to S\$12000 (US\$ 18120), and S\$12001(US\$ 18121) or above. Household monthly income was grouped into six categories of S\$2000 or less, S\$2001 to S\$4000, S\$4001 to S\$6000, S\$6001 to S\$8000, S\$8001 to S\$10000, and S\$10001 or above. Marital status was a binary variable i.e. married and unmarried. The available options for types of vehicle ownership were car, van, and small truck.

In the second, a third and fourth part of the RP component, information was gathered about the commuter's work/school trips. Work/school trips were considered because these trips are not discretionary and are usually made during the peak hours. Special consideration was given to capture the affects of the attributes that can influence the commuters' choice behaviour or have time and/or cost-related association. The information associated with trips was covered in the second part of the RP component. The trip variables provided knowledge about origin zone, destination zone, usual mode of transport, trip start time, acceptable delay, access to traveller information, and types of traveller information accessed. In the third part, information was gathered about availability of private mode option in the existing situation. The variables were total journey time, Electronic Road Pricing (ERP) charges, type of parking place (private or public), time required to find a parking place, parking charges, access and egress time to parking place, stoppage during the trip, nature of stoppage, and congestion experience. In the fourth part, information was gathered about the public mode option for work/school trips. The available modes were MRT, Bus, and Bus+MRT, which prevail in the existing transportation system of Singapore. The information was collected regarding access time, access

cost, egress time, egress cost, travel time, travel cost, total waiting time, number of transfers, and level of satisfaction.

The information gathered in RP component was then presented in the SP component while creating different hypothetical scenarios. The hypothetical scenarios were customised for each respondent based on the information provided in RP section. This approach resulted in creating realistic scenarios which can be conceptualised by the commuters as they reflected on certain parameters which existed in actual situations. There were five (5) SP scenarios, and the sequence of presenting these scenarios was designed in such a way that the respondent's willingness to access traveller information and to change his/her travel plan was first captured. In the second SP scenario, the impact of multimodal traveller information on his/her preferences to change travel mode was collected. In the third scenario, the respondents were presented with integrated traveller information and their willingness to change their travel choice decision in the previous i.e. second SP scenario was noted. The third scenario explicitly captured the commuters' mode switching propensity under the influence of integrated traveller information. The fourth scenario presented their work/school trips being commenced in a congested travel environment, and with the provision of integrated traveller information and certain incentives on public mode of transport, their mode choice behaviour was probed. Lastly, the fifth scenario captured the likelihood of commuters to access pre-trip or en-route information if they were to switch their mode of travel for their work/school trips.

Finally, based on the responses gathered from the sample of respondents, a database was developed. The captured data provided information regarding the influence of traveller information on commuters' travel behaviour. Approximately about two third of the available data was used to estimate the discrete choice models, while the remaining one third was utilised to evaluate the predictive performance of the estimated models.

3.2.3 Survey Instrument

Most of the travel behaviour surveys to date have been conducted as paper-pencil based interviews or telephone interviews (Dia, 2002; Khattak et al., 1995). These interviews have several advantages. They typically are less expensive to conduct because they reduce the cost associated with the failed attempts to find potential respondents. Telephone interview may also reduce the discomfort of respondents who are asked to describe sensitive behaviour. However, telephone interviews pose several problems with respect to the conceptualisation of SP scenarios. They do not offer visual aids, which can show the diagrammatic view of complete trips with dynamic changes in travel time, travel cost, etc.

Computer-aided surveys overcome the limitations of paper-pencil or telephone-based surveys. In the present survey, a computerised survey instrument was developed, which provided the respondents with the facility to read the survey questions from a computer screen. The computer programme automatically presented the respondent with the questions that were based on the information provided by the same respondent at the beginning of the survey. The dynamic interaction between the provided data and the generated scenarios greatly eased the administration of survey instruments, allowing complex scenario development, minimising the interviewer's error, and reducing the burden placed on the respondent by eliminating non-applicable questions. These features improved the internal consistency and overall quality of the collected data, both of which were further enhanced by the option of using built-in probes to internally reconcile any inconsistent responses.

The computer-aided survey instrument also allowed the interviewers to have respondents' answers to sensitive questions such as their age, personal income, or marital status. In these cases, the respondents had the privacy to enter his or her responses directly onto the computer keypad without directly sharing the information with any one else. There were, however some cases of illiteracy problem that limited the scope the interview as some respondents were not able to read the questions presented in English language. In such cases, the interviewer

would translate and read the question and corresponding categories to the respondent, as applicable.

The survey instrument was developed in Microsoft Visual Basic using Microsoft Access Database. Being a computer-aided survey instrument, it was able to simulate different scenarios and provide their complete graphical presentation with suitable images related to different mobility aspects. This helped the respondents to visualise and perceive the developed scenarios as per the research requirements, thus making it flexible in gathering existing travel characteristics and reproducing them in different scenarios at a later stage. The scenarios were designed to imitate the existing conditions for the RP component and the hypothetical travelling conditions for SP scenarios. With the help of such survey instrument, it was possible to find out the average level of congestion at which point travellers desire to access travel information, the mode or route switching propensity, and the level of acceptance to the provided or suggested travel options. On the whole, the survey instrument was a complete package to gather the commuters' travel behaviour with respect to the traveller information as provided.

3.2.4 Sampling Procedure

All travel surveys rely upon sampling techniques in which a part of total population is queried to make inference about the population as a whole. Sampling a population, rather than conducting a full population census, has the following advantages: economy, speediness and timeliness, feasibility and quality (Leslie, 1965). The broad types of sampling techniques in use include: simple random sampling, general stratified sampling, exogenous sampling choice based sampling, and enriched sampling. In this study, simple random sampling technique was used as there was no prior knowledge about the attributes of the commuters and every commuter had to be given same chance of selection. Random sampling technique ensures that bias is not introduced regarding the subject that is included in the survey (Stopher, 1993). The subjects were selected from the driver population located in the northern and eastern side of Singapore. The selected subjects were required to have their own valid driving license, and commute regularly by private mode of transport. The sampling procedure was such that the respondents did

broadly represent the targeted population of Singapore in terms of attributes like age, gender, education, income, nature of job, etc. The proportion of different strata corresponding to the commuters' socio-economic characteristics was based on data published by Department of Statistics, Singapore (DOS, 2004).

3.3 TRAVEL BEHAVIOUR MODELLING

The analysis of travel behaviour is typically disaggregated, meaning that the models represent the choice behaviour of individuals. Discrete choice models have been used to analyse and predict commuter's travel choice decisions (Ben-Akiva and Bierlaire, 1999). The data collected from the present travel behaviour survey were modelled to study the impact of different socio-economic and travel characteristics on commuters' travel behaviour. The analysis was based on the preferences of the commuters to change their travel plans or travelling modes corresponding to different hypothetical scenarios. These preferences provided the basis to stratify the sample and identify different behavioural groups within the sampled population.

The commuters' choice preferences were gathered in two different forms, such that either these preferences did not have any natural ordering, or they had some sequential ordering. The binary choice logit models were estimated to analyse the individual-specific outcomes corresponding to the non-ordered choices, and ordered probit models were estimated to analyse the ordered choices. Furthermore, a new dimension to the modelling effort was added by analysing the commuters' mode choice behaviour under the bounded rationality theory.

3.3.1 Mode Choice Logit Model

The mode choice logit model relates the dependent variable (Y) to the independent variables (X). The dependent variable is a discrete variable that represents the mode choice from a set of mutually exclusive choices i.e. public mode and private mode. The independent variables are presumed to affect the decision maker's choice, and represent a prior belief about the causal or associative elements important in the choice process (McFadden, 1987).

To calibrate the mode choice logit model, the observations on dependent variable (Y) have been randomly sampled from the commuter population of Singapore who take the private mode of transport as their usual commute mode. The choice decisions are based on a choice set with two alternatives i.e. public mode and private mode. These alternatives are characterised by a set of socio-economic and travel attributes, such that these attributes are independent in nature and are determined by the influences outside the model. Note that some attributes are generic to all alternatives, and some are specific to certain alternatives.

The decision rule is based on random utility theory, which assumes that the commuters have perfect discrimination capability, and their choice preference for an alternative is captured by a value called utility. The commuters select the alternative in the choice set with the highest utility. However, from the analytical point of view, it is difficult to have complete information and, therefore, uncertainty is taken into account for unobserved alternative attributes, unobserved taste variation, measurement errors, and proxy or instrumental variables (Hensher and Swait, 2000).

The utility is modelled as a random variable in order to reflect this uncertainty. More specifically, the utility that individual “ n ” associates with an alternative “ i ” in the choice set “ C_n ” is given by:

$$U_{in} = V_{in} + \varepsilon_{in} \quad (3.1)$$

where “ V_{in} ” is the deterministic (or systematic) part of the utility, “ ε_{in} ” is the random term, capturing uncertainty. The alternative with the highest utility is chosen. Therefore, the probability that alternative “ i ” is chosen by the commuter “ n ” from the choice set “ C_n ” is:

$$\Pr(i|C_n) = \Pr[U_{in} \geq U_{jn} \quad \forall j \in C_n] = \Pr\left[U_{in} = \max_{j \in C_n} U_{jn}\right] \quad (3.2a)$$

$$\Pr(i|C_n) = \Pr[V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}] \quad (3.2b)$$

$$\Pr(i|C_n) = \Pr[V_{in} - V_{jn} \geq \varepsilon_{jn} - \varepsilon_{in}] \quad (3.2c)$$

It should be noted that only the signs of the difference between utilities are relevant here, and not utilities themselves. The concept of ordinal utility is relevant here and not absolute. The means “ m_i ” and “ m_j ” of the random terms “ ε_{in} ” and “ ε_{jn} ” are assumed to be equal, such that $m_i = E[\varepsilon_{in}]$, then a new variable is defined:

$$e_{in} = \varepsilon_{in} - m_i + c, \text{ such that } E[e_{in}] = c \quad (3.3)$$

Now the probability that alternative “ i ” is chosen by the commuter “ n ” from the choice set “ C_n ” can be written as:

$$\Pr[U_{in} \geq U_{jn} \quad \forall j \in C_n] = \Pr[V_{in} + m_i + e_{in} \geq V_{jn} + m_j + e_{jn}] \quad (3.4)$$

In this model, the deterministic parts of the utilities are $V_{in} + m_i$ and the random terms are “ e_{in} ” (with mean c). The terms “ m_i ” are included as alternative specific constants that capture the means of the random terms. Thus, by including the alternative specific constants in the deterministic part of the utility function, without loss of generality, it can be assumed that the error terms of random utility models have a constant mean “ c ”. As only the differences between utilities are relevant, so the differences between alternative specific constants are relevant as well.

The deterministic term “ V_{in} ” of each alternative is a function of the attributes of the alternative itself and the characteristics of the commuter. That is:

$$V_{in} = V(z_{in}, S_n) \text{ and letting } x_{in} = (z_{in}, S_n) \text{ then} \quad (3.5a)$$

$$V_{in} = V(x_{in}) \quad (3.5b)$$

where “ z_{in} ” is the vector of attributes as perceived by the commuter “ n ” for alternative “ i ”, and “ S_n ” is the vector of characteristics of the commuters “ n ”. Linearity in parameters utility specification is assumed, thus the deterministic term is fully specified by the vector of parameters “ β ”, such that:

$$V_{in} = \sum_k \beta_k x_{ink} \text{ and } V_{jn} = \sum_k \beta_k x_{jnk} \quad (3.6a)$$

$$V_{in} = \beta x_{in} \text{ and } V_{jn} = \beta x_{jn} \quad (3.6b)$$

The logistic probability unit (logit model) is used in here. As the choice set has only two choices i.e. private mode (i) and public mode (j), the probability that alternative “ i ” is chosen by the commuter “ n ” from the choice set “ C_n ” is:

$$\Pr(i|C_n) = \frac{e^{\alpha V_{in}}}{\sum_{j \in C_n} e^{\alpha V_{jn}}} \quad (3.7a)$$

$$\text{or } \Pr(i|C_n) = \frac{e^{\alpha \beta x_{in}}}{\sum_{j \in C_n} e^{\alpha \beta x_{jn}}} \quad (3.7b)$$

where “ α ” is the scale parameter, which is usually assumed to be the one which makes “ β ” parameters equivalent to the coefficients for the independent variables.

Maximum likelihood method is used to solve for the “ β ” parameters in the mode choice model (Hensher and Swait, 2000). Consider the likelihood of a sample of “ n ” independent observations with probabilities $p_1, p_2, p_3, \dots, p_n$. The likelihood of the sample will simply be the product of the individual likelihoods. This product is maximum when the most likely set of p ’s is used, i.e.:

$$\text{Likelihood } L^* = p_1 p_2 p_3 \dots p_n = \prod_{i=1}^n p_i \quad (3.8)$$

In the case of mode choice logit model,

$$\text{Likelihood } L^* = (\beta_1, \beta_2, \beta_3, \dots, \beta_k) = \prod_{i=1}^n p_n(i)^{y_{in}} p_n(j)^{y_{jn}} \quad (3.9)$$

where “ p_n ” is a function of “ β ” parameters, and “ i ” and “ j ” are mode choice alternatives – private mode and public mode, respectively. Taking the log of the likelihood function, the above equation can be written as:

$$L = \log(L^*) = \log \left(\prod_{i=1}^n p_n(i)^{y_{in}} p_n(j)^{y_{jn}} \right) \quad (3.10a)$$

$$L = \log(L^*) = \sum_{i=1}^n \log \left[p_n(i)^{y_{in}} p_n(j)^{y_{jn}} \right] \quad (3.10b)$$

$$L = \log(L^*) = \sum_{i=1}^n y_{in} \log[p_n(i)] + (1 - y_{in}) \log[1 - p_n(i)] \quad (3.10c)$$

The maximum of “L” is solved by differentiating the function with respect to each of the “ β ” parameters and setting the partial derivatives to zero, or the values of “ β ” parameters that provide the maximum of “L”.

3.3.2 Ordered Probit Model

Some multinomial variables (e.g. mode switching propensity) are inherently ordered. In such cases, although the choice preferences are discrete, the multinomial logit model would fail to account for the ordinal nature of the dependent variables. The attitudinal outcomes, which are usually based on the Likert’s scale, generate data in the form of ordinal, or ordered responses. The scales to measure these responses do not have any natural unit of measurement, and the interval between the outcome choices cannot be assumed to be uniform (Abdel-Aty, 2001). Thus, analysing the ordered outcomes using linear regression technique is also not desirable. Ordered-response models e.g. Ordered Probit Model (OPM), recognises the indexed nature of the response variables, and have come into fairly wide use as a framework for analysing such responses (Abdel-Aty, 2003).

The OPM is based on the assumption that “ y_i^* ” depends linearly on “ x_i ”, according to the following equation: $y_i^* = x_i \beta' + u_i$; where “ y_i^* ” ($-\infty < y_i^* < +\infty$) is the underlying latent variable representing “ i ’s” propensity to agree with the statement advanced, “ i ” is the index of the respondent ($i = 1, \dots, n$), “ n ” is the sample size, “ x_i ” is a vector of characteristics relevant in explaining the attitude of a respondent, “ β ” is a vector of parameters, which will ultimately be interpretable in the same way as slope parameters in linear regression, and “ u_i ” a random error term assumed to follow a standard normal distribution $N(0,1)$.

“ y_i ” is the individual “ i ’s” response to the choice preferences, and assume that this can take one of the integer values 0, 1, 2, 3, ... , J . “ y^* ” is unobserved, but the relationship between “ y^* ” and the observed variable “ y ” is:

$$y = 0 \text{ if } -\infty < y^* < k_1 \tag{3.11a}$$

$$y = 1 \text{ if } k_1 < y^* < k_2 \tag{3.11b}$$

$$y = 2 \text{ if } k_2 < y^* < k_3 \tag{3.11c}$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$$

$$y = J \text{ if } k_J < y^* < \infty \tag{3.11d}$$

The parameters “ k_j ” ($j = 1, \dots, J$), are known as cut point or sometimes threshold parameters. Figure 3.2 shows the density function of the latent variable “ y^* ”, and illustrates the correspondence between the latent variable and the observed variable “ y ”. A set of threshold values for the case $J = 4$ is superimposed. Note that mean ($x_i\beta'$) of “ y^* ” depends on the explanatory variables contained in the vector “ x ”, and therefore the whole distribution shifts when the value of one such variable changes, in a direction dictated by the sign of corresponding “ β ” coefficient. It is obvious from the diagram that such a shift causes a change in the distribution of responses because the threshold values are fixed.

Figure 3.2 Probability Density Function of “ y^* ” and its Relationship to “ $y=0,1,2,3,4$ ”

The probabilities associated with the coded responses of an OPM are as follows:

$$\begin{aligned} P_i(0) &= Pr(y_i = 0) = Pr(y_i^* \leq k_l) = Pr(\mathbf{x}_i \boldsymbol{\beta}' + u_i \leq k_l) = Pr(u_i \leq k_l - \mathbf{x}_i \boldsymbol{\beta}') \\ &= \Phi(k_l - \mathbf{x}_i \boldsymbol{\beta}') \end{aligned} \quad (3.12a)$$

$$P_i(1) = \Phi(k_2 - x_i \beta') - \Phi(k_1 - x_i \beta') \quad (3.12b)$$

$$P_i(j) = \Phi(k_{j+1} - x_i \beta') - \Phi(k_j - x_i \beta') \quad (3.12c)$$

$$P_i(J) = 1 - \Phi(k_J - x_i \beta'), \quad (3.12d)$$

where “ i ” is an individual, “ j ” is the response alternative, $Pr(y_i = j)$ is the probability that individual “ i ” responds in manner “ j ”, and $\Phi(\cdot)$ is the standard normal cumulative distribution function. In this study, different OPMs were estimated, which provided information regarding the threshold levels corresponding to the commuters’ desire to access traveller information, desire to change their travel plan, and willingness to switch their usual mode of travel.

3.3.3 Bounded Rationality Model

The discrete choice models based on conventional microeconomic theory postulates economic and rational behaviour of the respondents. The concept of rational behaviour is used to describe a consistent and calculated decision process in which the individual follows his/her objectives – perfect rationality (Ben-Akiva and Lerman, 1985). Simon (1957) developed the distinction between perfect and bounded rationality. Unlike perfect rationality, bounded rationality recognises the constraints on the decision process that arise from the limitations of human beings as problem solvers with limited information-processing capabilities.

In this study, the bounded rational rule proposed by Simon (1957) is applied to model the decision makers' mode switching behaviour in the commuting context. The mode switching is from private mode i.e. Automobile (*Auto*) to public modes i.e., MRT (*M*), Bus (*B*), MRT and Bus (*M+B*). The mode switching between private

mode “ m ” and public mode “ k ” can be due to the travel time saving (TS_{im}^k) or travel cost saving (CS_{im}^k) by an individual “ i ”. The travel time saving (TS_{im}^k) is the time difference between travel time using private mode “ m ” (TT_i^m) and travel time using public mode “ k ” (TT_i^k), such that $TS_{im}^k = TT_i^m - TT_i^k$. Similarly, the travel cost saving (CS_{im}^k) is the cost difference between travel cost using private mode “ m ” (TC_i^m) and travel time using public mode “ k ” (TC_i^k), such that $CS_{im}^k = TC_i^m - TC_i^k$. The bounded rational notion here for mode switching is as follows. Commuter “ i ” does not switch his/her mode so long as the corresponding generalised travel cost saving (GTS_{im}^k), which is the weighted sum between travel time saving (TS_{im}^k) and cost saving (CS_{im}^k) on public mode “ k ” as compared to the current private mode “ m ”, remains within the decision maker’s mode indifference band (IBM_{im}^k), such that:

$$GTS_{im}^k = \alpha_m^k TS_{im}^k + \beta_m^k CS_{im}^k ; \quad (3.13a)$$

$$\phi_{im}^k = -1, \quad \text{if } GTS_{im}^k \leq IBM_{im}^k \quad (3.13b)$$

$$\phi_{im}^k = 1, \quad \text{otherwise} \quad (3.13c)$$

The variable ϕ_{im}^k is the mode switching decision indicator variable, which equals to 1 when user “ i ” switches his/her travel mode from the private travel mode “ m ” to the public transportation mode “ k ”, and equals -1 otherwise. IBM_{im}^k is the indifference band for mode switching corresponding to user “ i ”. The indifference band is a latent quantity modelled as random variables with systematic and random components given by:

$$IBM_{im}^k = g(X_i, Z_{im}^k, \theta_{im}^k) + \varepsilon_{im}^k ; \quad \text{where } \varepsilon_{im}^k \sim MVN(0, \Sigma_{\varepsilon, k}) \quad (3.14)$$

The systematic components of the indifference band for mode switching are $g(\bullet)$. The vector of user characteristics “ X_i ” and the vector of performance characteristics “ Z_{im}^k ” capture user’s inherent attributes and experience, respectively; “ θ_{im}^k ” is a

vector of parameters to be estimated. The random term “ ε_{im}^k ” is assumed to be normally distributed over different public transportation modes and across commuters with zero mean and general variance-covariance matrix “ $\Sigma_{\varepsilon,k}$ ”.

“ $\Sigma_{\varepsilon,k}$ ” represents $K \times K$ (where “ K ”, which is equal to 3 in this study, is the number of the public transportation modes that can be chosen by the respondents or sample of observations) variance-covariance matrix that captures serial correlation due to the persistence of unobservable attributes across the sequence of mode switching decisions made by the same user. Although the exact specification of the structure of this matrix is ultimately empirical, the general structure proposed for this study is specified in Table 3.1.

Table 3.1 General Expressions for Error Term Elements

$E[(\varepsilon_{im}^k)^2]$	$\sigma_{m,k}^2$	
$E[\varepsilon_{im}^k, \varepsilon_{im}^{k'}]$	γ_{m1}	$k \neq k'$
$E[\varepsilon_{im}^k, \varepsilon_{im}^{k''}]$	γ_{m2}	$k \neq k''$
$E[\varepsilon_{im}^{k'}, \varepsilon_{im}^{k''}]$	γ_{m3}	$k' \neq k''$

where m = automobile; k = rail transit; k' = bus, k'' = rail transit and bus. “ $\Sigma_{\varepsilon,k}$ ” can be rewritten in matrix form, as shown in Table 3.2.

Table 3.2 Variance-Covariance Matrix

	Automobile		
MRT	$\sigma_{M,Auto}^2$	γ_{Auto1}	γ_{Auto2}
Bus	γ_{Auto1}	$\sigma_{B,Auto}^2$	γ_{Auto3}
Bus and MRT	γ_{Auto2}	γ_{Auto3}	$\sigma_{B+M,Auto}^2$

The expression for “ $\Sigma_{\varepsilon,k}$ ” can be derived from the element of “ $\Sigma_{\varepsilon,Auto}$ ” as listed in Table 3.3.

Table 3.3 Expressions for Error Term Structure Elements

$E[(\varepsilon_{i \text{ Auto}}^{MRT})^2]$	σ_1^2
$E[(\varepsilon_{i \text{ Auto}}^{Bus})^2]$	σ_2^2
$E[(\varepsilon_{i \text{ Auto}}^{MRT+Bus})^2]$	σ_3^2
$E[\varepsilon_{i \text{ Auto}}^{MRT}, \varepsilon_{i \text{ Auto}}^{Bus}] = E[\varepsilon_{i \text{ Auto}}^{Bus}, \varepsilon_{i \text{ Auto}}^{MRT}]$	γ_1
$E[\varepsilon_{i \text{ Auto}}^{MRT}, \varepsilon_{i \text{ Auto}}^{Bus+MRT}] = E[\varepsilon_{i \text{ Auto}}^{Bus+MRT}, \varepsilon_{i \text{ Auto}}^{MRT}]$	γ_2
$E[\varepsilon_{i \text{ Auto}}^{Bus}, \varepsilon_{i \text{ Auto}}^{Bus+MRT}] = E[\varepsilon_{i \text{ Auto}}^{Bus+MRT}, \varepsilon_{i \text{ Auto}}^{Bus}]$	γ_3

“ $\Sigma_{\varepsilon, m}$ ” can now be rewritten in matrix form as follows:

Table 3.4 Expressions for Error in Matrix Form

	Auto		
MRT	σ_1^2	γ_1	γ_2
Bus	γ_1	σ_2^2	γ_3
Bus and MRT	γ_2	γ_3	σ_3^2

Given the specification for $g(\bullet)$, the available observations of the switching decisions made by “ N ” number of commuters in the sample provided a basis for the maximum likelihood estimation of the model parameters. A general approach to deal with the associated estimation was presented by Daganzo and Sheffi (1982) who showed that the probability of a sequence of decisions is essentially equivalent to a multinomial probit probability function. This approach was adopted by Mahmassani and co-workers to model the day-to-day switching decisions of departure time and route (Mahmassani and Herman, 1990; Jou and Mahmassani, 1996). The same approach was adopted here to model the mode choice switching under different levels of travel time saving and travel cost saving. The modelling approach is summarised hereafter.

The probability of an outcome (ϕ_{im}^k) for individual “i” at a specific decision switching from private mode “m” to public mode “k”, is given by:

$$\Pr(\phi_{im}^k) = \Pr[\phi_{im}^k (\alpha_m^k TS_{im}^k + \beta_m^k CS_{im}^k - IBM_{im}^k) \geq 0] \quad (3.15a)$$

$$\Pr(\phi_{im}^k) = \Pr[\phi_{im}^k \varepsilon_{im}^k \leq \phi_{im}^k (\alpha_m^k TS_{im}^k + \beta_m^k CS_{im}^k - g_k(X_i, Z_{im}^k, \theta_{im}^k))] \quad (3.15b)$$

The likelihood of three decisions ($\phi_{i, Auto}^M, \phi_{i, Auto}^B, \phi_{i, Auto}^{M+B}$) for individual “i” using private (i.e. Auto) mode is given by:

$$\Pr(\phi_{i, Auto}^M, \phi_{i, Auto}^B, \phi_{i, Auto}^{M+B}) = \Pr \left\{ \begin{array}{l} \phi_{i, Auto}^M \varepsilon_{i, Auto}^M \leq \phi_{i, Auto}^M (\alpha_{Auto}^M TS_{i, Auto}^M + \beta_{Auto}^M CS_{i, Auto}^M - g_k(X_i, Z_{i, Auto}^M, \theta_{i, Auto}^M)), \\ \phi_{i, Auto}^B \varepsilon_{i, Auto}^B \leq \phi_{i, Auto}^B (\alpha_{Auto}^B TS_{i, Auto}^B + \beta_{Auto}^B CS_{i, Auto}^B - g_k(X_i, Z_{i, Auto}^B, \theta_{i, Auto}^B)), \\ \phi_{i, Auto}^{B+M} \varepsilon_{i, Auto}^{B+M} \leq \phi_{i, Auto}^{B+M} (\alpha_{Auto}^{B+M} TS_{i, Auto}^{B+M} + \beta_{Auto}^{B+M} CS_{i, Auto}^{B+M} - g_k(X_i, Z_{i, Auto}^{B+M}, \theta_{i, Auto}^{B+M})) \end{array} \right\} \quad (3.16)$$

where M = MRT, B = Bus, and M+B = MRT and Bus.

The estimation of such a formulation, in which each possible decision is treated as a choice alternative, results in tremendous operational difficulty for any reasonable sample size (Daganzo, 1979). One approach to overcome the computational burden was developed by Daganzo and Sheffi (1982), who viewed an individual’s “n” consecutive decisions as a single choice among “n+1” hypothetical alternatives (of which one was viewed as an “auxiliary alternative”). This yields 3+1 alternatives with respective utilities “ u_i ”, as shown in the following:

$$\text{Auxiliary Alternative,} \quad u_0 = 0; \quad (3.17a)$$

Switch from Auto to MRT (M),

$$u_1 = \phi_{i, Auto}^M [g_M(X_i, Z_{i, Auto}^M, \theta_{i, Auto}^M) - \alpha_{i, Auto}^M TS_{i, Auto}^M - \beta_{i, Auto}^M CS_{i, Auto}^M + \varepsilon_{i, Auto}^M]; \quad (3.17b)$$

Switch from Auto to Bus (B),

$$u_2 = \phi_{i, Auto}^B \left[g_B(X_i, Z_{i, Auto}^B, \theta_{i, Auto}^B) - \alpha_{i, Auto}^B TS_{i, Auto}^B - \beta_{i, Auto}^B CS_{i, Auto}^B + \varepsilon_{i, Auto}^B \right]; \quad (3.17c)$$

Switch from Auto to MRT + Bus (M+B),

$$u_3 = \phi_{i, Auto}^{M+B} \left[g_{M+B}(X_i, Z_{i, Auto}^{M+B}, \theta_{i, Auto}^{M+B}) - \alpha_{i, Auto}^{M+B} TS_{i, Auto}^{M+B} - \beta_{i, Auto}^{M+B} CS_{i, Auto}^{M+B} + \varepsilon_{i, Auto}^{M+B} \right] \quad (3.17d)$$

Therefore, the probability of three decisions, $\Pr(\phi_{i, Auto}^M, \phi_{i, Auto}^B, \phi_{i, Auto}^{M+B})$, is identical to the probability of selecting the auxiliary alternative, P_0 :

$$P_0 = \Pr \left\{ \begin{array}{l} \phi_{i, Auto}^M \left[g_M(X_i, Z_{i, Auto}^M, \theta_{i, Auto}^M) - GTS_{i, Auto}^M + \varepsilon_{i, Auto}^M \right] \leq 0; \\ \phi_{i, Auto}^B \left[g_B(X_i, Z_{i, Auto}^B, \theta_{i, Auto}^B) - GTS_{i, Auto}^B + \varepsilon_{i, Auto}^B \right] \leq 0; \\ \phi_{i, Auto}^{M+B} \left[g_{M+B}(X_i, Z_{i, Auto}^{M+B}, \theta_{i, Auto}^{M+B}) - GTS_{i, Auto}^{M+B} + \varepsilon_{i, Auto}^{M+B} \right] \leq 0 \end{array} \right\} \quad (3.18)$$

The above equation can be rewritten as:

$$P_0 = \Pr \left\{ \begin{array}{l} \phi_{i, Auto}^M \left[IBM_{i, Auto}^M - GTS_{i, Auto}^M \right] \leq 0; \\ \phi_{i, Auto}^B \left[IBM_{i, Auto}^B - GTS_{i, Auto}^B \right] \leq 0; \\ \phi_{i, Auto}^{M+B} \left[IBM_{i, Auto}^{M+B} - GTS_{i, Auto}^{M+B} \right] \leq 0 \end{array} \right\} \quad (3.19)$$

The u 's are random variables with a multivariate normal distribution $MVN(V, \Sigma u)$. In deriving " Σu ", the error terms can be modelled as stand-alone alternative specific parameters and the variable " $\varepsilon_{i,m}^k$ " can be omitted with no loss of generality. Therefore, the expression of indifference band for mode switching can be rewritten as:

$$g'(X_i, Z_{im}^k, \theta_{im}^k) = g(X_i, Z_{im}^k, \theta_{im}^k) + \varepsilon_{im}^k, \quad (3.20a)$$

where;

$$\begin{aligned} g'_M(X_i, Z_{im}^M, \theta_{im}^M) &= g(X_i, Z_{im}^M, \theta_{im}^M) + \varepsilon_{im}^M \\ g'_B(X_i, Z_{im}^B, \theta_{im}^B) &= g(X_i, Z_{im}^B, \theta_{im}^B) + \varepsilon_{im}^B \\ g'_{M+B}(X_i, Z_{im}^{M+B}, \theta_{im}^{M+B}) &= g(X_i, Z_{im}^{M+B}, \theta_{im}^{M+B}) + \varepsilon_{im}^{M+B} \end{aligned} \quad (3.20b)$$

The utilities can also be rewritten as:

$$\text{Auxiliary Alternative, } u_0 = 0; \quad (3.21a)$$

Switch from Auto to MRT (M),

$$u_1 = \phi_{i, Auto}^M \left[g'_M(X_i, Z_{i, Auto}^M, \theta_{i, Auto}^M) - GST_{i, Auto}^M \right]; \quad (3.21b)$$

Switch from Auto to Bus (B),

$$u_2 = \phi_{i, Auto}^B \left[g'_B(X_i, Z_{i, Auto}^B, \theta_{i, Auto}^B) - GST_{i, Auto}^B \right]; \quad (3.21c)$$

Switch from Auto to MRT+ Bus (M+B),

$$u_3 = \phi_{i, Auto}^{M+B} \left[g'_{M+B}(X_i, Z_{i, Auto}^{M+B}, \theta_{i, Auto}^{M+B}) - GST_{i, Auto}^{M+B} \right] \quad (3.21d)$$

More generally, “ u ” and “ Σu ” can be written as follows (Daganzo and Sheffi, 1982):

$$u = \Theta^T [A], \text{ where } u \sim MVN[\Theta, \Sigma \theta[A]] \quad (3.22a)$$

$$\Sigma u = [A]^T \Sigma \theta[A]; \quad (3.22b)$$

where

$$u = [u_0, u_1, u_2, u_3] \quad (3.22c)$$

$$\Theta^T = [1, g_M(X_i, Z_{i, Auto}^M, \theta_{i, Auto}^M), g(X_i, Z_{i, Auto}^B, \theta_{i, Auto}^B), g(X_i, Z_{i, Auto}^{M+B}, \theta_{i, Auto}^{M+B})] \quad (3.22d)$$

$$A = \begin{bmatrix} 0 & H1 & H2 & H3 \\ 0 & \phi_{Auto}^M & 0 & 0 \\ 0 & 0 & \phi_{Auto}^B & 0 \\ 0 & 0 & 0 & \phi_{Auto}^{M+B} \end{bmatrix} \quad (3.22e)$$

where

$$\begin{aligned} H1 &= -\phi_{i, Auto}^M \left(\alpha_{Auto}^M TS_{i, Auto}^M + \beta_{Auto}^M CS_{i, Auto}^M \right) \\ H2 &= -\phi_{i, Auto}^B \left(\alpha_{Auto}^B TS_{i, Auto}^B + \beta_{Auto}^B CS_{i, Auto}^B \right) \\ H3 &= -\phi_{i, Auto}^{M+B} \left(\alpha_{Auto}^{M+B} TS_{i, Auto}^{M+B} + \beta_{Auto}^{M+B} CS_{i, Auto}^{M+B} \right) \end{aligned} \quad (3.22f)$$

Σu can be directly derived from the above equations, and written as:

$$\Sigma u = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & \phi_{i, Auto}^M \phi_{i, Auto}^B \gamma_1 & \phi_{i, Auto}^M \phi_{i, Auto}^{M+B} \gamma_2 \\ 0 & \phi_{i, Auto}^M \phi_{i, Auto}^B \gamma_1 & \sigma_2^2 & \phi_{i, Auto}^B \phi_{i, Auto}^{M+B} \gamma_3 \\ 0 & \phi_{i, Auto}^M \phi_{i, Auto}^{M+B} \gamma_2 & \phi_{i, Auto}^B \phi_{i, Auto}^{M+B} \gamma_3 & \sigma_3^2 \end{bmatrix} \quad (3.23)$$

The definitions of σ_1 , σ_2 , σ_3 , γ_1 , γ_2 and γ_3 are summarised in Table 3.3. Note that these parameters incorporate variance and covariance terms corresponding to the random error terms associated with both the relative indifference band and the absolute minimum threshold.

3.3.4 Statistical Characteristics of Mode Choice Model

There are several goodness-of-fit measures available for testing how well a discrete choice model fits the data on which it was estimated. The methods can be categorised as informal and formal tests. In the informal tests the sign, relative value, and ratio of coefficients are observed. In the formal tests the goodness-of-fit, *Student's t-test* and χ^2 test are performed.

(a) Sign of Coefficients

The model equation represents the association between a dependent variable (Y), which represents the probability of a particular choice being made, and one or more independent variables (X's) that reflect attributes of the choices and the choice-

maker. The coefficients in choice models are multiplicative on the response. The model parameters, or partial slope coefficients, represent the change in Y given a unit change in X, all else held constant. The sign convention of such model parameters expresses the increase or decrease in the utility of the alternative.

(b) *Relative Value of the Coefficients*

In this test the absolute values of the coefficients of same alternative are compared. For example, the absolute value of the in-vehicle travel time should be lower than the out-of-vehicle travel time, which shows the lesser preference to out-of-vehicle travel time.

(c) *Ratio of Two Coefficients*

Ratio of the two coefficients appearing in the same utility function provides a trade-off for marginal rate of substitution, between the two corresponding variables.

(d) *Overall Goodness-of-Fit Measures*

This test is used to describe the overall goodness-of-fit of the model. Everything else being equal, a specification with a higher maximum value of the likelihood function is considered to be better. It is known as:

$$\rho^2 = 1 - \left[\frac{L(\beta')}{L(0)} \right] \quad (3.24)$$

where “ $L(0)$ ” is the log likelihood of the restricted model, and “ $L(\beta')$ ” is the log likelihood of the unrestricted model. While for the same estimation of data set the ρ^2 of a model will always increase or at least stay the same whenever new variables are added to the utility function, so an adjusted likelihood ratio index ($\bar{\rho}^2$) is used with “ k ” (number of coefficients) degrees of freedom, such that:

$$\bar{\rho}^2 = 1 - \left[\frac{L(\beta') - k}{L(0)} \right] \quad (3.25)$$

(e) *Student's t-test*

As the model is estimated from the sample of relevant population, the magnitude of the sampling error in a parameter is given by the standard error associated with that parameter. A large standard error will imply a large sampling error and lower precision with which the corresponding parameter is estimated. The statistic used for testing the null hypothesis that is equal to a parameter “ β_k ” some estimated value “ β' ” is the asymptotic t -statistic that is given as:

$$t - statistic = \frac{\beta' - \beta_k}{S_k}, \quad (3.26)$$

where “ β' ” is the estimate for the k^{th} parameter and “ S_k ” is the standard error of the estimate. The critical values for the test statistic are percentiles of a standardized t -distribution which, for two-tailed tests at the frequently used significance levels of 0.10 and 0.05, correspond asymptotically to ± 1.65 and ± 1.96 , respectively. The statistic for an asymptotic t -test of a linear relationship can also be calculated by the covariance matrix information.

(f) *Likelihood Ratio Test*

The purpose of the likelihood ratio test is to compare models with different levels of complexity. Let “ $L(\beta')$ ” be the maximum log likelihood attained with the estimated parameter vector “ β' ”, on which no constraint has been imposed. Let “ $L(0)$ ” be the maximum log likelihood with constraints applied to a subset of coefficients in “ β' ”. Then, asymptotically $-2(L(0) - L(\beta'))$ has a χ^2 distribution with degrees of freedom equal to the number of constraint coefficients. Thus, the above statistic can be used to test the null hypothesis that two different models perform approximately the same in explaining the data. If there is insufficient evidence to support the more complex model, then the simpler model is preferred. For large differences in log likelihood value, there is evidence to support the more complex model as compared to the simpler one.

In the context of the calibrated model, two standard tests have been taken into consideration. The first test compares the model estimated with all the variables

postulated as influencing the choice process to a model that has no coefficients whatsoever – a model that predicts equal probability for all choices. The test statistic is given by: $-2(L(0) - L(\beta')) \approx \chi^2$, with degrees of freedom equal to total number of coefficients. The null hypothesis, “ H_0 ”, is that all the coefficients are zero i.e. $\beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$, or all the alternatives are equally likely to be chosen. “ $L(0)$ ” is the log likelihood computed when all coefficients including alternative specific constants are constrained to be zero, and “ $L(\beta')$ ” is the log likelihood computed with no constraints on the model.

The second test compares the complex model with another naïve model, which contains alternative specific constants for $n-1$ alternatives. This naïve model predicts choice probabilities based on the observed market share of the respective alternatives. The test statistic is given by: $-2(L(C) - L(\beta')) \approx \chi^2$, with degrees of freedom equal to total number of coefficients. The null hypothesis, “ H_0 ”, is that the coefficients are zero except the alternative specific constants. “ $L(C)$ ” is the log likelihood value computed when all slope coefficients are constrained to be equal to zero except alternative specific constants.

3.3.5 Sensitivity Analysis

The main intention of the travel behaviour modelling is to quantify the modal shares; find out the relationship between the trip makers and the available modes; and evaluate the impact of changes in the variables to the modal share. These changes in the variables may come with time or with changes in the policies. Policies such as change in the total travel cost and total travel time affect the modal shares. Hence, before implementing any such policies, it is very important to evaluate the possible impacts on the modal share. To analyse such impacts, sensitivity analysis is performed, which can explain the influence of the change in the dependent variable (X) on the independent choice variable (Y). It is known that the effect of each X on Y can be linear. Therefore interpreting parameter estimates as linear effects on Y must be common to all generalised linear models. Such interpretation, however, may not be intuitively appealing. Fortunately, interpreting

the linear effects is only one of the several possible ways of making sense of parameter estimates from probability models. In this study four different methods of interpreting parameter estimates from probability model have been adopted. These methods are discussed as follows.

(a) *Pivot Point Mode Choice Model*

According to Ben-Akiva and Lerman (1985), the pivot point model can be used to predict changes in behaviour with the existing choice probabilities of the alternatives and changes in variables. So it obviates the need to use the full set of independent variables to calculate the new choice probabilities. Derivation of the pivot point model is relatively straightforward. The linear-in-parameter logit predicts the probability that individual “ n ” will choose “ i ” from the set of alternatives “ C_n ”.

The revised choice probability resulting from a change in utilities is given by:

$$P_n(i) = \frac{e^{V_{in} + \Delta V_{in}}}{\sum_{j \in C_n} e^{V_{jn} + \Delta V_{jn}}}, \quad (3.27)$$

where “ ΔV_{in} ” is the change in utility for alternative “ i ”, and

$$\Delta V_{in} = \sum_{k=1}^K \beta_k \Delta x_{ink}, \quad (3.28)$$

where “ Δx_{ink} ” is the change in the k_{th} independent variable for alternative “ i ”, and individual “ n ”. Divide both the numerator and denominator by $\sum_{j \in C_n} e^{\Delta V_{jn}}$ to obtain:

$$P_n(i) = \frac{P_{n(i)} e^{\Delta V_{in}}}{\sum_{j \in C_n} P_n(j) e^{\Delta V_{jn}}} \quad (3.29)$$

Thus, to predict the changes with a linear-in-parameters choice model, it is needed to know the choice probabilities in the base and the changes in utilities due only to the affected variables.

(b) *Marginal Effect of the Independent Variables*

Regardless of the specific probability model, the estimated “ β_k ” gives a marginal effect of the corresponding “ x_k ” on Y . Thus, “ β_k ” represents the change in the probability associated with a unit change in the “ j_{th} ” parameter holding all other parameters constant. This change can be taken as linear, but it will not be useful because Y in a logit model is not interpretable as in the case of classical linear model. Often, it is more meaningful to interpret the effect of an estimated “ β_k ” on a transformed “ Y ” rather than on “ Y ” itself.

In case of logit model:

$$\text{Prob}(Y_i = 1) = \frac{\exp(\alpha + \beta X_k)}{1 + \exp(\alpha + \beta X_k)} \text{ and} \quad (3.30a)$$

$$1 - \text{Prob}(Y_i = 1) = \frac{1}{1 + \exp(\alpha + \beta X_k)} \quad (3.30b)$$

Using the above:

$$\frac{1 - \text{Prob}(Y_i = 1)}{\text{Prob}(Y_i = 1)} = \frac{\frac{1}{1 + \exp(\alpha + \beta X_k)}}{\frac{\exp(\alpha + \beta X_k)}{1 + \exp(\alpha + \beta X_k)}} = \frac{1}{\exp(\alpha + \beta X_k)} \quad (3.31a)$$

Re-arranging this expression, we obtain:

$$\frac{\text{Prob}(Y_i = 1)}{1 - \text{Prob}(Y_i = 1)} = \exp(\alpha + \beta X_k) \quad (3.31b)$$

Take natural logarithms of this last expression:

$$\ln \left[\frac{\text{Prob}(Y_i = 1)}{1 - \text{Prob}(Y_i = 1)} \right] = \alpha + \beta X_k \quad (3.31c)$$

where “ $\ln [\text{Prob}(Y_i = 1) / 1 - \text{Prob}(Y_i = 1)]$ ” is the logit or the log-odds ratio. The effect of a unit change in “ x_k ” on the log-odds of the event occurring is given by

the “ β_k ”. Thus taking the exponential of “ β_k ” can directly provide the effect of the independent variable on the odds ratio.

(c) *Predicted Probabilities Given a Set of Values in the Independent Variables*

The predicted probabilities are intuitively appealing because these probabilities give an idea of how likely certain types of commuters are going to take a certain type of mode. In generalised model framework, predicted probabilities are often derived by calculating the values $\sum_{k=1}^K (\alpha + \beta X_k)$, using certain values of “ x_k ” variables. In case of logit model, predicted probabilities of Y conditional upon set of “ x_k ” variables, the applicable equation is:

$$\text{Prob}(Y_i = 1) = \frac{\exp \sum_{k=1}^K (\alpha + \beta X_k)}{1 + \exp \sum_{k=1}^K (\alpha + \beta X_k)}, \quad (3.32)$$

which expresses event probability as a function of “ β_k ” and “ x_k ”.

(d) *Marginal Effect on the Probability of an Event*

The marginal effect on the probability of an event gives the combined output of the marginal effects of the independent variables and the predicted probabilities given a set of values in the explanatory variables. In case of logit model, the probability of an event is:

$$\text{Prob}(Y_i = 1) = \frac{\exp \sum_{k=1}^K (\alpha + \beta X_k)}{1 + \exp \sum_{k=1}^K (\alpha + \beta X_k)} \quad (3.33a)$$

The marginal effect of “ x_k ” on $\text{Prob}(Y_i = 1)$ is given by:

$$\frac{\partial \text{Prob}(y=1)}{\partial x} = \frac{e^{-\sum_{k=1}^K \beta_k x_k}}{(1 + e^{-\sum_{k=1}^K \beta_k x_k})^2} \beta_k \quad (3.33b)$$

where “ ∂ ” indicates partial derivative or marginal effect. Unlike the interpretation of the effect on odds, which is invariant to values of independent variables, the marginal effect on probability changes with the values of “ x ’s” and hence with the probability associated with the values of “ x ’s”.

(e) *Aggregate Elasticity*

Aggregate elasticity summarises the responsiveness of some group of decision makers rather than that of any individual when there is change in the value of the attribute. According to Ben-Akiva and Lerman (1985) the aggregate elasticity can be estimated as follows:

$$E_{X_{jk}}^{P(i)} = \frac{\sum_{n=1}^N P_n(i) E_{X_{jk}}^{P(i)}}{\sum_{n=1}^N P_n(i)} \quad (3.34)$$

3.4 RULE-BASED EXPERT SYSTEM

In this study a rule-based expert system i.e. INSIM Expert System (IES), has been designed to simulate the commuters’ travel choice behaviour. The source for the developed expert system is the knowledge and expertise gained from travel behaviour survey and modelling. Based on the commuters’ socio-economic characteristics and provided multimodal traveller information, the commuter travel choice behaviour is modelled and represented by cognitive rules.

The designed rules develop a knowledge-base that provides reasons to the inference engine of the Intelligent Agent (IA) to take required actions. At the macro level, these rules develop the capability of an IA to emulate the expertise of a travel information dissipating agent and at the micro level, they emulate the responses of commuters’ travel choice decisions with respect to the provided real-time multimodal traveller information. The rules help the IA to depict different beliefs and intentions and preferences of different commuters with respect to their objectives. With the provision of such rules the IA becomes sensitive to changes in

the travel environment and depicts a rational expert/human behaviour while making decisions. The IA processes autonomously, and is capable of making decisions and dictating the required commands to the simulation environment (i.e. traffic network), based on its perception of the overall situation.

The IA continuously observes the overall travel environment (problem-specific information) of the transportation network, via database (i.e. sensor). The traffic simulator (i.e. travel environment) generates travel attributes and updates the database at specific time intervals. Once the travel attributes are updated, the IA's inference engine compares the conditional part of the rules from the knowledge-base to the provided problem-specific information and takes actions that satisfy the required objectives, which are again defined as rules in the knowledge-base. Such actions allow the IA to perceive the existing condition of the travel environment, decide suitable actions with respect to the dissipation of traveller information and allow it to interact with the traffic simulator so as to improve the overall network conditions. The working mechanism of the rule-based IA is shown in Figure 3.3.

Figure 3.3 Working Mechanism of the Rule-Based Intelligent Agent
(Source: Russell and Norvig, 1995)

In this study, the implementation of rules is through the rule-based IA that is developed in commercially available software (Blaze Advisor). The model is based on an object-oriented design. The commuters, different travel characteristics, and travel mode are modelled as classes and objects. The knowledge about the travel preference of commuters, the transportation system operation rules and strategies, and the control heuristics, are implemented as rules.

3.4.1 Architecture and Working Mechanism

The main limitation of the existing transportation network simulation models is that the OD demand and the modal distribution are assumed to be constant over time, which does not correspond to the actual conditions in multimodal urban networks. Such a limitation can be relaxed if time-dependent modal distribution can be estimated and the resulting link flows over time can be determined. The IES has been designed to provide a simulation platform where a multimodal transportation network can be developed and the time-dependent modal distribution can be estimated. The IES has two main components i.e. Intelligent Expert (IE) and INSIM Commuter (IC) as shown in Figure 3.4.

Figure 3.4 Architecture of the INSIM Expert System

The IE provides knowledge regarding the commuters’ mode choice decisions based on a set of rules, and the logic to generate suitable decisions for the commuters. The IC has the capability to generate/simulate “commuters”, gather travel statistics from

the transportation network simulated in a microscopic traffic simulation model and estimate a time-dependent OD matrix. The IES initially activates the IC to generate commuters and gather travel statistics from the traffic simulator for every pre-specified time interval. The characteristics of the simulated commuters along with real-time travel statistics collected from the traffic simulator are then stored in the database as facts. The IES then activates the inference engine to determine mode decisions for each commuter based on the known facts and the prevailing traffic characteristics. The IES is designed in a way to provide the flexibility to adopt alternative mode choice model that can give prediction of commuters' mode choices.

After all the commuters are assigned a certain mode, the IES again activates the IC to estimate an OD matrix for the commuters in the database. Commuters assigned with public modes are technically removed from the traffic simulation as their movements will be simulated with buses and MRT. An OD matrix file for private transport is generated for the prevailing time interval and is loaded into the traffic simulator, which releases the assigned traffic demand into the network and generates the travel statistics after a fixed time stamp. For the commencing time stamp, the IC is again activated and another batch of commuters is generated and the newly generated travel statistics are imported by the IC. This cycle of activities continues for the entire simulation period. The flexibility, to specify time interval for travel statistics update, allows the system to predict the commuters' mode under the influence of real-time traveller information and also allows it to estimate a time-dependent OD matrix.

3.4.2 General Decision Framework

The intelligent agent approach adopted to simulate the decision process of a commuter prior to their trip making is the core of the IES, which facilitates the basic framework in allocating the appropriate mode to each commuter generated. Each commuter generated by the system can be considered as an intelligent agent, which takes the form of a utility agent in this study. The agent, i.e. each commuter, reads information from the environment, which is the virtual environment maintained by a microscopic traffic simulator, and reacts by either changing the commuter's mode

or route, to avoid congestion. The decision process within each agent is a utility-based process, which takes the form of rules or discrete choice based utilities. In the following sections, details of the decision process within the intelligent agent approach are presented.

(a) *The Basic Rule Structure*

The rule structure adopted in this study corresponds to a controlling sequence of subroutines in problem-specific information, where the outcome is a set of discrete alternatives. The basic rule template can be formulated as:

“if ... (the condition) ... is true

then select ... (the preferred option) ...

else select ... (the alternate option)”

In this study, the rules associated with the state of the system (i.e. transportation network) are based on the choice-related characteristics and/or preferences. These characteristics and/or preferences are the basis of the antecedent part of a rule. In the case of rule execution the consequent part of the rule dictates certain elements in the state of the system. A typical example of such a rule can be demonstrated as follows:

if the network wide travel time has increased to “x” minutes

then provide traveller information to all the commuters in the system

else do not provide traveller information

In the above mentioned rule, the premise has a proposition i.e. “network wide travel time has increased” and a predicate i.e. “x minutes”. It can be clearly observed that the antecedent part of the rule is associated with the state of the transportation network. When the network-wide travel time increases to “x” minutes, the antecedent part of the rule would become true. The rule is then executed and the action described in the consequent part, i.e. “provide traveller information to all commuters”, would be taken. Otherwise, the action taken in this condition will be

the consequent part of rule stated in the ‘else’ part of the rule. The rule performance at this level is goal oriented and due to the structural simplicity it is coherent with the characterisation of human attitude, and corresponds to common sense knowledge and intuitive behaviour. The use of such simple rules is also attractive because it allows the usage of heuristics and rules of thumb to reduce the amount of information processing when making decisions.

(b) *The Rule of Inference*

The developed rules define the decision(s) under a possible perception of travel attributes (problem-specific information) corresponding to the level of service of the transportation network. Each decision is based on the true or false conditions associated with the perceived problem-specific information. Depending upon the level of perception of the travel attributes and their association with the number of rule premise, a decision tree is formulated. During rule processing, if at any stage the antecedent is false, a value 0 is assigned to the consequent corresponding to the true condition of the antecedent, and a value 1 is assigned to the consequent for the false antecedent. From that point onwards, the rule process follows the decision tree structure associated with the false branch and vice versa. This process continues up till the point where all the antecedents are exhausted. At that stage, the rule process provides a solution to the given problem-specific information. The rule processing in this study is based on the designed rule of inference, which establishes syntactic relations between the provided problem-specific information (facts), the rule premise and conclusion. These syntactic relations are used in the process of inference.

The developed rule-base can be considered as a set of rules “ R ”, such that $R = \{r_1, r_2, \dots, r_m\}$, and each rule “ r_i ” represents a specific commuter’s travel choice decision “ d_i ”. The rule “ r_i ” can be decomposed into two parts i.e. $r_i = \{a_j, d_i\}$, where “ a_j ” is the rule premise (antecedent) and “ d_i ” is the rule conclusion (consequent). The antecedent part represents the socio-economic and travel characteristics, and the consequent part represents the decision taken by the commuter with respect to the antecedent part. To decide about the mode choice of

the generated commuter, a set of socio-economic and travel characteristics is taken into consideration and is regarded as problem-specific information. The problem-specific information for all such commuters is represented by a set of facts “ F ”, such that $F = \{f_1, f_2, \dots, f_n\}$, where each “ f_i ” represents a specific commuter’s socio-economic and travel characteristics i.e. $f_i = \{a_g\}$. The IE applies the rule of inference, which starts matching the “ a_g ” with all the available “ a_j ” in “ R ”. Once it finds a matching rule, the rule is marked $(r_i \{a_i, d_j\} \leftarrow \{r_1, r_2, \dots, r_m\}, d)$, and the decision “ d ” is generated as the mode of the commuter.

This general approach of finding a suitable solution in certain situations is very reasonable, specifically while modelling the commuters’ mode choice behaviour in an information-rich environment. The rule of inference also has the potential to generate a decision when problem-specific information is not complete. Thus, the rule of inference can model realistically the travel choice decisions, while new information regarding certain modes is being acquired dynamically.

3.5 RULE FLOW STRUCTURE

The IES imitates the expertise of an information provider on a macro level and the commuter’s mode choice behaviour at a micro level. As an information provider, the IES can change the traveller information update timing and the type of traveller information based on the traffic conditions and the travel environment. While imitating the commuter’s mode choice behaviour, it considers the provided traveller information and decides the mode of travel. Thus, the rule-base of the IES is developed in two different levels. Initially, the macro level rule flow executes followed by the micro level rule flow. The macro level rule flow structure, which is the basis for the emulation of traveller information provider, is shown in Figure 3.5.

It can be observed from Figure 3.5 that rules developed at the macro level can be generalised into the following categories:

- (a) rules dealing with the traveller information update-time strategies;
- (b) rules dealing with dissemination of traveller information;

Figure 3.5 Macro Level Rule Flow Structure Imitating the Expertise of a Traveller Information Provider

- (c) rules dealing with the type of traveller information; and
- (d) rules dealing with the automated knowledge acquisition.

The users' interaction with these models allows them to simulate different strategies regarding the dissipation of traveller information in a multimodal environment.

3.5.1 Execution of Rule Flow (Macro Level)

The IES after initialisation allows the user to decide the traveller information update timing. The traveller information in the system will then be updated accordingly with respect to the defined update time. The reason for such an input at the initial stage is necessary because the traffic simulator and IC have to coordinate with each other based on the update time interval. Once the traveller information update interval has been decided, the IES requires the specification of the dissemination of traveller information. This component allows the IES to simulate the travel environment with or without the provision of traveller information.

Depending upon the user needs, this component can simulate different types of traveller information strategies or calibrate the system without disseminating any traveller information or provide access to some run-time simulated traveller information. The calibration module provides the user with the flexibility to calibrate the complete system so that the results can be validated. The access info module provides the user with simulated real-time traveller information and allows the user to make some travel choice. The user also needs to specify the level of traveller information that is to be disseminated in the system. If the user specifies "Level 0" then no information would be dissipated. Similarly, if "Level 1" is specified then the information on travelling time on all the available modes would be provided, and if "Level 2" is specified then the information on travelling time and waiting time for public modes would be provided. If "Level 3" is specified then the information on travelling time for car and public modes of transport, waiting time and fare for public modes of transport would be provided. The IES can also assign the level of travel information on its own judgement, depending on the impacts of the traveller information on commuters' mode switching, which can be estimated by the number of commuters that change their mode from car to a public

mode of transport. The IES can estimate the mode switching percentage based on the type of information provided and adopt the one which provides the highest mode switching percentage.

Lastly, the user needs to specify that in case of conflicting decisions or missing rules, what approach should be adopted by the IES to estimate/generate the commuter's mode decision. There are three options available for the user to select: to prompt the user, to apply Discrete Choice Model (DCM) or to use heuristic model. The IES then activates the IC to generate commuters and applies commuter's mode choice model to estimate/generate commuters' mode decisions based on their socio-economic characteristics and perceived travel environment.

3.5.2 Execution of Rule Flow (Micro Level)

The IES has three different models to simulate the commuter mode choice decision. These models are: Pure Rule-Based Model (PRB), Discrete Choice Model (DCM) and Probabilistic Model (COM). The user needs to specify any one model that is desired to simulate the commuter's mode choice behaviour. All the three models are based on the commuters' responses gathered from the travel behaviour survey. The PRB model provides a crisp set of rules and each rule represents a single commuter, such that the commuter's socio-economic and travel characteristics is the rule premise and the commuter's mode choice decision is the consequent part of the rule. The DCM is a mode choice logit model which is estimated based on the commuters, responses gathered during the same travel behaviour survey. The COM model applies the Bayesian technique to estimate probability of car and public mode of transport based on the commuters, socio-economic and travel characteristics. The details regarding these three models are presented in the following section. These models function at a higher level of sophistication, where the level of performance is goal-controlled, as the developed rules have to deal with unfamiliar situation. The rules of this level can capture more complex reasoning and interaction, implicit preferences, as well as apparently non-intuitive behaviour.

Once all the basic requirements are specified, the IES processes different rules, and specific actions are taken so as to get the desired output. The designed IES is

sophisticated enough to allow the user to redefine or incorporate changes in all the above mentioned conditional models. So, in the later stage, user can change the information update time strategy, traveller information type and dissemination strategy, traveller information access model, or mode choice models and incorporate his/her own developed models.

3.6 THE MODE CHOICE MODEL

The general rule structure described in Section 3.5 allows a great flexibility in developing the rules. Depending on the complexity of commuters' decision-making process, three different mode choice models have been developed. Each model is capable of predicting the commuters' mode choice based on their socio-economic and travel characteristics.

3.6.1 Pure Rule-Based Model

The Pure Rule-Based (PRB) model is based on the rules designed from the knowledge and expertise gained from the travel behaviour survey. The design of the rules for the PRB model follows the production rule architecture. The rule structure of PRB model is simple and can clearly represent recommendations or directives. The form of rules in the PRB model follows either the situation-action type or the premise-conclusion type such that the action or conclusion in the THEN part is reached if the situation or the premise in the IF part is true. The rule set for PRB model consists of crisp rules, each of which represents a single commuter's socio-economic and travel characteristics and the decision of taking car or public mode of transport. The rule set is defined in a database and allows dynamic update for new rules. To access the rules from the database a generic rule pattern is designed, which is shown in Figure 3.6. In the PRB model, the commuter's socio-economic and travel characteristics considered are age, gender, education, income, stoppage during trip, ride sharing, travel time by car, ERP charges, parking fees, travel time by public mode of transport, level of comfort and transit fare. The PRB model is capable of finding the exact crisp rule(s) from the rule-base for problem-specific information in the database. As there are only two mode choices in this study i.e. private or public, each choice is given a weight for every occurrence.

Figure 3.6 Generic Rule Pattern for PRB Model

The choice with the higher weight is taken as the final decision and the mode choice is generated. If there is no matching rule or conflicting decisions, the PRB triggers the automated knowledge acquisition model, which gets the decision as mentioned earlier.

3.6.2 Discrete Choice Model

To develop a DCM Model a mode choice logit model has been estimated based on the data collected in the travel behaviour survey. Initially, different discrete choice models are estimated and analysed and the most suitable model with lowest prediction error and highest goodness-of-fit index (ρ^2) value is taken into consideration. The estimated DCM model is incorporated as a rule in the Knowledge-Base (KB). The Frame Technique is adopted to represent the knowledge. This technique allows the IE to execute all the commands defined in the frame and format the output as a rule premise. For example, one frame represents the utility by private mode of transport and the other represents the utility by public mode of transport. The IE will execute the first frame and estimate the utility by private mode and form a rule premise for private mode. In a similar manner, it will execute the second frame and estimate the utility for public mode and form a rule premise. Then, considering the developed rule premises, a decision will be given. Finally a decision regarding the commuter’s travel mode will be generated. The generic rule pattern for DCM model is shown in Figure 3.7.

Figure 3.7 Generic Rule Pattern for DCM Model

The frames shown in Figure 3.7 represent the modules to calculate the utility function and probability of private and public modes of transport. The rules are represented by R1 and R2, which follow a sequence, such that the next rule is executed after the execution of the previous rule. The IES simulates mode choice of each commuter based on this model and assigns a certain mode of travel.

3.6.3 Probabilistic Model

The COM model is based on the Bayes’ Theorem, which fits in well with the decision-making process. The theorem is based on the estimated prior probability of an event that is, the known probability of an event to occur. In the decision-making process for generated commuters, there are two events out of which one has to occur for each commuter. These two events are the commuter taking the private mode, or the commuter taking the public mode.

The prior probabilities for each of the events are associated with the socio-economic and travel characteristics of the commuter. These prior probabilities can be estimated form the data gathered during the travel behaviour survey. Every socio-economic and travel characteristic is classified as a separate set of variables which are segregated into different classes to give a limited choice set. A description of such sets is presented in the Venn diagram shown in Figure 3.8.

Figure 3.8 Venn diagram showing the Partitioning of Different Variables

According to Bayes’ Theorem, if the events Age1, Age2, Age3, and Age4 constitute a partition of the sample space, where $P(Age_i) \neq 0$ for $i = 1,2,3,4$ then the probability of the event Car to occur, such that $P(Car) \neq 0$, can be estimated by:

$$P(Age_i|Car) = \frac{P(Age_i)P(Car|Age_i)}{\sum_{i=1}^4 P(Age_i)P(Car|Age_i)} \tag{3.35}$$

Once the prior probability for each class of each variables is estimated then, based on the facts (socio-economic and travel characteristics) provided in the database, a new partitioning of the sample space is developed as shown in Figure 3.9.

Figure 3.9 Partitioning of the Sample Space Showing the Union of Probabilistic Distribution of Different Mutually Exclusive Variables

This partitioning provides details about the conditional probability of each variable. By adding up all these probabilities, the probability of the event that a commuter would choose the private mode based on his/her socio-economic and travel characteristics can be estimated. If the socio-economic and travel characteristics are represented by B_i and $P(B_i) \neq 0$ for $i = 1, 2 \dots k$, then the event private mode (Car) to occur can be estimated by:

$$P(Car) = \sum_1^k P(B_i)P(Car|B_i) \quad (3.36)$$

Similarly, the probability for the commuter to choose bus can also be estimated. Then, based on estimated probabilities a mode choice decision is generated in favour of the mode which has the higher probability of occurrence.

3.7 INTELLIGENT TRANSPORTATION NETWORK SIMULATION MODEL

In this study, an Intelligent Network Simulation Model (INSIM) is developed by interfacing the INSIM Expert System with a transportation network simulation model. The simulation model used for this study is PARAMICS, which is an advanced suit of software tools for microscopic traffic simulation. It can model the individual vehicles in fine details for the duration of their entire trip. It also allows access to a library of simulation functions for advanced users. This unique feature provides the opportunity to the user to customise many features of the underlying simulation model through an Application Programming Interface (API). In this study, the API is used for interfacing the INSIM Expert System with the transportation network simulation model. The architectural design of INSIM is presented in Figure 3.10.

3.7.1 Working Mechanism of INSIM

In general, travellers are faced with a choice set that includes destination, mode, departure/arrival time, route and any of their combinations. Commuter work-related trips usually have the destination fixed, and they can choose departure/arrival time, modes, and routes based on traffic conditions. Such choice dynamics and consequently traffic dynamics can vary in scope depending on ‘within-day’ or ‘day-to-day’ dynamics. From the modelling perspective, simplification is usually made by fixing the departure/arrival time and mode, and limiting the choice set to route choice only. INSIM relaxes such modelling limitation by expanding the choice set to mode and route choices. The mode choice is dependent on pre-trip information

and route choice is based on en-route information. The INSIM components model the mode choice and the route choice.

Figure 3.10 Architectural Design of Intelligent Network Simulation Model

INSIM has four main components; INSIM expert system (IES), transportation network simulation model (PARAMICS), application programming interface (API), and data transfer interface (DTI). The IES, as described further in Chapter 6, is responsible for generating mode choice decision. The transportation network simulation model simulates the traffic and generates travel statistics as output files. These output text files are collected by the API, and transformed to Microsoft Access Database (MDB) files and exported to DTI, which is a database that stores the PARAMICS output files in MDB format while retaining the original time-dependent updated OD matrix file intact. The API takes in the information in the OD matrix file and releases the vehicles in the simulated network according to the

pattern specified in the file. The complete functional cycle of INSIM is shown in Figure 3.11.

Figure 3.11 Complete Functional Cycle of INSIM

As illustrated in Figure 3.11, the components of INSIM are activated in a sequential manner, which repeats itself after every fixed time stamp. The time stamp is a user defined function, which can be any duration of time periods between 1 and 20 minutes. Initially IES generates an OD matrix and exports it to the network simulation model. Once the OD matrix is generated, the simulation model is activated. The model simulates the traffic and generates the output file for the specified time stamp. At the end of each time stamp the commuters are generated and are assigned travel characteristics. Based on the commuters' characteristics, modes are assigned and a new OD matrix is estimated and exported to the simulator, which is simulated in the commencing time stamp.

The functionality of INSIM allows it to simulate traffic for every time stamp, and gather real-time traveller information on the available modes existing in the network. In this study, the simulated scenario imitates a multimodal transportation network, where by integrated traveller information is being disseminated. INSIM

dynamically generates OD matrix for each specified time stamp, relaxing a very common limitation of simulation models i.e. to simulate the travel demand based on the predetermined modal split. It is also necessary to note that at the end of every time stamp the network simulation is paused till the new OD matrix for the commencing time stamp is generated. The functionality of IES has been explained in detail in Sections 3.4 and 3.5. In the following sections, only certain relevant functionalities of INSIM Commuter related to COMGEN and ODGEN are further discussed along with the detailed discussion on transportation network model, API, and DTI.

3.7.2 INSIM Commuter

To perform the functionality of assigning modes, the IES needs to be activated and then provided with certain required traveller information. The IES is activated after receiving the information about network-wide (i.e. overall) travel time and/or travelling speed. Once the IES is activated it can assign modes in response to the provided traveller information. To assign modes IES needs commuters' socio-economic and travel characteristics. The commuters and their socio-economic characteristics are generated by COMGEN, and the travel statistics are provided by PARAMICS through an output file. After every time stamp TRAFSTAT gathers (reads) the output file generated by PARAMICS via an API.

The travel statistics are sorted according to origins and destinations. TRAFSTAT matches the origins and destinations of commuters with the origins and destinations of the travel statistics and assigns the travel characteristics to each commuter accordingly. The generation of commuters is based on three user-defined parameters i.e. total number of commuters to be generated, trip release profile, and zone based percentile trip distribution. The user can define any number of commuters (trips) to be generated, and control their release to simulate any desired level of congestion at any time during the entire simulation period. The number of commuters to be generated is the number of trips that will be assigned with either public or private mode of transport for their work/school trips. This number represents the travel demand that is to be simulated.

If the total number of commuters to be generated during the simulation period is “ T_C ” and the percentage of commuters to be released in the commencing time stamp “ x ” is “ P_{Cx} ” then the commuters to be released in that time stamp will be “ t_x ”, such that $t_x = T_C \times P_{Cx}$. It is necessary that the sum of all the commuters release percentage “ P_{Cx} ”, for the entire simulation period to be equal to 100 i.e. $\sum P_{Cx} = 100$, so that the total desired number of commuters “ T_C ” are generated by the end of the simulation. Once the number of commuters to be generated in the commencing time stamp has been estimated i.e. “ t_x ”, it is necessary to assign the generated commuters certain origin and destination zones. This OD distribution of commuters is obtained by taking the product of “ t_x ” and the percentage of commuters travelling from certain origin to destination “ OD_{ij} ”, e.g. if the percentage of trips commencing from zone 1 to 2 is “ OD_{12} ”, then the number of trips commuting between 1 to 2 will be $tr_{12} = t_x \times OD_{12}$. At the end of the every time stamp $t_x = \sum (t_x \times OD_{ij})$, and at the end of simulation $T_C = \sum t_x$.

The allocation of socio-economic characteristics to the generated commuters is dependent on the percentage of each socio-economic variable defined by the user. In this study, the socio-economic variables under consideration are: gender, age, education, income, car ownership, stoppage, ERP charges, parking fees, and access and compliance to traveller information.

3.7.3 API Functionalities

An API has been developed to integrate and perform the required data transfer functionalities between the IES and transportation network simulation model. To develop the API three kinds of functions have been used. These functions are used for changing/overwriting the standard code, which is developed to run the core PARAMICS model. The functions are as follows:

- (a) QPX: Extend Standard Code - define a function in the plug-in that adds to the functionality in PARAMICS. It can be triggered by one of a large number of events e.g. when the network is loaded, saved, refreshed, or at the start/end of each time step etc;

- (b) QPG: Get a value from the Standard Code - retrieve data from an external source or within the PARAMICS simulation/graphics engines; and
- (c) QPS: Set a value in the Standard Code – set/change/add a data value to the view displayed.

The API also performs the functions to control the types of vehicles, assign aggressiveness and awareness to drivers, gather the required performance measures, and convert the output text files into Microsoft Access Database files. It should be noted that in the designed simulation environment the private mode refers to the simulated Driver Vehicle Unit (DVU), and the public mode to the simulated fixed route facilities i.e. buses and trains.

(a) Types of Vehicles

The API provides the functionality to change the vehicle composition (%) that the user desires to simulate. In the developed API, there is the provision to incorporate 18 different types of small (e.g. car or small trucks) and large (e.g. Buses or articulated trailers) vehicles. It also allows the user to define their release percentage in the vehicle population.

(b) Aggressiveness and Awareness

The developed API can override the default values for the aggressiveness and awareness of the simulated drivers.

(c) Estimation of Zone to Zone Travel Time

The API performs the functionality to estimate the minimum travel time for public and private modes of transport. In the case of public mode of transport, initially the API reads all the zones in the network, and then estimates the average travelling time from zone to zone, which is based on the travel times of last three trip arrivals in the specified time stamp between each pair of origin and destination zones. If the arriving trips in the specified time stamp are less than three then the last two arrivals are taken into consideration, but if no arrival takes place then the data from the previous time stamp is considered. Once the travel times between all the zones is

estimated, the Dijkstra's algorithm is used to find the minimum travel time path for each trip. The minimum travel time for the private mode is extracted from the output file generated by PARAMICS. In this case only the average of the last three arrivals is taken into consideration, as the arriving DVU are assumed to be following the shortest path algorithm defined within the PARAMICS model.

(d) Performance Measures Related to Public Mode of Transport

The average waiting time and seat availability for public mode of transport are extracted from the output files generated by PARAMICS. The PARAMICS output file provides details about the mean waiting time at each bus stop or transit station, which are then directly extracted by the API. The statistics regarding the seat availability are estimated by the API, from the details about the capacity, passenger queue, occupancy, number of alighting passengers and number of boarding passengers at the previous bus stop.

(e) Network-wide Performance Measures

The API also extracts the network-wide performance measures such as overall time spent by private modes in the network and overall DVU travelling speed. These performance measures are provided by the PARAMICS model in the output files and are directly extracted by the API.

3.7.4 Data Transfer Interface

The API and IES can access (read and/or write) data from the generated output files, data record files and OD matrix file, via the Data Transfer Interface (DTI), in which two common files have been developed i.e. the time-dependent OD matrix file (text file) and travel statistics file (Microsoft Access database files). The IES develops the time-dependent OD matrix and writes in it. At the same time all the relevant information regarding the generated commuters' socio-economic and travel characteristics, their mode choice decisions, and mode switching are also stored in the database file. The API reads that OD matrix file and releases the vehicles to the network accordingly. The network simulation model simulates the traffic and

generates the output files. The API then reads the output files and writes (transfers) the desired output data into the respective database files.

3.7.5 Transportation Network Simulation Model

The transportation network simulation model, i.e. PARAMICS, simulates the private as well as the public modes of transport. The private vehicles are considered as driving vehicle units (DVU) and are released onto the road network based on the generated time-dependent OD matrix. The public mode services such as buses and MRT trains are modelled as fixed route vehicles. At the end of every time stamp PARAMICS generates travel statistics as an output file and provides the desired measures of effectiveness. The PARAMICS also models the route choice phenomenon. To model the dynamic route choice behaviour the commuters are divided as familiar and unfamiliar drivers. The unfamiliar drivers basically choose routes based on the perceived static cost to their destination, on familiar (major) routes only, whereas familiar drivers have access to dynamically updated costs to the destination that reflect the time varying congestion patterns, and use them to make turning decisions at each decision point or junction in the network. Travel costs are provided at each junction in the form of cost-to-destinations table that gets updated on a regular basis. The informed population is assumed to comply in such a way that drivers always follow their perception of the provided information to make routing decisions. Perception error or variation in perceiving true travel cost is introduced by adding random perturbation (noise) to the true cost, distributed across the driver population. The PARAMICS model simulates buses, bus stops, transit (MRT), and MRT stations. The release of each public mode vehicle follows a user-defined schedule, and moves along a predefined route. The public mode services are associated with their respective stops/stations. The PARAMICS simulates the arrival rate of passengers, and number of alighting and boarding passengers at every stop/station. The PARAMICS model generates an output file at the end of every time stamp (e.g. 1 minute). The output file provides the details about the required performance measures such as the overall travel time spent by private mode in the system, average travelling speed of DVU, travel time by private and public modes

of transport, average transit waiting time, seat availability in transit services, and total number of generated vehicles.

3.7.6 Calibration of the INSIM Network

The calibration of the simulated network was done from two different perspectives. The first one was to adjust the traffic counts and the OD matrix in the simulation as compared with the existing situations, such that the simulated network within the complete simulation framework can be close to the real traffic conditions. The second one was the calibration of behavioural parameters of the PARAMICS simulation model, such that they imitate the local commuters' travel behaviour.

The calibration of the behavioural parameters was necessary because the IES was developed based on the commuters' travel behaviour observed/estimated from the travel behaviour survey conducted in Singapore. To integrate IES with PARAMICS and to allow PARAMICS to simulate the local traffic environment in accordance with IES, both models should have the same behavioural assumptions, so that the behavioural parameters that govern the PARAMICS simulation model can realistically represent the behaviour of local commuters.

(a) Calibration Process

The output parameters expected from the PARAMICS model are travel times by DVU and public mode of transport. The INSIM model is to be calibrated so as to get reliable and realistic information with respect to these parameters. As travel times reflect the performance of a network, they are affected by the travel demand and the driver behaviour. To calibrate the driver behaviour, traffic counts observed at the real network are to be compared with the simulated traffic counts. If the traffic counts are well calibrated then the parameters associated with the driver behaviour can be examined by comparing the simulated and real travel time data.

In calibrating the PARAMICS model, a sequential approach is adopted to calibrate IES within the PARAMICS environment. Therefore the time-dependent OD matrix is generated by IES, which also controls the release of commuter trips, whereas the

resulting traffic flows are simulated by PARAMICS. The calibration process is based on the following steps:

- (a) Estimation of OD matrix;
- (b) Calibration of route choice model;
- (c) Calibration of driving behaviour model; and
- (d) Calibration of public mode of transport.

The calibration process starts from an un-calibrated model, and the coding errors are checked before addressing the above mentioned calibration steps. It should be noted that the network coding errors are a major source of abnormal vehicular movements. Such errors can be found at any time during the process of calibration. The fixing of network coding errors remains an important task throughout the whole calibration process.

(b) Number of Simulation Runs

The results obtained from the PARAMICS model are random, as the model generates random numbers to release vehicles, assign vehicle types, select their destination and their route, and to determine their behaviour as they move through the network. To address this issue of randomness, multiple simulation runs with different seed number are to be conducted, and the average values for the desired performance measures are estimated. The required number of simulation runs is estimated based on:

$$n = \left(\frac{ts}{\mu\epsilon} \times 100 \right)^2 \quad (3.37)$$

where “ n ” is the required number of runs, “ μ ” is the mean travel time in the runs, and “ s ” is the standard deviation, “ ϵ ” is desired margin of error (percent of “ μ ”), and “ t ” (Student’s t-statistic) is the confidence co-efficient.

All performance measures of interest are obtained, and the highest value of “ n ” is taken as the required number of runs. If the current number of runs is already larger than this value, the simulation of that scenario is ended. Otherwise, an additional

run is performed and then the required number of runs needed is recalculated. At the beginning of each calibration step the required number of simulation runs for that calibration is determined. A 95% confidence interval and a 2.5% allowable error are used in the calculation. The system level measure i.e. generated number of vehicles (NV) is taken as the criterion to select the required number of simulation runs. The number of simulation runs that resulted in the median NV is selected as the representative traffic conditions for calibration.

(c) *Goodness-of-Fit Measures*

The purpose of the calibration is to minimise the deviation between the observed and corresponding simulated traffic counts at selected measurement locations for the simulation period. The objective function can be stated as:

$$\min \sum_{n=0}^N (V_{obs_n} - V_{sim_n})^2 \quad (3.38)$$

where “ N ” is the total number of data collection points, “ V_{obs_n} ” and “ V_{sim_n} ” are the observed and simulated traffic counts at data collection point “ n ”, respectively. There are 18 data collection points selected for this study. The GEH criterion used by British engineers (UK Design Manual for Roads and Bridges, 1996) is applied here:

$$GEH = \sqrt{\frac{(E - V)^2}{(E + V)/2}} \quad (3.39)$$

where “ E ” is the candidate data and “ V ” is the average data. If the GEH values for more than 85% of the data collection points are less than 5, the objective function is satisfied.

In order for the calibration methodology to be efficient and robust the goodness-of-fit tests used should not just provide a metric describing the fit, but they should include information as to what is the nature of discrepancy between reality and simulation, so the user can target specific parameters for calibration. A widely used error measure that can provide a fairly good initial estimates of the degree of fit

between the simulated and the actual traffic measurements is the Root Means Squared Percent Error (RMSP), defined as:

$$RMSP = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - y_i}{y_i} \right)^2} \quad (3.40)$$

where “*RMSP*” is the root mean squared percent error, “*x_i*” is the simulated traffic measurement at time “*i*”, and “*y_i*” is the actual traffic measurement at time “*i*”. It is used to compare the deviation between the actual traffic data and the simulated output.

3.8 TRAVELLER INFORMATION SYSTEMS

The designed INSIM setup, which consists of the INSIM Expert System and Microscopic Traffic Simulation Model, imitates a virtual transportation network with the provision of traveller information systems. This simulation model provides a flexible platform to dynamically analyse the performance of multimodal networks with a variety of elements that are subjected to external uncertainties or correspond to system design parameters. Thus, the controlled interactive experiments conducted in a simulated environment are proposed in this research. These experiments can provide the platform to analyse the commuters’ mode choice behaviour under the influence of AMTIS. The designed approach provides a higher degree of experimental control, systematic investigation of network performance and its determinants, and a wide range of scenarios that are not practically available through observation. The prime objectives of conducting these experiments are:

- (a) to analyse the significance of commuters’ mode choice model in an intelligent transportation network simulation framework;
- (b) to quantify the impacts of various information provision strategies, and the traveller information update timing on commuters’ mode switching propensity; and

- (c) to evaluate the influence of traveller information on the network level-of-service by incorporating certain changes in the transportation infrastructure/network.

3.8.1 Experimental Factors

The purpose of conducting the experiments is to evaluate the impacts of different schemes in disseminating traveller information with the aim to improve the overall level of service of the transportation network. The factors related to traveller information that are studied within such scenarios can be broadly segregated into four categories: type of traveller information, transit operations, development of transportation infrastructure, and incident management. These factors in certain ways can influence the commuters' mode choice behaviour resulting in certain changes that can be reflected in the overall level of service. The following discussion provides the details with respect to these factors.

(a) *Traveller Information*

The effect of different strategies to supply traveller information is examined. There are two factors that may affect the commuters' mode switching behaviour (from private to public mode). These factors are type of information, and quality of information. To analyse the impact of different types of traveller information, four different cases are simulated. In each case different amount of information is provided to commuters and their mode choice behaviour is studied. The other factor which is the quality of information is assumed to be dependent on the update time, such that if the information is updated within a short interval of time it may be more reliable (or more updated) as compared to when the information is updated after a longer interval of time, e.g., 5 minutes. The effect of uncertainty and accuracy of the disseminated information can also be analysed with the help of this factor. To analyse the impact of information update time (or frequency), four different cases are simulated with different update time intervals.

There are two factors that are related to the commuters. These factors are the number of commuters with access to traveller information and the commuters' level of compliance to the provided information. These factors can also be used to

mitigate the congestion. The increase in the number of commuters accessing traveller information and the higher level of compliance may result in a higher mode switching propensity, which can help to mitigate the level of congestion by reducing the number of private vehicles on the surface street network. To analyse the effects of access to traveller information, four scenarios are simulated with different proportion of commuters having access to traveller information. The objective of analysing the access to traveller information is to find an optimal number of commuters (market share), which can improve the overall level of service of the network.

The impact of commuters' level of compliance with the traveller information is studied by simulating four different cases in which different percentages of commuters are assumed to comply with the provided information. The commuters who do not have access to information or do not comply with information are assumed to take their usual modes for travel. By analysing the commuters' level of compliance, the information regarding the proportion of commuters among those who have access to traveller information can be estimated. Such an estimate can provide the knowledge about the minimum level of compliance that can improve the overall network performance.

(b) Transit Operations

The factors that are analysed with respect to the transit (MRT) operations are the frequency of feeder buses and the transit comfort level. The frequency of feeder buses directly influences the waiting time. To make the transit mode more attractive the feeder bus frequency is improved, and three different cases are simulated in this regard. The impact of such improvement can be analysed by observing the number of commuters who switch from private to public mode.

(c) Development of Transportation Infrastructure

The effect of adding or removing new Light Rapid Transit (LRT) service is considered in this experiment. The provision of LRT service provides a seamless integration with MRT service, reduces the walking time, and the transfer time. The provision of such facility can attract the commuters and increase the transit

ridership. In this regard an experiment is conducted with an LRT service, and commuters' mode choice behaviour is observed along with the overall network level of service.

(d) Incident Management

Another aspect of analysing the dissemination of traveller information is to study its affects on the network level of service in congestion due to any incident/accident. To analyse such a scenario the transportation network is simulated with the occurrence of an incident on one of the expressways. It is expected that, with the provision of traveller information, some commuters may switch their modes of travel – resulting in certain changes that can be observed from the overall network performance.

3.8.2 Measures of Effectiveness

To analyse the impact of the experimental factors that will be discussed in Section 8.3 Measures of Effectiveness (MOE) are collected from the simulated scenarios. These MOE provide the basis to study and draw conclusion regarding the experimental factors. The MOE are segregated into two categories; the first category captures the overall network performance, and the second category is about the transit corridor. A total of 6 zones (zones 3, 8, 9, 10, 12 and 13) are classified as transit corridor, because in these zones the number of MRT stations is more than one, and the MRT stations are within the walking distance for most of the generated trips. The MOE considered are:

- (a) Percentage of commuters taking private mode of transport;
- (b) Percentage of commuters taking public mode of transport;
- (c) Percentage of commuters switching their modes from private to public – it should be noted that mode switching percentage is calculated based on the figures provided in the base case;
- (d) Average vehicle travel time by private mode;
- (e) Average vehicle travelling speed of private mode;
- (f) Average vehicle travel time by public mode;

- (g) Average vehicle travelling speed of public mode; and
- (h) Average transit waiting time.

The public or private mode is the mode of travel that a commuter has chosen for his/her journey. The average vehicle travel time refers to the average time spent by a commuter in the network from the point of origin to destination. The procedure to estimate the average travel time for public or private mode of transport shall be discussed in Section 7.2. The average travel time for the transit corridor trips is only estimated for the trips that originate and are destined within the transit corridor. All the other trips with one or both ends outside the transit corridor are considered as network-wide trips.

The average transit waiting time ($atwt$) is the overall average of the estimated waiting times (wt) corresponding to the number of transit stops (ts) within the network, such that $atwt = wt/ts$. The average waiting time (wt) at a transit stop is estimated by dividing the total time lapsed (T) by the total number of arrived passengers (np) during the lapsed time (T) at the stop i.e. $wt = T/np$, where the lapsed time (T) is the difference between the departure time (t_0) of last transit service and the arrival time (t_1) of the next transit. It should be noted that mode switching (in this study) refers to the change of mode from private to public only. The mode switching percentage is estimated as the difference between the base-case modal split and the experimental case being considered. All the above mentioned MOE provide the details about the network travel environment and the impact of traveller information on commuters' mode choice behaviour. These MOE can be a direct measurement to assess the changes occurring in the network with respect to the changes in experimental factors.

3.8.3 Experimental Setup

The experimental setup designed to achieve the mentioned objectives consists of eight sets of experiments, each of which simulates the same transportation network and the number of generated commuters discussed in Sections 8.2.2 and 8.2.3 respectively. The following gives a brief discussion on the eight experiments.

- (a) The first experiment is taken as a base case which provides the details about the network conditions without availability of traveller information.
- (b) In the second set of experiments, the impact of the different levels of detail of traveller information on commuters' mode switching propensity is analysed.
- (c) In the third set of experiments, the impact of update frequency (or the reliability of information) is analysed.
- (d) The fourth set of experiments focuses on the number of commuters with access to traveller information.
- (e) The fifth set of experiments analyses the impact of commuters' level of compliance on the network performance.
- (f) The sixth set of experiments provides the grounds to analyse the improvement in public transport level of service, i.e., waiting times.
- (g) The seventh experiment evaluates the improvement in the overall network level of service by introducing the LRT service.
- (h) The eighth experiment provides the insights to evaluate the influence of traveller information in a congested environment due to the occurrence of an accident.

Each simulation is performed for one hour time period (i.e. 0800 hours to 0900 hours) with three different seed numbers. In the simulation, there are three different modes of transport i.e. car, bus and transit, available for each commuter to "choose". They can also choose any combination of routes comprising bus and transit modes depending on the level of comfort and travel time. The level of comfort is assumed to be based on the seat availability and number of transfers. It is assumed that the commuters all have access to pre-trip traveller information on all the available modes. Furthermore, it is also assumed that there is no significant capacity constraint with respect to the public mode of transport. The overall network traffic data are collected after every minute. There are loop detector stations, which are imaginary data collection points designated on the simulated

network, placed across the entire network. In a similar way, the traffic data are also collected for the transit corridor.

3.9 SUMMARY

In this chapter, an overall conceptual framework is provided for the development of INSIM that is capable of evaluating the impacts of integrated real-time traveller information on commuters' behaviour in different scenarios. The INSIM consists of two main components i.e. IES and a transportation network simulation model. The mode choice decision-making process of IES is based on the PRB model and depends on rules that are gathered from the travel behaviour survey. To analyse the working mechanism of INSIM a set of simulation experiments have been designed. These experiments have a two-fold impact; one is to validate the working mechanism of INSIM, and the other is to recommend traveller information supply strategies that can improve the transportation system performance. The use of INSIM to simulate traffic system performance and to evaluate different integrated traveller information schemes can overcome the limitations that are related to the analytical formulation of the prime objective.

CHAPTER 4

TRAVEL BEHAVIOUR SURVEY

4.1 BACKGROUND

This study aims to investigate the use of the rule-based intelligent agent to capture commuters' mode choice behaviour, and to analyse the effects of real-time multimodal traveller information on the performance of a multimodal transportation network in Singapore. The overall modelling framework consists of a decision component, and a traffic simulation component. The decision component determines the commuters' responses to the disseminated information. An intelligent agent is used for this purpose, which can determine individual commuter's mode choice preferences in response to the disseminated information.

The development of the intelligent agent is based on the knowledge about the commuters' mode decisions with respect to the provided traveller information. The commuters' mode decisions are usually dependent on their socio-economic characteristics and travel attributes. It is thus important to understand the sensitivity of these attributes and their influence on individual's mode choice behaviour. To achieve such an understanding, a travel behaviour survey was conducted. The goal of the survey was to analyse the commuters' travel behaviour under the influence of traveller information.

The survey captured commuters' travel behaviour and provided the knowledge regarding different factors that can influence commuters' mode choice behaviour. The gathered knowledge provided insights about the commuters' behaviour in response to two different aspects: the real-time multimodal traveller information, and the congestion occurring in the transportation network. The knowledge gathered from this survey is based on the responses of commuters who make regular work/school trips from their homes to office/school destinations, and have experienced some congestion during commuting.

This chapter provides details about the design, development and conduct of the travel behaviour survey. Section 4.2 provides details about the travel behaviour framework. In Section 4.3, a discussion about survey administration, technique, and conduct is presented. The empirical findings are detailed in Section 4.4. Lastly, a discussion on the attributes that significantly influenced the mode choice decision is presented in Section 4.5.

4.2 TRAVEL BEHAVIOUR FRAMEWORK

The utility in providing traveller information, as a tool in transport policy, is that the provision of information about various modal options will allow the commuters to access information, compare information about various modes, and make rational mode choice decisions (Lyons et al., 2001; Grotenhuis, 2005). It is assumed that commuters shall attempt to find the best alternative mode using the information available to them (Kenyon and Lyons, 2003; Grotenhuis et al., 2007). Under the influence of such information, the commuters' mode choice decisions are then related to their socio-economic characteristics, the nature of journey, and the transport facility characteristics.

The effect of information on commuter's decisions depends on the information's content, format, nature, and reliability. This information can be qualitative, quantitative or predictive, as well as prescriptive or descriptive (Jou et al., 2005). Prescriptive information is normally used to advise commuters to make a specific choice. Descriptive information can be either qualitative or quantitative and usually provides information on travel times and saving in time and cost. The relevancy in the traveller information comes from the level of sophistication with which the AMTIS captures the transport facility characteristics and provides them to commuters.

The purpose of this survey was to gather perception and information, which can be reorganised to create a knowledge-base about the level of significance, and the impact of different attributes, that can be associated within the rules in order to understand individual commuter's mode choice decisions. In this study, the impacts of congestion on commuters' desire to access traveller information, the commuters'

mode switching propensity, and the commuters' personal and travel characteristics that can significantly influence their mode choice decisions were studied.

4.3 TRAVEL BEHAVIOUR SURVEY OF COMMUTERS

Travel behavioural surveys are best suited for developing behavioural models associated with traveller information (Polydoropoulou et al., 1996). This approach provides the platform to assess the potential demand and the user needs of multimodal traveller information, and its effects on commuters' travel behaviour. In this study, the focus was on travellers' mode choice behaviour under the influence of disseminated traveller information.

4.3.1 Survey Technique

To estimate mode choice models for mandatory and discretionary trips, two types of data were used: revealed preference and stated preference data. These data were gathered from the representative population through a cross-sectional survey at a single point in time. The technique adopted for this survey involved a combined strategy, which centred on combining revealed preferences (RP) and stated preferences (SP) data. The designed questionnaire that was used to gather data was quantitative in nature and presented the SP scenarios with limited choices, which were based on the data gathered from the RP section of the survey. The combination of RP and SP data can increase the usefulness of RP data and result in more realistic SP scenarios, such that both data sources jointly allow one to exploit their advantages and to overcome their limitations (Louviere et al., 2000; Lam and Xie, 2002; Bradley 2006).

4.3.2 Instrument Design

The survey questions were grouped into two sections: RP and SP. In the RP section, data on actual choices were obtained from the respondents, so as to facilitate the selection of a reference trip for subsequent SP questions. In the SP section, the selected reference trip was customised in different SP scenarios for each individual respondent. These SP scenarios were then presented to the respondent for which the responses were recorded. Implementing a valid and reliable RP and SP study

requires precise definitions of attributes, attention to presentation of preferential information (ratings, rankings, or semantic), efficient experimental design, and rigorous statistical analysis. Given the researcher's intention to use the RP and SP data jointly, the desired data were provided in semantic format, and the choice experiment appeared the most suitable because: (a) preferences are expressed in a context similar to that of an RP survey; (b) choices are perceived to be more realistic than ratings or rankings and (c) SP method allows in principle to test any discrete choice model structure. In spite of the fact that the choice context was quite typical, in order to ensure greater realism and reliability, a computer-assisted survey instrument was designed. Such an approach provided the facility to customise each SP scenario for every specific respondent, according to their information provided in the RP section during the surveys.

4.3.3 Pilot Survey

A pilot survey was conducted within the premises of Nanyang Technological University campus before carrying out the main survey. The purpose of conducting the pilot survey was to ensure that the survey questionnaire was well designed and to provide training to the survey staff. During the pilot survey, the focus group, which comprised car users/drivers, was given the survey forms and their responses were gathered. A total of 23 samples were collected. The main problem that was observed during the survey was the wording of the questions. This problem was resolved by taking the opinions of respondents about different questions and its wording. These questions were then rephrased accordingly. After the questionnaire was amended to satisfaction, the main survey was conducted.

4.3.4 Survey Administration

A computer-assisted travel behaviour survey was conducted between 13th and 21st June 2003. The participants were randomly selected within a continuous nine-day period, which served to cover all differential influences (by days of the week inclusive of weekend) of external events. A broad demographic mix of participants was selected to ensure that commuters in all major categories were represented.

(a) *Sample Size*

If the subjects are chosen randomly and assuming that the distribution of the characteristics in the population as being normally distributed, sample size (N) can be determined from the following formula (Walpole, 2006):

$$N = \left[\left(Z_{(1-\alpha/2)} \sqrt{pq} \right) / e \right]^2 \quad (4.1)$$

where “Z” is the standard normal variate; “(1- α)” is the desired confidence level; “e” is the tolerance error; and “p” is the population proportion.

The focus group in this study was the commuters who use private mode of transport. In Singapore context, the proportion of such focus group was about 41.6% (Singstat, 2004) within the total commuter population. Thus, taking the worst case when $p=0.5$, the level of confidence at 95% or $Z_{(1-0.5\alpha)}$ equal to 1.96, and a tolerance error (e) of 0.05, the sample size was estimated to be 385 respondents. This sample size was considered as the minimum requirement; therefore the sample size was chosen to be greater than the calculated value of 385 respondents. For the case herein, a sample of 479 individuals across Singapore was undertaken to estimate the models. Another sample of 200 subjects was sampled across Singapore for validation of models.

(b) *Sampling Procedure and Data Collection*

The sampling of subjects was carried out in the central and northern parts of Singapore. The central and northern parts were selected due to the reason that the transportation network simulation model mainly covered these parts of Singapore. The participants were contacted by the interviewers at various locations such as petrol stations, car parks of shopping centres, and food centres. The surveys took place when the subjects drove to these locations to fill up petrol, to do shopping or to have their meals. The discussions with the participants were firstly aimed at highlighting the modal choice and travel behaviour to give an understanding of the decision-making process for the current modal choices. Then the use and influence of multimodal traveller information on the travel behaviour and modal choice was

discussed. After which they were required to provide their personal characteristics, and details about their usual travel plans. Later, hypothetical scenarios representing different traveller information schemes were presented to them and their preferences were gathered. No gifts were offered to the respondents, and it emerged during the survey that they were genuinely interested in the experience, and placed confidence in the organisation (Nanyang Technological University) that carried out the survey.

4.3.5 Data Requirements

The data gathered in the RP and SP survey can be grouped into the following categories:

(a) Personal information

A person's socio-economic characteristics can have strong influences on his/her mode choice decisions (Ortuzar, 2001). To analyse the possible impact of socio-economic characteristics, information was gathered about the respondent's gender, age, education, nature of job, income level, and car ownership status.

(b) General trip information

Information was required about the respondents' usual/normal travel pattern, as it can help in the design of SP scenarios, and enhance its degree of realism. Such information can be helpful in estimating the realistic level of attributes that can be easily conceptualised by the respondents. During the survey, respondents were asked about their usual mode of travel, journey start time, total journey time, acceptable delay, access of traveller information, and sources of traveller information.

(c) Private-mode trip information

This information was obtained from the respondents regarding their attributes related to work/school trip, if private mode of transport was taken. The gathered information was about car usage per week, Electronic Road Pricing (ERP), number and nature of stoppage, access and egress time to and from parking lots, parking fee, and average travelling speed or distance.

(d) Public-mode trip information

Information was gathered on the availability of public mode of transport for respondent's work/school trip. This information was helpful in developing the public mode options in hypothetical scenarios. The information focused mainly on the access and egress time to and from Mass Rapid Transit (MRT) station, access and egress mode to and from MRT station, usage of MRT per week, total journey fare, total journey time, total waiting time, total number of transfers, and respondent's opinion about transit service accessibility.

(e) Pre-trip response to congestion information

The knowledge that the network is congested may influence commuters' mode choice decision. In such situation, the commuters may desire to access more information on other available modes and may change their travel plan. The respondent's preferences under such situations were duly gathered.

(f) Existing travel pattern

The respondent's preference to existing travel conditions was obtained. Such information was utilised to develop the RP models. These models were used in analysing the reliability of the preferences that were gathered from SP scenarios.

(g) Willingness to change existing mode choice

In a congested environment, commuters may be willing to change their mode of travel if certain incentives (e.g. reduction in travel time) can be achieved on alternative modes. Respondents' behaviour was thus noted under the influence of congestion in a multimodal and information-rich environment, with the incentives being provided on the public mode of transport.

4.3.6 RP Section

The RP section was designed to obtain the behavioural responses of commuters in the existing travel environment, and to have some grounds to establish a platform where new alternatives concerning the impact of traveller information could be

analysed in the SP section. The data collected in RP section focused on four categories:

- (a) Personal information;
- (b) General trip information;
- (c) Private-mode trip information; and
- (d) Public-mode trip information.

4.3.7 SP Section

The main aims of the SP survey were to gather information about commuters' desire to access traveller information, the impact of traveller information on commuters' mode switching propensity, and commuters' mode choice behaviour in an information-rich environment. The commuters' responses were to be observed under the influence of congestion and in the presence of real-time multimodal traveller information. Five scenarios were presented to the respondents.

Scenario 1 (SP1) on Accessing Traveller Information: the details about a congested work/school trip were first presented. Then the respondents' preferences about their desire to access traveller information and to change their travel plan under the influences of such information were collected.

Scenario 2 (SP2) on Impact of Multimodal Traveller Information: the details on two travel plans for the same congested work/school trip (used in SP1) were presented. The first travel plan provided information regarding the private mode of transport, whereas the second travel plan provided the details regarding the public mode of transport. In this scenario, respondents were required to give their preferences regarding their choice of travel plan based on the given multimodal traveller information.

Scenario 3 (SP3) on Commuters' Mode Switching Propensity: respondents were asked to detail their mode switching propensity based on the given integrated multimodal traveller information. In this scenario, the same congested work/school

trip was considered as in SP2, and integrated traveller information was provided on the public and private mode travel plans.

Scenario 4 (SP4) on Impact of Integrated Traveller Information: respondents were presented with the same congested work/school trip as in SP2, along with the integrated multimodal traveller information on the public and private modes of transport. In this scenario, some travel time incentive on public mode of transport was given, and the respondents were required to choose between the public and the private travel plans.

Scenario 5 (SP5) on Pre-Trip and En-Route Traveller Information: respondents were required to state their preferences about the likelihood of receiving pre-trip or en-route information. The following sections give detailed discussions on each of the five scenarios.

(a) *Scenario 1: Accessing Traveller Information (SP1)*

Figure 4.1 presents a generic template for the hypothetical scenario (SP1). In this scenario, the respondent was provided with the information on his/her work/school trip. The provided details were about usual travel time, expected travel time due to congestion and the amount of delay. The respondent was then asked about his/her desire to access traveller information, and the desire to switch from his/her regular travel plan.

All the transport facility characteristics and the attributes related to the school/work trip were based on the information provided by the respondent during the RP survey. The usual travel time was the time that was provided by the respondent as his/her usual travel time for work/school trip. The congested travel time was based on the respondent's acceptable delay, and it was estimated by taking the sum of usual travel time and the generated delay. The generated delay is the product of Acceptable Delay (AD) and a uniform Random Number (RN). The RN was generated between 0.5 and 1.5, such that due to RN the delay can be randomly increased or decreased.

This exercise can influence the commuters' travel behaviour. Thus, the congested travel time can be estimated as:

$$\text{Congested Travel Time} = \text{Usual Travel Time} + \text{Delay} \quad (4.2)$$

where $\text{Delay} = AD \times RN$, ($0.5 \leq RN \leq 1.5$). The range of RN was fixed between 0.5 and 1.5, so that the respondent can conceptualise certain realistic occurrences of delay in his/her trip. This range can also prevent the generation of such SP scenarios where the respondent may show any extreme behaviour. For example, if the lower range of RN was to be kept at 0.0, then there would have been no delay, and could probably result in no significant change in respondents' behaviour. The given range of RN simplifies the design of choice set, and provides an opportunity to analyse the effects of marginal delays, where the socio-economic factors can play a vital role along with the delay factor, and influence the commuters' mode choice behaviour.

SCENARIO (SP1)

Usual travel time (min)

Expected travel time due to congestion (min)

Amount of delay due to congestion (min)

Percentage increase in travel time (%)

As regular travel time of your work trip has been increased due to congestion, and you are facing some delay. Use preferences given below to determine;

How likely would you desire to receive travel information from any source?

☐ Strongly do not desire

☐ Do not desire

☒ Neutral

☐ Do desire

☐ Strongly do desire

How likely would you desire to switch your regular travel plan with the provision of real travel information?

☐ Strongly do not switch

☐ Do not switch

☐ Neutral

☐ Do switch

☐ Strongly do switch

Next Page

Record: 14

Figure 4.1 Generic Template for Scenario (SP1)

Once the respondents were provided with the required information they were asked to choose their preferences about their access to traveller information. They were given five choices based on a semantic scale (0: Strongly do not desire, 1: Do not

desire, 2: Neutral, 3: Do desire, and 4: Strongly do desire). These responses correspond with the usual discrete choice RP approach, except for the fact that both alternatives and choices are hypothetical. In a similar manner, the respondents were also asked about their level of desire to change their usual travel plan under the influence of the provided information and travelling conditions. This scenario permits a wider range of analysing the attributes that affect the commuters' desire to access traveller information and to change their usual travel plans.

(b) Scenario 2: Impacts of Multimodal Traveller Information (SP2)

Figure 4.2 presents a generic template for the hypothetical scenario (SP2). In this scenario, the impact of multimodal (private and public modes of transport) traveller information on the commuters' travel behaviour, represented in terms alternative travel plans, was analysed. The respondents were presented with the same details (as in SP1) about their congested work/school trip. The traveller information was provided on two different travel plans.

Microsoft Access - [1SPa]

SCENARIO

Usual travel time (min)

Expected travel time due to congestion (min)

Amount of delay due to congestion (min)

Percentage increase in travel time (%)

In such congestion, with the provision of following real time travel information, which travel plan would you like to choose for your trip?

☐ Travel Plan 1 ☐ Travel Plan 2 ☐ Do not select any plan

TRAVEL PLAN NO.1

Expected travel time is: min

Total Journey Cost: S\$

TRAVEL PLAN NO.2

Expected travel time is: min

Total Journey Cost: S\$

Suggested Travel Plan **Next Page**

Record: 14 of 1

Figure 4.2 Generic Template for Scenario (SP2)

The Travel Plan No.1 gave information about travelling conditions on expressway, and Travel Plan No.2 provided information about the travelling conditions on public mode of transport. In both plans the provided information included the travel time and cost on each facility. The delay and congested travel time were estimated as described in the previous section. The information about the travel time provided in the travel plans was based on the information given by the respondents regarding their usual travel time and average travelling speed for their work/school trip. Initially, the respondents work/school trip distance was estimated ($Distance = Speed \times Time$), based on their average travelling speed and usual travel time.

The auto cost included the fuel cost, parking fee and ERP charges. The parking fee and ERP charges were included in total auto cost only when the respondent had declared in the RP section that he/she paid the parking fee and ERP charges. The parking fee and ERP charges were kept the same as that mentioned by the respondent. The fuel cost was estimated as the product of trip distance (km), fuel cost (S\$/l), and average fuel consumption of 9 km/l (Singapore LTA Statistics, 2004 and 2005). The public mode cost was provided by the respondents in the RP section. The respondents were requested to choose amongst three travel plans which were based on the above information. Travel Plan 1 was to take the expressway, Travel Plan 2 was to take public mode of transport, while the third one was to retain their usual/regular travel plan. The provision of such choice option provided the insights about the respondent's habitual behaviour. The respondents were then asked to choose any one plan out of the given three choices.

(c) *Scenario 3: Commuters' Mode Switching Propensity (SP3)*

Figure 4.3 presents the generic template for the scenario (SP3) in which integrated multimodal traveller information was provided to the respondents by comparing the travel attributes of the available modes. The respondents were provided with the same congested work/school trip as shown in SP2 with the same travel plan options. In order to analyse the impact of traveller information on commuter's mode switching behaviour, extra information on both travel options (i.e. private and public) was provided.

Furthermore, the cost and travel times were also compared so that the respondents can easily analyse which travel plan is better and why. The respondents were then asked to mark their willingness to change their previously chosen travel plan, which they had selected in SP2. The respondents were given five choices based on a semantic scale (0: Absolutely No Change, 1: No Change, 2: Neutral, 3: Would Change, and 4: Would Absolutely Change), and were asked to choose any one based on their preferences.

The information for auto cost and travel time was estimated in the similar manner as in the scenario SP2. The access time and waiting time for public mode of transport were provided by the respondent in the RP section. The availability of seats was randomly varied and made available to 90% of the respondents, assuming that there was some small constraint in capacity constraint in the public mode option. It is necessary to note that the seat availability means that a seat will be available for the commuter to sit throughout his journey from point of boarding to point of alighting. Seat availability does not correspond to the standing option, standing place, or overall capacity.

Microsoft Access - [15Pa]

SCENARIO

Usual travel time (min)

Expected travel time due to congestion (min)

Amount of delay due to congestion (min)

Percentage increase in travel time (%)

With the provision of integrated traveller information regarding the time and cost saving on each option, will the preference of taking your previously selected mode change?

☐ Absolutely No Change ☐ No Change ☒ Neutral ☐ Will Change ☐ Will Absolutely Change

TRAVEL PLAN NO. 1	TRAVEL PLAN NO. 2
Fuel cost: <input type="text"/> S\$	Access time: <input type="text"/> min
ERP charges: <input type="text"/> S\$	Waiting time: <input type="text"/> min
Parking fees: <input type="text"/> S\$	Expected journey time: <input type="text"/> min
Total cost: <input type="text"/> S\$	Expected total cost: <input type="text"/> S\$
Expected travel time is: <input type="text"/> min	Seat Availability: <input type="text"/>
Estimated time saved is: <input type="text"/> min	Estimated time saved: <input type="text"/> min
Estimated cost saved is: <input type="text"/> S\$	Estimated cost saved: <input type="text"/> S\$

Next Page

Record: 14 of 1

Figure 4.3 Generic Template for Scenario (SP3)

(d) Scenario 4: Impact of Integrated Traveller Information (SP4)

Figure 4.4 presents a generic template for the scenario (SP4). In this scenario the respondents were provided with integrated traveller information, and certain incentives were given on the public mode of transport. This scenario captured the impact of traveller information on commuters' mode switching propensity in a congested travel environment.

To imitate the real travelling situation in a multimodal and information-rich environment, two different options with varying travel and comfort characteristics were developed. Integrated traveller information comparing both the options was also provided. In this scenario, the respondent was given the details about his/her usual travel time, expected travel time due to congestion and the amount of delay. The respondent was also presented with two different travel plans. Travel Plan 1 provided information regarding the travelling conditions on expressway, and Travel Plan 2, provided information about the transit mode of transport. The respondent was then asked to choose a travel plan with respect to his/her preferences.

Microsoft Access - [TSPa]

SCENARIO

Usual travel time (min)

Expected travel time due to congestion (min)

Amount of delay due to congestion (min)

Percentage increase in travel time (%)

In such congestion, with the provision of following real time travel information, which travel plan would you like to choose for work/school-trip?

☐ Travel Plan 1 ☐ Travel Plan 2 ☐ Do not select any plan

TRAVEL PLAN NO.1

Fuel cost: S\$

ERP charges: S\$

Parking fees: S\$

Total cost: S\$

Expected travel time is: min

Estimated time saved is: min

Estimated cost saved is: S\$

TRAVEL PLAN NO.2

Access mode: Walk

Access time: min

Waiting time: min

Expected journey time: min

Expected total cost: S\$

Seat Availability:

Estimated time saved: min

Estimated cost saved: S\$

Main mode: MRT

Egress mode: Bus

Next Page

Record: 4 of 1

Figure 4.4 Generic Template for Scenario (SP4)

All the transport facility characteristics and the attributes related to the school/work trip were based on the information provided by the respondents in the RP section. The usual travel time was provided by the respondent as his/her usual travel time for work/school trip. The congested travel time was based on the respondent's acceptable delay in his/her trip, as estimated in the previous scenarios.

To analyse the impact of integrated traveller information on mode switching, the scenario was generated in such a manner that the public transport provided the travel facility at par with the private mode of transport. The congested travel time on expressway, and travel time on public mode of transport were estimated, except that the mean travelling speed on expressway was reduced from 45 kph to 40 kph, at a level more comparable with the public transport.

In the same context, lower waiting time and higher rate of seat availability were also generated for transit mode. The generation of waiting times followed the normal distribution with a mean of 2 minutes and standard deviation of 1 minute. The availability of seats on MRT trains was assumed at 95% of time. With the provision of such information the respondents were presented three choices. The first choice was to opt for private mode and drive on expressway. The second choice was to switch to public transport, and the third choice was not to change their travel plan. The third option allowed the respondents to commence their trip as per their usual travel plan.

(e) *Scenario 5: Pre-Trip and En-Route Traveller Information (SP5)*

To analyse the respondents' likelihood towards accessing the pre-trip and en-route information, specifically considering their mode switching propensity, the respondents were presented with the Scenario (SP5) as shown in Figure 4.5. The purpose of this scenario was to make the commuters conceptualise the influence of information on their mode switching propensity, and to study their attitude in considering the access event (i.e. pre-trip or en-route) of information. The respondents were asked to choose any one among the given five choices based on a semantic scale (0: Strongly do not desire, 1: Do not desire, 2: Neutral, 3: Do desire, and 4: Strongly do desire).

The screenshot shows a Microsoft Access window titled 'Microsoft Access - [1SPa]'. It displays two forms side-by-side. The left form is titled 'Before your journey' and the right form is titled 'During your journey'. Both forms have a question: 'If real time travel information is provided [before/during] your journey, how likely would you desire to switch your mode or route?'. Below the question are five radio button options: 'Strongly do not desire', 'Do not desire', 'Neutral', 'Do desire', and 'Strongly do desire'. The forms are color-coded: yellow for 'Before your journey' and green for 'During your journey'. The 'Before your journey' form has a blue header box, and the 'During your journey' form has a blue header box. The status bar at the bottom indicates 'Records: 14 of 1'.

Figure 4.5 Generic Template for Scenario (SP5)

4.3.8 The Survey

The survey instrument was designed and developed using Visual Basic and Microsoft Access. Final year civil engineering students were given the task to collect the data. These students were trained during the pilot survey that was conducted within the premises of NTU. Transportation Laboratory at NTU provided the portable notebooks, which were loaded with the developed survey instrument and the surveyors were assigned different locations within Singapore to collect data. Due to the computer-aided approach, the time taken in generating any form (scenario) was almost negligible. The time taken to fill the forms was also minimum as the respondents only had to read and click the desired option.

The average time taken by each respondent to complete all the RP and SP sections was 8 minutes and 48 seconds with a standard deviation of 1 minute and 6 seconds. Data were collected from a total number of 479 respondents. All data samples were useful and contained complete information. Another set of data was collected for a separate sample of 200 respondents for model validation with similar attributes.

There was no rejection of samples neither was there error in collecting data as these data were collected from the respondents who did own a car and used their car on a frequent basis. Furthermore, the survey questions were quantitative in nature and the surveyors were present to guide and help the respondents throughout the survey duration.

The participants were presented with 9 forms, out of which 4 forms were designed for RP section and 5 forms for SP section. The presentation of the RP and SP forms was such that initially all the RP forms were shown to the respondents. Then based on their responses in the RP section the SP scenarios were developed and presented. In the RP section, information about their personal characteristics, work/school trip characteristics, car usage, and MRT usage was gathered. In the SP section, the information was gathered about the respondent's desire: to access traveller information, to change travel plan under the influence of the traveller information, to choose travel mode under the influence of the multimodal traveller information, to express the mode switching propensity under the influence of integrated multimodal traveller information, to switch travel mode under the influence of the integrated traveller information, and the likelihood to access pre-trip or en-route traveller information. The tree diagram shown in Figure 4.6 illustrates the presentation sequence of the RP and SP forms. The details of all the RP and SP forms are provided in Appendix A.

The visual presentation of all the choice-related forms (i.e. SP forms) were personalised for each individual. These survey forms were developed in a way that allowed dynamic customisation according to the respondent's provided information in the RP section. All the SP scenarios had generic templates, and based on the provided information these templates were changed accordingly. The changes that could dynamically be incorporated in these templates were related to the multimodal or integrated traveller information schemes, access mode to MRT station, egress mode from MRT station, and all the information related to travel and transport characteristics. The information presented within the visual aids was factually correct for each individual as it was based on his/her provided information.

Figure 4.6 Tree Diagram Illustrating the Presentation Style of RP and SP Slides

4.4 EMPIRICAL FINDINGS OF TRAVEL BEHAVIOUR SURVEY

The main objective of this survey was to analyse the factors that influence the commuters’ mode choice behaviour, such as the personal/socio-economic characteristics, travel characteristics, and the attributes related to traveller information. In this section, the reported behaviour of the respondents and findings related to their mode choice decisions are presented.

4.4.1 Sample Representativeness

As part of this survey, a rich source of individual data has been collected, which will be important in modelling the impact of integrated traveller information. Table 4.1 presents the comparison of the socio-economic attributes of the 479 respondents versus the Singapore census (Singstat, 2004).

Table 4.1 Summary Report on Personal Attributes of the Respondents

Personal/Trip Attributes	Travel Behaviour Survey		Singapore Census	
	Market Segment	(%)	Market Segment	(%)
Gender	Male	52.19	Male	51.14
	Female	47.81	Female	48.86
Age (years)	18 - 35	9.39		
	36 - 45	38.2		
	46 - 55	50.31		
	Above 55	2.09		
Education	A Level and Below	14.15	Primary and Below	52.17
	Bachelors	42.25	Secondary	15.94
	Masters	33.11	Post Secondary	11.6
	PhD	10.49	University	20.28
Job	Student	7.93	Student/Academic	4.3
	Academic	1.88		
	Administrative	25.26	Administrative	42.1
	Technical	32.57	Technical	35.18
Income (S\$) per Month	Professional	32.36	Professional	18.4
	Less than 1500	3.55	Less than 1999	3.35
	1501 - 3000	13.36	1999 2999	9.4
	3001 - 6000	39.67	3000 5999	29.23
	6001 - 12000	38.83	6000 7999	16.83
Car Ownership	Above 12000	4.59	Above 7999	41.19
	Yes	88.51		
Access to Transit Station	No	11.49		
	Yes	94.36		
Usual Travel Time (min)	No	5.64		
	Average	31.3		
	Standard Deviation	9		

It should be noted that the issue of sample representativeness is critical in impact evaluations. Biases can arise from a sample which does not represent the participant population, or a non-participant sample which does not represent the participant population. Representativeness expresses the degree to which sample data accurately and precisely represents a characteristic of a population's parameter variations at a sampling point. Representativeness is a qualitative parameter which is most concerned with the proper design of the sampling programme. The representativeness criterion is best satisfied by making certain that sampling locations are selected properly and a sufficient number of samples are collected.

The strength of a statistical inference is determined by the degree to which the sample is representative of the population that is, how similar in the relevant respects the sample and the population are. Desirable characteristics of sample statistics are:

- Unbiased: if the arithmetic mean of the statistic calculated for all possible samples of a given size n exactly equals its population parameter.
- Sufficient: summarises all relevant information about the parent population contained in the sample, while ignoring any sample-specific information.
- Efficient: the more the statistic values for various samples cluster around the true parameter value, the lower the sampling error and the greater the efficiency.
- Consistent: the larger the sample, the closer the statistic should be to its parameter value.

When the target population is known, it can be tested if the sample is likely to have been drawn from that population. For this test one hopes to retain H_0 :

- H_0 : The sample matches or represents the target population for the variable of interest.
- H_1 : Otherwise

If the sample is representative in some important target variables, then we have increased confidence in the validity of the sample to represent the target population. We use two different tests (i.e. Z-test and χ^2 -test) to check for sample representativeness.

Comparing the Male and Female proportions in the collected survey sample with the Singapore Census, it can be concluded that based on the Z-test for proportion that the survey sample did represent the population in terms of gender variable as Z value is 0.46 for Male proportion and -0.46 for Female proportion, which is within $Z_{0.05}$ critical value of ± 1.96 at 5% significance level. Thus, one fails to reject H_0 that the respective Male and Female proportion in the survey sample is the same as that of the Singapore population.

However, upon testing the Job variable based on the χ^2 -test, one cannot accept H_0 as the χ^2 value of 24.59 is greater than the decision criterion of 7.81 for 5% significance level with 3 degrees of freedom. This is due to the nature of the focus group, which in this survey was the driver population only, whereas the Singapore Census covers details all jobs in the entire population and not just the driver population segment.

4.4.2 Socio-Economic and Travel Attributes

The socio-economic and travel attributes of respondents have an important effect on travel behaviour. Table 4.1 presents the comparison of the socio-economic attributes of the 479 respondents versus the Singapore census (Singstat, 2004). The average respondent was middle-aged and regularly used car for commuting. The respondents were divided fairly equally between males and females, with 52% of the respondents being males. The average monthly income was in the range of S\$3001 to S\$6000, and 75% of respondents had either bachelors or masters degree qualifications. The primary occupations were technical, professional, or administrative. A large number of respondents (88%) owned a car, while almost 95% of them had access to MRT service at both their trip origins and destinations. The selected sample of respondents represented mainly the middle age, upper income groups, and well-educated individuals. There were also higher proportions of respondents belonging to the job categories of technical, professional and

administrative grades. The profile of the sample was not unexpected given that the target population was the car drivers (owners).

4.4.3 Traveller Information Sources

The respondents reported that they received information from a variety of sources. Figure 4.7 shows that about 91% of commuters accessed different sources of traveller information.

Figure 4.7 Commuters’ Access to Different Traveller Information Sources

Most of the respondents (42%) indicated that their primary source of traffic information was radio traffic reports. About 24% indicated that they relied on Expressway Monitoring and Advisory System (EMAS) or variable message signs (VMS) as a source of traffic information on their usual route. This clearly indicates that the implementation of strategically located EMAS or VMS has the potential to influence commuters’ travel choice behaviour. The commonly used pre-trip information system was internet whereby about 25% of the subjects used internet to gain knowledge about traffic conditions before commencing their journey.

4.4.4 Respondent’s Desire to Access Traveller Information

The respondents’ attitude towards accessing traveller information was examined by asking them to provide details on their access of traveller information. They were

provided with two options i.e. pre-trip information and en-route information. Of the 479 respondents, 25% of them accessed information before starting their journeys; about 66% accessed information during their journeys, while 9% did not access traveller information on a regular basis. The findings show that many commuters access traveller information before or during their journeys.

In order to gain further insight into commuters’ access to traveller information, the respondents were presented with the hypothetical scenario (SP1). The experimental sample included all respondents, regardless of their revealed preferences about regular access to pre-trip or en-route traveller information. The respondent’s desire to access traveller information was examined under different levels of congestion. The respondents were given five choices: “Strongly do not desire”, “Do not desire”, “Neutral”, “Do desire”, and “Strongly do desire”. They were allowed to choose their best suited preference. The distribution of respondents’ desire to access traveller information in congested environment is presented in Figure 4.8.

Figure 4.8 Distribution of Respondents’ Desire to Access Traveller Information in Congested Environment

It can be observed from Figure 4.8 that during congestion the commuters exhibited a stronger desire to access traveller information, especially when delay was longer than 17 minutes. To estimate the changes occurring in the commuters’ desire to

access traveller information with increasing or decreasing level of delay, the commuters’ desire was modelled against the delay variable. As the respondents’ level of desire was in the form of 5 ordered choice preferences, an ordered logit model was estimated. The results of the estimated model are presented in Table 4.2.

Table 4.2 Influence of Delay on Commuters’ Desire to Access Traveller Information

Commute Variables	Coefficients	t-Statistics
Constant	0.327	1.82
Delay (min)	0.229	15.33
Threshold parameters for index		
μ_1 “Do not desire”	0.922	5.84
μ_2 “Neutral”	2.040	10.96
μ_3 “Do desire”	4.409	16.82
Summary Statistics		
Number of observations	479	
Log likelihood function	-330.43	
Restricted log likelihood	-470.59	
Degrees of freedom	478	
Chi-squared	280.3194	
ρ^2	0.297	

In the estimated model, the commuters’ desire to access traveller information is the dependent variable (y), whereas the delay is the independent variable (x). The threshold parameters μ_1 , μ_2 , and μ_3 correspond to the commuters’ level of desire to access traveller information, such that if the estimated value of “ y ” is greater than

4.409, it will fall in the category of “Strongly do desire”, and if it is less than 0.922, but greater than 0.00, it will fall in the category of “Do not desire”. Any estimated value of “y” less than 0.00 will fall in the “Strongly do not desire” category.

It can be observed from the positive value of the “constant” variable that even if there was no delay, the commuters may still have some desire to access the traveller information. At a delay of 17.63 minutes the estimated “y” value will be 4.41, which corresponds to the threshold value of “ μ_3 ”. It means that at the delay of 17.63 minutes the commuters may fall in the category of “Do desire” to access traveller information, and if the delay further increases from 17.63 minutes the commuters may “Strongly do desire” to access traveller information.

4.4.5 Influence of Traveller Information on Respondents’ Travel Plan

One of the objectives of this study is to analyse the impact of traveller information on commuters’ desire to change their usual travel plans under different levels of delay. To analyse such behaviour, the respondents were asked to indicate their preferences when presented with the hypothetical scenario (SP1). These preferences included: “Strongly do not change travel plan”, “Do not change travel plan”, “Neutral”, “Do change travel plan”, and “Strongly do change travel plan”. The respondents were allowed to choose the best-suited preference with respect to their choices. The distribution of respondents’ desire to change their usual travel plan is presented in Figure 4.9.

Figure 4.9 Distribution of Respondents’ Desire to Change their Usual Travel Plan under the Influence of Traveller Information

It can be observed from Figure 4.9 that with an increase in delay, traveller information can influence the commuters’ desire to change their regular travel plans. To analyse the commuters’ desire to change their usual travel plans, an ordered logit model was estimated and the results are provided in Table 4.3. It can be observed from the negative value of the “constant” variable that if there was no delay, the commuters may not have any desire to change their usual travel plans under the influence of the traveller information. Furthermore, the delay of 20.16 minutes would result in an estimated value of “y” equal to 5.14. It means that the delay of 20.16 minutes may bring the commuters in the category of “Do desire” to change their usual travel plans. Any delay greater than 20.16 minutes can shift the commuters’ level of desire to the category of “Strongly do desire”.

On the other hand, a delay of 11.28 minutes ($y=2.32$) seems to be tolerable as the commuters still fall in the “Neutral” category, but any delay more than 11.28 minutes can shift their category from “Neutral” to “Do desire”. From the estimated model, it can be concluded that the commuters are willing to change their usual travel plans in response to the delay and provided traveller information. The commuters expressed the desire to change their usual travel plans at the delay of

11.28 minutes and longer. This desire becomes more concrete as the delay goes to 20.16 minutes and beyond.

Table 4.3 Influence of Delay on Commuters’ Desire to Change their Usual Travel Plan

Commute Variables	Coefficients	t-Statistics
Constant	-1.267	-7.64
Delay (min)	0.317	20.42
Threshold parameters for index		
μ_1 “Do not desire”	1.00	7.74
μ_2 “Neutral”	2.32	13.89
μ_3 “Do desire”	5.14	20.63
Summary Statistics		
Number of observations	479	
Log likelihood function	-323.01	
Restricted log likelihood	-533.95	
Degrees of freedom	478	
Chi-squared	421.89	
ρ^2	0.356	

4.4.6 Pre-Trip and En-Route Traveller Information

In this study the focus was also given to analyse the commuters’ desire for pre-trip and en-route traveller information, when considering their plan to change their usual travel mode. Should they feel that the pre-trip information is more desirable when changing their regular travel plans, then such a conclusion can enhance the reliability of the mode choice model.

To analyse the acceptance of pre-trip and en-route traveller information, the respondents were provided with scenario SP5 at the end of the SP section. They were asked to give their opinions about the impact of pre-trip and en-route information on their desire to change their regular travel plans. The respondents were given five choices: “Strongly do not desire”, “Do not desire”, “Neutral”, “Do desire”, and “Strongly do desire”. The respondents were allowed to choose the best suited preference. The respondents’ attitude towards pre-trip and en-route traveller information is presented in Figure 4.10.

Figure 4.10 Respondents’ Attitude towards Pre-Trip and En-Route Traveller Information

Figure 4.10 provides information on three different aspects. Firstly, it can be observed that respondents did not show severe negative attitude towards the access of traveller information. It could be that this SP scenario (SP5) was presented to the respondents at the end of the SP section, by which time respondents had a clear idea about the importance and utilisation of traveller information. Secondly, as the information was enquired about the influence of traveller information on their desire to change travel plans, a higher percentage of likelihood (i.e. 52% = 19% + 33%) for en-route information can be observed in the situation when the respondents were not willing to, or were indecisive about, changing their regular

travel plans. On the other hand, only 26% (10% + 16%) of the commuters showed their likelihood to access pre-trip traveller information, when they were not willing to change their regular travel plan or were indecisive. Thirdly, the respondents showed higher likelihood (about 74%) towards pre-trip traveller information if they were willing to change their regular travel plans. The results clearly indicate that the respondents' desire to access pre-trip info was higher as compared to en-route info when they were willing to change their usual travel plans.

4.4.7 Respondents' Mode Switching Behaviour

The empirical findings discussed in the previous sections provide evidence that commuters may desire to change their travel plans if they are provided with pre-trip integrated traveller information and are subjected to certain level of congestion. To have detailed knowledge about the commuters' mode choice behaviour, the influence of commuters' socio-economic variables, travel characteristics, and multimodal traveller information attributes were analysed. The analysis can provide the grounds to evaluate the significance of these factors in influencing the commuters' mode choice behaviour.

During the survey, respondents were presented with a hypothetical scenario (SP4). In this scenario, they were presented with a congested work/school trip and were provided with integrated multimodal traveller information. The information gave details regarding the travel environment on private and public modes of transport. The respondents were then asked to choose their preferred mode of travel and results were obtained from the sample of 479 respondents. It should be note that all the respondents used car as their usual mode of transport for their work/school trips.

(a) Commuters' Socio-Economic Characteristics

The respondents' socio-economic characteristics that were considered in the analysis are: gender, age, level of education, level of income, car ownership, stoppage, Electronic Road Pricing (ERP), and acceptable delay. The impact of each variable on respondent's mode choice was studied and is presented in Table 4.4.

**Table 4.4 Influence of Socio-Economic Characteristics of Car Users on
Mode Choice Behaviour**

Socio-Economic Variables (Code)		Market Segment (Car Users)	Mode Choice (Number)		Mode Choice (%)	
			Switch to Public	Continue with Private	Switch to Public	Continue with Private
Gender	Female (0)	220	106	114	48.18	51.82
	Male (1)	259	45	214	17.37	82.63
Age (years)	18 – 35 (0)	48	8	40	16.67	83.33
	36 – 45 (1)	186	104	82	55.91	44.09
	46 – 55 (2)	235	37	198	15.74	84.26
	Above 55 (3)	9	1	8	11.11	88.89
Education	A Level (0)	75	55	20	73.33	26.67
	Bachelors (1)	142	67	75	47.18	52.82
	Masters (2)	162	24	138	14.81	85.19
	PhD (3)	100	5	95	5	95
Income (S\$) / Month)	Below 1501 (0)	23	4	19	17.39	82.61
	1501 – 3000 (1)	69	25	44	36.23	63.77
	3001 – 6000 (2)	174	56	118	32.18	67.82
	6001 – 12000 (3)	186	59	127	31.72	68.28
	Above 12000 (4)	27	7	20	25.93	74.07
Car Ownership	Yes (1)	418	128	290	30.62	69.38
	No (0)	60	23	37	38.33	61.67
Stoppage	Yes (1)	4	0	4	0	100
	No (0)	475	151	324	31.79	68.21
Pay ERP	Yes (1)	285	82	203	28.77	71.23
	No (0)	194	70	124	36.08	63.92
Acceptable Delay (min)	5 (0)	161	43	118	26.71	73.29
	10 (1)	229	82	147	35.81	64.19
	15 (2)	90	27	63	30	70

In Table 4.4 the information is provided about the market segment of each socio-economic variable, distribution with respect to each mode of transport, and the details of respondents' mode switching behaviour. It should be noted that in this study the "stoppage" refers to a regular compulsory stop that the car driver makes to

drop off kids at school, spouse at work, or have breakfast before reaching his/her destination.

The mode choice data provide the details about the distribution of the respondents' public or private mode choice decisions. It is important to note that all the respondents took private mode as their usual mode of transport. Thus, the figures shown in the public mode column are the number of respondents who chose the public mode of transport, i.e. they would switch their mode of travel from private to public mode of transport.

The mode switching column reflects the comparative percentage of mode switching respondents within the specific category, e.g. based on the gender category, about 48% of the females exhibited certain degrees of mode switching behaviour, whereas for males only 17% of them had any intention of switching their modes of travel. This may imply that females have higher mode switching propensity as compared to males. The age factor also seems to influence the mode choice behaviour. It can be observed that respondents in the age group of 36 to 45 years had the highest inclination in switching their usual mode, as compared to other respondents. As for educational level, the highest mode switching percentage can be observed in respondents having "A" Level. It can be seen that respondents with middle income had a higher mode switching propensity.

Car ownership also showed some influence on mode choice behaviour. Respondents who did not own cars and drove to work were more inclined to change their mode as compared to those who did own cars. On the other hand, respondents who made regular stops were less willing to change their mode of transport. Such behaviour can be due to the commitment that they might have such as dropping their kids at school or dropping their spouses at work. Surprisingly, those who did not have to pay ERP were more inclined towards mode switching. It can be observed that acceptable delay did influence the mode switching behaviour, but not in direct relation to the magnitude of delay. By aggregating the data into two groups based on gender i.e. male and female, the influence of age, education, and income on mode choice behaviour can be further analysed.

Figure 4.11 Mode Choice Behaviour of Male Respondents Based on their Age, Education and Income

Figure 4.11 presents the mode choice behaviour of male respondents based on their age, education and income. It should be noted that for each socio-economic variable such as gender, age, education or income, the corresponding class categories were coded (see Table 4.4) as 0, 1, 2, 3, 4, e.g. income variable corresponding to class category of S\$ 6001 – 12000 was represented by code 3. Based on the coding for each class category for corresponding socio-economic variable, it can be observed that respondents in older-age groups with higher education and income levels were more inclined towards staying with the private mode of transport. Similarly, respondents in the younger age group with higher education were also inclined towards private mode of transport. As for respondents falling in middle age groups (i.e. 36 to 45 and 45 to 55 years), those with higher level of education, and higher level of income were inclined towards private mode of transport, while those with lower level of education and income were more willing to switch to public mode of transport. The inclination of respondents towards private mode seems to be more dependent on their levels of income and education, and to some extent their age

groups. Basically, respondents with higher income and higher level of education were more inclined towards private mode.

Figure 4.12 Mode Choice Behaviour of Female Respondents Based on their Age, Education and Income

Figure 4.12 presents the mode choice behaviour of female respondents based on their age, education and income. As noted earlier, female respondents were more inclined towards using private mode of transport as compared to male respondents. Younger females with higher education and income levels tended to prefer private mode of transport, which was similar for elderly females regardless of their income and education level. Females in middle age group (i.e. 36 to 45 years) with a lower level of education were more willing to switch to public mode of transport. In some cases, females with higher education but lower level of income were also willing to switch to public mode of transport.

(b) Travel Characteristics

Three factors namely delay, estimated time saving and estimated cost saving were analysed and their impacts on commuters' mode choice decision were studied. The

information on delayed travel times was provided to the respondents to analyse the effects of delay on their mode switching propensity. The information on estimated time saving and cost saving was provided to study the attractiveness of the respective mode of transport. The empirical findings regarding the delay, time saving, and cost saving are presented in Figures 4.13, 4.14, and 4.15, respectively.

Figure 4.13 Impact of Delay on Commuters’ Mode Switching Propensity

It can be observed from Figure 4.13 that information about travel time delay played a significant role in mode choice decisions. A linear relationship is observed between mode switching and delay, such that an increase in delayed travel time increased the mode switching propensity (from private to public mode). On the average about 36% of the respondents were willing to change to public mode of transport in case of 40 minutes of delayed travel time. The cross-over point, where modal split is 50%, occurred at delayed travel time of 61.34 minutes and tendency of mode switching continued with increase in delayed travel time. Such trend reflects significance of delayed travel time on commuters’ mode switching propensity.

The information regarding the estimated time saving on public mode attracted the respondents to switch from their usual mode. Figure 4.14 illustrates the impact of

estimated time saving by public (transit) mode transport on commuters' mode switching propensity. The information on estimated time saving shows a similar kind of impact as delayed travel time. It shows that information on increase in estimated time saving attracted the commuters towards higher degree of mode switching propensity. A 20-minute time saving resulted in 37% of mode switching. The cross-over point is estimated at a time saving of 27.25 minutes whence the public and private mode proportions were equal at 50%.

Figure 4.14 Impact of Estimated Time Saving on Commuters' Mode Switching Propensity

The increase in cost saving also attracted the respondents to change their mode of transport. It can be observed from Figure 4.15 that respondents who used car as their usual mode of commute for their work/school trip were influenced by cost saving. At a cost saving of S\$2.00, only a small proportion of the respondents would switch their usual mode of travel from car to public transport. This mode switching propensity continued beyond the cross-over point of a cost saving of around S\$21.

Figure 4.15 Impact of Estimated Cost Saving on Commuters' Mode Switching Propensity

(c) Information on Public (Transit) Mode Facility

To further explore the impact of public (transit) mode facility, respondents were given information on transit waiting time, seat availability, and transit fare. As the objective was to analyse the respondents' mode switching propensity, it was assumed that such information can elicit some mode switching attitude in the respondents. The impact of waiting time and seat availability are presented in Figure 4.16. It can be observed from Figure 4.16 that information on waiting times had significant influence on the mode switching propensity. A high proportion of the respondents i.e. 51% choose transit mode when they were given the information that the waiting time was 1 minute and seats were available, whereas 50% of the respondents chose transit mode with same waiting time, but without the availability of seats.

Figure 4.16 Impact of Waiting Time (min) and Seat Availability (1 if seat is available, else 0) on Mode Switching Propensity

The increase in waiting time had an inverse effect on mode switching propensity, such that at 4 minutes of waiting time, 12% respondents chose transit mode with the availability of seats, and only 9% respondents chose transit mode without the availability of seats. Information on seat availability had a significant influence on mode switching propensity when the waiting time was higher. At a lower waiting time, the seat availability did not show any significant impact on mode switching propensity. It reflects that commuters expect a higher level of service when waiting times are longer, and make a trade off between the transit level of service and their usual mode of travel, while making any mode switching decision.

4.5 SUMMARY

The review of existing literature suggests that travel behaviour surveys can provide an effective method to capture the impacts of traveller information on commuters’ travel behaviour. The designed survey technique for this study combined the revealed preference and stated preference data collection methods. In the RP section, the respondents’ socio-economic characteristics and their existing travelling conditions were captured. In the SP section, different hypothetical scenarios were presented to the respondents and their perceived travel preferences were collected.

The hypothetical scenarios were customised according to the information provided by the respondents in the RP section. This approach enabled the examination of commuters' behaviour, under specified conditions and with provision of traveller information, at a temporal resolution that is appropriate for dynamic modelling purpose. The advantages of this approach include adequate experimental control over factors of interest, and moderate cost when compared to full scale operational tests. The main objective of the travel behaviour survey was to determine the factors that influence commuters' mode choice decisions. It is clear from the empirical findings that commuters showed mode switching propensity in congested environment, and under the influence of integrated multimodal traveller information. A strong mode switching propensity was observed in the commuters when they were provided with the integrated traveller information.

The socio-economic characteristics that influence the mode choice decision include gender, age, level of education, and level of income. The mode switching propensity was prominent among middle-age male respondents with a low level of income. Females with a higher level of education were less willing to switch travel mode from private to public mode of transport. Respondents who regularly made a stop during their work trip were less willing to change their mode of travel. The key information content provided by an integrated multimodal traveller information system was tested. Information regarding travel time delay instilled the desire in commuters to access traveller information, and information on estimated time saving led to commuters to analyse different modes of travel. Information about transit facility also influenced the mode choice decision such that an improved transit level of service can increase transit ridership. From the empirical results of this survey, it appears that respondents may be willing to access pre-trip traveller information, change their usual travel plans under the influence of such information, and given certain incentives on public mode of transport may switch their usual mode of travel from private to public in congested travel environment.

CHAPTER 5

TRAVEL BEHAVIOUR MODELLING AND ANALYSIS

5.1 INTRODUCTION

The development of an intelligent agent for travel behaviour modelling is based on knowledge about the commuters' mode choice decisions with respect to the provided traveller information. These decisions are usually related to commuters' socio-economic characteristics and travel attributes. The travel behaviour survey and associated empirical findings (as reported in Chapter 4) have captured commuters' travel behaviour and provided knowledge regarding different factors that can influence commuters' mode choice decisions. The results provided insights on two different areas namely: the provided real-time traveller information, and the mode choice behaviour in a congested environment.

The effects of information on commuter's travel decisions depend on the content, format, and nature of the disseminated information. Gaining knowledge about these variables can help in the development of the mode choice model, which in this study, is a rule-based expert system. The motivation of the proposed rule-based expert system is drawn from earlier research work in developing cognitive (mental model-based) agents (e.g. Shoham, 1993; Suh and Trabasso, 1993). Cognitive agents possess a mental state (rule-base) which is composed of various mental elements: beliefs, capabilities, commitment; as well as behavioural rules. The agent then uses the rule-base to achieve certain outcomes. The rules in the rule-base are designed in such a way that each rule has a conditional part and an action part. If the conditional part is satisfied the corresponding action will be executed.

Each commuter's decision can be represented as rules, and by collecting a variety of such rules (which imitate different types of individuals), a rule-base can be formed. The rule-base can be utilised to simulate the decisions of commuters whose preferences and characteristics are known. The conditional part of a rule is based on the commuter's (respondent's) socio-economic and travel attributes, and the action part corresponds to his/her decision regarding the chosen mode. The purpose of

modelling the commuters' mode choice behaviour is to gain expertise about the level of significance and the impact of different attributes that can be associated with the rules in order to model individual commuter's mode choice decisions. Specifically, the impacts of congestion on commuters' desire to access traveller information, the commuters' mode switching propensity, and the commuters' personal and travel characteristics that may significantly influence their mode choice decisions are modelled.

This chapter is concerned with the calibration of the travel behavioural models. The following, Section 5.2, provides details about the impact of traveller information on commuters' travel behaviour. In Section 5.3, the commuters' mode switching propensity is analysed. A detailed discussion and analysis regarding the influence on integrated traveller information is provided in Section 5.4. Lastly, a discussion summarising the integration of traveller information with the socio-economic and travel attributes on commuters' mode choice behaviour is outlined in Section 5.5.

5.2 IMPACT OF TRAVELLER INFORMATION ON COMMUTERS' TRAVEL BEHAVIOUR

To gain an understanding of the commuters' travel behaviour, two different models were adopted, namely the ordered probit model and multinomial logit model. The ordered probit model focused on the impact of congestion on the commuters' desire to access traveller information, the commuters' desire to change their travel plans under the influence of traveller information, and the commuters' mode switching propensity under the influence of integrated multimodal traveller information. The multinomial logit model examined the impact of different socio-economic and travel characteristics on commuters' mode choice behaviour in a multimodal environment.

5.2.1 Commuters' Desire to Access Traveller Information

To analyse the impact of congestion on commuters' desire to access traveller information or to change travel plans, the commuter's desire was taken as a dependent variable. This dependent variable could take any value between 0 and 4

depending on the choice preferences of respondents (i.e. 0: “Strongly do not desire”, 1: “Do not desire”, 2: “Neutral”, 3: “Do desire”, and 4: “Strongly do desire”). These choice values did have a natural order, and in such case the ordered probit model is readily applicable. It would be inappropriate to use the multinomial logit model because it cannot capture the order of the dependent variables. On a similar note, an ordinary regression model would not be appropriate because it assumes differences between ordered categories of the dependent variables to be equal, whereas the data are only ordinal and the dependent variables are discrete. The results would be substantially different if ordered dependent variables are analysed using regression instead of using the ordered probit (Zavoina and McKelvey, 1975; Khattak et al., 1992).

To model the commuters’ likelihood to access traveller information, respondents were provided with the hypothetical scenario (SP1) as explained in Section 4.3.5. In the estimation of this model, the variables associated with the respondents’ usual work/school trip were entered. The final model was estimated based on the respondents’ preferences and the commute variables, and is presented in Table 5.1. The summary statistics support the reliability of the model. The lower absolute log-likelihood value at convergence makes it better from a statistical point of view. The results show that all the coefficients were significant (significance at $p = 0.10$), except the one for usual travel time. It is also clear from the estimated model that delay is a very important factor affecting a traveller’s decision in accessing traveller information.

There were 5 choices given to the respondent, out of which two choices i.e. “do desire” and “strongly do desire”, depicted the respondents’ behaviour to access traveller information. These choices correspond to the threshold parameter μ_2 ($y = 3.832$) and μ_3 ($y = 8.532$), respectively. Considering only the delay variable in the estimated model, it can be observed that the commuters’ level of desire to access traveller information can fall into the category of “do desire”, if a delay of 12 minutes is estimated to occur. Similarly, their level of desire can fall in the category of “strongly do desire” if the estimated delay is 23 minutes. Based on different levels of delay, the corresponding regions showing the commuter’s level of desire is

presented in Figure 5.1. The results showed that commuters experiencing longer travel times were more likely to access traveller information as compared to those experiencing short travel times.

Table 5.1 Ordered Probit Model Estimating the Likelihood that the Respondent would access Traveller Information under the Influence of Delay

Commute Variables	Coefficients	t-Statistics
Constant	-1.159	-2.101
Delay (min)	0.416	13.064
Usual Travel Time (min)	0.112	1.043
Acceptable Delay (min)	0.816	1.905
Car Ownership	0.825	2.702
Pay ERP	0.36	1.779
Regularly Access to Traveller Information	0.22	1.98
Threshold parameters for index		
μ_1	1.845	5.223
μ_2	3.832	9.393
μ_3	8.532	15.659
Summary Statistics		
Number of observations	479	
Log likelihood function	-406.14	
Restricted log likelihood function	-610.18	
Degrees of freedom	473	
Chi-squared	408.07	
ρ^2	0.334	

The increasing delay can motivate commuters to access traveller information to gain better knowledge about their travel plans. In the estimated model, the variable “acceptable delay” has a positive sign, which means longer acceptable delay may enhance the commuters’ likelihood to access traveller information. It can be observed from the gathered data that the commuters usually had acceptable delays of 5 minutes or 10 minutes. Thus, if only acceptable delay is considered (assuming all else to be equal), then in the case of 5 and 10 minutes acceptable delay, the

probability that the commuters would access traveller information falls in the category “do desire” or “strongly do desire”, respectively. It means that the commuters, who cannot afford to be late more than 10 minutes, had a very high probability to access traveller information. It is also evident from the model that when delay did occur, the commuters who own cars, pay ERP, and regularly access traveller information were more inclined to access traveller information.

Figure 5.1 Different Regions Illustrating the Commuter’s Desire to Access Travel Information corresponding to the Delay

5.2.2 Commuters’ Desire to Change Travel Plans

To model the commuters’ desire to change their usual travel plan, respondents were provided with the hypothetical scenario (SP1), and were asked to provide their preferences from a set of choices. Based on their preferences and delays experienced, the model was estimated. As the choices were; “Strongly do not desire”, “Do not desire”, “Neutral”, “Do desire” and “Strongly do desire”, with an order of 0, 1, 2, 3, and 4, respectively, an ordered probit model was estimated. The estimated model parameters are presented in Table 5.2.

The results show that the coefficients of all the variables were significant (at $p = 0.10$), except the coefficient for “regularly access to traveller information”. The

respondents generally presented a rather habitual behaviour and did not show high propensity in changing regular travel plans, except during congestion which may delay the trip. Such travel behaviour depicts that, if commuters are facing delay and are provided with information on the expected delays/congestion, they may change their usual travel plans.

Table 5.2 Ordered Probit Model Estimating the Likelihood of Commuters’ Desire to Change their Usual Travel Plans

Commute Variables	Coefficients	t-Statistics
Constant	-2.258	-4.312
Delay (min)	0.585	19.188
Usual Travel Time (min)	0.245	1.851
Regularly Access to Traveller Information	0.206	1.006
Car Ownership	0.306	1.937
Threshold parameters for index		
μ_1	1.736	8.561
μ_2	4.017	14.849
μ_3	9.886	19.913
Summary Statistics		
Number of observations	479	
Log likelihood function	-417.8467	
Restricted log likelihood	-699.9609	
Degrees of freedom	475	
Chi-squared	564.2293	
ρ^2	0.403	

In other words, any increase in delays enhances the mode switching propensity in commuters. If only delay is considered in the estimated model, it can be observed that at about 7 minutes of delays the commuters’ desire to change their usual travel plans under the influence of traveller information can fall in the category of “do desire”, and at 17 minutes of delays in the category of “strongly do desire”. Respondents with longer journey times are more likely to change their travel plans; it may be due to the reason that respondents like to minimise their driving effort in case of long journeys. Another reason could be the availability of options, i.e. for

longer journeys there may be more travel options to reach a certain destination. The estimated model also shows that respondents who own cars and access traveller information regularly have a higher tendency of switching their usual travel plans.

5.2.3 Impact of Multimodal Traveller Information on Commuters' Mode Choice Behaviour

The estimated models provide information about the commuters' desire to access traveller information and to change their usual travel plans, subject to the conditions that they are provided with traveller information, and that they are going to face some estimated delays during their trips. But these models do not provide any knowledge about the commuters' mode choice behaviour, specifically considering the multimodal nature of traveller information. It is also interesting to identify the commuters' personal and travel characteristics that may influence their mode choice decisions, under the influence of multimodal traveller information. In this regard, binary logit models were thus estimated to analyse the influence of different socio-economic and travel characteristics.

The scenario SP2, which refers to a congested work/school trip with information on three different travel options, was administered and the respondents were given a choice between private, public, and their usual modes of transport. The total number of useable responses was 400, out of which 245 selected private mode of transport, and 115 selected public (transit) mode of transport, and 40 opted for their usual mode of travel. Excluding the 40 responses, the remaining 360 responses were used for estimating the model. With only public and private mode data in consideration, binary logit models were developed based on two types of input parameters; the first was related to the commuters' socio-economic characteristics, and the second was related to transport facility characteristics. Different socio-economic and transport facility related variables were considered while estimating the model, and based on their levels of significance some of them were incorporated in the final model, in which the socio-economic variables used were: gender (1 for Male and 0 for Female), age (0 for age group 18 to 35, 1 for age group 36 to 45, 2 for age group 46 to 55, and 3 for age group above 55), level of education (0 for 'A' Level and Below, 1 for Bachelors, 2 for Masters, and 3 for PhD), and level of income (0 for

income group less than S\$1500, 1 for income group S\$1501 to S\$3000, 2 for income group S\$3001 to S\$6000, 3 for income group S\$6000 to S\$12000, and 4 for income group above S\$12000). The transport facility characteristics were total travel time, and total travel cost. The estimated mode choice model is presented in Table 5.3.

Table 5.3 Binary Logit Model Estimating the Respondent’s Mode Choice Behaviour, given Multimodal Traveller Information

Socio-Economic Variables	Coefficients	t-Statistics
Constant specific to private mode	11.531	1.941
Gender	2.472	1.863
Age	1.921	1.958
Level of Education	2.076	2.735
Level of Income (S\$)	0.503	1.832
Transport Facility Variables		
Travel Time (min)	-3.36	-2.464
Cost (S\$)	-0.985	-1.937
Summary Statistics		
Number of observations		360
Log likelihood function		-11.39
Restricted log likelihood		-249.53
Degrees of freedom		354
Chi-squared		100.665
ρ^2		0.815

All the estimated coefficients were significant at the 10% significance level. The value of ρ^2 was very high, and it implied that the model could describe the mode choice process well. The utility functions for the private (U_{PR}) and public (U_{PB}) modes of transport were:

$$\begin{aligned} U_{PR} = & 11.531 - 3.361 (Travel\ Time)_{PR} \\ & - 0.985 (Cost)_{PR} + 2.472 (Gender) + 1.921 (Age) \\ & + 2.076 (Education) + 0.503 (Income) \end{aligned} \tag{5.1a}$$

$$U_{PB} = - 3.361 (Travel\ Time)_{PB} - 0.985 (Cost)_{PB} \tag{5.1b}$$

The results as presented in Table 5.3 showed that in a non-congested travel environment, commuters preferred to take private mode of transport, but with increase in travel time the likelihood of taking the specific mode decreased. The alternative specific constant reflected that, if all the variables were equal the commuters would choose private mode of transport. It may be due to the higher level of comfort and freedom in mobility. Among the socio-economic variables, the coefficient for gender had a positive sign, which reflected that males were more inclined to use private mode as compared to females, probably due to the higher availability of vehicles. Similarly, the estimated coefficients for age, income and education had positive sign, which reflected that with the increase in age, income and education the probability of choosing private mode became higher. Thus, it can be concluded that in a congested environment and with higher auto cost, commuters who belonged to younger age groups, lower income groups and lower level of education might show higher likelihood towards public mode of transport. Commuters with higher income were less likely to switch their modes (from private to public), which might be attributed to safety consideration, ready availability of a car and lesser consideration of operational cost. The public mode of transport has two main attributes: travel times and travel costs. In a congested environment the transit mode may be attractive if it has a lower travel time and lower travelling cost as compared to the private mode of transport. The public mode could be made attractive by increasing the travelling cost of private mode. For example, increasing the ERP or parking charges could directly influence the attraction of public mode of transport. The impact of increase in travel cost on the existing travel demand can be estimated by pivot point modelling method, such that:

$$p'_k = \frac{p_k^0 e^{(V_k - V_k^0)}}{\sum_j p_j^0 e^{(V_j - V_j^0)}} \quad (5.2)$$

where p'_k is the new proportion of trips using mode k , p_k^0 is the original proportion of trips by mode k , and $(V_k - V_k^0)$ is the change in utility of using mode k . The impact of increasing private mode cost is presented in Figure 5.2, and it can be

observed that each 10% increase in the travel cost of private mode can increase public mode ridership by 2.12%.

Figure 5.2 Change in Modal Share due to Increase in Private Mode Travel Cost.

The estimated model shown in Table 5.3 indicates that respondents do show some likelihood of taking public mode of transport under the influence of congestion, and in the presence of limited multimodal travel information. This indicates a potential for multimodal traveller information to influence commuters mode switching propensity from private to public mode of transport. The model also indicates several significant attributes that affect the commuters’ mode switching propensity. These attributes include: gender, age, education, income, travel time and travel cost.

5.3 COMMUTERS’ MODE SWITCHING PROPENSITY

The commuters’ mode choice behaviour was analysed in the previous section by providing limited (i.e. travel time and travel cost) multimodal traveller information. To further explore the commuters’ mode switching propensity, the hypothetical scenario (SP3) was presented and the likelihood of changing their previous decision (i.e. their decision in scenario SP2) was examined. In scenario SP3, the commuters were presented with the same situation as in SP2, but with provision of integrated

traveller information on public and private modes of transport. The integrated information was provided by comparing the value of time and cost saving for each mode of transport. The information regarding fuel cost, ERP charges, parking fees, access time, waiting time, and seat availability for public mode of transport was also provided. The respondents were then asked to indicate their preferences regarding their willingness to change their previous mode decisions upon the provision of integrated traveller information. Due to the ordering of the choices, an ordered probit model was considered appropriate. Thus, an ordered probit model was estimated to analyse the commuters' mode switching propensity under the influence of integrated multimodal traveller information.

The socio-economic variables taken into consideration were gender, age, level of education, and level of income. The transport facility characteristics were access mode to MRT station (0 if the mode was walk and 1 if the mode was feeder bus), access time from home to the nearest MRT station, waiting time, seat availability (1 if the seat was available, else 0), travel time difference between private and public modes of transport, and travel cost difference between private and public modes of transport. The estimated ordered probit model is presented in Table 5.4. The estimated model has a lower absolute log-likelihood at convergence, which makes it a better model from the statistical point of view. All the estimated variables are significant at 10% significance level. The results show that commuters would be willing to switch their mode of travel if they are given comparative (integrated) information. For example, if the given information provides the comparison on the time saving by public mode, then the increase in time saving by public mode can increase the commuters' willingness to switch their mode from private to public. Similarly, the increase in cost saving can also enhance the willingness of commuters to switch their mode of travel.

From the results, information of transport facility on access mode, access time, and seat availability has significant influence on the commuters' mode switching propensity. If the MRT station is within walking distance (i.e. less access time value) and commuters are informed that they would get a seat in the MRT, then they would be more willing to switch their mode of travel. Walking mode to MRT

station increases the probability of using transit. The information regarding waiting time at MRT station also influences commuters' willingness to change their usual modes of travel. Commuters prefer lower waiting times whereas higher waiting times can reduce their willingness to switch their mode from private to public mode of transport.

**Table 5.4 Ordered Probit Model Estimating the Willingness of the
Commuter to Switch His/Her Mode of Travel, given Integrated
Multimodal Traveller Information**

Socio-Economic Characteristics	Coefficients(β)	t-Statistics	Exp(β)
Constant	2.435	2.991	
Gender	0.814	3.434	2.256
Age	0.346	2.16	1.413
Level of education	0.291	1.974	1.337
Level of income	-0.251	-1.838	0.778
Transport Facility Attributes	Coefficients(β)	t-Statistics	Exp(β)
Access mode to MRT station	-0.45	-1.701	0.637
Access time to MRT station	-0.513	-3.42	0.598
Waiting time at MRT station	-0.139	-1.832	0.87
Seat availability	0.479	1.815	1.614
Travel time difference	0.158	2.229	1.171
Travel cost difference	0.13	1.641	1.138
Threshold parameters for index	Coefficients(β)	t-Statistics	
μ_1	1.512	5.015	
μ_2	2.812	7.968	
μ_3	4.1	8.935	
Summary Statistics			
Number of observations	360		
Log likelihood function	-127.222		
Restricted log likelihood function	-175.6669		
Degrees of freedom	350		
χ^2	96.88975		
ρ^2	0.274		

The socio-economic variables all appear to be significant at 95% confidence level, except for the level of income (at 90% confidence level). Male respondents are

more likely to switch to public mode of travel. Individuals in higher age groups are more willing to switch, whereas respondents with higher income are less willing to switch, which might be attributed to their high job profile with higher level of income, where a car may be readily available and there may be no consideration of out-of-pocket costs such as parking cost or ERP charges.

5.3.1 Marginal Effect of the Estimated Probit Variables

The marginal effect of any variable can be interpreted as the expected change in the odds of belonging to any category of the dependent variable. In this case it is equivalent to the willingness to change their usual mode of travel, which has five categories i.e. “Absolutely will not change”, “Will not change”, “Neutral”, “Will change” and “Will absolutely change”, with an order of 0, 1, 2, 3, and 4, respectively. The expected change in the odds is the multiplicative effect of the exponential value of estimated coefficient i.e. $\exp(\beta)$, given a unit change in the variable under consideration.

By examining the estimated socio-economic variables, it can be observed that the effect of gender on the odds is 2.256. This marginal effect suggests that the odds for males to be in category 4 “Will absolutely change” instead of category 3 “Will change” are about 2.256 times higher as compared to females. Similarly, the odds for the males to be in category 3 “Will change” instead of category 2 “Neutral” are about 2.256 times as high as respondents who are females.

The effect of the level of income can be interpreted in a similar way. The probit estimate for level of income is -0.251, and the corresponding effect of odds after taking exponentiation is 0.778. Other things being equal, the odds of being in category 4 “Will absolutely change” versus category 3 “Will change” would be increasing by 0.778 times with each unit increase in the level of income. By observing the transport facility attributes it can be concluded that the marginal effect of waiting time at MRT station (i.e. 0.87) is higher than that of the access time to MRT station (i.e. 0.598). Similarly, the marginal effect of travel time difference (i.e. 1.171) is more effective than the marginal effect of cost difference (i.e. 1.138). This interpretation provides a naive method of evaluating the marginal

effects on the choice variable, because it estimates a linear growth for each category. In reality it may not be true.

5.3.2 Predicted Probabilities based on the Explanatory Variables

To have a better understanding of the marginal effects on the expected change in the odds of belonging to any category of the dependent variable, the predicted probability for a given set of values of the estimated variables is computed. The predicted probability with respect to any variable can be computed as:

$$\text{Prob}(y = j) = \Phi \left[\mu_j - \sum_{k=1}^K \beta_k x_k \right] - \Phi \left[\mu_{j-1} - \sum_{k=1}^K \beta_k x_k \right] \tag{5.3}$$

where “y” is the stated choice falling in “J” number of ordered categories, “ $\Phi(\square)$ ” gives the standard normal cumulative probability, “ μ ” is the threshold parameter separating the adjacent categories, “ β ” is the parameter estimate for the interaction term, and “ x_k ” is the k^{th} variable involved in the interaction.

Figure 5.3 Predicted Probabilities at Different Levels of Access Time

To facilitate the interpretation, all the predicted probabilities at different levels of the access time, waiting time, travel time difference, and cost difference are

estimated based on this equation. During the estimation of predicted probabilities all the other variables are valued at their means except the variables for which the probabilities are being estimated. Figure 5.3 presents the predicted probabilities at different levels of access time. It can be observed that increasing access time decreases the likelihood of mode switching propensity.

The mode switching willingness can be categorised as maximum when the choice variable i.e. “y” falls in the category y=3 or y=4, which corresponds to the decision 3: “Will change” and 4: “Will absolutely change”. The predicted probabilities of y=3 and y=4, at an access time of 10 minutes are 48% and 25%, respectively. These are the maximum values, because any value of access time greater than 10 minutes will result in lower values of predicted probabilities for category y=3 and y=4. It can be concluded from these estimated values that the maximum allowable access time associated with mode switching propensity can be 10 minutes, after which the predicted probabilities can change from categories y=3 and y=4 to y=2, y=1, and y=0.

Figure 5.4 Predicted Probabilities at Different Levels of Waiting Time

Figure 5.4 presents the predicted probabilities at different levels of waiting time. It can be observed that the predicted probabilities falling in categories y=3 and y=4

are decreasing monotonically. It means that the likelihood of mode switching from private to public mode of transport decreases with each minute increase in waiting time. It is estimated that the sum of predicted probabilities for categories $y=3$ and $y=4$ remains above 50% for any value of waiting time that is less than 4.5 minutes, after which the predicted probabilities can shift from categories $y=3$ and $y=4$ to $y=1$, $y=2$, and $y=3$. It means that any values of waiting time greater than 4.5 minutes may weaken the commuters' mode switching propensity.

To get good results in terms of an appreciable proportion (i.e. 63%) of the predicted probabilities falling in categories $y=3$ and $y=4$, the waiting time is to be kept below 2 minutes. The estimated probability for waiting time of 9 minutes show a cross-over point, where the predicted probability for category of $y=3$ is 25%, and the predicted probability for $y=3$ is equal to $y=1$, and that of $y=4$ equal to $y=0$. It is clear from Figure 5.4 that the estimated values of predicted probability for category $y=1$ show an increasing trend with a steep slope. It indicates that any increase in waiting time rapidly shifts large proportions of predicted probabilities from other categories to category $y=1$.

Figure 5.5 Predicted Probabilities at Different Levels of Travel Time Difference

The impact of increasing travel time difference (in terms of reduced travelling time) on predicted probabilities is presented in Figure 5.5. It is observed that the sum of the predicted probabilities for categories $y=3$ and $y=4$ will always be greater than 75% for any travel time difference of 6 or more minutes. It indicates that if the time saving on public mode is 6 minutes and above, the commuters' mode switching propensity increases rapidly, and the likelihood of taking public mode is predominant. The travel time difference is an estimate that reflects the time saving on public mode of transport.

Figure 5.6 Predicted Probabilities at Different Levels of Cost Difference

Figure 5.6 presents the predicted probabilities of mode switching at different levels of cost difference in terms of cost saving on public mode of transport. It is estimated that a cost difference up to S\$10 would result in a shift of predicted probabilities from various categories to categories $y=3$ and $y=4$. At a cost difference of S\$10 or more, the predicted probabilities shift to $y=4$. About 51% of predicted probabilities fall in the category of $y=4$ at a cost difference of S\$15. Thus, it can be inferred that mode switching propensity becomes significant at a cost difference of S\$10 and becomes predominant from S\$13 onwards. The category $y=4$ shows a continuous increasing trend with a steep slope, and capturing a major share of predicted probabilities from categories $y=1$, $y=2$, and $y=3$.

5.3.3 Marginal Effect on the Probability of Commuter's Mode Switching Decision

The impact of a unit change in the value of one variable on the overall probability of the outcome is analysed by estimating the marginal effect of that specific variable on the probability of commuter's mode switching decision. There are two kinds of variables that are under consideration, one being the continuous variables and the other dummy variables. The marginal effect on the event probability, in an ordered probit model, is expressed as the partial derivative of the probability function with respect to the continuous variable "x". In general, it can be estimated as:

$$\frac{\partial \text{Prob}(y = j)}{\partial x_k} = \left[\phi \left[\mu_j - \sum_{k=1}^K \beta_k x_k \right] - \phi \left[\mu_{j-1} - \sum_{k=1}^K \beta_k x_k \right] \right] \beta_k \quad (5.4)$$

where " $\phi(\cdot)$ " gives the probability density function. The marginal effect of a dummy variable is analysed by estimating the change in probabilities. The estimated marginal effects of probit variables on the commuters' mode switching decision are presented in Table 5.5.

Table 5.5 shows that the subject being a male would increase the likelihood of making the decision of changing the previous mode choice decision under the influence of the provided integrated traveller information. Using the change in probability method, the increase in probability of being in category y=3 is 22% and y=4 is 7.8%, while the reduction in the probability of being in category y=2 is 8.3%. The effect of age is such that with one unit increase, the probability of being in category y=3 increases by 7% and y=4 increases by 6%. A similar kind of effect is observed for education level i.e. one unit increase in education increases the probability of being in category y=3 by 6% and y=4 by 5%. This shows that increasing age and level of education increases mode switching behaviour. The effect of income is different from the other socio-economic variables. One unit increase in income decreases the probability of being in category y=3 by 5% and y=4 by 4%, while it increases the probability of being in category y=1 by 4% and

y=2 by 6%. This shows that mode switching propensity decreases by increasing level of income.

Table 5.5 Marginal Effects on the Probabilities of Commuters’ Mode Switching Decision

Variable	<i>prob</i> (y=0)	<i>prob</i> (y=1)	<i>prob</i> (y=2)	<i>prob</i> (y=3)	<i>prob</i> (y=4)
Gender	-1.763	-19.870	-8.362	22.156	7.839
Age (years)	-0.238	-5.407	-8.149	7.474	6.320
Level of education	-0.200	-4.547	-6.854	6.286	5.315
Level of income (S\$/month)	0.173	3.922	5.912	-5.422	-4.584
Access time to MRT station	0.035	0.797	1.201	-1.102	-0.931
Waiting time at MRT station	0.096	2.172	3.274	-3.003	-2.539
Travel time difference (min)	-0.109	-2.469	-3.721	3.413	2.886
Travel cost difference (S\$)	-0.089	-2.031	-3.062	2.808	2.374

Figure 5.7 profiles all the marginal effects of the respective socio-economic variables. It is clear from Figure 5.7 that the marginal effect of gender is most significant as compared to age, education, and income, which have similar degrees of marginal effects, while noting the effect of income is inverse to that of others. From this analysis, it can be concluded that males with higher level of education and lower level of income are more willing to change their mode decisions under the influence of integrated traveller information. In Table 5.5, by observing the transport facility variables, it can be seen that every unit (of 1 minute) increase in access time decreases the probability of category y=3 by 1% and y=4 by 1%. Such effect can give the understanding that improving, i.e. reducing, the access time by 5 minutes can result in increasing the probability by 5% of categories y=3 and y=4. Similarly when waiting time is improved by 1 minute, it would increase the probability by 3% for category y=3, and by 2.5% for category y=4.

Figure 5.7 Marginal Effects of the Socio-Economic Variables

The travel time difference shows a different pattern. Increasing time difference increases the mode switching propensity. One unit (of 1 minute) increase in travel time difference increases the probability of category $y=3$ by 3.4% and $y=4$ by 2.9%. Similarly, S\$1 increase in cost difference increases the probability of category $y=3$ by 2.8% and $y=4$ by 2.3%, while it reduces the probability of category $y=1$ by 2.03% and $y=2$ by 3.06%. It means that the integrated information that provides details on time saving and cost saving for public mode transport can enhance commuters' mode switching propensity. When the marginal effects of transport facility variables are sorted accordingly, it can be observed that the marginal effect of travel time saving is more significant as compared to the rest. The least influence is access time, with marginal effect in the range of 1%. This can also be observed from Figure 5.8, which consolidates the marginal effects of the respective transport facility variables.

Figure 5.8 shows that the effect of travel time difference is the most significant. This may be due to the context of the conducted survey, whereby all the respondents were regular private mode users, and from their perspectives, time saving is generally more valuable than cost saving. The model presented in Table 5.5 and the conducted sensitivity analysis indicate that survey subjects did respond

to the contents of the integrated traveller information in stating their mode choice decision. This indicates a potential usage of AMTIS. The model also indicates several attributes that significantly affect the commuters' mode switching propensity in terms of influencing their choice to use the public mode of transport. These attributes were found to be: access time to MRT station, waiting time at MRT station, seat availability, travel time saving and travel cost saving.

Figure 5.8 Marginal Effects of the Transport Facility Variables

5.4 IMPACT OF INTEGRATED TRAVELLER INFORMATION ON
COMMUTERS' MODE CHOICE BEHAVIOUR

It is observed from the estimated ordered probit model (See Section 5.2.3) that integrated traveller information can influence commuters' willingness to switch their mode of travel. It is interesting to also examine the impacts of integrated traveller information on commuters' mode choice behaviour in a congested travel environment. In this regard, the respondents were presented with a hypothetical scenario (SP4), which presented the same delayed work/school trips along with the integrated traveller information as in SP2. But in this scenario (SP4), certain incentives were given to public mode of transport. The information provided by the respondents regarding the access time for public mode of transport was randomly

reduced by 10%, 30%, and 50%. Similarly, the parking cost and ERP charges were randomly increased by 50%, 75%, and 100%. The respondents were given a choice between private mode, public mode, and their usual mode of transport. A total of 400 respondents participated, and they were those respondents who had either chosen private mode or public mode of transport in the previous questions. A binary logit model was developed to capture the mode choices based on two types of input parameters; the first one was related to the commuters' socio-economic characteristics, while the second one was related to transport facility characteristics.

Table 5.6 Mode Choice Logit Model Estimating the Respondent's Mode Choice Behaviour, given Integrated Multimodal Traveller Information

Socio-Economic	β	t-Stat	SE	P
Constant	5.731	2.253	2.544	0.024
Gender	-1.97	-1.833	1.075	0.067
Age	1.929	3.016	1.303	0.003
Level of education	1.137	2.066	0.55	0.039
Level of income	1.402	2.287	0.613	0.022
Transport Facility Attributes	β	t-Stat	SE	P
Access mode to MRT station	2.478	2.108	1.176	0.035
Access time to MRT station	1.759	1.65	1.087	0.105
Waiting time at MRT station	2.514	2.212	1.137	0.027
Seat availability	-3.161	-2.163	1.462	0.031
Travel time difference	-0.655	-2.172	0.302	0.03
Travel cost difference	-0.33	-2.585	0.128	0.01
Summary Statistics				
Number of observations	400			
Log likelihood function	-56.343			
Restricted log likelihood	-225.522			
Degrees of freedom	390			
χ^2	74.808			
ρ^2	0.663			

Different socio-economic and transport facility related variables were considered while estimating the model, and based on their levels of significance some of them

were incorporated in the final model. The levels of each transport facility attribute were entered into the model and were assumed to be provided by an AMTIS according to the SP design. This effort would enable the identification of two important aspects of the disseminated integrated traveller information. First, the usage of the provided information by the respondents, and second the significant attributes that are considered important by the respondents. The sign convention, estimated coefficient values, and their corresponding significance level are presented in Table 5.6. All the included variables were significant at 90% confidence level. The likelihood static ratio shows that the model was significantly different from the null or intercept-only (or know-nothing) model by the χ^2 test (74.8 with 10 degrees of freedom). All the variables have an estimate significantly different from 0, as judged by the size of β relative to its estimated asymptotic standard error, and further indicated by the column labelled p , which gives the upper bound of the probability of making Type 1 error. The lower value of log likelihood function and the estimated value of ρ^2 reflect the robustness of the estimated model. These test statistics show that the model can describe the mode choice process well.

The utility functions for the private (U_{PR}) and public (U_{PB}) modes of transport are:

$$\begin{aligned}
 U_{PR} = & 5.731 - 0.655 (\text{Travel Time Difference}) \\
 & - 0.330 (\text{Travel cost difference}) \\
 & - 1.970 (\text{Gender}) \quad 3.929 (\text{Age}) \\
 & + 1.137 (\text{Education}) + 1.402 (\text{Income})
 \end{aligned} \tag{5.5a}$$

$$\begin{aligned}
 U_{PB} = & - 2.478 (\text{Access Mode}) \\
 & - 1.759 (\text{Access Time}) \\
 & - 2.514 (\text{Waiting Time}) + 3.161 (\text{Seat Availability})
 \end{aligned} \tag{5.5b}$$

The estimated model consists of one alternative specific constant, four socio-economic variables, and six transport facility attributes. The sign convention of each variable provides the information from which the commuters' mode switching behaviour can be analysed. The positive value of alternative specific constant shows that the commuters' tend more towards the private mode of transport as compared

to public mode. The model results show that with the provision of information female commuters are more likely to continue using their usual modes of transport, as compared to male commuters. This may be due to a lower tolerance level of congestion among males. Higher age group commuters with a high level of education show higher tendency to use private mode. It can be due to the level of comfort and/or privacy that they may need with respect to their socio-economic status. Commuters with lower income level are more likely to switch from private mode to public mode. This may be attributed in part to cost saving, as the parking cost and ERP charges were increased in the scenario presented to respondents.

Among the transport facility attributes, commuters prefer to have MRT stations nearby their residence, within their walking distance, as walk is the preferred access mode. It can be due to the reason that travelling by bus to MRT stations increases the overall waiting time and the number of transfers. Lower access time is preferred, as it can save the effort utilised in travelling. Commuters prefer shorter waiting times. Increased waiting time at MRT stations decreases the tendency to take public mode of transport. The negative signs of Travel Time Difference (TTD) and Travel Cost Difference (TCD) show that commuters' likelihood of mode switching can be increased by increasing the absolute values of these variables. The TTD and TCD can be estimated as:

$$TTD = TT_{PR} - TT_{PB} \quad (5.6a)$$

$$TCD = TC_{PR} - TC_{PB} \quad (5.6b)$$

where TT_{PR} is the travel time by private mode of transport, TT_{PB} is the travel time by public mode of transport, TC_{PR} is the travel cost by private mode of transport, and TC_{PB} is the travel cost by public mode of transport. Increasing the value of travel time on private mode of transport as compared to public mode of transport would increase the TTD, resulting in less likelihood for the private mode of transport. In other words, increasing the value of TTD would result in higher mode switching as compared to a lower value TTD. Thus, by analysing the TTD and TCD, it can be inferred that increasing travel time or cost on any mode will decrease its utility.

Furthermore, increase in travel time or cost of private mode would enhance commuters' mode switching propensity.

5.4.1 Marginal Effect of the Estimated Logit Variables

The exponential values of β give the odds of having an event occurring versus not occurring, per unit change in the explanatory variables, other things being equal. The same interpretation applies to both the dummy and the continuous variables. From Table 5.6, it can be observed that the estimate for gender is -1.97. The resulting exponential value of β , which is 7.17, indicates that the odds for males are 7.171 times as high as females, to switch their usual mode of travel. Hence, males have higher probability of switching their modes of travel as compared to females.

The level of income has a negative impact on mode switching propensity. Increasing the level of income decreases mode switching, and vice versa. The effect of each stage of income (i.e. 0, 1, 2, 3 and 4) on the odds of taking usual mode under the influence of integrated traveller information is -1.402. Thus, other things being equal, a decrease in income would increase the mode switching by a factor of $\exp(-1.402) = 0.246$. In other words, the odds for a person belonging to the income group one stage lower, to switch his/her usual mode of travel are higher to the person who is one stage above that person. Among the transport facility characteristics the travel time difference and travel cost difference have the most significant effect on the commuters' mode switching propensity under the influence of the integrated traveller information. It is estimated that each minute increase in travel time difference increases the mode switching propensity by a factor of 1.927. Similarly, each dollar increase in cost difference increases mode switching by a factor of 1.392.

5.4.2 Predicted Probabilities based on the Explanatory Variables

The predicted probabilities based on the explanatory variables can provide estimates, upon which different policies can be designed and analysed. It should be noted that variables such as travel time and travel cost can be utilised to design policies which can cause diversion effect on shifting the ridership from private to

public mode of transport. There are several measures to reduce travel time by public mode of transport, such as improvement in accessibility or service frequency etc. Similarly, increasing parking cost and ERP charges can directly influence the public mode ridership. To analyse the impact of such policy sensitive variables, once again pivot point modelling approach was adopted. The results are presented in Table 5.7.

Table 5.7 Impact on Modal Split by Increasing Travel Time and Travel Cost of Private Mode of Transport

UNIT CHANGE (%)	TRAVEL TIME			TRAVEL COST		
	Private Mode (%)	Public Mode (%)	Diff (%)	Private Mode (%)	Public Mode (%)	Diff (%)
10	80.01	19.99	1.07	80.52	19.48	0.52
20	78.94	21.06	1.07	80	20	0.52
30	77.82	22.18	1.11	79.46	20.54	0.53
40	76.67	23.33	1.15	78.92	21.08	0.54
50	75.48	24.52	1.19	78.36	21.64	0.56
60	74.24	25.76	1.23	77.8	22.2	0.57
70	72.97	27.03	1.27	77.22	22.78	0.58
80	71.66	28.34	1.31	76.64	23.36	0.59
90	70.31	29.69	1.35	76.04	23.96	0.6
100	68.92	31.08	1.39	75.43	24.57	0.61
Average			1.22			0.56

The results in Table 5.7 reveal that every 10% increase in travel time in private mode, on average, decreased the modal share of private mode by 1%. Similarly, every 10% increase in travel cost of private mode decreased its modal share by 0.56%.

5.4.3 Marginal Effect on the Probability of Commuters’ Mode Switching Propensity

Instead of examining the marginal effect of an *x* variable on the odds, it is better to look at the marginal effect of the variable on the probability of the event. Such marginal effect is given by the following equation:

$$\frac{\partial \text{Prob}(y = 0)}{\partial x_k} = \frac{e^{-\sum_{k=1}^K \beta_k x_k}}{\left(1 + e^{-\sum_{k=1}^K \beta_k x_k}\right)^2} \beta_k \tag{5.6}$$

Table 5.8 presents the marginal effects of the variables on the probability of commuters’ mode switching propensity. Observing the marginal effect column, it can be seen that among the socio-economic variables, the most significant effect is for gender and age, followed by income and education respectively.

Table 5.8 Marginal Effects of the Variables on the Probability of Commuters’ Mode Switching Propensity

Variable	Coefficient (β)	Mean	Marginal Effect
Gender	1.970	0.543	0.725
Age (years)	-1.930	1.491	-0.710
Level of education	-1.137	1.750	-0.418
Level of income (S\$)	-1.402	2.241	-0.516
Access mode to MRT station	-2.479	0.578	-0.912
Access time to MRT station	-1.760	0.750	-0.647
Waiting time at MRT station	-2.514	0.664	-0.925
Seat availability	3.162	0.690	1.163
Travel time difference (min)	0.656	2.414	0.241
Travel cost difference (S\$)	0.330	2.202	0.122

Among the transport facility, the sensitive variable is the waiting time for transit service. Each single minute increase in travel time reduces the mode switching propensity by a factor of 0.925. This analysis can also be viewed in another perspective i.e. any improvement in the waiting time reduction can cause a significant impact on transit ridership, such that 1 minute reduction in waiting time will increase the mode switching propensity by a factor of 0.925. A similar kind of effect can be observed for the access time variable. It can be observed that 1 minute reduction in access time can influence the mode switching propensity by an

additional factor of 0.647. The time difference and cost difference, which in this case represent the time savings and cost savings on public mode of transport respectively, also show positive effect on commuters' mode switching propensity. It is estimated that a single minute time difference will increase the mode switching propensity by a factor of 0.241, whereas a single dollar cost saving will increase the mode switching by a factor of 0.122.

It can be concluded from this analysis that commuters show a certain level of mode switching propensity under the influence of integrated traveller information. Younger male commuters with lower level of income are more willing to switch their usual mode as compared to richer and better educated commuters. The cost factor also influences the mode choice, such that if the commuter belongs to a lower level income group he/she may show higher mode switching propensity. Richer commuters with a higher level of education are less willing to switch. The time and cost saving is appreciated by all commuters. In time context, commuters are more sensitive to waiting time than travelling time. They choose modes with overall shorter journey times. The cost saving enhances the attraction of public mode, which the commuters compare with higher private mode cost. Thus, the provision of integrated information regarding the transport facility variables can enhance the mode switching propensity in the commuters.

5.5 DISCUSSION

The main objective of the travel behaviour survey was to determine the factors that influence the commuters' mode choice decision. It is clear from the estimated travel behaviour models that commuters do show mode switching propensity in a congested environment, and under the influence of integrated multimodal traveller information. Commuters are willing to change their mode choice decisions when they are provided with integrated traveller information. A higher mode switching propensity was observed in the commuters once they were provided with (simulated) integrated traveller information. The socio-economic characteristics that significantly influence the mode choice decision are: gender, age, level of education, and level of income. A higher mode switching propensity was found for middle aged male respondents having a lower level of income. Females with higher

level of education were also less willing to switch their mode from private to public mode of transport. The respondents who regularly made a stop during their work trip were not less willing to change their private mode of travel.

The attributes related to multimodal traveller information that significantly influence the mode choice decision are: estimated time saving, estimated cost saving, access mode to MRT station, access time to MRT station, transit waiting time, and transit seat availability. The information regarding delayed travel time generated the desire in commuters to access multimodal traveller information, and information on estimated time saving allowed the commuters to analyse different modes of travel. Information about transit facility, also significantly influenced the mode choice decision, such that improved transit level of service can increase the transit ridership. The knowledge gathered from the travel behaviour survey provides the expertise to select variables that significantly influence the mode choice decision. The selected variables can then be incorporated in the rules as their conditional part, and the corresponding respondent's decisions as the action part. The selected variables for the conditional part of the rules are: gender, age, education, income, stoppage, car ownership, travel time delays, estimated time saving, estimated cost saving, access time to MRT station, transit waiting time, and transit seat availability.

CHAPTER 6

DEVELOPMENT OF RULE-BASED MODE CHOICE MODEL – INSIM EXPERT SYSTEM

6.1 INTRODUCTION

Existing transportation network simulation models focus mainly on private mode of transport i.e. car. Some of these models do simulate public mode of transport like bus and transit, but do not provide any dynamic mode choice model that simulates commuters' mode choice decisions based on real-time travelling conditions. Due to such limitations these models are not capable of simulating the impacts of transportation planning and operational strategies designed to mitigate congestion in multimodal urban networks where an appreciable number of transit and intermodal trips exist. These models are also limited in their applicability in analysing the impacts of Intelligent Transportation Systems (ITS) user services and Advanced Multimodal Traveller Information Systems (AMTIS), which can become an important factor in mitigating congestion in metropolitan areas. To develop a simulation model that can capture the dynamic interaction between commuters' mode choice behaviour, and the evolving travel environment, it is necessary to have a simulation platform that can simulate a multimodal transportation network and can capture the interaction between the existing modes of travel. It is also necessary for such a model to have a certain level of intelligence to imitate the commuters' mode choice decision-making process dynamically, by perceiving and analysing the evolving travel environment. The integration of artificial intelligent agent with a microscopic traffic simulator is therefore an appropriate approach to provide a breakthrough in overcoming this limitation. The intelligent agent is to have the function of monitoring the system performance and provide necessary actions, through the effects of switching modes by commuters, to reduce the network demand and relieve the traffic congestion in the network.

This chapter presents the framework of the development of a rule-based Intelligent Network Simulation Model known as the INSIM Expert System (IES). The IES is

an artificial intelligent agent, which is specifically designed to induce intelligence in the transportation network simulation model. The IES developed for this research consolidates parts of the experience and knowledge gathered from the travel behaviour surveys, which were carried out to analyse the commuters' mode choice behaviour in a multimodal transportation system. Based on the gathered knowledge about the commuters' mode choice behaviour a variety of rules were developed. The developed rules equip the IES with a certain level of sophistication to simulate/imitate commuters' mode choice behaviour on a micro level, and the expertise of a traveller information provider on a macro level. As a whole, with the integration of a microscopic traffic simulator, the IES allows one to analyse the impacts of different multimodal traveller information strategies on the level of service of the transportation network.

This chapter focuses on the architecture and functionality of the IES and is organised as follows. Section 6.2 describes details about the architecture and working mechanism of the IES. An overview about the characteristics of the IES is provided in Section 6.3. The description about the reliability analysis of the mode choice models is presented in Section 6.4. Lastly, the concluding remarks are presented in Section 6.5. The calibration and application of IES shall be covered in the following chapters.

6.2 THE ARCHITECTURE AND WORKING MECHANISM

The main limitation of existing transportation network simulation models is that the OD demand and the modal distribution are assumed to be constant over time, which does not correspond to the actual conditions in multimodal urban networks. Such a limitation can be overcome if the time-dependent modal distribution can be estimated and the resulting link flows over time can be determined. The IES has been designed to provide a simulation platform where a multimodal transportation network can be established and the time-dependent modal distribution can be estimated. The model is developed to capture explicitly the dynamic interaction between mode choice and traffic assignment as network conditions evolve with changing events.

The IES has two main components namely INSIM Expert (IE) and INSIM Commuter (IC) as shown in Figure 6.1. The IE provides knowledge regarding commuters’ mode choice decisions based on a set of rules, and the logic to generate suitable decisions for the commuters. The IC has the capability to generate simulated “commuters” (called “commuters” in the following sections), gather travel statistics from the network via a microscopic traffic simulator and estimate a time-dependent OD matrix.

Figure 6.1 Architecture of the INSIM Expert System

The IES initially activates the IC to generate commuters and gather travel statistics from the traffic simulator for every pre-specified time interval. The characteristics of the simulated commuters along with the real-time travel statistics collected from the traffic simulator are then stored in the database as facts. The IES then activates the inference engine to determine mode decisions for each commuter based on known facts and prevailing traffic characteristics. The IES is designed in a way to provide the flexibility to adopt alternative mode choice model that can give prediction of commuters’ mode choices.

After all the commuters are assigned certain modes, the IES again activates the IC to estimate an OD matrix for the commuters in the database. Commuters assigned with public modes are technically removed from the traffic simulation as their movements will be simulated with buses and MRT. An OD matrix file for private transport is generated for the prevailing time interval and is loaded into the traffic simulator, which releases the assigned traffic demand into the network and generates the travel statistics after a fixed time stamp. For the commencing time stamp the IC is again activated and another batch of commuters is generated and the newly generated travel statistics are imported by the IC. This cycle of activities continues for the entire simulation period. The flexibility to specify the travel statistics update time interval allows the system to predict the commuters' mode under the influence of real-time travel information and estimate a time-dependent OD matrix.

6.2.1 The INSIM Expert (IE)

The IE is the intelligent part of the IES and provides rule-based knowledge and logic to determine a suitable mode choice decision for the commuters. The inference engine in IE utilises the rules and facts to generate the commuters' mode choice decision. The IC generates commuters without any pre-determined mode choice. The IE imitates the commuters' mode choice behaviour, and allocates travel modes to them. The commuters are allocated with modes based on the rules present in the rule-base and the prevailing travel environment.

The intelligence in the form of rules is gained from the travel behaviour surveys. It was observed from the survey results that the commuters' socio-economic characteristics and the travel attributes associated with the existing travel environment influence the commuters' mode choice decision. The rules are designed in a way that each of them represents a typical commuter's decision. Each rule has a premise and a conclusion. The premise contains information about the commuter's socio-economic and travel characteristics. The conclusion provides the decision made by the commuter with respect to his/her socio-economic characteristics and perceived travel attributes.

To make a decision based on the rules developed, it is necessary that the premise should match some facts, i.e. the premise of the rule needs certain facts to become true so that certain action/decision can be taken. The facts are provided in the database and represent the problem-specific information, which can be commuters' socio-economic characteristics and/or the travel characteristics.

The matching of facts with the rule premise is done by the inference engine, which adopts a logical reasoning process and applies the rules to draw conclusions from the matched premises. For example, for a rule such as $P \Rightarrow A$, if “ P ” occurs and is found to be true, then “ A ” can be concluded. Thus, once the inference engine finds a rule with such a premise which matches the provided fact then the rule is said to fire and the action part is executed imitating the commuter's mode choice decision.

6.2.2 The INSIM Commuter (IC)

The IC is the medium of communication between the IE and the traffic simulator. It allows the IE to perceive the travel environment, makes decision about commuters' mode choice and then updates the travel demand. To assign modes to commuters and estimate a dynamic modal distribution, the detailed information about the commuters, their socio-economic characteristics and travel plan indicating their origin and destination, are required. The generation of commuters, along with their socio-economic characteristics are based on the findings of travel behaviour survey. The details of this survey have been presented in Chapter 4.

The commuters should also follow a certain time-dependent zonal demand pattern and should be released in a certain fixed time intervals over the simulated time period. The details about the time-dependent distribution of commuter trips over the study area were studied and OD matrices for each time stamp were estimated. The details of the traffic simulation are discussed in Chapter 7.

Based on the gathered information from the traffic survey, the COMGEN component of the IC generates the commuters. A specified number of commuters is generated for every fixed time stamp. The number of commuters to be generated is defined in a commuter release profile, in which the time stamp of generation is

taken as 1 to 5 minutes and the corresponding number (or percentage) of commuters to be generated are prescribed. The COMGEN initially generates the commuters. Based on the percentile zonal demand, it distributes the total number of generated commuters within certain origin to destination zones. Such functionality allows the user to have control over the total generation of commuters, the zonal demand distribution and the time-dependent release of commuters in the system. The IES helps to simulate any level of congestion with respect to an OD distribution, and analyses its effects within the transportation network. The COMGEN, as a component of IC, assigns the origin and destination zones to commuters and does not assign any mode of travel to the generated commuters.

It was found from the travel behaviour survey as discussed in Chapters 4 and 5 that socio-economic characteristics that significantly influence the mode choice decisions include age, gender, education, income, car ownership, ride sharing, regular stoppage, access to traveller information and reliability of the traveller information. The travel characteristics that were found to influence the mode choice decision are travel time, travel cost, access time, transit waiting time, transit fare and transit level of service.

COMGEN, after assigning origins and destinations to the commuters, categorises them into different groups based on age, gender, level of education, level of income, fuel cost, parking cost, ERP, and OD-based transit fare. Each commuter is also assigned four binary variables that indicate the regular mode of travel (car or public mode), access to traveller information, level of compliance to the provided information, and regular stoppage during their work trip.

All the associated socio-economic variables can be defined in a configuration file and the desired generation percentage of any variable can be specified. For example, by specifying the percentage value for the specific variable in the configuration file, the impacts of commuters' access to traveller information can be studied. Accordingly, COMGEN will allocate access to traveller information to the specified percentage of generated commuters. In this study, the distribution of socio-economic characteristics of the generated commuters is based on the

distribution of those characteristics observed in the travel behaviour survey, as discussed in Chapter 4.

The INSIM Expert (IE) imitates the commuter's mode choice decision, which is not only based on the commuter's socio-economic characteristics but also influenced by the travel attributes. To generate a rational decision it is necessary for the IE to analyse the travel characteristics corresponding to the origin and destination of the commuter. In this regard TRAFSTAT is developed as a component of IC. The functionality of TRAFSTAT is to gather the transport statistics from the traffic simulator. The transport statistics are collected from the output files generated by the traffic simulator and updated in the database with respect to each commuter's origin and destination. The simulated traffic data provide information regarding the minimum travel times by public mode, i.e. bus and/or transit, the minimum travel times by car, waiting times for bus and/or transit, numbers of transfers and seat availability in public mode. This information is zone based and provides details in the form of OD pairs. TRAFSTAT matches the OD of each commuter with those in the transport statistics and updates the commuters travel characteristics in the database. Once all the generated commuters have been updated with the corresponding travel attributes based on their OD, the IE activates the inference engine and the rules start to execute.

The inference engine activates according to the update time specified by the user and correspondingly the simulator also provides the travel statistics for the same time interval. The update time option allows the user to define any information update time scheme ranging from 1 to 20 minutes. This option develops the dynamism in the interaction between the commuters' mode choice decisions and the evolving network condition. It also provides a platform to analyse the impacts of different update timings on the reliability of the traveller information. Thus, after every specific time stamp the inference engine activates and executes the rules. The provided travel attributes represents the traveller information, and corresponding to the commuter access and level of compliance to the provided traveller information. IE makes a decision resulting in certain mode switching from the regular mode of travel. The updated mode choices change the modal split, and the component

ODGEN of IC gathers all the information with respect to the new modal split. It sorts all the commuters based on their OD and mode of travel, and estimates an OD matrix in that specific time stamp.

6.2.3 Microscopic Traffic Simulator

The simulation environment can be used to imitate the scenario of a multimodal transportation network, where real time traveller information is provided and commuters decide their mode of travel before they start their journeys. The microscopic traffic simulator provides a multimodal transportation network consisting of nodes and directed arcs with multiple origins and destinations. Different public and private modes are available for commuting from different origins to different destinations. Every public mode is described in terms of its routes, stop location, schedule and fare structure. The MRT has a separate right of way and is not allowed to mix with other traffic. The trips generated in the traffic simulator are based on the estimated OD matrix. The ODGEN component of IC estimates the OD matrix, and sends it to the microscopic traffic simulator, which loads the corresponding traffic into the network and the trips are then released accordingly in the commencing time stamp.

The total simulation time period can be any peak or off peak period of interest, which can be segregated into small intervals (time stamps) ranging between 1 to 20 minutes for the entire simulation period. The percentile distribution of time-dependent zonal demand over the simulated area is assumed to be known. Similarly, the percentile distribution of trips being released into the network is also assumed to be known. Real-time multimodal traveller information on the available modes in the network is provided by the microscopic traffic simulator.

Based on the traveller information the commuters decide about their modes of travel and undertake their trips using private cars, buses or MRT or combination of bus and MRT. The commuters evaluate different travel alternatives and make their choices based on a range of travel attributes. Once they decide about their travel mode they start their journey. Depending upon the interval of each time stamp a certain number of commuters are generated and released into the system.

6.3 CHARACTERISTICS OF THE IES

The most important characteristics of the IES are its high-quality performance and adaptive learning. The high-quality performance assures system reliability and timeliness. The adaptive learning allows the IES to update dynamically and learn on its own through its experience.

6.3.1 Reliability

The reliability of the IES depends upon its decision-making process, which can be observed by two specific features i.e. the capability to explain about the decision taken, and to deal with uncertainty.

(a) Capability to Explain

The IES is capable to explain how and why a decision was reached. The execution of rules tending towards certain decisions can be explained by the inference chain or the rule flow. A simple illustration is presented in Figure 6.2. The rule flow shown in Figure 6.2 has a rule R1 with two distinct IF conditions i.e. C1 and C2. If these two conditions are true then the following action A1 will be concluded and the actions will be taken. In case these two conditions or either one of these two conditions is not true then the concluding action will not be taken.

The rule flow structure is designed in such a way that every condition tends to a certain action, which provides the reason why such an action is taken. Similarly, every action is dependent on certain conditions and because of the occurrence of those conditions the action is taken. For example the action A3 will only be taken if the conditions C1 and C2 conclude action A1, conditions C3 and C4 conclude action A2 and condition C5 is true. It implies that the action A3 is implicitly dependent on conditions C1, C2, C3, C4, and C5 to be true. The design of rule flow structure is such that every action taken proves that the preceding conditions have been met. Thus, in rule flow the forward inference will provide the details on why a decision is taken and the backward inference will indicate how a decision is concluded.

(b) Dealing with Uncertainty

In an ideal condition, the perfect decision about any situation can be based on the knowledge i.e. facts and rules to be strictly true or false. It can then be concluded with certainty that the derived solution is correct. Unfortunately, in reality often incomplete information is observed. To overcome such situations the rule of inference has also been designed to deal with knowledge that is incomplete or not completely certain. A heuristic algorithm is developed to solve the uncertainty problem and generate a suitable decision for the provided facts.

This ability of the IES allows it to provide a solution even when the provided information is not 100 percent accurate or is incomplete. In this regard it is also necessary to analyse the level of accuracy of the provided solution. To check the accuracy of the designed algorithm an experiment was conducted and its level of mode prediction accuracy was analysed. The details of the conducted experiments are discussed in Section 6.4.

The heuristic algorithm is designed to find the nearest matching rule(s) and generate a decision. The provided fact (i.e. the commuters' socio-economic and travel characteristics) are compared with all the antecedent parts (premise) of all the available rules in the rule-base. Then the rules whose antecedent parts closely match with the fact are taken into consideration. The possibility distribution of the resulting decision is related to the possibility distributions of the crisp decision in the rule consequence, and the similarity between the provided fact(s) and the corresponding premise of the rule. The degree of similarity between the facts and the rule premise, often called the firing of the rule, is estimated by calculating the difference between the fact(s) and the rule premise. Let $\mathbf{R}$ and $\mathbf{F}$ be two matrices such that:

$$\mathbf{R} = \{R_{ij}\}, \text{ where } i = 1, 2, \dots, N \text{ and } j = 1, 2, \dots, M \quad (6.1a)$$

$$\mathbf{F} = \{F_{kj}\}, \text{ where } k = 1, 2, \dots, P \text{ and } j = 1, 2, \dots, M \quad (6.1b)$$

Figure 6.2 Rule Flow Illustration

In R every “ i ” represents a rule with “ j ” variables in its rule premise and in F every “ k ” represents the commuter’s under consideration with “ j ” variables (fact)

associated with the problem-specific information corresponding to each commuters. To find the nearest match a comparison search between the variables associated with the rules premise and the fact is performed, and the Root Sum Square Difference (RSSD) amongst these variables is observed such that:

$$D = \underset{i=1}{\overset{N}{\text{Min}}} \left\{ \sqrt{\sum_{j=1}^m (R_{ij} - F_{kj})^2} \right\} \quad (6.2)$$

The variables in the rule premise that give the minimum difference “ D ” is marked and the decision associated to that rule is taken into consideration.

This approach will always give the best possible match for the fact(s) which is incomplete or does not have any matching rule. But it may not be necessary that the decision generated based on this approach may be the real choice of the commuter. In that case it is necessary that the possibility distribution of the resulting decision may be estimated by considering all the rules which fall within a certain range “ Rg ” of the estimated “ D ”. Thus,

$$Rg = \text{rule selected} \begin{cases} 1, & \text{if the } \text{MinD} \leq \text{RSSD} \leq \text{MinD} \times Rg, \quad (0 \leq Rg \leq 1) \\ 0, & \text{if the } \text{RSSD} > \text{MinD} \times Rg \end{cases} \quad (6.3)$$

The range can be defined as the extra allowance (percent relaxation) which widens the choice of rule selection. It allows the search algorithm not only to choose the rule with the minimum value of “ D ”, but also takes into consideration the rules that fall within the specified range. In this situation, the provided fact may overlap with the premise of many rules. Hence, more than one rules may be fired (processed) simultaneously, each of them to a degree reflecting the similarity between the individual fact and the rule premise. The decision of each fired rule is then composed, and based on the occurrence of a similar decision an increment in its weight is assigned to improve its firing strength. The composition mechanism then

applies the deterministic choice rule and assumes that the alternative with the highest weight is the decision to be generated as the commuter’s mode choice.

A representation of the general rule flow model to search the mode decision in case of missing data is illustrated in Figure 6.3. The fact(s) F^* represents the commuter’s socio-economic and travel characteristics which are matched with the premises of each rule present in the rule-base. The heuristic algorithm is used to perform the search. The rules are processed simultaneously and a composition mechanism aggregates all the decisions corresponding to the matching rules. The rules having the same decisions are aggregated, and an increment for each similar decision is added, which enhances the weight of that decision. Finally, based on the comparison of the weights allocated to each decision, the decision with the highest weight is generated.

Figure 6.3 Mode Choice Decision Rule Flow for Missing Data

To evaluate the mode prediction accuracy of the developed algorithm, an experiment was conducted. The evaluation criterion was based on the level of prediction error, which provides the estimate of error between the actual and the predicted modes. The details are provided in Section 6.4.

6.3.2 Adaptive Learning

The adaptive learning feature involves the ability of the IES to learn i.e. add or modify the rules on its own through experience. An automated knowledge acquisition module is developed that allows the IES to provide decisions on its own

and update the rule-base. This feature helps in resolving the issue when the provided facts are not compatible with the set of available rules i.e. the facts are new and do not have any matching rule premise, or the generated mode decision is conflicting in its nature. The conflicting mode decision occurs in a situation where the number of matching rules favouring private mode decision becomes equal to the number of matching rules for the public mode decision.

The flow chart shown in Figure 6.4 presents the working mechanism of the automated knowledge acquisition module. The automated knowledge acquisition module activates when the inference engine is unable to find any suitable rule for the problem-specific information. Once the module is activated, then based on the user's input parameters it automatically acquires the decision for the existing situation. It has three different options to get the decision: prompt the user to provide the decision, apply the discrete choice model to estimate the decision, or run the heuristic algorithm to find the nearest match and generate the decision for the existing situation.

Figure 6.4 Automated Knowledge Acquisition Module

The automated knowledge acquisition module also consists of a consistency checking tool, which refines the knowledge base to eliminate any contradiction or

confusion that occurred during the decision-making process. Such confusion can happen when the commuter's choice probability for public mode becomes equal to the private mode. In that situation, any mode choice decision would be contradicting, thus the automated knowledge acquisition module would be activated and the users would be prompted for decisions.

6.4 RELIABILITY ANALYSIS

The developed mode choice models have their own functionality and associated limitations. In this study, a common goal in the application of these models is to estimate the mode choice for each commuter. The estimation of mode choice is based on the provided facts that are the socio-economic characteristics and the observed travel environment. Thus, to evaluate the reliability of all these models a common criterion i.e. estimated Prediction Error (PE) is selected. An experiment is conducted to estimate the PE associated with the models and based on the PE the model's prediction accuracy is analysed.

6.4.1 Evaluation Criterion

The PE provides details about the error in predicting the travel mode decisions of commuters. The PE gives the percentage of wrong predictions for commuters who choose private mode but were not assigned with one, or similarly for the public mode. An overall error is estimated by calculating the total number of wrong predictions with respect to the provided number of choice decisions. Considering the private mode to be Car and public mode to be Transit, the PE can be estimated as follows:

$$PE(\text{Car}) = \frac{\sum \text{Commuter Choosing Car} - \sum \text{Commuter Assigned Car}}{\sum \text{Commuter Choosing Car}}, \quad (6.4a)$$

$$PE(\text{Transit}) = \frac{\sum \text{Commuter Choosing Transit} - \sum \text{Commuter Assigned Transit}}{\sum \text{Commuter Choosing Transit}} \quad (6.4b)$$

$$PE(\text{Total}) = \frac{\sum \text{Wrong Predictions}}{\sum \text{Commuters Choice Decisions}} \quad (6.4c)$$

It is assumed that the socio-economic and travel characteristics can influence the mode decision of any commuter. Thus, the capability of a mode choice model depends on its sophistication to capture the influence of each variable (associated with the socio-economic and travel characteristics) on the mode decision. If the variables are independent and identically distributed then based on the distribution of associated error term the logit model can be used to estimate the mode choice. Therefore, all the three mode choice models have their own methodology of handling the error term and generating a decision.

6.4.2 Experimental Design

The objective of the experiment was to evaluate the PE related to each mode choice model, and find the model which gives the minimum PE. The experimental design was such that a new sample of 200 respondents was gathered, and presented with the questionnaire SP4 designed for the travel behaviour survey, as detailed in Section 4.3. The data collected from this survey provided the socio-economic and travel characteristics of the 200 respondents, along with their mode choice decision. This data set was considered as the facts or problem-specific information. The designed mode choice models were then simulated to predict the mode decisions for these 200 respondents. Based on each mode choice model's prediction capabilities the PE was estimated. The details of the developed mode choice models have been presented in Section 3.6.

6.4.3 Results and Discussion

In the PRB model there are two controlled variables which can influence the prediction capability of the model. One is the range “*Rg*” that allows a certain number of matching rules to be considered, and the other is the number of rules in the rule-base. The impact of range is two-fold. An increase in range means a larger number of matching rules can be incorporated, but at the same time the error would also increase during the composition of the decisions corresponding to those rules.

Thus, there should be a trade-off bringing the prediction error to a minimum level. The range can be assigned any value between 0 and 1 based on which the PE can be estimated. The range that gives the minimum PE can be taken into consideration.

In this experiment, two sets of rule-base were used to analyse the impact of different range values on the PE. A rule set consisting of 30 rules and another of 50 rules were stored in rule-base. The range values considered were 0.00, 0.10, 0.20, 0.22, 0.25, 0.30, and 0.40. The impact of these range values on the rule-base consisting of the 30 rules is presented in Figure 6.5.

Figure 6.5 Impact of Range on PE Considering a Rule-Set of 30 Rules

It can be observed from Figure 6.5 that the minimum value of overall PE is at the range value of 0.20. It means that rules that have matching facts within a 20% margin allowance provide a better prediction estimate for the PRB model. No matter whether the range is increased or decreased, in both the cases the overall PE value increases from the minimum value of the range of 0.20. It can also be observed that with the decreases in range the PE for transit decrease but the PE for car increases. If the range is 0.00 which implies that only the exact match of rule(s) is taken into consideration, then the PE for Car is 44% and the PE for transit is 26%. Whereas with the increase in range, the PE for transit increases and PE for car decreases, such that at a range value of 0.40 the PE for transit is 52% and the PE for

car is 21%. The optimum value of PE can be obtained at a range value of 0.20 where the PE for transit is 27% and the PE for car is 30%.

Figure 6.6 Impact of Range on PE Considering a Rule-Set of 50 Rules

Similarly, the impact of range values on the PE for a rule-set of 50 rules is presented in Figure 6.6. It can again be observed that the minimum value of overall PE at 18% is attained at a range value of 0.20. The increase or decrease in the range value respectively increases or decreases the overall PE. The impact on the PE for car and transit shows similar kind of behaviour i.e. increasing range value increases the PE for transit and decreases the PE for car and vice versa. Thus, it can be concluded that the flexibility provided by the range allows the model to incorporate and compose the decision of more than one rule. Such flexibility also restricts the composition of the remaining rules whose premise does not match with the provided facts and it reduces the induction of unnecessary error in predicting the mode choice. Hence, it is necessary that an optimum value of range that, in this case is 0.20, should be taken into consideration before initialising the IES.

While analysing the range value, it was observed that the increasing number of rules can also improve the level of accuracy of the PRB model. A variety of rules can be incorporated and their impact on PE can be evaluated. But it will not be feasible that a very large rule-base may be developed. Thus, to analyse the impact of the size of

rule-base, five different sets of rule-base were taken into consideration, such that the number of rules in each set was 10, 20, 30, 40, and 50, respectively. The range value was taken as 0.20 for all the sets. Figure 6.6 provides the details about the impact of size of a rule-base on the PE.

Figure 6.7 Impact of the Number of Rules on the Overall PE

It can be observed from Figure 6.7 that the size of the rule-base is inversely proportional to the overall PE. The increase in number of rules increases the domain of the matching rules, which reduces the prediction error. It can be observed that with only 10 rules the overall prediction error is 39% and with the increase in number of rules this error reduces to an extent and becomes 18% with the set having 50 rules. Thus, it can be concluded that the prediction accuracy of the PRB model can be further improved by increasing the size of the rule-base and incorporating all the 479 rules.

In case of the DCM, the choice variability is defined in terms of the distribution of the associated error term. In the developed DCM model the error term is assumed to be logistically distributed. Based on this distribution the choices are predicted and the PE is estimated. To check the reliability of the DCM the same database consisting of 200 commuters (facts) was analysed, and the DCM was applied to predict the mode choices. The estimated overall PE was observed to be 22%, the PE

for car was 13% and the PE for transit was 33%. It can be seen that the level of PE for DCM is higher as compared to the PE observed in PRB model. It is due to the reason that while estimating the DCM all the data points are taken into account, which increased the overall error in the model. The PRB model only considers the matching rules within the specified range, which reduces the error that could be incorporated due to the non-matching rules. This is an associated limitation of the DCM, and cannot be avoided during its estimation. The after-effects of this limitation can be observed in its application for a data set that is different from the data set on which the model was estimated.

To check the reliability of the COM model the same set of 200 commuters was taken into consideration. After allocating modes to the provided commuters (facts) in database the PE is estimated. The overall PE is found to be 42%, the PE for car is 0% and the PE for transit is 92%. Such kind of error is observed due to the prior probabilities of the socio-economic variables, which tend more towards the car as compared to public mode of transport. To estimate the mode choice the general multiplication law was applied and it was assumed that all the commuters' characteristics are independent. As most of the respondents did choose the mode car, all prior probabilities of socio-economic variables were tending more towards the car. Thus, the COM model did correctly predict the number of commuters travelling by car and the error was 0%, but while predicting the commuters' mode choice for public mode it predicted a large proportion of commuters who selected car. As a result of such prediction, a very high value of PE for the commuters travelling by transit is observed.

6.5 SUMMARY

The design and development of the INSIM Expert System, which can simulate the mode choice decision-making process of commuters in real-time environment, is presented. The IES captures the interaction between different available modes, and then based on the associated travel characteristics it decides on the commuters' mode. The source for the development of IES is the knowledge and expertise gained from travel behaviour surveys. Based on the commuters' socio-economic characteristics and provided multimodal traveller information the commuter mode

choice behaviour was modelled and represented by cognitive rules in the rule-base of the IES.

A general decision framework is also presented, in which the details about the decision-making rules and the rule of inference are presented. The incorporation of such rules allows the IES to become sensitive to changes in the travel environment and depicts a rational expert/human behaviour while making decisions. At the macro level, these rules possess the capability to emulate the expertise of a travel information provider and at the micro level they imitate the responses of commuters' mode choice decisions in the form of an intelligent agent.

Two important characteristics of the IES i.e. the reliability and the adaptive learning were also discussed. These allow the IES to generate decision on its own experience and update the knowledge base. In case of missing data or conflicting outputs the IES activates the heuristic algorithm that searches for the nearest match and generates the mode choice decision. The IES processes autonomously, and is capable of making mode choice decisions and dictating the required commands in the form of time-dependent OD matrix to the simulation environment (i.e. traffic network), based on its perception of the overall situation.

The mode choice decisions can be generated based on any one of the three different models i.e. PRB, DCM, or COM models. The development of these models is dependent on the commuters' mode choice decisions gathered from the travel behaviour survey. The PRB model constitutes crisp rules, the DCM is based on the estimated mode choice logit model, and the COM model adopts the Bayesian approach. An experiment was conducted to check the reliability of these models. It was found that highest level of accuracy can be achieved by applying the PRB model to generate mode choice decisions.

In this study, the implementation of rules is through the IES that is developed in a commercially available software expert system shell (Blaze Advisor). The model is based on an object-oriented design. The commuters, their associated travel characteristics, and the travel modes are modelled as classes and objects. The

knowledge about the travel preference of commuters, the transportation system operation rules and strategies, and the control heuristics, are implemented as rules.

CHAPTER 7

INTELLIGENT NETWORK SIMULATION MODEL

7.1 INTRODUCTION

The empirical findings of the travel behaviour survey and the travel behaviour models (presented in Chapter 4) provided evidence that if commuters are presented with integrated multimodal traveller information in congested travel environment, they would show a certain degree of mode switching. Based on the knowledge and experience gained from the travel behaviour survey regarding the significant variables that influence the commuters' mode choice decisions, a rule-based mode choice model (INSIM Expert System or IES) was developed, as presented in Chapter 6. The IES has the sophistication to imitate the commuters' mode choice behaviour in an information-rich environment.

Transportation network simulation modelling is one of the most suitable approaches to model and analyse different scenarios with respect to a particular travel environment. The IES developed in this study can interface with a network simulation model to enhance its simulation capabilities by providing a platform where the impacts of real-time integrated multimodal traveller information on the overall performance of the transportation network can be analysed. The simulation component is used to represent individual vehicular movements on the surface street network, and provides the details about the travel environment after every fixed time stamp. The IES allocates modes to the commuters in response to the supplied traveller information. The mode choice decisions are generated based on the commuter's goals, socio-economic characteristics and the perception and knowledge of prevailing travel environment.

This chapter shall cover the calibration of the IES and is organised as follows. The study network and network coding is presented in Section 7.2. A detailed discussion regarding the calibration of INSIM is provided in Section 7.3 followed by discussion in Section 7.4.

7.2 STUDY NETWORK

To assess the capabilities of the integrated IES with the transportation network simulation model, the transportation network for the central and western part of Singapore was selected for the simulation. The simulated network covers the area bounded by 5 major expressways and numerous arterial streets. The expressways are: Pan-Island Expressway (18 km), Ayer Rajah Expressway (19 km), Central Expressway (9 km), Bukit Timah Expressway (5 km), and Kranji Expressway (9 km). In the multimodal network established for this study, there are a total 21 residential and commercial zones, 2 Mass Rapid Transit (MRT) Lines, 23 MRT stations, 2 MRT interchanges, 38 selected signalised road junctions, 14 flyovers, and 1 roundabout. Every MRT station has feeder bus services, which traverse between nearby zones and MRT stations. Figure 7.1 shows the overall view of the simulated multimodal transportation network.

7.2.1 Data Collection and Network Coding

The physical data of the selected transportation network was collected through a survey of the study area, and also from several web-based reliable sources. All the selected expressways and the two arterial streets (i.e. Jalan Boon Lay and Pioneer North Road) were covered. The design criteria provided by the Land Transport Authority of Singapore (Civil Design Criteria for Roads and Rail Transit System, 2002) was also consulted. The road geometrics and details of the signalised junctions were observed and sketched during site survey. The details of public mode of transport were collected from Singapore Mass Rapid Transit (SMRT, 2005) website.

(a) *Geometric Design*

The imported map of the study area provides all the essential details regarding the road network layout; location of flyovers, roundabouts and road junctions; MRT network, MRT stations, and bus interchanges. It also gives the details about residential, recreational, and commercial zones. This map was loaded and displayed as the background, over which nodes, links, bus stops and MRT stations were placed.

Figure 7.1 Overall View of the Simulated Transportation Network

It was observed during the survey that the expressways are mostly 3-lane dual carriageways, except at certain locations where they are upgraded to 4-lane dual carriageways, depending upon the physical network conditions. The expressways have ramps and slip lanes that allow the vehicles to enter or exit, respectively. All the arterial roads are 2-lane dual carriageways. The lane width is 3.7 metres for expressways and 3.4 metres for other streets. The speed limits varied from 80 kph to 90 kph for expressways and 60 kph to 70 kph for arterial roads.

The road network in PARAMICS was coded according to the field-observed road geometrics, including the intersections, roundabouts, and flyovers. The type of links for expressways and arterial streets were specifically designed and provided in the “category” file, where details about number of lanes and speed limits were defined. The speed limit for expressways was standardised as 85 kph and for arterial streets as 55 kph, given that the PARAMICS models simulates the average travelling speed at 10 kph higher than the defined values.

(b) Surveillance Equipment

In the real network, traffic streams are observed via video cameras and loop detectors, and the statistics regarding traffic volumes and speeds are estimated and provided to certain users through web-based services. The Transportation Lab at Nanyang Technological University is one of the locations that receives such information. In this study, the data for the selected network were extracted from the archived database of such records.

To collect the simulated traffic statistics, 18 loop detectors were coded at points where surveillance cameras and loop detectors were present in the physical network. The loop detectors provided information about traffic counts and average travelling speeds on specific links. The gathered traffic data were also utilised in the calibration and validation process of the developed transportation network model.

(c) Signalised Junctions

There are 38 signalised junctions, which were selected for the simulation model. These junctions are either cross-junctions or T-junctions, with slip lanes for left

turning vehicles. The signal phasing patterns are fixed with varying signal cycle lengths. All the signalised junctions provide pedestrian crossing. The signal phasing pattern is such that it has 3 seconds amber time and 2 seconds all red time. The maximum cycle time is 120 seconds and maximum green time is 50 seconds. The details regarding the coded phasing pattern schemes are shown in Figure 7.2. The un-signalised junctions are set as follows: no control (57 junctions), yield sign control (27 junctions) and stop sign control (33 junctions).

Figure 7.2 Coded Phasing Pattern Schemes

(d) Public Mode of Transport

The public mode of transport consists of two modes; one is the bus service and the other is the MRT. It was found from the travel behaviour survey that more than 90% of commuters in the sample reported that they have access to MRT, and they are willing to travel by public mode of transport if it provides better service. One of the service criteria was the journey time, which is assumed to be lower in the case of MRT as compared to bus service. Thus, the bus service is modelled as a feeder service to MRT, and MRT is taken as the major mode of public transport for the work/school trips.

The bus routes that serve as feeder service to MRT and the bus stops associated to these bus routes are coded and modelled by location. Similarly, the MRT lines and stations were also coded. The MRT lines were modelled as single-lane dual carriageways, with right of way dedicated to MRT services only. Each MRT station was designated a zone, and explicitly simulated in such a way that the residential and commercial zones were connected to the MRT zones (stations) by feeder bus services. The simulated buses were 12 metres long with a seating capacity of 47 passengers and a top speed of 20 kph. The simulated trains have 6 cars with a total length of 138 metres and a top speed of 45 kph. Each train has a seating capacity of 372 passengers.

(e) Traffic Information and Route Guidance

In this study, it is assumed that travellers commute in an information-rich environment, where pre-trip integrated multimodal traveller information and en-route traffic information is available. All the Driver Vehicle Units (DVU) are coded as controlled by familiar drivers and the dynamic feedback assignment is adopted. The PARAMICS model disseminates traffic information to all the familiar drivers and information is updated every minute.

(f) OD matrix

To calibrate the developed transportation network, an OD matrix was estimated based on the traffic counts. A traffic survey was conducted and the statistics such as traffic counts, average travelling speed, and average travel time were gathered. The details about the traffic survey and the estimated OD matrix will be discussed in Section 7.3.4.

7.3 CALIBRATION OF THE INSIM NETWORK

The calibration of the simulated network was done from two different perspectives. The first one was to adjust the traffic counts and the OD matrix in the simulation as compared with the existing situations, such that the simulated network within the complete simulation framework can be close to the real traffic conditions. The second one was the calibration of behavioural parameters of the PARAMICS

simulation model, such that they imitate the local commuters' travel behaviour. The calibration of the behavioural parameters was necessary because the IES was developed based on the commuters' travel behaviour observed/estimated from the travel behaviour survey conducted in Singapore. To integrate IES with PARAMICS to simulate the local traffic environment in accordance with IES, both models should have the same behavioural assumptions, so that the behavioural parameters that govern the PARAMICS simulation model can realistically represent the behaviour of local commuters.

7.3.1 Study Site

To adjust the traffic counts and the OD matrix, the network chosen for the simulation model is shown in Figure 7.3. To calibrate the behavioural parameters, a portion of the selected study network (marked with blue broken line) in Figure 7.3, was considered. This portion of the network was selected because it covers an area bounded by two expressways (i.e. Pan-Island Expressway and Ayer Rajah Expressway), four arterials (i.e. Jalan Boon Lay, Boon Lay Way, International Road, Pioneer Road North), and one collector (i.e. Jurong West Street 75).

Secondly, there is one major bus interchange integrated with the MRT station. Such an arrangement facilitates possible transfer between feeder buses and the transit mode. Thus, this portion of the network provides the opportunity to collect the travel data for both the public and private modes, for similar trips within the same corridor. The details and overview of the simulated network are presented in Figures 7.3 and 7.4.

7.3.2 Basic Input Data

The basic input data include network geometry, driver behaviour, vehicle characteristics, transportation zones, travel demands, traffic control systems, traffic detection systems, bus routes, and bus stops. The road geometry and infrastructure locations were obtained from field surveys. The transportation zones were located at the entrance and exit nodes of all the major roads.

Figure 7.3 Detailed Map and the Overview of the Simulated Network

Figure 7.4 Detailed Simulated Network showing Zones, Links and Nodes

Since PARAMICS regards each vehicle in the simulation as a Driver Vehicle Unit (DVU), therefore the driver behaviour data, vehicle mix by type, vehicle characteristics data are considered as basic inputs of the established PARAMICS simulation model. The vehicle mix by type was determined by the statistical analysis of traffic flow data collected during the survey at two locations (i.e. Nanyang Flyover and Pioneer Road North Junction). The driver behaviour data, represented by aggressiveness and awareness factors in PARAMICS, were assumed to be the default normal. The details of public transportation system were obtained from SMRT website (SMRT, 2005). The public transport system in the simulated network consists of different bus routes that connect all the zones with bus interchange. There are 36 bus routes, 127 bus stops, and 1 bus interchange. The frequencies for different bus routes range from 5 to 20 minutes.

The simulated network has the same locations of loop detector stations and traffic control operations as that the traffic control systems and traffic detection systems on the physical network. A total of 18 loop detectors were coded within this network which provided information about traffic counts and average travelling speed on specific links. Altogether, 38 signalised junctions were selected for simulation purpose. These junctions were operated on a fixed-time phasing pattern with an amber time of 3 seconds and an all red time of 2 seconds. The details of signal cycle timing and phasing patterns can be found in Section 7.2.1.

7.3.3 Required Data for Calibration

The calibration involves the checking of model results against observed data and adjusting the values of the parameters until the model results fall within an acceptable threshold of error. The collected data for model calibration included traffic volume, average travelling speed, and travel time data. The traffic volume and average travelling speed data on expressways and arterial streets were obtained from an archived database in NTU. The travel time data were obtained by floating car method for Pioneer North Road and Boon Lay Way. To construct the real-world traffic variations in the simulation, a typical traffic variation can be represented by the traffic conditions of a typical day, of which the traffic data are the target of calibration process. The selection of a typical day can be implemented based on the

comparison of traffic volume (at any selected station) of a candidate day with the average traffic volume of all candidate days using the GEH statistic, used by British engineers (UK Design Manual for Roads and Bridges, 1996):

$$GEH = \sqrt{\frac{(E - V)^2}{(E + V)/2}} \quad (7.1)$$

where “ E ” is the candidate data and “ V ” is the average data. If the GEH values for more than 85% of the selected stations are less than 5, the traffic condition and the demand pattern of the candidate day are typical. There were 35 stations for which data were available. These stations are located on the two expressways i.e. PIE and AYE, and various arterial streets within the network. The data were collected for 1 hour during the morning peak period from 8 to 9 AM. It was found that traffic condition on Wednesday (i.e. 1st October 2004) was typical with respect to all weekdays. Therefore, the traffic volume and average travelling speed data of 1st October 2004 were chosen for calibration purpose. It was observed from the extracted data that the aggregated average traffic volume on PIE was 3600 vehicles per hour with an average travelling speed of 94 kph, whereas on AYE the aggregated traffic volume was 5300 vehicles per hour with an average travelling speed of 87 kph.

In this study, the travel time data were collected by floating car method. On the typical day (Wednesday) during the peak period (8 to 9 AM) several runs were made between Nanyang Technological University (NTU) and Boon Lay Interchange (BLI). The length of the selected route was 4.2 km. The number of floating car runs was obtained from the following equation:

$$n = \left(\frac{ts}{\mu\epsilon} \times 100 \right)^2 \quad (7.2)$$

where “ n ” is the required number of runs, “ μ ” is the mean travel time in the runs, and “ s ” is the standard deviation, “ ϵ ” is the desired margin of error (percent of “ μ ”), and “ t ” (Student’s t-statistic) is the confidence coefficient.

The standard deviation was unknown prior to runs being conducted. A total of 12 runs were performed. It was observed from the floating car runs that the average travel time was 314 seconds and the standard deviation was 22.3 seconds. The estimated number of runs with margin of error at 5% was $9.75 \approx 10$. The 95% confidence interval at 12 runs was estimated to be ± 14.2 seconds, and it was inferred that the true mean lies between 300 and 328 seconds. The corresponding average travelling speed on this route was estimated to be between 46.1 and 50.4 kph. The average travelling speed on the typical day for this route was observed to be 50.4 kph with standard deviation of 14 kph. These results provided evidence that the estimated travel time statistics by floating car method give the details regarding the basic travelling conditions.

Similarly, the travel time data for the public mode of transport were collected. Two bus routes were considered; the first one (SBS 199) served between NTU and Boonlay Interchange (BLI), and the other (SBS 242) served between Jurong West area and BLI. The route length of SBS 199 is 4.2 km, and it is associated with 9 bus stops out of which 3 have bus bays. The route length of SBS 242 is 3 km with 13 bus stops, out of which 3 have bus bays. The frequency of SBS 199 is 4 minutes during peak hours and 7 minutes during off-peak hours and for SBS 242 it is 3 minutes during peak hours and 5 minutes during off-peak hours on week days, respectively. On each route, there are 4 signalised junctions and 2 mid-block pedestrian signals.

Initially, 30 runs were conducted on each bus route, and the travel time data were collected. For SBS 199 route, it was found that the average travelling time was 752 seconds with a standard deviation of 51.4 seconds. With an error margin of 5% the estimated number of runs was $7.8 \approx 8$. The resulting 95% confidence interval was ± 19.5 seconds for 30 runs, which means that the true mean travel time can be between 732 and 772 seconds. For SBS 242 route, the average travelling time was 556 seconds and the standard deviation was 46.2 seconds. With an error margin of 5% the estimated number of runs was $11.6 \approx 12$. The resulting 95% confidence interval was estimated to be 17.5 seconds, which means that the true mean travel time can be in-between 538 and 574 seconds. The average travelling speeds based

on the observed travel time values on routes SBS 199 and SBS 242 were 23.9 kph and 32.4 kph, respectively.

7.3.4 OD Matrix

The PARAMICS simulation model needs an origin and destination (OD) matrix as a starting point of the calibration process. In this study the time-dependent OD matrix was generated by IES. The OD estimation process was dependent on commuters release profile and zone-based trip distribution. It is necessary that the generated OD matrix should match with some reference OD matrix, so that the PARAMICS model can simulate realistic traffic conditions. The reference OD matrix was estimated from the traffic count data obtained for the typical day.

The cordon points of the network were taken as zones, and 15-minute interval traffic counts at all cordon points were gathered from the data. The total traffic attractions and generations of each zone were then assumed to be the distribution of traffic with respect to the OD matrix. The available loop detector data were then compared for checking the network-wide equilibrium of traffic volumes, such that the traffic entering the network should be equal to the traffic leaving the network.

To estimate the OD matrix, the observed traffic counts at the cordon points were taken into consideration. These traffic counts were assumed to be the total attraction and generation from the respective zones corresponding to the cordon points. Once the traffic volume being generated and attracted to each zone was estimated, it was then compared with a reference OD matrix. This reference OD matrix was obtained from the study done by Ang (2003), which provided the details regarding the traffic volumes being generated during the morning peak period.

The data gathered from the traffic survey provided details about the traffic volumes for 7 zones. The traffic volume data for these zones were compared with those provided in the reference OD matrix, and the growth factor was estimated. Based on the estimated growth factor, which was found to be 5.69% per annum, the total productions and attractions for each zone of the reference OD matrix were revised. The adopted expansion procedure was such that:

- (a) the total production/attraction for each zone of the reference OD was expanded, excluding the total productions /attractions for the zones that were obtained from the traffic survey,
- (b) all the expanded total productions/attractions were summed up, and compared with the expanded total traffic volume, to estimate the overall difference,
- (c) the estimated overall difference was distributed among all the zone-based total productions/attractions depending upon their share of traffic volume aggregated in the total traffic volume,
- (d) each zone-based total productions/attractions were then adjusted according to the allocated share of the estimated difference,
- (e) the adjusted zone-based total productions/attractions were again summed up, such that their sum was now equal to the expanded total traffic volume, and
- (f) the finalised zone-based total productions/attractions were taken as the target values to be used in the final version of the estimated OD matrix.

Each cell value of the reference OD matrix was also expanded, corresponding to the expanded total traffic volume. The zone-based total productions/attractions for these cell values were estimated and taken as the base case values. The Furness technique (Ortuzar, 2002) was then used for balancing the base case values with the target values of the zone-based total productions/attractions. In this study only the expanded values of the zone-based total productions/attractions were changed, and the total productions/attractions estimated from the traffic survey were unchanged. Based on the estimated OD matrix, the percentages of trips being productions and attractions for each zone were calculated. The zone-based distribution of trips was estimated and is presented in Table 7.1. The zone based trip distribution was utilised by IES to generate the time-dependent OD matrix.

Table 7.1 Zone-Based Trip Distribution (%)

O/D	1	2	3	4	5	6	7	8	9	10	11	12	21	22	23	24	25	26
1		0.1	1	0.2	0.3		0.2		0.1	0.4	0	0.5	0.2	0.2	0.6		0.1	0.7
2	0		0.3	0.1	0.2		0		0	0.1		0.1	0.1	0.1	0.1		0.1	
3	0.6	0.3		0.8	2.1	0.2	0.3	0.8	0.6	0.7	0.2	0.5	0.2	0.5	0.8		0.3	0.1
4	0.1	0.1	0.7		0.3		0	0.2	0.1	0.3	0.1	0.1	0.1	0.1	0.2		0.1	0.2
5	0.2	0.3	2	0.3		0.1	0.1	0.7	0.4	1.1	0.1	0.4	0.8	0.6	1.4	0	0.1	0.2
6			0.2		0.1		0		0	0.2	0	0.2	0.2	0	0.2		0.1	0.1
7	0.1		0.4	0	0.1	0		0.3	0.1	0.2	0	0.4	0.1	0.1	0.2		0.1	0.3
8	0.1		1.5	0.3	0.4		0.5		0.1	1.4	0.1	2.1	0.8	0.3	1.4		0.2	
9	0.1	0	0.8	0.1	0.5	0	0	0.1		0.8	0.1	0.6	0.4	0.2	0.9		0.2	0.6
10	0.2	0.2	0.7	0.3	1	0.2	0.2	0.6	0.8		0.2	0.4	0.7	0.3	0.5		0.1	0.1
11	0		0.1	0.1	0	0	0	0.3	0.1	0.2		0.1	0.1	0	0.1		0	0.2
12	0.4	0.2	0.7	0.3	0.9	0.2	0.5	3	0.9	0.6	0.2		0.2	0.3	0.6		0.6	0.8
21	0.1	0	0.1	0.1	0.8	0	0.1	0.7	0.3	0.6	0.1	0.1		0.1	0.5		0	0.1
22	0.3	0.2	1.9	0.3	2.3	0.1	0.1	0.7	0.7	1	0.1	0.8	0.5		3.6	0.1	0.8	0.3
23	0.3	0.2	0.9	0.2	1.5	0.2	0.2	1.2	0.7	0.5	0.2	0.2	0.6	1.2		0.1	0.9	0.2
24					0									0.1	0.1			
25	0.1	0.1	0.4	0	0.2	0.1	0.1		0.2	0.3	0	0.5	0.1	0.3	1.2			0.2
26	1		0.2	0.3	0.5	0.6	0.9		0.7		0.4	1.2		0.2	0.4		0.6	

**Empty cells denote ZERO value*

7.3.5 Preliminary Testing

In INSIM, three types of demand-related data are necessary: zone-based percentile distribution of trips, percentage of total traffic flow to be released onto the network in each time interval, and the total number of trips to be generated. Field data were processed to gather the demand-related data along with the required performance measures. The transportation network model was simulated based on the demand-related data, and was calibrated by comparing the simulated output results with the gathered performance measures. It was observed from the field data that on a typical day the peak period was from 8 to 9 AM, and 33,120 private mode trips were generated within the physical geography of the simulated network during this period. The peak period was defined by observing the traffic volumes on a typical day from 7 to 10 AM. It was found that the traffic volumes were highest during the time period of 8 to 9 AM. Available information (Singstat, 2005) revealed that

41.6% commuters took private mode for their work trip, 52.4% took public mode, and 6.1% did not require any motorised transport. Based on these statistics, a total of 79,877 trips were simulated for a 1-hour time period. The estimated zone-based percentile trip distribution (Table 7.1), and a uniform trip release profile was also defined. Default values for vehicle population composition, vehicle characteristics, and driver aggressiveness and awareness distributions were not changed. The dynamic feedback assignment model provided by PARAMICS was adopted to replicate route choice behaviour.

Initially, 10 simulation runs were conducted in order to identify critical variables that may significantly influence the performance of PARAMICS. In this procedure, the variables related to driver behaviour (e.g. the distributions of aggressiveness and awareness), compositions of vehicles, driver's familiarity etc. were examined. The number of simulation runs was based on number of vehicles (NV). The average NV value was 32,580 and the standard deviation was 335. The estimated number of runs with a margin of error of 1% was $5.12 \approx 6$. The required simulation runs were conducted with different seed numbers. The numerical outputs of the median simulation run were then compared to corresponding real data.

Table 7.2 Output Data Comparison with the Un-calibrated Model

	Simulated	Actual	Error (%)
Total Demand (Production)	28782	33120	-13.09
Total Outflow (Attraction)	25288	33120	-23.64

Outputs used for comparison were related to total demand generated and total flow reaching the destination. The outputs were aggregated and analysed using Microsoft Excel. Outputs aggregated in this step for further comparison included total generations (total vehicles released in 1 hour), and total attractions (total number of vehicles completing their journey in 1 hour). Results in Table 7.2 indicate a shortage in the number of vehicles being released into the network, and the number of vehicles reaching their destination. The difference in productions and attractions is due to vehicles which were released in the network but have not reached their

destinations. Thus, the difference between the productions and attractions is equal to the total number of vehicles which are still in the network. Therefore, parameter values other than the default needed to be explored. In order to improve the accuracy of the simulation at this stage, the types of vehicles using the network and their respective percentages in the vehicle population were adjusted. There was no systematic way for adjusting the percentage of different types of vehicles except to experiment different combinations so as to get acceptable results. After several combinations the final percentages of different types of vehicles used in the simulation are presented in Table 7.3.

Table 7.3 Calibrated Percentages of Different Types of Vehicles

Type		Percentage
Serial Number	Shape	
1	Car	50
2	Car	15
3	LGV	9
4	OGV 1	8
5	OGV 2	8
6	Mini Bus	7
7	Coach	3
Total		100

It was observed during the simulation that the DVU traversing on expressways did not use the outermost lanes, specifically at the merging points at on-ramps along expressway lanes. At such locations, the vehicles slowed down for longer time periods, and such behaviour resulted in heavy congestion. To analyse such issue, three parameters concerned with driver behaviour were investigated. Two parameter sets in particular, aggressiveness and awareness, were examined – since these two significantly influence the behaviour of drivers. There are four default types of distributions associated with each of these two parameters. Numerous runs with different combinations of distributions were conducted until the outputs of lane use were improved further. The third parameter was the ramp awareness distance. It was observed from the PARAMICS graphical interface of the simulated network

that increasing the length of ramp awareness distance increased the flow on expressways and improved the outermost lane utilisation. The ramp awareness distance was adjusted one by one for all the ramps, and the minimum value was adjusted to 150 metres. Using combination of the above adjustments, 10 simulation runs were conducted and the improved results are summarised in Table 7.4.

Table 7.4 Output Data Comparison Based on the Calibrated Model

	Simulated	Actual	Error (%)
Total Demand (Generation)	32428	33120	-2.09
Total Outflow (Attraction)	31914	33120	-3.64

7.3.6 Trip Release Profile

Theoretically, it is a dynamic OD demand estimation problem. To date, there is yet an effective method to be defined that can solve this problem for a corridor network (Cascetta et al., 1993; Ashok and Ben-Akiva, 1993). There are some OD estimation tools e.g. Estimator of PARAMICS, but their capabilities and potentials are still unclear. In this study, the time-dependent OD matrix was estimated on the basis of zone-based percentile trip distribution. The proposed dynamic OD estimation process could be regarded as a process that assigns the total OD to a series of consecutive time stamps. The time-dependent OD demand for each time stamp was then extracted by reconstructing the dynamic OD matrix based on a set of demand profiles for every time stamp e.g. 1 minute or 5 minutes. INSIM allows the user to define the time-dependent releases of demand (trips) and the zone-based trip distribution.

As discussed in Section 7.3.3, the field data provided 15-minute interval counts for all the cordon points in the network. Based on these data, the profile of vehicle production from any origin zone and that of vehicle attraction to any destination zone were thus estimated. Four different trip release profiles for different time stamps were developed, (shown in Figure 7.5), such that:

- (a) A uniform distribution was assumed for the demand release profile i.e. RP1. The total demand was uniformly spread over 60 minutes of the simulation period with a time stamp value of 5 minutes. For the first four time stamps 9% of the total volume was released, and in the remaining 8 time stamps 8% were released.

Figure 7.5 Different Trips Release Profiles

- (b) A non uniform distribution was assumed for the demand profile RP2. The 15-minute counts were segregated into 5-minute counts, by uniformly distributing the 15-minute counts over 3 periods of 5-minutes duration, such that 7% of demand was released in each 5-minute time stamp for the first 15-minute time interval, 9% of demand was released in each 5-minute time stamp for second 15-minute time interval, and so on.
- (c) A normal distribution was assumed for the demand profile RP3, such that the maximum release of trips occurred during the 2nd and 3rd quarter of the simulation period. In this profile the demand increased uniformly and arrived at the peak discharge of 12%, after which it decreased uniformly.

- (d) A normal distribution skewed towards the right was assumed for the demand profile RP4. In this profile the peak discharge of 12% occurred in the 3rd quarter of the simulation period. This profile imitated the real traffic flow condition.

7.3.7 Calibration of Route Choice Model

In the OD estimation process, the network, traffic and route choice behaviour parameters were to be fixed initially, as this process is based on the traffic assignment matrix that is affected by any change in the simulation input parameters.

Due to the existence of expressways and parallel streets in the study network, the routing algorithm adopted in the PARAMICS simulation was important. The network was calibrated using the dynamic feedback assignment model provided by PARAMICS. Dynamic feedback assignment in PARAMICS assumes that different drivers perceive different costs from a decision node to the destination. The perceived cost is calculated, and the perceived shortest route is chosen at the decision node. At this stage the parameter to be calibrated for the route choice model is the number of drivers who are familiar with the road network. Since there was no data to calibrate it, it was assumed that most drivers in the morning peak period were familiar drivers who had the knowledge of road network and traffic conditions. So it was assumed that 95% of drivers were familiar drivers, who could choose their route from the available options.

7.3.8 Secondary Testing

Multiple simulation runs for each parameter combined with the each trip release profile were also performed. During these simulations, certain changes in the cell values (i.e. specific origin to destination traffic volumes) of the estimated OD matrix were incorporated. After multiple iterations, the calibration criteria were satisfied. The highest GEH value was obtained by adopting RP4 trip release profiles with 95% of familiar drivers. The calibration results of this step are shown in Table 7.5, which presents the 15-minute intervals traffic counts at the selected measurement stations. It shows that for more than 87.5% of all measurement

locations, their GEH values are lower than 5, which satisfies the calibration acceptance criteria of this step.

7.3.9 Calibration of Driving Behaviour Model

The last step was to calibrate the driver behaviour model to reflect local driver characteristics. The local driver characteristics can be examined through the comparison of the simulated and observed volume-occupancy curves drawn based on aggregated point detector data, and the point-to-point travel time measurements. Thus, the network was further calibrated under the following conditions: the network has been coded and partially calibrated for the parameters mentioned in Sections 7.3.7 and 7.3.8; there is no data that can support the calibration of route choice model; the dynamic assignment model in the microscopic simulator can be accepted; and the zone based percentile trip distribution and trip release profile have been obtained.

The objective function is to minimise the deviations among the observed and the simulated volume-occupancy curves, and the point-to-point travel time measurements, subject to the condition that the traffic counts, which have already been calibrated, do not change. But if the traffic counts do change then they must satisfy the GEH calibration acceptance criteria. In this study, the point-to-point travel time match was performed only for one route between NTU and Boon Lay Interchange (BLI), as described in Section 7.3.3. The driver behaviour model can be calibrated by adjusting the car-following and lane-changing models. The global parameters for the car-following and lane changing models are mean headway and driver reaction time, which can drastically influence the simulated driver behaviour. These two parameters were fine-tuned by trial-and-error. The purpose was to reconstruct traffic variations and match the congestion pattern of the study network, So that the resulting simulated travel time measurement matched with the real life estimated travel time values. Five different combinations of mean headway and driver reaction time were analysed. For each combination 5 simulation runs were performed and the median NV value was taken into consideration. It was observed that higher values of mean headway and reaction time resulted in lower travelling speeds and vice versa.

Table 7.5 Traffic Count Calibration Results for Release Profile RP4

Detector ID	Traffic Count			Traffic Count			Traffic Count			Traffic Count		
	(0800 - 0815)			(0815 - 0830)			(0830 - 0845)			(0845 - 0900)		
	Actual	Sim	GEH	Actual	Sim	GEH	Actual	Sim	GEH	Actual	Sim	GEH
80002851	793	778	0.54	947	875	2.39	961	1037	2.4	853	870	0.58
80006958	161	158	0.24	192	193	0.07	195	231	2.47	173	200	1.98
80006966	68	57	1.39	81	71	1.15	82	68	1.62	73	82	1.02
80006967	122	96	2.49	146	115	2.71	148	118	2.6	131	137	0.52
80006975	575	564	0.46	687	674	0.5	697	701	0.15	619	581	1.55
80007818	219	195	1.67	262	268	0.37	265	250	0.93	236	176	4.18
80007819	194	149	3.44	232	183	3.4	235	184	3.52	209	252	2.83
80007822	192	144	3.7	229	276	2.96	233	203	2.03	207	210	0.21
80007823	140	145	0.42	167	117	4.2	170	124	3.79	151	181	2.33
80007842	234	244	0.65	279	330	2.92	284	292	0.47	252	266	0.87
80007843	192	156	2.73	229	252	1.48	233	323	5.4	207	283	4.86
80010570	73	64	1.09	87	93	0.63	88	101	1.34	79	82	0.33
80010571	219	173	3.29	262	287	1.51	265	306	2.43	236	187	3.37
80010574	170	97	6.32	203	214	0.76	206	255	3.23	183	183	0
80010578	176	131	3.63	210	162	3.52	213	236	1.54	189	193	0.29
80010579	315	252	3.74	376	359	0.89	382	395	0.66	339	323	0.88
80010586	648	592	2.25	774	797	0.82	785	849	2.24	697	745	1.79
80010587	352	321	1.69	420	475	2.6	427	550	5.57	379	386	0.36

The final calibrated mean target headway and driver reaction time were 0.85 and 0.75, respectively. The calibrated demand profile for the simulated peak period is shown in Figure 7.6.

Figure 7.6 Calibrated Demand Profile

Figure 7.7 Observed and Simulated Point to Point Travel Time

The Root Mean Square Percent Error (RMSP) was estimated to examine the error between the simulated traffic demand and the actual observed traffic demand, and it was found to be 8.29%. Figure 7.7 shows the comparison of observed and simulated point-to-point travel time for the selected route, which has an RMSP value of 2.49%. Based on the lower value of RMSP, it can be stated that calibration process has been effective as indicated by the test statistics.

7.3.10 Calibration of Public Mode of Transport

The last stage of calibration process was to calibrate the public mode of transport. In this study, two public modes of transport i.e. train and bus services were simulated. As the bus shares the right of way on the surface street network, its journey time can be affected by the congestion occurring in the network. The bus journey time is dependent on two parameters i.e. the average travelling speed of buses and the number of passengers waiting at bus stops.

The number of waiting passengers can influence the travel time by affecting the bus dwell time at the bus stops, which is an integral part of travel time. Increasing the number of passengers can increase the boarding passengers' queue length resulting in longer bus dwell time at the bus stops and vice versa. Thus, the bus dwell time can be a function of number of passengers arriving at the bus stop. The public mode of transport was calibrated by adjusting the bus travel time. The selected parameters were the average travelling speed by bus and the arrival rate of passengers at the bus stop. The objective function was to minimise the difference between the simulated and observed bus travel times. The field data were available for two bus routes i.e. SBS 242 and SBS 199, as described in Section 7.3.3. Along with other bus services these two routes were simulated.

Initially the bus travel time was calibrated by adjusting the travelling speed. Three different speeds i.e. 15 kph, 20 kph, and 25 kph were simulated. Each speed was simulated for 5 times with different seed number. Then the bus travel time was fine-tuned by adjusting the arrival rate of passengers at each bus stop. After multiple simulation runs, the bus journey time was found to be acceptable. The average travel time measured for buses from simulation were compared with the average

travel time for buses measured in the field. The comparison of observed and simulated point-to-point travel times on bus route SBS 242 and SBS 199 are presented in Figures 7.8 and 7.9, respectively. The RMSP was estimated to be 9.54% for SBS 242, and 9.01% for SBS 199 route.

Figure 7.8 Observed and Simulated Travel Time by Bus Route SBS 242

Figure 7.9 Observed and Simulated Travel Time by Bus Route SBS 199

7.4 DISCUSSION

The working mechanism of INSIM and details about the generation of commuters and the transportation network simulation model are discussed in this chapter. The INSIM commuter generates commuters based on three user-defined parameters i.e. total number of commuters to be generated, trip release profile, and zone-based percentile trip distribution. The transportation network simulation model i.e. PARAMICS, simulates the private as well as the public modes of transport and generates the desired measures of effectiveness. An API has been developed to integrate and perform the required data transfer functionalities between the IES and transportation network simulation model. To analyse the feasibility of the integrated IES with the transportation network simulation model, a real life transportation network was selected, which reflects the transportation network for the central and western part of Singapore. The network simulation model was calibrated with the observed field data.

A major concern in model calibration/validation is error inherent in the collection of input data. In spite of the methodology adopted to calibrate the simulation model, some of the calibration errors might have been derived from problems in observed data, such as data discrepancy, poor quality, or missing data. Problems with input data can propagate to erroneous corrections to models that will damage model performance, credibility and results.

The calibration of microscopic simulation models depends on the quality of the observed data. One important reason for the calibration error was that the data used for calibration was not of good quality. It was one of the reasons that several measurement locations had GEH values larger than 5. Another reason was the missing data, which had to be estimated based on the traffic patterns of other days. The assumption that the traffic patterns under all data collection days were the same seemed to underestimate the variations of the traffic patterns, and some potential inconsistencies among the traffic flows were thus inevitable.

The calibration of microscopic simulation models also depends on the quantity of the observed data. The observed data need to cover every part of the network.

However, due to unavailability of data at certain locations, some parts of the network are still un-calibrated, which would be another source of the calibration errors. The completeness and quality of the observed data are especially important for the calibration of a simulation model. It is also important to recognise that uncertainty is inevitable in microscopic simulation model calibration and validation due to variability of traffic conditions as well as the quality of data.

Summarising the efforts made to calibrate the transportation network model, it is concluded that the improved model and the generated outputs are acceptable. The network coding errors were rectified and a reliable time-dependent OD matrix was estimated. The INSIM was adjusted, such that the commuter release profile and the percentile trip distribution resulted in simulated traffic counts that were acceptable according to the GEH statistics. The PARAMICS route choice model was adopted and 95% of the driver population was assumed to be familiar with the road network. The driver behaviour model was calibrated, and it reflects the local driver behaviour. The mean headway and driver reaction time were adjusted to achieve the calibration of driver behaviour model. Finally, the public mode of transport was calibrated by adjusting the average travelling speed of buses and the arrival rate of passengers at the bus stops.

The calibrated network simulated model reflects the local travel environment. It can simulate the travel time and travelling speed parameters, within the statistically acceptable limits, as compared to the observed data. These parameters shall be utilised by the IES to generate commuters' travel mode choice under the influence of provided integrated traveller information.