Takeshi Takama

Stochastic agent-based modelling for reality:

Dynamic discrete choice analysis with interaction

A Thesis for
The Degree of Doctor of Philosophy

Oxford Centre for Environment

University of Oxford

To

my wife Rumi,

my family in hometown,

and

in the memory of Jyon

Abstract

Title:	Stochastic agent-based modelling for reality: Dynamic discrete choice analysis with interaction
Name:	Takeshi Takama
Society:	St. Catherine's College
Degree:	Doctor of Philosophy
Submission:	Trinity term 2005

This D.Phil. thesis develops a new agent-based simulation model to improve the results of analysis, which solely uses discrete choice modelling, as well as to analyse the effects of a road user charging scheme for the Upper Derwent Valley in the Peak District National Park. The advantages of discrete choice analysis are well known. However, results with these conventional conventional approaches, which conduct analysis solely with discrete choice models, can be biased if interaction and learning effects are significant. The Minority Game, in which agents try to choose the option of the minority side, is an appropriate tool to deal with these problems. The situation in the Upper Derwent Valley can be explained with economic game theories and the Minority Game. The two approaches mutually help to analyse the situation in the Upper Derwent Valley leading to the development of a stochastic Minority Game. The stochastic Minority Game was tested with an online game (questionnaire), which was played 3,886 times by response in all around the world.

The practical part of this thesis examines the components of the stochastic Minority Game with the data collected around the Upper Derwent Valley. The main data was collected using a stated preference survey. Overall, 700 questionnaires were distributed and 323 of them were returned (i.e. a return rate of 46.1 %). In the practical part, the agent-based model has four sub modules: 1) Multinomial mixed logit model for mode choice, 2) Binary logit model for parking location choice, 3) Markov queue model for parking network, and 4) the Minority Game for parking congestion and learning.

This simulation model produces comprehensive outputs including mode choices, congestion levels, and user utilities. The results show that the road user charging scheme reduces car demand in the Upper Derwent Valley and ensures a reduction in congestion at the parking areas. The model also shows that an exemption will increase the utilities of elderly visitors without substantially sacrificing those of younger visitors.

In conclusion, the simulation model demonstrated that oversimplification in conventional approaches solely using discrete choice models gave significant biases when real world problems were analysed.

Extended abstract

Title:	Stochastic agent-based modelling for reality: Dynamic discrete choice analysis with interaction
Name:	Takeshi Takama
Society:	St. Catherine's College
Degree:	Doctor of Philosophy
Submission:	Trinity term 2005

This D.Phil. thesis develops a new agent-based simulation model to improve the results of analysis, which solely uses discrete choice modelling, as well as to analyse the effects of a road user charging scheme for the Upper Derwent Valley in the Peak District National Park. This thesis starts with confirming the statement of problem below:

> ***"Conventional approaches which conduct analysis solely with discrete choice models have the advantage of simplicity, but severe biases exist due to the neglect of some interaction and learning effects, which might be seen as oversimplification"***

The proposed solution to the problem in this thesis is:

> ***"Innovative interaction and learning are added to the conventional approaches by agent-based modelling together with discrete choice analysis"***

To achieve this target, the road user charging and park & ride schemes at the Upper Derwent Valley in the Peak District National Park are used as a testing ground. Therefore, this thesis is the combination of theoretical and practical work. The advantages of conventional approaches using discrete choice analysis and stochastic processes are confirmed in the introduction. The main advantage of these methods is the efficiency in terms of computational power,

simulation time, data collection, and so on. However, results with these conventional approaches, which conduct analysis solely with discrete choice models, are biased if interaction and learning effects are significant. The ignorance of interaction effects are strongly related with the problems of imperfect information, dynamic congestion, and transforming probability to the proportion of population. The Minority Game, in which agents try to choose the option of the minority side, is an appropriate tool to deal with these problems.

The theoretical part of this thesis starts by discussing the advantages and disadvantages of the Minority Game to analyse real world problems. Following this, the situation in the Upper Derwent Valley is explained with economic game theories and the Minority Game. The two approaches mutually allow analysis of the situation in the Upper Derwent Valley leading to the development of a stochastic Minority Game. In addition, the strategies of the stochastic Minority Game in the Valley, as the combination of thought patterns and memories, are defined. The thought patterns are deductively discussed. The distribution of memories is figured out while the mechanism of the stochastic Minority Game is briefly tested by an online game (questionnaire), which was played 3,886 times by response all around the world.

Even after the theoretical part ends, the theoretical discussion appears repeatedly in the practical part of this thesis due to the generative approach of this thesis. Some theoretical arguments are meaningless without practical examples and justification in the generative approach. The theoretical arguments are cleared more by the end of the practical part.

The practical part of this thesis examines the components of the stochastic Minority Game with the data collected around the Upper Derwent Valley. The behavioural data used for practical analysis were collected during the summer of 2003. The behavioural survey on mode and parking location choices was a destination survey with stated preference survey methods. The survey was conducted at the parking areas around the Upper Derwent Valley. After a pilot survey, 700 questionnaires were distributed with face-to-face interviews and 323 were returned (i.e. a return rate of 46.1 %). The traffic related data used in this project are car movements, the arrival rate of private cars, parking hours, and annual traffic flow on the A57. The arrival rate was collected in front of a parking area during the survey period mentioned above. The parking hours at each parking area were obtained from the parking beat surveys, which were undertaken over three days during August 2001 by the Transport Office of Derbyshire County Council. In addition, the annual traffic flow on the A57 was collected by an automated system during 2003. The flow was westbound and eastbound from 7:00 to 19:00, and the average flow was used for this project.

In the practical part, the agent-based model is combined with stochastic discrete choice analysis such as multinomial discrete choice models to analyse the situation in the Upper Derwent Valley. Therefore, this simulation model is termed as stochastic agent-based modelling in this thesis. The simulation model is created by Java programming language and the RePast agent-based modelling toolkit. The discrete choice analysis and statistical analysis are carried out with BIOGEME and R. The agent-based model has four sub modules: 1) Multinomial mixed logit model for mode choice, 2) Binary logit model for parking location choice, 3) Markov queue model for parking network, and 4) the Minority Game for parking congestion and learning. Multinomial mixed logit model analyses travel behaviours with heterogeneous taste variation. A binary logit model is used to analyse parking location choice. A Markov queue model simulates the movement of cars and this module is considered as a microsimulation. Agent-based modelling combines all four models including the Minority Game, which is discussed in the theoretical part.

In this case study, the Minority Game is about choosing a less congested travel mode in the dynamic situation in the parking areas of the Upper Derwent

Valley. This simulation model produces comprehensive outputs including mode choices, congestion levels, and user utilities at both a micro and macro level. The results show that the road user charging scheme reduces car demand in the Upper Derwent Valley and ensures a reduction in congestion at the parking areas. The model also shows that an exemption will increase the utilities of elderly visitors without substantially sacrificing those of younger visitors.

In conclusion, the bias resulting from the neglect of interaction effects was observed and confirmed. The proposed solution was achieved by improving the approaches solely using discrete choice models in this thesis. The simulation model developed in this thesis demonstrated that oversimplification in conventional analysis gave significant biases when real world problems were analysed. For the Upper Derwent Valley, the oversimplification was the ignorance of dynamic interaction among visitors, which was represented as congestion in a parking network. Agent-based models have the advantages of incorporating dynamic modelling and connecting different components in the model. Therefore, the agent-based model simulated the situation of the Upper Derwent Valley more realistically. This thesis established a new type of agent-based model to examine the inter-relationships between a road user charging and park & ride schemes and socio-economic and physical factors simultaneously.

Acknowledgments

First, I am particularly grateful to my supervisor, Dr. John Preston, for the best combination of freedom and support. This was what I actually wanted for my D.Phil. and I would not have enjoyed my D.Phil. as much otherwise. Additionally, the most frequent and dedicated meetings in the final several weeks were invaluable to my thesis as I did not think I could have finished my thesis without his advice from these.

Second, I would like to thank to the Oxford Kobe scholarship for fully funding my D.Phil. as well as my M.Sc. Without their financial support, I would not have been able to even start my postgraduate education at the University of Oxford. Conjointly, I would like to thank Dr. Francis M. Mburu for encouraging me to apply for this opportunity. After graduation, I hope to contribute to society by pledging my work, knowledge, and skills, gained during this period, to show my appreciation to those who have invested in me.

Third, I would like to acknowledge the following people: Dr. Edmund Chattoe for introducing the Minority Game, the park rangers and Derbyshire County Council for personal interviews, Nikolaos Thomopoulos for collecting data as his M.Sc. project, Dr. Kenichi Ishibashi for funding the M.Sc. project, Fiona Rajé, Graham Sinden, James Keirstead, and S. E. Thornbush for proof readings, Homero Neves da Silva for lending many economics books, and the 323 travellers for responding to my questionnaires. Unfortunately, I cannot mention all contributors' names here due to the lack of space, but I will never forget the many contributions to this thesis.

Very lastly, I would like to thank my wife, Mrs. Rumi Takama, in my mother tongue - ありがとう - for really everything. I do not often express my appreciation to her, but I have never taken her support for granted from the first day I started my D.Phil. to the 999th day I had my viva.

Glossary

Only important keywords, which appear in more than two separate sections, are presented here.

Terminology

Agent	(Usually) A visitor to the Upper Derwent Valley in computer models
ASC	Alternative Specific Constant
Auto	Going to the Upper Derwent Valley by car
Bridge End Pasture	3rd parking area from the Information Centre
Bus	Going to the Upper Derwent Valley by bus
Bus fare	Fare for a bus service to the Information Centre
Cancel	Cancelling a trip to the Upper Derwent Valley
CKR	Common knowledge of rationality
Derwent Overlook	2nd parking area from the Information Centre
FIFO	First In First Out in network models
Horizon	See H below
Hurst Clough	4th parking area from the Information Centre
Headway	Period between departure times of buses
IIA	Independence of Irrelevant Alternatives (See footnote on page 89)
IID	Independent and Identical Distribution (See footnote on page 89)
Log-likelihood ratio	See footnote on page 89
Memory	Discrete experience each agent uses to play a game
Parking fee	Parking fee for Bus is the fee visitors pay when they park their car before getting on a bus. Parking fee for Auto is the fee visitors pay when they park at the Information Centre.
SIRO	System In Random Order in network models
Strategy	In the models developed in this thesis, it is the combination of a thought pattern and a memory. Its symbol is either s or TH1M1.
Searching time	Minute to find a parking space
The Information Centre	The Upper Derwent Information Centre
The Valley	The Upper Derwent Valley
Toll	£ to enter Derwent Lane from the A57
Visitor	A visitor to the Upper Derwent Valley
Walking time	Minute from a parking space to the Information Centre on foot
WTP	Willingness to pay for road user charging

Symbol

$\rightarrow \mathbb{R}$	Associating to a real value
α	Constant coefficient
β	Nonconstant coefficient
ε	Error term
θ_t^s	Success score of a given strategy s at a time step t within the scope of horizon H. Its equation is shown on page 78
λ	Departure rate of car per minute from the Valley
μ	Arrival rate of car per minute at the Valley
$\bar{\rho}$	Adjusted rho-square for log-likelihood ratio test
Σ	Summation
σ	Sample standard deviation
A	Auto option (Side A in Chapter 4)
Ac	a set of actions leading to the next unknown future state
a	Action
B	Bus option (Side B in Chapter 4)
C	Cancel option
e	Environment state
$\exp(x)$	Expositional transformed x
H	Length of the horizon which represents the horizon for which each strategy records its score
H0	Null hypothesis
H1	Alternative hypothesis
hour	Time of day and its interval is [10:00, 15:00]
$L(0)$	Null log-likelihood ratio for multinomial logit models
$L(\hat{\beta})$	Final log-likelihood ratio for multinomial logit models
m	Sample mean
$max(x)$ ...	Maximise x
N	Population size, the number of agents
n^x	Number of individuals choosing x
$P(x)$	Probability of choosing choice x
r	Run, which is a sequence of environmental states and actions, i.e. $e_0 \xrightarrow{a_0} e_1 \xrightarrow{a_1} e_2 \xrightarrow{a_2} e_3 \ldots \xrightarrow{a_{t-1}} e_t$
$R_i^{x_i^s}$	Return from the selected choice at time step i where x is selected choice by strategy s at i. Generally simplified to R^x.
s	Strategy. Same as TP1M1 below
TP1M1 ...	Strategy, which is the combination of **T**hought **P**attern and **M**emory. In this case, this strategy uses thought pattern 1 with memory 1.
t	Time step, e.g. $t-1$ is the last time step
U	Unobserved utility including error term
V	Observed utility excluding error term

Contents

List of Figures

List of Tables

Part I

Introduction

Chapter 1

Framework of thesis and the background on choices and interaction

1.1 Research aims

This D.Phil. thesis develops a new agent-based simulation model to improve the results of analysis, which solely uses discrete choice modelling, as well as to analyse the effects of a road user charging scheme for the Upper Derwent Valley in the Peak District National Park. The purposes of this chapter are to clarify the research aims and to provide overall background information for this thesis. The statement of a key problem throughout this thesis is:

> ***"Conventional approaches which conduct analysis solely with discrete choice models have the advantage of simplicity, but severe biases exist due to the neglect of some interaction and learning effects, which might be seen as oversimplification"***

The proposed solution to the problem in this thesis is:

> ***"Innovative interaction and learning are added to the conventional approaches by agent-based modelling together with discrete choice analysis"***

Based on the problem and proposed solution above, this thesis is subdivided into two parts, theoretical and practical, so that there is also a sub-aim in each part.

1.1.1 Aim at theoretical level

At a theoretical level, this thesis tests the traditional unidirectional cause-and-effect relationships, learning, and consequently some degrees of rationality. These assumptions are widely known; however, they are often ignored for the sake of the methodological usage in mathematical programming. The theoretical part discusses the implication of the ignorance from the interdisciplinary viewpoint of economics, social simulation, and statistics.

In particular, two methodologies are combined to analyse congestion: 1) traditional economic approaches, including game theories and probabilistic discrete choice analysis and 2) agent-based modelling. Agent-based modelling is the simulation of society from the bottom-up with the interaction of autonomous agents and tries to duplicate real world phenomena through computerisation (Epstein and Axtell, 1996, pp.153–162). In other words, the agent-based model tries to use the detailed decision making mechanism from lower-level society to analyse upper-level society. In contrast, one of the state-of-the-art econometric models, discrete choice analysis, produces a choice from a discrete set by treating unknown factors as random components (Ben-Akiva and Lerman, 1985). The latter has advantages over the former in terms of the simplification of problems. Both models are analysed with computational power although their approaches are different. Once parameters are calculated by a computer software package, the behaviours of individuals are represented by mathematical formulas in discrete choice analysis. Therefore, afterward, the choice of individuals can be calculated manually with any input data. In contrast, the conventional agent-based model is non-parametric, which means all behaviours are expressed as IF-THEN rules. This means that the choice of individuals has to be simulated to calculate each result in an agent-based modelling analysis. Moreover, the recent studies on human genomic show that the human mind works as IF-THEN rules even at a level of genes and proteins (Marcus, 2004); therefore, this approach could be a better representation of the real human decision making.

The comparison of the two types of models can be illustrated as:

$$\begin{array}{rlll}
\textbf{Discrete choice} & \Rightarrow \text{reductive} & \Rightarrow \text{simplistic} & \underbrace{\Rightarrow \text{☑ more-feasible}}_{\text{efficiency of model}} \\
 & & \Rightarrow \text{parametric} & \Rightarrow \square \text{ theoretic} \\
\textbf{Agent-based} & \Rightarrow \text{holistic} & \Rightarrow \text{complex} & \Rightarrow \square \text{ less-feasible} \\
 & & \Rightarrow \text{non-parametric} & \underbrace{\Rightarrow \text{☑ realistic}}_{\text{reality of model}}
\end{array}$$

The analysis of real world problems such as parking congestion may have been compromised in reductive models for methodological purposes since it was more feasible in terms of research time and cost. However, the compromise may be unnecessary if holistic models, which represent the world clearer and more realistically, are feasible. Therefore, while integrating the two approaches, this project proposes a new model, which combines the advantages of modelling efficiency and modelling reality.

1.1.2 Aim at practical level

The practical aim is to analyse real congestion problems with the arguments explained in the theoretical part. It has been argued that methodological studies have not been motivated enough to make theories reflect the real world in social sciences compared to physical sciences in the past centuries (Moss, 2002). This is a strong viewpoint; however, this issue has been continuously criticised since 1970s (Siegfried, 1970), and the problem is true especially in transportation modelling (Mackie and Preston, 1998). A panel discussion was held during the ninth World Conference on Transport Research 2001 to improve the relevance of transportation modelling to practice. One of the conclusions from the discussion was that the current transport modelling was too theoretically oriented and did not look at model outcomes and realities (Ben-Akiva and Bonsall, 2004, p.102).

This study is particularly interested in the multinomial discrete choice model, which has been widely practised in transport modelling. Unlike agent-based simulation models, conventional approaches which conduct analysis solely with discrete choice models have no mechanism to consider the interaction of people and consequently cannot formulate the dynamic process of congestion (Takama, 2004b). Hence, conventional approaches which conduct analysis solely with discrete choice models are expected to produce biased outputs. The comparison and combination of the two approaches were conducted at the Upper Derwent Valley, the Peak District National Park (Figure 1.1). Moreover, both approaches are used to forecast the effect of a road user charging and park & ride schemes on congestion by day-trip visitors at the four car parks. The forecasting and evaluating of the transport policies are carried out in terms of 1) the level of congestion and 2) user utility.

One of the practical aims in this thesis is to improve the relevancy of

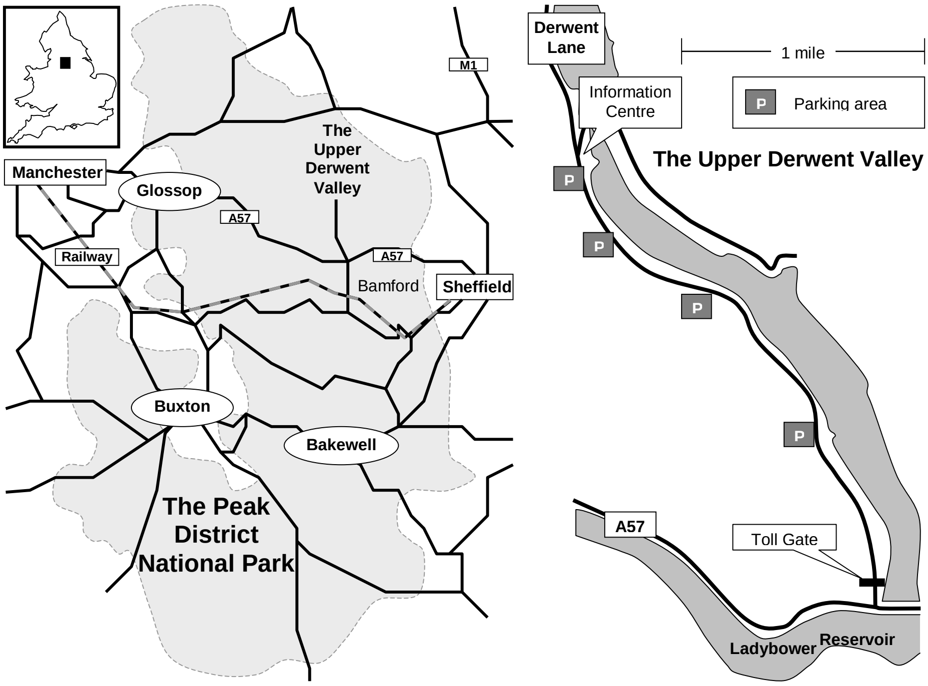


Figure 1.1: Map of Derwent Valley

transportation modelling to reality. Having said that, many concepts in this project are at an experimental level and the proposed scheme may not be implemented in the Upper Derwent Valley. In other words, this D.Phil. thesis can be seen as a relatively large pilot study for the future research, and the Upper Derwent Valley was used as a testing ground for a new concept of simulation modelling.

It is important to emphasise that although the theoretical and practical aims are separated at different levels, they are strongly interrelated. For example, the accuracy in the results between the two different analyses is not testable without the case study of the road user charging scheme at the Upper Derwent Valley. For example, one of the traditional neo-classical economics assumptions, *"the summation of local optima equals a general optimum"*, is testable by comparing the results from a conventional approach using solely discrete choice models and an agent-based simulation model with same input data.

1.1.3 Conceptual framework and research approach

The research is conducted by combining various approaches - empirical, deductive, and inductive. Particularly, the agent-based simulation model requires iteration between a deductive[1] and inductive[2] strategy as the researcher develops the model. In other words, one simulation starts with a set of assumptions, and then uses an experimental method to *generate* data which can be analysed inductively (Axelrod, 1997, pp4–5). Therefore, Epstein and Axtell (1996, p.177) identify this new category as the generative approach, which is the combination of the deductive and inductive approaches.

The final model of this D.Phil. project is the combination of four sub modules, namely multinomial discrete choice models of mode choice and parking location choice, Markov queue parking network, and the Minority Game (Figure 1.2). Multinomial logit model is a type of discrete choice models and these terms are interchangeably used in this thesis. Also, this thesis is largely divided into two parts, theory and practice. As Table 1.1 shows, the four sub models

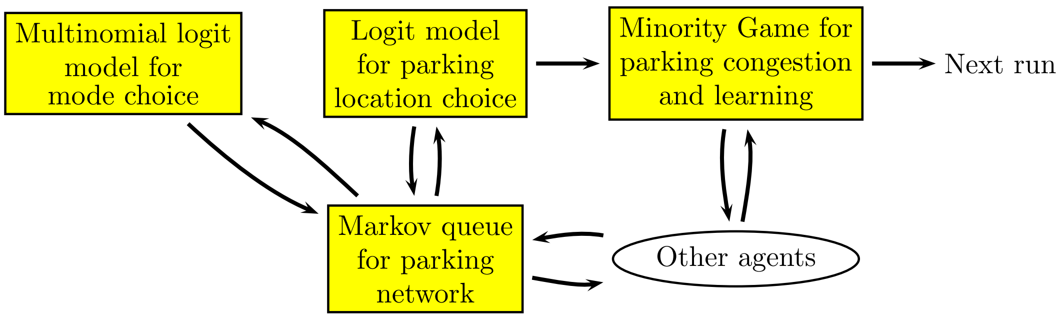


Figure 1.2: Structure of agent-based simulation

are discussed and implemented in both parts of this thesis. This monograph thesis has to be presented in a linear format, i.e. from the first page to the final page, but the process of writing was generative, i.e. iterative and interactive

[1] loosely, testing of sets of assumptions and their consequences (Gilbert and Troitzsch, 1999a, p.25), i.e. general to specific

[2] the development of theories by generalisation of observations, i.e. specific to general

Table 1.1: Relationship between chapters and models of the thesis

		Theoretical part			Practical part		
Purpose of chapter		Ch.2	Ch.3	Ch.4	Ch.6	Ch.7	Ch.8
		Explains the background of the Minority Game	Discusses congestion as a game and combine game theory and the Minority Game	Justifies the combined model and figures out distribution of memories	Analyses the visitor characteristics and travel behaviour with discrete choice models	Analyses and simulates the movement of cars with Markov queue network model	Analyses and simulates the parking situation in the Upper Derwent Valley
Type	**Model**	**Relationship between chapters and models**					
Econometrics	**Mode choice**		Congestion as a mode choice game	Justification in terms of game theory	Uses a multinomial discrete choice model to analyse travel behaviour		Handling visitors' mode choice
Econometrics	**Parking location**		Parking congestion as a source of game		Uses a discrete choice model to analyse parking behaviour		Handling visitors' parking location choice
Statistical	**Parking Network**					Uses Markov queue model to simulate car movements	Handling car movements
Multi-agent	**Minority Game**	Minority Game as a concept of congestion	Introduces the concept of thought patterns and strategies into congestion problem	Justification in the real world and finding memory distribution			Uses concept of thought patterns and strategies and combines all components

modification of theoretical and practical parts. This means that theoretical discussion contributes to the construction and evaluation of models, and practical analysis contributes to the formulation of theoretical thoughts. The theoretical arguments and models were developed by co-evolutional processes. Due to this co-evolutional process and generative approach, the final products are different from my work presented in two conferences in Florida (Takama, 2004b) and Bristol (Takama, 2005) and a workshop in Oxford (Takama, 2004a).

Furthermore, it is important to clarify that although this thesis has a theoretical part, this does not aim to develop any kinds of social theories. The theoretical part is used only to discuss my theoretical arguments. That is, this thesis uses conventional theories to understand the mechanism of empirical phenomenon or to check how the theories reflet reality in some early chapters; however, this thesis does not extract any new theory from the empirical results. In other words, this thesis tries not to be theory-oriented since as Clark (1998, p. 77) says; "*we seem to have too many theories for the empirical observations available and too little theory that makes sense of the scope and diversity of the world*". This thesis uses the term 'theoretical argument', but it can be paraphrased to 'anti-theoretical argument'.

1.2 Background for theoretical arguments

This section shows the problems of conventional assumptions, explains what agent-based modelling is, and then, proposes how the simulation modelling works as an alternative.

1.2.1 Limitations of traditional assumptions

Current social science modelling, such as neoclassical economics with mathematical programming methods, is constrained by some significant underlying assumptions. One of the main ones is that the summation of local optima forms a general optimum (Epstein and Axtell, 1996, pp.10-12). In other words, a general optimum can be achieved if, and only if, the interaction of agents does not produce any sub-effects to the society and environment. Therefore, it ignores the co-evolutionary interactions of agent-agent and agent-environment that are observed in real life. The omission of these interactions can be interpreted as zero transaction and external costs in the society in economic term (Berger, 2001, pp.245-246). These costs exist in real-life, for example in the transport sector, and are determined by the quality of the system. For example, a better transport system improves information availability and reduces transaction costs; therefore, visitors can reduce the risk of travelling to congested areas. At the same time, society can reduce the external costs of visible pollution and traffic noise. Because of the assumptions, the decision making and forecasting abilities of the conventional modelling approach are limited when it is applied to real world problems including traffic congestion (Tolley and Turton, 1995, p.20).

Traditional analysing approaches using discrete choice analysis have the underlying assumption of fixed tastes. This is problematic if people have heterogeneous characteristics. Transport and tourism economics involves heterogeneous agents since visitor behaviours are often discrete choices. In other words, visitors and commuters cannot select a part of a mode while using another type of a mode (McCarthy, 2001, p.93). If all consumers have homogeneous taste as defined by neo-classical economics, only one optimal mode of transport will survive. However, more than one mode exists in the real transport system so the underlying assumption of the identical preference for all agents is implausible.

The problem of homogeneous taste is solved in multinomial mixed logit models (Train, 2003) and recently in the agent-based simulation model (Page, 1999, p.11), recently. Therefore, the problem can be solved by using a multinomial mixed logit model, an agent-based simulation model, or the combination of both. This project proposes the combination approach since it copes with the co-evolutionary interactions of heterogeneous factors. One of the crucial questions to be asked by this research is:

> ***"How do the results from conventional analysis, which solely uses discrete choice models, change when dynamic interaction and learning are integrated by agent-based modelling?"***.

1.2.2 Agent-based modelling as a solution to conventional assumptions

Before explaining why agent-based modelling can be one of the solutions to the problem of the traditional assumptions, this section briefly explains the basics of agent-based modelling.

History and basic concept of agent-oriented program, multi-agent system (MAS) and agent-based modelling

Agent-based simulation originated from the next-generation-computer-paradigm, agent-oriented programming following object-oriented programming (e.g. Java and Object-C) and its computer system, multi-agent system (MAS) (Wooldridge, 2002, pp.303–315). The concept of an agent did not emerge until the mid 1980s and the research in MAS seriously started in the late 1990s.

Although the definition of an agent may not yet be agreed upon, Wooldridge and Jennings's (1995) statement is widely accepted:

> An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment

in order to meet its design objects.

The definition above is still rather abstract and Wooldridge and Jennings (1995) state the four typical properties of computer agents are: autonomy[3], social ability[4], reactivity[5], and proactivity[6]. It is not necessary to understand how these four properties of an agent work. Rather than showing the mechanics of the complex system, many researchers use an intentional stance, such as "belief" and "hope", to explain and programme the system (Wooldridge, 2002, p.30). The intentional stance is merely a new abstraction tool to manage the complexity of the computer system. For example, the history of programming languages moves from low-level machine-oriented languages (e.g. `0100110101...`) to abstract languages similar to human perception (e.g. `for, while, class, switch`). The intentional stance used for many researchers is human-like mental attributes such as belief, desire, and intentions (Bratman et al., 1988), and the agent system with the three attributes is the so-called **B**elief-**D**esire-**I**ntention (BDI) model.

It is important to bear in mind that beliefs are not always true[7], so that the information an agent has could be wrong, particularly, for contingency truths in the beliefs of social simulation (e.g. you believe it will rain tomorrow). This is expressed by a modal logic:

$$M_{\text{Belief}} \models_w \varphi \qquad \text{where} \quad \begin{aligned} &W = \{w, w', \dots\} \\ &W \neq \emptyset \end{aligned}$$

[3] Agents operate without others having direct control of their actions and internal state.

[4] Agents interact with other agents through some kind of 'language' (a computer language, rather than natural language).

[5] Agents are able to perceive their environment (which may be the physical world, a virtual world of electronic networks, or a simulated world including other agents) and respond to it.

[6] As well as reacting to their environment, agents are also able to take the initiative, engaging in goal-directed behaviour.

[7] This attribute is known as "no reflexive" (Meyer and Lomuscio, 2003)

The formula is read as φ is true at w in a model of belief M_{Belief} (i.e. M_{Belief} captures φ in w), and w is a world in a non empty set of possible worlds[8] W. From this formula, the two properties are deduced; a belief is "possible true", which means true at some worlds, and perhaps false at others (Meyer and Lomuscio, 2003), and an agent believes φ if φ is present in the agent's belief model (Konolige, 1986). This means that an agent in a BDI model may have localised information from experience, but not have perfect information in the world. The desires could be a set of listed choices an agent has and the intention is a chosen desire.

Moreover, it should be stated that agent-based modelling and MAS use not only computer science concepts, but also economic and social science concepts. This trend is likely to become much more widespread over the coming years. For example, although conventional multi-agent system is based on IF-THEN rule base, the idea of utility with a classical decision theory is sometimes used to determine the intensity amongst the listed desires (Rao and Georgeff, 1995, p.4).

Simulation of social phenomenon with agent-based modelling

Concurrently with the emergence of agent-oriented programming and multi agent system, the computer simulation of social phenomena started in the 1990s (Gilbert and Troitzsch, 1999a, p.1). The combination of the two trends brought prosperity to the agent-based modelling of social simulation.

In social simulation, there are the environments of artificial societies and computer agents. After the agents are placed into an environment, they need 'sensors' to perceive their local neighbourhood and some means with which to affect the environment (Gilbert and Troitzsch, 1999a, p.167). Communication

[8]The idea of possible worlds is developed by Kripke in the 1950s (Kripke, 1980). Each world represents one state of affairs considered possible, given what is known.

between agents is likely to be conducted through the environment, so agents need to be able to 'hear' messages coming from the environment and to send messages to the environment for the communication between other agents. The reason to route the communication between agents through the environment is to make the effect of sequential running of simulation less serious [9] (Gilbert and Troitzsch, 1999a, p.168).

Agent-based modelling is used from the theoretical modelling of the human settlement in the prehistoric Europe - EOS project (Doran et al., 1994) to the practical decision making tools of policy analysis such as freshwater management in England - FIRMA project (Downing et al., 2001). One of the most common misunderstandings is that "the goal of simulation research is to fully simulate the real situation". The initial aim can be the duplication of real data in computer simulation (Moss and Edmonds, 2004), but the researchers in this field are more interested in understanding factors and processes of simulation models. For example, in the EOS project, the goal of simulation is to try to understand some of the factors involved in the emergence of social complexity rather than to fully simulate these ancient societies (Doran et al., 1994). Similarly, in the FIRMA project, the aim is to understand the impact of governments exhorting water consumers to exercise care and caution in water usage during times of drought (Downing et al., 2001, p.206). The final target of agent-based modelling in this thesis is not to reproduce the real transport system in a computer, but to show how the conventional approach, which use solely discrete choice models, can be improved by agent-based modelling in terms of model reality and model efficiency.

[9] The order of running simulation is sequential since programme code is written sequentially; if the programme of agent `A` is written before agent `B`, the agent `A` communicates to agent `B` before agent `B` communicates to agent `A`. However, through the communication between agents in the environment, the sequential problem is nullified; messages from agents are collected and stored in the environment. At the beginning of the next time step, all the stored messages are delivered to their recipients.

Agent-based modelling as a solution

Agent-based modelling could be a solution to the problem of the assumptions made by discrete choice analysis. Agents in the model interact with the environment and other agents through socialisation, reaction and pro-action. In other words, agents do not need to know all consequence of actions since agents interact through their experience and current situation of local environment – *There is no assumption of perfect information.*

In reality, it is unlikely that individuals can gather perfect information because of shortsightedness and bounded rationality (Varian, 1993, pp.602-662). Therefore, when utility theory is combined with agent-based modelling, agents calculate their utility based on their belief and experience. In other words, an agent chooses the best option which has the maximum utility in the current local environment including the reflection from own experience[10] and the actions[11] that lead to the next environment situation. The maximum utility function of an agent at a particular environment is:

$$\max\{U : e \xrightarrow{Ac}\} \rightarrow \mathbb{R} \qquad \text{where} \quad \begin{aligned} E &= \{e_0, e_1, \dots\} \\ Ac &= \{a, a', \dots\} \end{aligned} \tag{1.1}$$

This function shows that the maximum utility, $\max\{U\}$, of an agent associates a real value[12] with its current environmental state, e and a set of actions leading to the next unknown future state, Ac. This is a rather shortsighted optimisation since the value assigns utilities to a local and short-term state; it is difficult to specify a long-term view.

In contrast, if an agent has perfect information as assumed in conventional

[10] The environment may be in any of finite set E of discrete instantaneous states (Wooldridge, 2002, p.31). For example, an environment state changes from e_0 to e_1 and so on.

[11] An agent has a repertoire of possible actions, which transform the state of the environment

[12] '$\rightarrow \mathbb{R}$' means associating to a real value

modelling, the agent could have the utility function based on its run (r), which is a sequence of entire environment states (e) and actions (a):

$$\max\{U : r\} \to \mathbb{R} \qquad \text{where} \quad r : e_0 \xrightarrow{a_0} e_1 \xrightarrow{a_1} e_2 \xrightarrow{a_2} e_3 \ldots \xrightarrow{a_{t-1}} e_t \tag{1.2}$$

For the decision making process with a utility function based on a run (function 1.2), an agent must know all future environmental states, actions, and their interactions, i.e. *the assumption of perfect information.* Such information is impossibly achieved in the real world, so conventional models tend to ignore the interaction effects at a practical level. Therefore, the agent-based model in this research uses agents with random utility theory; however, the rationality is bounded by beliefs, experience, and shortsighted interactions (Simon, 1957). By comparing the agent-based modelling of bounded rational agents with the perfect rational agents, this research may answer how the imperfection of information, such as interaction effects, affects the result of analysis. The formal hypotheses in this question are:

H0: Model output$|\max\{U : e \xrightarrow{Ac}, \text{imperfect information}\} \to \mathbb{R} \quad =$
Model output$|\max\{U : r, \text{perfect information}\} \to \mathbb{R}$

H1: Model output$|\max\{U : e \xrightarrow{Ac}, \text{imperfect information}\} \to \mathbb{R} \quad \neq$
Model output$|\max\{U : r, \text{perfect information}\} \to \mathbb{R}$

In brief, this hypothetical question tests if a model based on an iterative process with imperfect information and another model based on a single calculation with perfect information produce the same output.

This section attempted to explain the problems of traditional economic modelling and the potential advantage of agent-based modelling. Then, the conventional approach and improved approach with agent-based modelling are compared with a real case study of the Upper Derwent Valley.

1.3 Background on the transport situation of England and the Upper Derwent Valley

This section starts with the issues related to the transport sector in England and then the story is narrowed into the case study of the thesis project. Tourism is one of the key transport components, especially in the areas of outstanding natural beauty. Transport policies, such as private car access regulations or road user charging, can improve values of natural beauty by reducing visual intrusion, air pollution, and traffic noise (Dillon, 2002). Ecological conservation values, including the preservation of biodiversity and rare species, may also improve directly and indirectly by reducing emissions of gases and suppressing travel demand, which cuts the need for road construction.

Also, travellers may directly improve their user utilities after congestion levels are reduced by these transport policies. Car drivers may shorten their travel times and pedestrians and cyclists may increase their travel safety around the tourist destinations. In this research, visitor demand and congestion are focused on since they are the major transport related concerns in the area of British natural beauty.

1.3.1 Traffic congestion and road user charging

Road congestion has been seriously affecting the tourism sector in England and will do so for the next decade:

> Traffic congestion and associated air pollution are now perceived by some as the biggest threat to maintaining the economic utilities from tourism in many of the UK's prime destinations. Traffic congestion also impacts negatively on the visitor experience, the health of local people, etc. (Department for Culture, Media and

Sport, 1998, p.6).

Visitors to the National Parks are heavily dependent on their private cars. According to underlying economic theory, road user charging scheme is a suitable method to make sure that the road users (e.g. car drivers) pay for the external costs generated from their travel (Steiner and Bristow, 2000, p.95). Currently, one of the major objectives to install road user charging scheme is to reduce traffic congestion. In addition, road user charging scheme is planned where there are high levels of seasonal traffic in rural areas such as the Lake District National Park and the Peak District National Park (Eckton, 2003; Department of Environment, Transport and Regions, 1998, p.115). Therefore, it is inevitable to consider the implementation of a road user charging scheme during a summer holiday period around the Upper Derwent Valley, which is one of the most popular destinations for visitors to the Peak District National Park.

1.3.2 Case study site description

The case study site is the Upper Derwent Valley, the Peak District National Park, during the summer holiday period. It is estimated that 60% of the population of England and Wales lives within two hours travel distance by car to the Peak District National Park. Moreover, the National Park is the second most visited national park in the world after Mt. Fuji in Japan (Derbyshire County Council, 2002). The Upper Derwent Valley is visited by around two million visitors a year, approximating a travel of 500,000 cars in use. For example, on the 2001 August Bank Holiday Monday, 3,044 cars travelled along the Upper Derwent Valley. There is only one practical transport mode to get to the Valley namely the automobile. The survey was conducted in the Upper Derwent Valley in the summer of 2003. In totally, 700 questionnaires were distributed and 323 were returned.

The Upper Derwent Valley is located between two large cities, Manchester and Sheffield. The access to the Valley by private car is easy not only from local towns but also from these nearby cities via the A57 (Figure 1.1). The entrance to the Upper Derwent Valley by car is only from the A57 and only through Derwent Lane, which comes to a dead-end. Access by public transport to the Upper Derwent Valley is revealed to be unpopular from the interviews during this survey and comments on the questionnaires. The reasons for unpopularity include the low frequency of bus services and the inflexible regulations (e.g. visitors cannot travel by the public buses with their bicycles). In fact, most of the buses observed during this survey were almost empty. There are four parking areas on Derwent Lane: (from the Information Centre to the A57) 1) Upper Derwent Information Centre parking area, 2) Derwent Overlook, 3) Bridge End Pasture, and 4) Hurst Clough. The approximate parking capacity of each parking area is 134, 77, 58, and 18 cars respectively[13]. Only the first parking area requires a parking ticket, which costs £2.50 for one day parking or 50 pence per hour.

There are two main reasons why the Upper Derwent Valley was chosen for this research. First, the Upper Derwent Valley receives a large number of visitors by cars during the summer period. It is important to note that even on the busiest days, the congestion on the roads, such as on the A57 and Derwent Lane, are minimal, but severe congestion occurs around and in the parking area of the Upper Derwent Information Centre (the Information Centre). The most scenic area starts from the Information Centre (The area north from the Information Centre in Figure 1.1) and the access to the area by private cars is restricted during weekends, holidays, and high seasons. Therefore, visitors

[13]The first parking area at the Information Centre has two sections, but these are considered as one parking area in this research. Also, the third parking area includes a parking space on road shoulder around the parking area.

to the Valley try to park as close to the Information Centre as possible to reduce walking time before they start hiking in the scenic area. Additionally, the Information Centre also has extra attractions such as a souvenir shop and a take away shop. These situations make the competition of parking at the Information Centre severe. If the parking area is full, the visitors move to the next parking area, and so on. The movements of cars in the Upper Derwent Valley are discussed in Chapter 7.

Second, currently the implementation of a road user charging scheme is being considered for Derwent Lane (Derbyshire County Council, 2000). A proposed, but not confirmed, tollgate for the road user charging scheme is shown in Figure 1.1. A park & ride scheme is also planned in this area as the complementary policy tool of the road user charging scheme. There are three proposed parking areas for the park & ride scheme around the A57 and Bamford, but these locations have not been confirmed yet. The road user charging scheme is still under consideration and may take some time to be implemented or could be withdrawn. However, the Upper Derwent Valley is still one of the best case study sites in the UK to analyse the effects of a road user charging scheme around the area of natural beauty. The road user charging and park & ride schemes are further explained in Chapter 6.

1.4 Background for modelling discrete choice behaviours

1.4.1 Random utility theory and discrete choice analysis

The random utility theory was formulated in the late 1950s by Luce (1959) and implemented statistically in 1970s (Manski, 1977; McFadden, 1981). Today, it is adapted in many human behavioural models including transport demand models. In this theory, individuals tend to select the alternative, which maximises their utilities; however, the decision is still bounded to random components coming from the uncertainty in utility functions (Ben-Akiva and Lerman, 1985). In 1944, Neumann and Morgenstern (1953) developed the formal theory of risk and uncertainty, which was applied to the random utility theory. Although the two concepts make the expected decision of agents differ between agents, the agents try to maximise their utilities (Sandholm, 1999, p.214). For example, although an agent may visit a national park at a 50% congestion level, although another agent my prefer to stay at home. This is because of the risk, which agents have different feelings about. On the other hand, they may make different decisions even within the same individuals due to uncertainty in the utility functions. The utility of a decision maker can be estimated either by assuming mathematical functions or by fitting a curve empirically among a set of a discrete utility distribution. No matter what the utility function is, there is always an estimation problem, and the functions above are likely to be theoretical idealisations, but not reality. Therefore, the unobserved disturbance term, ε, should be associated on the right-hand side of the equations, i.e. $U(x) = V(x) + \varepsilon$. Also, the disturbance term, ε, is a random variable that is usually distributed with mean zero and some form of variance. This is called the random utility theory, which is an important baseline of econometric discrete choice analysis.

Random utility models have been extensively used in the field of transportation research since the emergence of the model in the last 30 years. All applications were likely to be based on the multinomial logit model for discrete choice analysis (McFadden, 1974). Discrete choice analysis examines individual choice between discrete alternatives, such as the choices of travel mode based on individual behavioural data, including travel origin and the frequencies of trip (Spear, 1977). Therefore, its models are often called disaggregate travel demand models. In this thesis, the mixed logit model is used, which is one of the state-of-the-art discrete choice models (Bolduc and Ben-Akiva, 1991; Ben-Akiva and Bierlaire, 2003). The mixed logit model considers heterogeneous individual tastes, which are assumed to have some defined distribution as well as overcoming the problem of independence of irrelevant alternatives (IIA)[14] in conventional logit models by taking the covariance between choices into account (Train, 2003).

The mixed logit model and any other discrete choice models cannot be calibrated by using standard curve-fitting techniques, such as least squares estimation, because their dependent variable is an unobserved probability (between 0 and 1) and the observations are the individual choices, which are either 0 or 1. Therefore, the mathematical transformation of the utility values is required to get probability values between 0 and 1. A basic logit function for three alternatives is:

$$P(A) = \frac{\exp(U^A)}{\exp(U^A) + \exp(U^B) + \exp(V^C)}$$

where U^x is utility and $P(x)$ is an unobserved probability, and the parameters of the utility function are not fixed in the mixed logit model due to taste variation. The factors considered in this thesis include parking fee, bus fare, toll fee, the interspaces between two buses (headway), and searching time to find a car park and walking time to distinction in alternative travel modes.

As mentioned earlier, the mixed logit model is based on individuals like the agent-based simulation model. However, the analysis at the aggregated level is different from that of an agent-based simulation model. The disaggregated econometric model is probabilistic; therefore, its usage has to be made in a proportional term as they yield the probability of choosing each alternative and

[14]The probability of choosing one option between two is not affected by adding any other third option (Arrow, 1951). Also see footnote on page 89

do not indicate which one is selected (Ortúzar and Willumsen, 2001, p.221). For example, the expected number of people using a certain travel option equals the sum over each individual's probabilities of choosing the alternative:

$$n^x = \sum_N P^x \tag{1.3}$$

where n^x is the number of individuals choosing x and P^x is the probability of choosing x in N population. This is the transformation of probability to proportion, and population is treated as a continuous variable. However, in reality, the population is the summation of discrete entities, which is determined stochastically and hence not a continuous variable. If the interaction effect between individuals is insignificant, the outcomes from the mixed logit can be the same as those of agent-based modelling, but if it is significant, the mixed logit model is expected to produce biased outcomes. Under these conditions, conventional approaches which conduct analysis solely with discrete choice models have the neo-classical economics assumption that *"the summation of local optima forms a general optimum"*.

1.4.2 Four stage model

The case study of this thesis is about transportation modelling; therefore, it is important to mention the four stage model, which is most commonly practiced in this sector. The four stage model was first used in the US in the 1950s (Oi and Shuldiner, 1962) and is composed of 1) trip generation (and attraction), 2) trip distribution, 3) mode split, and 4) assignment (McNally, 2000) (Figure 1.3)[15].

The first stage measures trip frequencies and provides the propensity to

[15] Although this is the most common order, the variants of the four stage model exist in which mode split is undertaken prior to trip distribution or the mode split and trip distribution stages are undertaken simultaneously

travel, i.e. trip production (the number of trips from origin i, T_i) and trip attraction (the number of trips to destination j, T_j). This will typically be a function of the characteristics of origin zone i (O_i), the characteristics of destination zone j (D_j) and the overall cost of travel (C). The second stage distributes the trip production to match the trip attraction, thus linking trips between origin i and destination j (T_{ij}). This distribution is based on underlying travel impedances including travel time and costs (C_{ij}). The third stage is the mode split model which determines the number of trips between i and j by mode m (T_{ijm}). The fourth stage assigns the mode split to specific routes k (T_{ijmk}).

Originally, four stage models had a sequential structure although the mode split stage could be either before or after the distribution stage (as in Figure 1.3). Today such models have feedback loops (as illustrated by Figure 1.3) and hence the four stages can be estimated and applied simultaneously. The feedback loops in the four stage model aims to achieve convergence or an equilibrium status between demand and supply as in a system dynamic model (Gilbert and Troitzsch, 1999a). Although convergence is not guaranteed or takes days, four stage models are feasible even in the largest of the UK multi-modal study areas, including the Central Scotland Transport Model[16] (Department for Transport, 2003, §1.4.7). Moreover, the model in Figure 1.3 is a basic model, but more complex models may consider the importance of land-use, travel time, incremental change to the base, etc. (Bates et al., 1991).

The conceptual diagram for the agent-based model in Figure 1.2 can be partially compatible to the four stage model in Figure 1.3. The mixed logit model covers the third stage (mode split) and partially covers the first stage (trip generation). The Markov queue model covers the fourth stage (assignment) since it assigns the demand of Auto to the parking networks of the Valley. The agent

[16] This covers much of Scotland from the English border to north of Perth

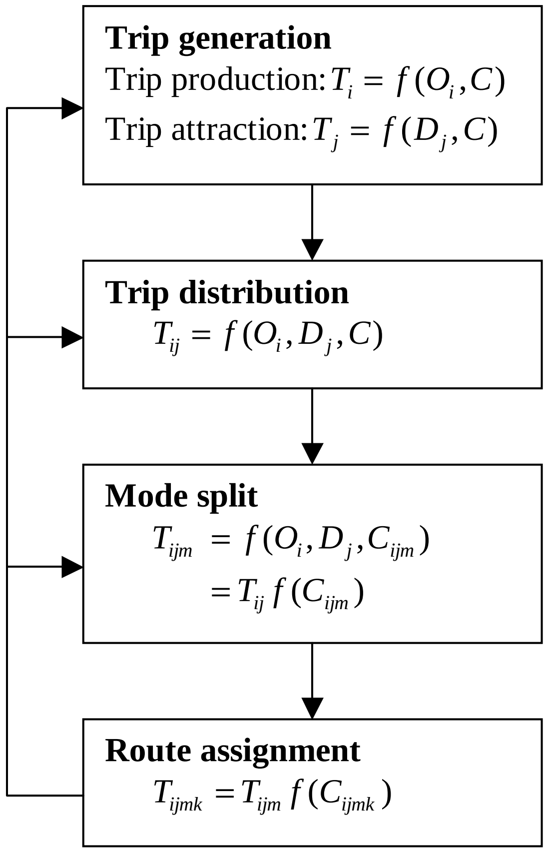


Figure 1.3: Four stage model with feedback loops

based model does not have a second stage (trip distribution) as the destination is assumed to be fixed. However, the two feedback loops of the agent-based model and the aggregated four stage model are essentially different.

Generally, the feedback loops in the four stage model occur at a system level[17], so that this feedback loops the interactions of modelling stages, but not the interactions amongst agents. Moreover, this system dynamic feedback loop cannot show emergent phenomena. In contrast, the feedback loop in the agent-based simulation occurs at the individual level, but not at the system level, i.e. many discrete choices and feedback loops are conducted in one cycle.

Also, the feedback loop of the agent-based simulation model aims at presenting emergent phenomena rather than convergence (Epstein and Axtell, 1996). For example, in the case of this thesis, the convergence is not expected due to unforeseeable interaction amongst agents (See page 15). More specifically, the unforeseeability is handled with a Minority Game algorithm, so that convergence is not expected. In a similar manner when a stock market is represented as a Minority Game it does not achieve an equilibrium state due to uncertainly (Arthur, 1994).

The difference between the feedback loop of system dynamics and that of agent-based simulation is illustrated in Parunak et al. (1998) and Gilbert and Troitzsch (1999a, pp.12–3) and summarised in Table 1.2. Besides these differences in the feedback loops, the case study of this thesis does not require all four stages. The agent-based modelling at the Derwent Valley focuses only on the congestion levels in parking areas and the number of private car visitors. In other words, the case study is not interested in visitors' other destinations if they do not come to the Valley. Trip distribution is not discussed and

[17]TRANSIMS (`http://www.ccs.lanl.gov/transims/`) can be considered as a four stage model carrying out feedback loops at disaggregated levels. However, this type of models is different from the conventional four stage model. Actually, TRANSIMS claims that it is an agent-based simulation system in its web page.

Table 1.2: Comparison between feedback loops in system dynamics and agent-based modelling

Simulation type	Feedback levels	Communication in agents	Number of agents	Aim
System dynamics	System	No	1	Static equilibrium
Agent-based model	individual	Yes	Many	Dynamic phenomena

only a fraction of trip generation is discussed as a part of discrete choice. In conclusion, the mechanism of the four stage model is different from that of the agent-based model. The scope of the four stage model is too large for the case study; therefore, the four stage model is not discussed after this section.

1.4.3 Minority Game

This study uses an agent-based simulation of discrete choice known as the Minority Game. Arthur (1994) started problem with the El Farol model, which is also known as the Minority Game today, after he was inspired by the phenomenon of the bar "El Farol". El Farol is located in Santa Fe, the U.S.A, and offers Irish music on Thursday nights. However, the offer is held only if the bar is not crowded[18]. Therefore, people who go to El Farol on Thursday nights are playing the Minority Game - i.e. they may be better off coming to El Farol if the bar is not crowded, otherwise, they will be better off staying at home. As you see, the problem described above is a discrete choice, i.e. go or not go to the bar, and it is about congestion. Thus, this problem is applicable to the situation in the Upper Derwent Valley.

In the situation of the Upper Derwent Valley, visitors are better off travelling by private automobile when they can park at their target parking areas since it is their initial intention. Also, visitors arriving by bus will be glad

[18] He set the crowdedness is the condition that 60% of regular customers go to the bar in his model

that they chose the bus option if there are no empty spaces in the parking areas. These two situations indicate that the visitors are playing the Minority Game, i.e. visitors are better off choosing the less congested option. On the other hand, from the interviews conducted in the summer of 2003 and 2004, some people said that they would not bother to come to the Valley and would go elsewhere if the new policy tools were put into effect. Therefore, the visitors are able to choose the third option, which is not going to the Valley, in addition to the two Minority Game choices. Overall, the Minority Game in the Upper Derwent Valley is a multinomial discrete choice of Auto, Bus, and Cancel. Further description of the Minority Game is given in Chapter 2.

1.4.4 Learning model

Although the Minority Game algorithm contains some aspects of learning and perception, this thesis does not focus on these issues. Thus, learning models for transportation mode choice and some aspects of learning and perception in the Minority Game are briefly reviewed here.

The learning model is generally divided into two sub-categories namely explicit updating models and non explicit updating models (Jotisankasa and Polak, 2005). The explicit models have the modules of perception updating and decision making. The decision making component in the explicit models are often implemented from the discrete choice analysis, which is discussed in this thesis. For example, van Berkum and van Der Mede (1998) use a logit model to analyse a route choice and Jha et al. (1998) use a nested logit model to analyse a combined route and departure time choice.

Several perception updating approaches have been considered in the explicit models for transport analysis. First, the weighted average approach assumes that travellers formulate the perception of the travel time based on the weighted average of past travel times (Jotisankasa and Polak, 2005). Second, the adaptive expectation approach assumes that travellers update their perception of travel time based on the difference between the actual and the perceived travel times from the previous trip. Third, the Bayesian updating approach was successfully implemented in some studies (Jha et al., 1998; Ben-Akiva et al., 1999). This approach focuses more on random variables of the mean travel time and the experienced travel time, which indicate the confidence levels about the travel information. Therefore, the Bayesian updating approach can consider the confidence about information, i.e. variance, more than the other two approaches.

Since this stochastic agent-based modelling is based on discrete choice analysis as shown in Figure 1.2, the discrete choice model is considered as an explicit decision making component and another module namely the Minority Game can be considered as a learning component. The adaptive expectation approach is most popular amongst the three; however, the agent-based model developed in this thesis may be best characterised by the weighted average approach. One of the major criticisms of this approach is the ambiguous assumptions on the length of the distribution of memories (Jotisankasa and Polak, 2005). These assumptions are briefly tested with an on-line game in Chapter 4.

In contrast, the non explicit models do not have an explicit division between perception and decision making parts. Moreover, they usually use more innovative approaches instead of relying on discrete choice analysis. For example, Arentze and Timmermans (2003) use the reinforcement learning, which controls and adjusts an agent's action to the environment through trial-and-error processes without supervision, showing a correct action to the specified environment state (Sutton and Barto, 1998). Also, Nakayama and Kitamura (2000) propose a model based on IF-THEN inductive psychology with genetic algorithm and the concept of memory, which is described in Chapters 2 and 3. These approaches are similar to the Minority Game component of this thesis. Therefore, the learning process of this Minority Game can be also categorised as a non explicit approach. As it will be explained further in the next two chapters, the best strategy of agents to play the Minority Game are updated by reinforcement learning while the perception as selection of memory is formulated in the Minority Game.

In conclusion, this agent-based model has a learning and perception mechanism. The innovative Minority Game handles the learning and perception mechanism and most decision making processes arise from the discrete choice model. The best explanation of the learning model in this thesis is that the discrete choice analysis is explicitly separated from the learning component in the Minority Game while the results of the discrete choice models are feedback to the Minority Game together with other results from the parking network and parking choice models (i.e. explicit approach) (Figure 1.4). The Minority Game passes the selected memories and strategies in the Minority Game run discrete choice models (Chapter 3), so whole Minority Game algorithm is considered as a learning and perception mechanism.

However, choosing the best strategy is, in fact, a decision making phenomenon, so that the Minority Game component can be still viewed as a part of decision making component (i.e. non explicit approach). Although the model is initially based on a weighted average of the last n memories, reinforcement learning and inductive psychology in the Minority Game mean that choosing the best strategy involves decision making processes in a non explicit manner. Therefore, the double loop with the label 'Choosing the best strategy' in Figure 1.4 is the non explicit learning model mechanisms.

The discussion between explicit and non explicit learning models is difficult in this case due to the innovative approach of Minority Game and the nature

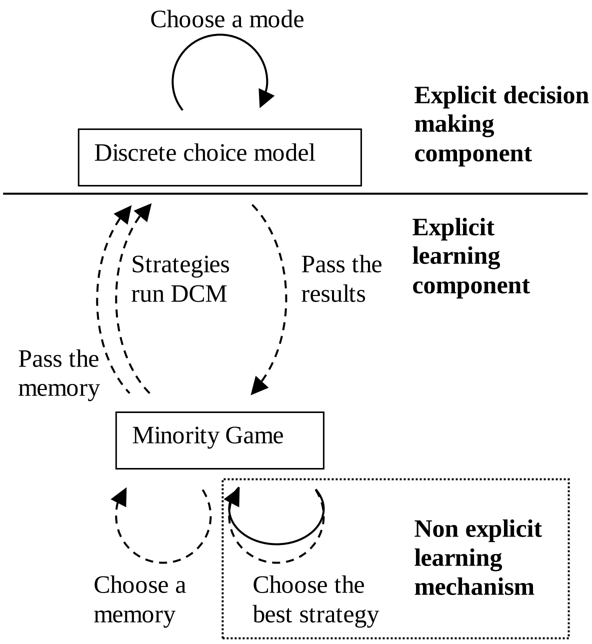


Figure 1.4: Leaning model with discrete choice model and Minority Game. Normal lines are decision making processes and dotted lines are learning and perception processes. Choosing the best strategy can be considered as the decision making as well as the learning mechanisms, i.e. non explicit learning.

of agent-based modelling, i.e. combining many modules together. Nonetheless, the Minority Game mechanisms are explained further in Chapters 2 and 3.

1.4.5 Modelling platform for agent-based modelling

The agent-based simulation model is programming oriented, and usually has to be written in one of the object-oriented programme language, such as Java, C++, Python, etc.. However, there are various toolkits (i.e. collection of classes) available to help the development of agent-based models. Some personal criteria for the platform and toolkit selection were:

- Flexibility for non-abstract and interdisciplinary modelling
- Based on a popular programme language for better learning materials and development tools
- Widely distributed platform for reliability and efficiency

Table 1.3: Sources and developers of the selected simulation toolkits

Name	Web site	Organisation
RePast	`http://repast.sourceforge.net/`	University of Chicago
Swarm	`http://www.swarm.org/`	Santa Fe Institute

- Under open source project for long-term development and easy cooperation with other researchers

These criteria are for long-term usability. Therefore, short-term criteria, such as non-steep learning carve, are not seriously considered. Most of the simulation platforms and toolkits were excluded because of the first and second criteria. For example, GUI (e.g. drag-and-drop interface) based platforms are easier to start, but they are less flexible so the platform may not handle all concepts needed to be modelled on the unique case of the Upper Derwent Valley, i.e. the model is too abstract. Some popular platforms use less popular programming languages, for example, CORMAS[19] uses Smalltalk. After these investigation, Java[20], which is one of the most popular object-oriented language, was selected. Then, the two most popular Java toolkits were considered: RePast (REcursive Porous Agent Simulation Toolkit) and Swarm (Table 1.3).

Swarm is originally written in Objective-C, but researchers can also model simulation with Java language by using Java Swarm, which is a Java layer running on top of the Swarm kernel (Tobias and Hofmann, 2004, § 2.4). Moreover, Swarm is probably the most popular, most widely distributed, and most flexible platform. RePast is a Swarm like multi-agent platform, but written purely in Java and used more in the field of social simulation than artificial intelligence (Collier et al., 2003). After an evaluation period, RePast was se-

[19] `http://cormas.cirad.fr/`
[20] `http://www.java.com/`

lected as the platform of all simulation processes due to social science oriented features and potential to future development.

1.5 Outline of the thesis

The outline of thesis is presented in Table 1.1 and the rest of this thesis is organised in the following way. The theoretical part starts with the additional information about the Minority Game, which is the kernel of this project (Chapter 2). The next chapter in this part discusses how congestion can be described as a game and why it is necessary to include the idea of the Minority Game into the problem (Chapter 3). Then, this chapter defines a new stochastic Minority Game as the combination of discrete choice analysis and conventional Minority Game. Following this, The distribution of memories is figured out while the mechanism of the stochastic Minority Game is briefly tested by an online game (Chapter 4).

The practical part starts with the background information about practical modelling issues (Chapter 5). The second chapter in this part is about the econometric analysis on the parking congestion in the Upper Derwent Valley and develops multinomial discrete choice models for a mode choice and a parking location choice (Chapter 6). This chapter also analyses the characteristics of visitors. The third chapter explains the Markov queue model about the car movements, which is the source of dynamic congestion process. Then, Markov-queue-like microsimulation is developed (Chapter 7). The last chapter of the practical part combines all four models including the Minority Game, which is discussed in the theoretical part by agent-based modelling (Chapter 8). In the final chapter, this thesis finishes with the main conclusions including discussion on the question raised in the introduction, some recommendations for policy makers, thoughts about future research, and closing comments (Chapter 9).

Part II

Theoretical part

Chapter 2

Background of the Minority Game and the El Farol Problem

2.1 Introduction

Rather than simple article reviews, this background chapter about the Minority Game explains the two replica models of the previous Minority Game while former works in this field are discussed. Since the Minority Game is a core module of this thesis, it is important to explain this concept in detail.

2.2 Main references and source codes to replicate the El Farol / Minority Game model

This chapter tries to replicate two El Farol / Minority Game models with currently available information. The first duplicated model is the original El Farol of Arthur's (1994) model and the second model is Bazzan et al.'s (2000) model, which took into account agents' characteristics in the transport sector.

To start to replicate Arthur's (1994) model, free Swarm source codes of the

El Farol[1] were examined. These source codes were distributed by Prof. Paul Johnson, Department of Political Science, University of Kansas. The codes had not been modified for the last four years, and not having been tested properly contained significant bugs; furthermore, they had been programmed without the RePast tool kit[2]. Therefore, these codes were not used at all for this thesis; however, they were useful to understand the structure and mechanism of the Minority Game simulation, which were not revealed in the two articles mentioned above. In addition, Lofton's (2000) working paper was also valuable to understand the source codes since the codes were not documented at all. Unfortunately, the working paper is no longer available from the URL link as stated in the bibliography.

2.3 Original El Farol model (Arthur, 1994)

As mentioned in Section 1.4.3, the regular customers to the El Farol bar are better off if the bar is not crowded since the bar offers Irish music only in this condition. It is impossible to find out the exact number of people coming to the bar in advance. So, regular customers to the bar have to guess if they are better off going to the bar or staying at home from their experience - i.e. inductive but not deductive rationality. In this paper, the author mentions two reasons why perfect / deductive rationality does not work, and how bounded / inductive rationality works in Minority Game situations (Arthur, 1994, p.406):

1. *"Beyond a certain complicatedness, our logical apparatus ceases to cope, our rationality is bounded"*, and

2. *"In interactive situations of complication, agents cannot rely upon the other agents they are dealing with to behave under perfect rationality,*

[1] `http://lark.cc.ku.edu/~pauljohn/Swarm/MySwarmCode/El_Farol/`

[2] `http://repast.sourceforge.net/`

and so they are forced to guess their behaviour"

He explained inductive reasoning, which could be understood as "localised deductive reasoning". This reasoning is applicable not only for real customers in the El Farol bar, but also for computer agents in the Minority Game. In a given environment of the El Farol simulation, computer agents always test hypotheses if their prediction is correct. If the prediction is incorrect, the agents choose an alternative strategy to predict the number of customers coming to the bar on the following Thursday night. Without deductive methods, the author showed that agents were reaching close to the optimal ratio, which was 40/60 = 'Stay at home' / 'Go to bar', at a macro level. On the other hand, the agent number coming to the bar never converged into one point and always oscillated around the optimal ratio.

2.3.1 Explanation of key classes in this simulation

In this section, the main classes, which were used for the duplicated simulation, are explained. It is not necessary to understand the relationships of these classes in detail, but understanding these will help readers to understand the successive chapters. In brief, object-oriented programmers do not actually programme objects, but programme classes, which define the structure of objects. A class has its own data, methods, and other classes (objects) and its objects are produced from the class. Therefore, if a class of agents is designed, the millions of agent objects are easily duplicated from the class, i.e. the relationship of class and object is like a foundry mould to produce many plastic dishes. One class may contain other classes and it is not necessary to know how the methods of the other classes work. For example, an agent class may have strategy classes and the agent class does not need to know how each strategy class proceeds in decision making, but needs to know what the result of the decision making is. In other words, each class encapsulates its own problems for itself to solve larger and more complex problems at a macro level.

The duplicated El Farol simulation model has 49 classes, so only four key classes are explained here (Table 2.1). It should be kept in mind that these classes may not be the same as those of the original El Farol model since the information was limited. The main idea of the duplication is to understand the concept of El Farol model to develop my own Minority Game model rather than just to rebuild the model in Java and RePast. Even if some differences exist, the concepts of key classes should remain the same. As shown in Table 2.1, Model is the root class of the simulation model and has a number of agents and a bar. Then, each agent has a number of strategies.

A simulation process of the El Farol model is described as a pseudo-code in Figure 2.1. Each agent first generates a subset of the possible strategies individually, and then chooses the best strategy to make a decision. In other words, an agent just follows the decision of the best strategy and the best strategy is determined by trial-and-error processes. The bar is notified how many agents come there at every time step, and all processes are controlled by the model. There are 21 types of strategies in the model (Table 2.2), so that this is assumed as the possible strategies for the agents in this model[3]

2.3.2 Results and discussion of original El Farol

The important parameters set for this simulation were:

- The number of agents is 100

[3]Even without considering trends and average attendances, all possible strategies are immense, i.e. 2^{2^M} where M is the size of agent's memory. Read Challet and Zhang (1997, p.1) for more explanation. For example, if an agent remembers the past three attendances to the bar, the number of the possible strategies is 256 ($2^{2^3} = 256$). This number is very fast increasing with the size of memory. Therefore, the pre-selected 21 strategies were used for this example.

Table 2.1: Key classes in duplicated El Farol model. N^A is the number of agents in the model and N_i^S is the number of strategies in an agent i.

Name	No. objects	Possess	Description
Model (M)	1	A, B	Root class managing the schedule of simulation
Agent (A)	N^A	S	Regular customer making a decision if it goes to bar
Bar (B)	1	–	Class storing the history of agent attendance to the bar
Strategy (S)	$\sum_{i=1}^{N^A} N_i^S$	–	Strategy each agent use to make a decision

Table 2.2: 21 strategies in duplicated El Farol model

ID	Description
0	Go to the bar randomly without considering the history
1	Assume to be the mirror image around mean of last week's
2	Assume to be the mirror image of attendance in two weeks ago
3	Assume the attendance is the same as the trend in the last five weeks
4	Assume trend in last three weeks bounded by 0 and total number of agents
5	Assume the attendance is the same as the trend in last eight weeks
6	Assume the attendance to be 110 minus last week's attendance
7	Assume the attendance is the same as the attendance five weeks ago
8	Go to the bar whatever the history is
9	Assume the attendance is the same as the overall average in whole history
10	Assume the attendance is the same as the moving average of attendance
11	Assume the attendance is the same as the average of past five weeks
12	Assume the attendance is the same as the average of past two weeks
13	Assume to be the mirror image of moving average
14	Go to the bar randomly, but different from Rule ID 0
15	Assume the attendance is the same as the average of past four weeks
16	Assume the attendance is the same as eight weeks ago
17	Do not go to the bar whatever the history is
18	Assume to be the mirror image of the average of past three weeks
19	Assume same as ten weeks ago
20	Assume the attendance is the same as the attendance two weeks ago

```
01 Start initialising a model
02     Setting up the initial values of parameters
03 End the initialisation

04 Start building the model
05     Building the Bar
06         Building history of attendance
07     Building strategies
08     Building agents
09 End building

10 Repeat for each time step
11     Start setting a step
12         Check the last attendance number from Bar's memory
13         Check all strategies if they are successful
14         Make sure no one is at the bar, yet
15     End setting a step

16     Start the step of agents
17         For each agent from population
18             Choose the best active strategy to predict
19             Make decision if he goes to the bar
20             Bar notice the agent arrives
21         End 'for' loop
22     End the step of agents

23     Start closing this step
24         the Bar add the attendance number in its memory
25     End closing this step
26 End repeat for each time step
```

Figure 2.1: Pseudo-code of simulation procedure in duplicated original El Farol. In the main simulation step of agents (lines 16-22), each agent first chooses the best locally available strategy from a set of active strategies (line 18), which are randomly assigned to each agents at the beginning of simulation (line 08) and fixed during the simulation process. Then, the agent decides if they go to the Bar El Farol or not based on their past experience (line 19). The attendance of the agent is informed to the Bar (line 20). At the end of each time step, the total number of agents arriving at the bar is reported to the main class of the model (i.e. Model), which produce the graph of the output (line 24).

- The number of active strategies that each agent possesses is 5
- The history length, which stores the number of past attendances, is set to 20
- The threshold level of crowdedness is 0.6, i.e. when the bar is filled by 60 agents, the bar does not offer the music.

The result of this duplicated model was the same as the original result and showed the emergent property Arthur revealed. The attendance number of agents to the bar oscillates randomly around the target attendance of 60 (Figure 2.2). There are two important messages from this result. The first one is that the system achieved a Nash equilibrium[4] at macro level:

> *"a mixed strategy of forecasting above 60 with probability 0.4 and below with probability 0.6, which would engender a mean attendance of 60 individuals, is a Nash equilibrium when the situation is viewed in terms of game theory."* (Fogel et al., 1999, p.143)

The second one is that although the system seems to have a Nash equilibrium at the target attendance of 60, this is not because of agents' co-optation (Challet and Zhang, 1997, p.3). Each agent acts individually and selfishly and tries to select the best strategy to win each game.

When an economist hears about El Farol problem, s/he may say that "*The optimal strategy agents choose is pseud-optimal. So, I will just stay home to maximise my utility.*" His/Her answer is correct if one uses the deductive reasoning as a conventional Nash's game theory. Each agent has the best strategy to predict the attendance number, but there is no real optimal or best strategy to win this game. For example, the random attendance strategy to the bar (Strategy 0 in Table 2.2) will be successful about 60% of the time, i.e. that

[4]A situation is a Nash equilibrium if no agent has incentive to deviate from its choice given the other players do not deviate (Rasmusen, 2001, p.26).

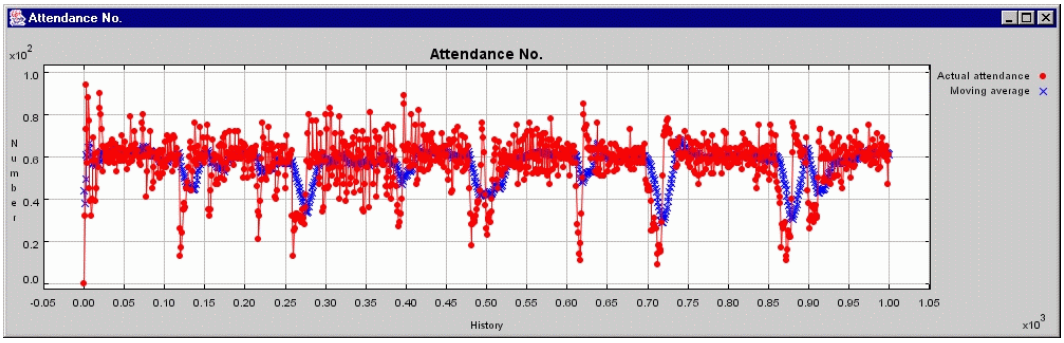


Figure 2.2: Attendance number of Agents

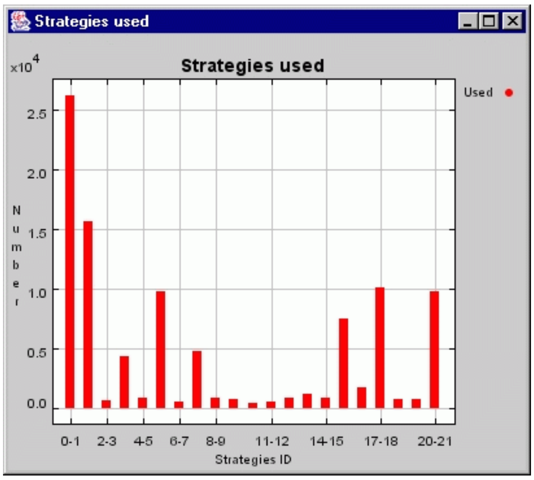


Figure 2.3: Strategies used by all agents through the simulation

Figure 2.4: Successful Strategies through the simulation

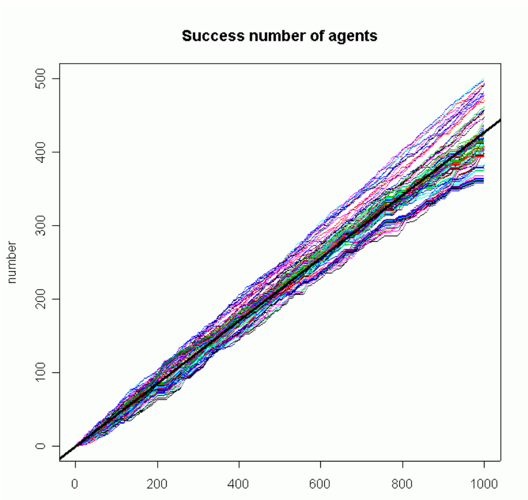


Figure 2.5: Success number of Agents

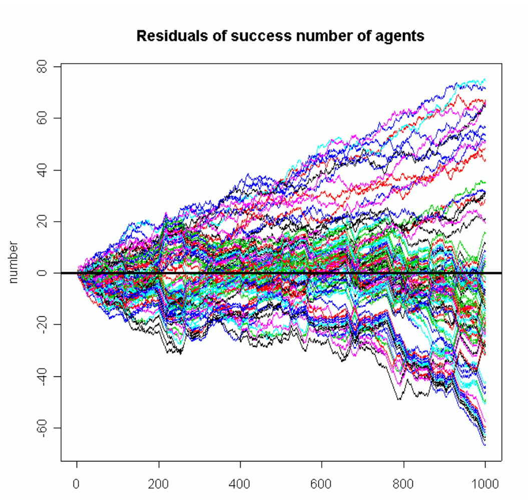


Figure 2.6: Residual of success number of Agents

is the probability when attendance is over 60 as mentioned in the quotation above. The rest of the strategies will also achieve similar success rates (Lofton, 2000, p.11). The result of this simulation process proves the pseud-optimality. Figure 2.4 shows that the numbers of strategies that successeded[5] at the end of a simulation process[6] are not varied as much as the numbers of strategies used by agents (Figure 2.3). The different patterns between the two figures are due to the bounded rationality and the problem of the local and global optimal, i.e. agents choose the best local strategy from their subset of the strategies, but this best strategy may not be the same as the best possible strategy in the simulation world. Because of this trial-and-error reasoning, the successful numbers of agents are not equally distributed (Figure 2.5). Some agents win the game more than some other agents. This movement is more clearly shown when we look at residuals from the trend (mean) of the successful numbers (Figure 2.6). The trend of residuals indicates that successful agents will be more likely to be successful. Similarly, unsuccessful agents will be more likely to be unsuccessful. In other words, the Minority Game including El Farol is symmetric as a Nash equilibrium shown at a macro level, but the games are not symmetric for agents (Savit et al., 1998), i.e. the distribution of wins is not equal between agents. The symmetrical nature of the Minority Game could be unclear in this original El Farol game since this game did not have any payoff function, and the threshold of the congestion level was 60%, but not 50% of the population. However, this symmetric nature remains in any Minority Game regardless of the threshold level (Challet and Zhang, 1997, p.4). In conclusion, the original El Farol was the beginning of the paradigm 'Minority Game', and the game showed how perfect rationality was not necessary to win this game.

[5]If the predicted choice (i.e. either go to the bar or not) results in the minority side, the strategy succusseded.

[6]Successful strategies are countered whether or not the strategies are used by an agent.

2.3.3 Other work after Arthur's original model (1994)

After the original work, there were several improvements in the Minority Game. In fact, the original model was not called 'the Minority Game' when Arthur started this study in fact it was Challet and Zhang (1997) who re-phrased this problem as the Minority Game. Having in mind a general outline of the application to stock markets, they shifted the threshold of binary choice to 50/50 and focused on the oscillation of the attendance number. In the finance sector, buyers can be sellers, simultaneously and the minority side is the winning side. If there are more sellers than buyers, the buyers can buy stocks at lower prices. Similarly, if there are more buyers than sellers, the sellers can sell the stocks at higher price. Moreover, they introduced the concept of memory, which was the length at which each agent remembered past experiences.

Following this, de Cara et al. (2000) studied the Minority Game with personalised memory rather than global information. Edmonds (1999a) and Bazzan et al. (2000) studied characteristics of agents in different approaches. Liu et al. (2004) introduced the horizon of strategy successfulness. The horizon is related with the adaptability of agents, since a long horizon makes agents consider too much historical information, which may not be relevant to the current situation. More works have been done in this field, but these are the ones that are most relevant to the Minority Game in this thesis. In the next section, the model of Bazzan et al. (2000) is duplicated. It is not because this study was the most important work after the original El Farol problem, but because this study considers the application to the transportation sector.

2.4 Minority Game with characteristics (Bazzan et al., 2000)

In the original El Farol model, distributing subset-strategies to each agent was purely random, so that the agent did not have a revealed characteristic. The Minority Game with characteristics is one step toward the phenomena observed in the real world. Bazzan et al. (2000) categorised strategies into higher-levelled rules and named them as characteristics or personalities of agents. Although the Minority Game in this thesis has a different approach to define the characteristics of agents, it was worth to replicate Bazzan et al.'s (2000) model to think about the practicability of his approach.

2.4.1 Features of the model

Most parts of classes in Table 2.1 remain the same for this duplicated model. The major change was that strategies were categorised into characteristics. In consequence, the difference between the model in the last section and the current model are as follows:

- The strategies of agents are broadly categorised into nine characteristics (Table 2.3).
- The characteristics of each agent are assigned at the beginning of a simulation process and do not change throughout the process.

Besides these features, the backdrop of the simulation was changed to the transportation sector. Their motivation to use the Minority Game was to forecast a binary route choice based on congestion on the roads. Therefore, 'Go' means 'going via the main route' and 'Not' means 'not going via the main route', i.e. 'going via an alternative route'.

The first feature means that some strategies are categorised into one of nine characteristics although other strategies represent their own characteristics. For example, in the case of two memories (i.e. agents remember the number of agents choosing main route last two times), there are four possible outcomes (two memories and two possible choices, Go / Not, $= 2^2$) and there are 16 possible agent's strategies (four outcomes and two agents' actions, Go / Not, $= 4^2$). Then, if a characteristic is $\mathsf{G}\vee$ in Table 2.3, the agent uses 15 strategies. So, an agent with the characteristic of $\mathsf{G}\vee$ chooses "Go" when the choice "Go" is unsuccessful in the last two periods. On the other hand, if its characteristic is N, the agent's strategy is only one, which is the one rejected by the characteristic of $\mathsf{G}\vee$. The number of strategies used by other characteristics is shown in Table 2.3.

Table 2.3: Description of the characteristics

ID	Description of characteristics	No. of strategies
G	Go regardless of the history	1
N	NOT go regardless of the history	1
$\mathsf{G}\vee$	Go if it has won at least one game	15
$\mathsf{G}\wedge$	Go if it has always won	8
$\mathsf{N}\vee$	NOT Go if it has won at least one game	14
$\mathsf{N}\wedge$	NOT Go if it has always won	13
P	choose the route that won in the previous game	1
W	choose the route that lost in the previous game	1
R	choose a route randomly	15

The second feature was the result of simplifying the problem and the background of the current model. Edmonds' model (1999a) tried to describe actual human behaviour by using communication and learning to observe the emergence of heterogeneous characteristics in the society. This approach was complex and needed large computational power, Edmonds' model could imitate this with solely ten agents. However, ten agents were not suitable in the context of Bazzan et al.'s model, i.e. congestion on road traffic. Moreover, ten agents are not large enough to reflect the situation of the Upper Derwent Valley.

2.4.2 Results and discussion of the Minority Game with characteristics

The important parameters set for this simulation are:

- The number of agents is 901,
- Agents remember the last two events, i.e. the memory length is 2.
- The threshold of the congestion level is 50%, so the game is now clearly symmetrical.

The parameters of active strategies disappear in this model since the numbers of strategies agents possess are different and it depends on their characteristics. Agents sizes are evenly distributed among nine characteristics[7] (i.e. $100 \times 9 + 1$).

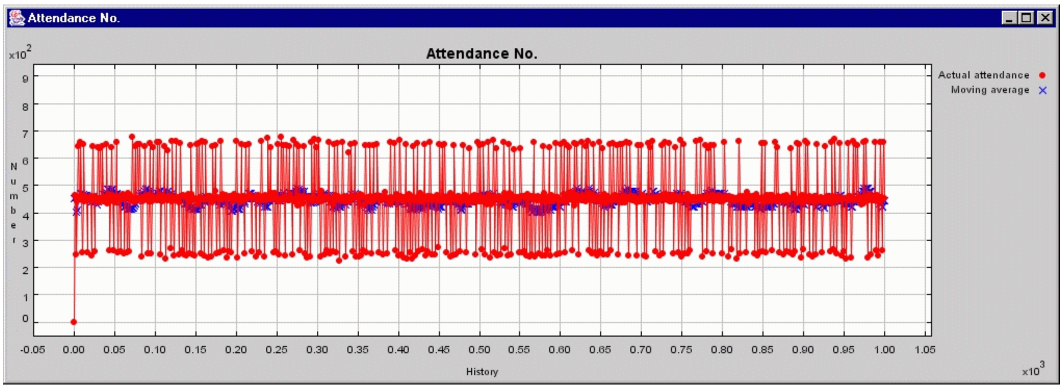


Figure 2.7: Number of Agents on the main route

The number of agents on the main route oscillates randomly around the threshold congestion level of 450 (Figure 2.7). This oscillation shows that this Minority Game also possesses a Nash equilibrium at a macro level like

[7]Only G has 101 to make odd number of agents

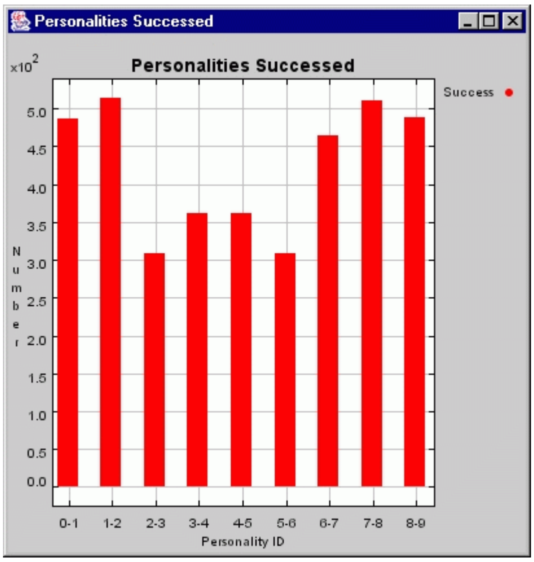


Figure 2.8: Successful characteristics through the simulation

the original El Farol model. However, the oscillation has a discovered pattern, which is due to the smaller memory size and the variation of strategies. Agents with the same characteristics react in the same manner since the decision making is based on conventional IF-THEN rules, but not fuzzy or stochastic approaches[8]. For example, if going via the main route is successful in the past two periods, all 100 agents with characteristics G∧ go via the main route this time. There are similar manners for other characteristics. Therefore, the trend in agent movements is the repetition around a few levels rather than a random walk[9]. Focusing on the successfulness of characteristics, some characteristics and consequently some agents are more successful than others (Figure 2.8); therefore, the game is not symmetrical among agents as expected from the results of the previous Minority Game. In conclusion, this duplicated Minority Game with characteristics showed the possibility of the Minority Game in transportation research although this research was only at a theoretical level.

[8]These approaches are explained in detail in Section 5.4.2 (Chapter 5).
[9]The random walk is a process produced by the sequence of countless discrete levels.

2.5 Minority Game in this thesis

As mentioned in the last chapter, the Minority Game in this thesis is the mode choice to visit the Upper Derwent Valley based on the congestion level in the parking areas. Since this Minority Game is used to analyse a real world problem, the setting of the game, such as characteristics and decision making processes, has to be realistic. The characteristics of agents were implemented in the last section; however, this approach is unlikely to be implemented at a practical level. Those travel-related characteristics (e.g. $\mathsf{G}\wedge$ etc.) are usually unobservable directly (Ben-Akiva et al., 2002). The more practical alternative approach is to use observable characteristics such as the age of the traveller as an indicator of the travel characteristics. These characteristics of visitors to the Upper Derwent Valley are revealed in Chapter 6 and integrated into the Minority Game by an agent-based simulation model in Chapter 8.

Similarly, the IF-THEN rules as a decision making mechanism and the strategy of agents have not been tested to justify their mechanisms and usage in the previous studies. For example, in Arthur's (1994) agent-based model, the behaviours of regular customers to the El Farol bar were defined by IF-THEN rules, but the rules were not supported by any real world observation. In this thesis, the Minority Game is not a conceptual model, so that the behaviour of agents and strategies have to be tested and reflected reality. The problem of the IF-THEN rules is altered by implementing the random utility theory and discrete choice analysis. The justification of the implementation is explained in Chapter 5. The rest of the chapters in the theoretical part discuss the possible and practical strategies for the Minority Game in the Upper Derwent Valley.

Chapter 3

Congestion as a game in the Upper Derwent Valley

3.1 Introduction

This chapter describes the situation of the Upper Derwent Valley from a game theoretic viewpoint as well as that of the Minority Game. The main purpose of this chapter is to find suitable strategies of the Minority Game, which reflect the situation of the Upper Derwent Valley. The complexity of the parking congestion in the Upper Derwent Valley is first discussed in terms of conventional game theories and then the discussion moves to the Minority Game. This transition shows how the conventional game theories are connected with the Minority Game. Additionally, the Minority Game is eventually combined with another economic concept namely discrete choice analysis.

Theoretical arguments, which do not reflect reality, are not the purposes of this thesis. However, some situations in this chapter are not supported by any observation. These non-supported situations are purely to introduce real complex situations and will not be used in the subsequent chapters. Moreover, this chapter starts using the terms, 'agents' and 'visitors to the Upper Derwent Valley', interchangeably. The term 'visitors' is still dominating, but they describe the same subjects.

3.2 Simplistic deterministic game

Transport mode choice has been studied for many decades (Baumol and Vinod, 1970; Winston, 1985; Hurley and Petersen, 1994); however, most studies have considered only the situation with a single agent, i.e. no interaction among transport users. In reality, competition among users exists. For instance, Haugen and Hervik (2004) mention an example of competitive game between an automobile transport and a boat transport for Norway's salted cod exporters (game players). The situation is a prisoner's dilemma[1] (Fudenberg and Tirole, 1991, p.9). As a road becomes crowded, the road cannot be fully utilised; therefore, the cost of road transport rises. More importantly, these two modes are interdependent of each other since choosing the boat transport will make other exporters choose the auto transport as the road becomes less congested. Similarly, choosing the auto transport will make other exporters choose the boat transport as the road becomes crowded. This situation is similar to the current situation in the Upper Derwent Valley, and its utility matrix could be like Table 3.1. As explained in Section 1.4.3, there are three possible options for visitors to the Upper Derwent Valley, namely Auto, Bus, and Cancel. To simplify the problem, the third choice, 'Cancelling the trip' is omitted at this stage. If both visitors (Table 3.1) choose Auto, the parking areas are more crowded so that the utilities of both visitors are equally small. If both visitors choose Bus, the visitors enjoy less traffic, noise, and emissions in the Upper

[1]Prisoners are two suspects charged with complicity relation, but are not strict prisoners. If one confesses and the other does not confess, the crime becomes heavy only for the latter. If both keep negating, the crime becomes light for both. If both are confessed, the crime becomes medium. Because the crime becomes heavy if the other confesses, both suspects think that the other party might have confessed mutually. Therefore, both suspects confess and do not end up choosing the optimal choice, 'keep negating'.

Table 3.1: Current utility matrix between Auto and Bus

		Visitor 2	
		Auto	Bus
Visitor 1	Auto	3 / 3	5 / 2
	Bus	2 / 5	4 / 4

Derwent Valley. In addition, the bus service as a public transportation service can be improved as more visitors choose the option due to the economies of scale (Mohring, 1972). Therefore, the summed utilities can be the highest when both visitors choose Bus. However, any visitor has a possibility of increasing his / her utility by changing the option from Bus to Auto in the situation of Bus & Bus. If not many visitors come by Auto, a visitor coming by Auto does not need to search for a parking space or to walk to the Information Centre as the visitor is likely to park at the Information Centre. On the other hand, in this case, there is a side effect on innocent bystanders coming to the Valley by bus, i.e. the Upper Derwent Valley is experiencing more traffic, which generates noise and emissions. Moreover, the bus service may not be as good as it used to be in terms of costs and operation frequencies due to the smaller economies of scale. Therefore, the utility of visitors by Auto is higher. In contrast, the utility of visitors by Bus is lower.

Therefore, the current situation in the Valley can be a game:

- Go by auto = defect
- Go by bus = cooperate
- non-zero-sum game

The Nash equilibrium in the current situation seems to be ‘All defect’ since regardless of what others choose, the equilibrium ends up to be Auto & Auto.

As Steiner and Bristow (2000) assumed, the size of a travel group is usually considered an attribute to an agent, i.e. trip leader owns this attribute. Therefore, if we do not consider car sharing, the situation cannot be cooperative since visitors to the Upper Derwent Valley are complete strangers to one another. More importantly, the problem in this situation is not to gain the highest social welfare, which can be Bus & Bus. This could be a reason why there is no practical bus service running through the Upper Derwent Valley today. The utility of Auto is high because of external costs, which are generated by the option, are not paid by the users of the option. Therefore, a policy tool is needed to consider the external costs into the utility of Auto and eventually alter the equilibrium point.

A road user charging scheme and a park & ride scheme will reduce the utility of Auto and increase the utility of Bus, so the Nash equilibrium could be altered as in Table 3.2. A toll fee reduces the utility of Auto and a frequent

Table 3.2: Utility matrix between Auto and Bus after implementing a road user charging

		Visitor 2: Auto	Visitor 2: Bus
Visitor 1	Auto	1, 1	2, 3
	Bus	3, 2	4, 4

bus service improves the utility of Bus. Therefore, no matter what other visitor chooses, the equilibrium will end up as Bus & Bus, which is known as a dominant strategy[2] (Fudenberg and Tirole, 1991, pp6–11).

This section showed how the current situation and the situation after implementing schemes were expressed as a game. However, this is an optimistic

[2]A dominated strategy is a strategy, that yields a higher payoff than the other strategies regardless of the opponents' strategies.

way to view the game. In reality, the situation of the Upper Derwent Valley cannot be expressed through this simple game since the situation is more like a mixed game, which is discussed in Section 3.4. Moreover, this game in the parking congestion of the Upper Derwent Valley involves more than two visitors, so that it is an N-person game, not a two-person game. Even with agent-based modelling, a real world N-person game is difficult since it is costly to achieve the behavioural data about agents against specific opponents. Moreover, the N-person game is likely to be infeasible from a computational viewpoint since it eventually leads to the big-O or NP-hard problem[3] as an agent size increases (Karp, 1972; Miller and Lung Shaw, 2002, pp.200-201). Also, although the conventional game theory has the concept of mistake-and-imperfect-information condition as 'Trembling-hand perfection[4]' (Selten, 1975; van Damme, 1983), this concept cannot be used in the N-person game (Mas-Colell et al., 1995, p.259). Consequently, previous studies in the agent-based modelling of an N-person game are not conducted at a practical level (Epstein and Axtell, 1996; Cohen et al., 1999; Gilbert and Troitzsch, 1999b; Sunitiyoso and Matsumoto, 2005). This is also a serious problem in the case of the Upper Derwent Valley since perfect information cannot be assumed in the Valley (Thompson and Richardson, 1998; Klügl and Bazzan, 2004, p.162), and the situation is an N-person game. Therefore, it is sensible to give agents inductive reasoning to play the game in the Upper Derwent Valley.

[3] A problem which is both NP (verifiable in nondeterministic polynomial time that requires no more than a polynomial function of the problem size) and NP-hard (Weisstein, 2003). A problem is NP-hard if an algorithm for solving it can be translated into one for solving any other NP-problem (nondeterministic polynomial time). NP-hard therefore means 'at least as hard as any NP-problem', although it might, in fact, be harder.

[4] The notion of trembling hand perfection assumes that an agent, who wants to select a choice, might select another by the slip of a hand. Therefore, agents could make minor mistakes (tremble) and the mistakes lead to unexpected events. In a trembling hand perfect equilibrium, an agent's action is not only based on equilibrium beliefs but also based on perturbed beliefs. Hence, the trembling hand perfection excludes strategies that are 'unsafe' given the risk of slight mistakes.

3.3 Introducing the concepts of the Minority Game

It is necessary to add the idea of the Minority Game into the conventional game-theoretic model in order to analyse real world problems more realistically, as in the case of the Upper Derwent Valley. From a different perspective, the N-person game in the Upper Derwent Valley is ‘indirect’, but not an one-to-one game. This means that the game is similar to that of the El Farol bar or the Minority Game - i.e. opponents will be an unspecified group of visitors coming to the Upper Derwent Valley.

In the situation of the Upper Derwent Valley, visitors should not come to the Valley by car when they cannot park at their target parking areas. Also, visitors arriving by bus will be glad that they chose the Bus option if they find out there are no empty spaces in the parking areas, where the buses pass and stop. These two situations indicate that the visitors are indirectly playing the Minority Game.

For example, in the all-defect-situation, both visitors in Table 3.1 choose Auto and this situation indirectly implies that Visitor 1 chooses Auto when the car park is congested and *vice versa*. Similarly, in the all-cooperate-situation, Visitor 1 chooses Bus when a number of other visitors go by bus and so the parking areas are not congested.

The viewpoint of the game in the Upper Derwent Valley is changed from a one-to-one competition to an implicit competition after introducing the Minority Game. However, this integration is not completed since there are many ambiguous settings in the utility matrixes above. Hence, the utility matrixes are stopped being used from here since these real utilities are not known and visitors are competing only indirectly, but not on the one-to-one basis. The concept of defect and cooperate strategies are not used in this situation since visitors to the Valley are playing the Minority Game without the concept of co-operation as explained in Section 2.3.2. By exchanging the concepts of game theories and the Minority Game, visitors in the Upper Derwent Valley are equipped with the strategies of the Minority Game.

The strategies of visitors use the information that the visitors experienced. The visitors to the Valley experience toll, bus fare, and parking fees as well as the searching time to find a parking space and walking time to the Information Centre. It is difficult to measure congestion levels directly and it is more robust to measure time, which is a congestion indicator (Fortin and Rousseau, 1998; Bugmann and Coventry, 2004). Searching time and walking time can be used as indicators to measure congestion in the transport sector since time is a standard unit to analyse transportation problems, e.g. 'value of time[5]' (Fowkes, 2000).

This ability of guessing and indicative reasoning are implemented as thought patterns and consequently as strategies, which is explained further in the following chapter. The choice of memories, which store experience to calculate the current and local best choice for a visitor, can be considered as a part of strategies. Therefore, the strategies in the Minority Game in the Upper Derwent Valley are the combination of the thought patterns and the choice of memories. Before developing the Minority Game of the Upper Derwent Valley, the Minority Game is compared with more sophisticated game theoretic concepts in the next section.

[5] This is the ratio of weights on time and money (Mackie et al., 2003, §1.2)

3.4 Comparing the Minority Game with more sophisticated game theories

The situation at the Upper Derwent Valley can be explained by mixed strategies[6] of game theories since the situation has an equilibrium point where all visitors should decide to go according to a random distribution, which is weighted towards going a fixed proportion of the time at a theoretical level (Edmonds, 1999b). However, the agent actions never converge to such equilibrium in the Minority Game simulation due to the dynamics of the problem and imperfect information, which is explained in Section 8.3.1. The same phenomena are observed in everyday life, e.g. the stock market is always fluctuating and does not converge toward an equilibrium. Therefore, the focus of the Minority Game is generally more on fluctuation or variability rather than on equilibrium, so that this problem should not be confused with the problems that traditional game-theoretic mixed strategy focuses on (Challet et al., 2004).

The Minority Game is more applicable to the situation of the Upper Derwent Valley than is the game-theoretic mixed strategy. The traffic demand in the Valley has seasonality so the situation is dynamic. Visitors are not acquainted with one another, so perfect information cannot be assumed. Also, this study is interested in the variation of user utility and mode choice amongst the visitors. This is the reason the Minority Game rather than game-theoretic mixed strategy, is implemented into this simulation model.

Moreover, the common knowledge[7] of rationality (CKR) (Aumann, 1995)

---

[6] A strategy consisting of possible choices with a probability distribution (collection of weights). This corresponds to how frequently each choice is selected (Rasmusen, 2001, pp.66–81).

[7] Two people, 1 and 2, are said to have common knowledge of an event E if both know it, 1 knows that 2 knows it, 2 knows that 1 knows it, 1 knows that 2 knows that 1 knows it, and so on. THEOREM. If two people have the same priors, and their posteriors for an event A are common knowledge, then these posteriors are equal (Aumann, 1976, p.1236).

cannot be assumed in the Minority Game situation, since if every visitor has CKR, the result would be the same as playing the game with uncertainty. To win the Minority Game, a visitor has to know which option the majority of other players choose. If the visitor obtains this information by using CKR, the visitor wins the game. However, why can the other visitors not obtain the same information by using their CKR? In this case, this logic fails. Having said that, the different degrees of CKR among visitors will be an interesting topic for future research.

In terms of data collection, the Minority Game is more straightforward than game theories in practice since it is not necessary to collect the behaviours on one-to-one interaction. Instead of the one-to-one behavioural data, the frequency of bus and the searching time and walking time are asked in the stated preference questions. Therefore, this approach uses the behavioural data of agents against an unspecified group of visitors. The utility for each agent is calculated on each simulation time step while a new set of the behavioural data is generated. Additionally, each agent has a different set of data since the agents' utility function is based on belief and experience.

3.5 Re-combination of economic concept from multinomial discrete choice model

The sections above show how the game theoretic situation in the Upper Derwent Valley can be described in terms of the Minority Game to simplify the problem. However, the decision making mechanism of visitors has not been cleared and it is difficult to achieve this mechanism by using the conventional Minority Game. The conventional Minority Game uses strategies to make agents select their choices as described in the previous chapter. This simple approach works at a conceptual level, but not at a practical level. The strategies of visitors could be numerous in the Upper Derwent Valley situation, and there is no well-defined methodology to estimate the strategies of visitors even if we find out small feasible choice sets for visitors (Manski, 1977). For example, the IF-THEN rule of a strategy may be as follows: choose Auto if the searching time was 10 minutes and toll fee was 1.5 pounds in the most recent trip. In this case, how can we justify this IF-THEN rule? Fortunately, econometrics has established an alternative approach to the IF-THEN rule, namely stated preference analysis surveys and multinomial discrete choice analysis. The stated preference survey can be expressed in response to hypothetical scenarios of non-existing alternatives (Fowkes, 2000; Ortúzar and Willumsen, 2001), so that this is a suitable approach to analyse the road user charging and park & ride schemes, which have not been implemented in the Upper Derwent Valley. The stated preference analysis is explained further in Chapter 6.

The biggest transformation from the conventional Minority Game approach to this approach is the conversion of a deterministic form to a probabilistic form. So, this new Minority Game can be termed identified as stochastic Minority Game in this thesis. For example, the probability of an agent going to the Upper Derwent Valley by car (i.e. Auto) is:

$$P(\text{Auto}) = \frac{\exp(U^{\text{Auto}})}{\exp(U^{\text{Auto}}) + \exp(U^{\text{Bus}}) + \exp(U^{\text{Cancel}})} \tag{3.1}$$

The utility function of U^{Auto} in equation (3.1) is expected to be partially composed of searching time for a parking space and walking time to the destination area. These time factors are expected to have negative effects on travel utilities in equilibrium models such as multinomial discrete choice models (Hess and Polak, 2004), but these negative effects may not be correct in dynamic models.

According to the Minority Game, the two thought patterns are deductively considered to deal with the negative effects from the time factors. For example, if the parking areas are severely congested at the time of travel, the searching time and walking time tend to be long for the visitors by car. Therefore, these visitors may think: 1) the parking area will be congested so I will not go to the Valley by car next time, or 2) many visitors will be discouraged to come to the Valley by car and then parking areas will be empty so I will go to the Valley by car next time. Therefore, searching time and walking time can affect the utility of Auto negatively as well as positively. One advantage of the discrete choice model is that values of utilities and probabilities are relative and have no meaning in absolute terms (Ben-Akiva and Lerman, 1985). This means that the utilities generated by searching time and walking time can be added to alternatives to express the thought patterns of the visitors. From the description above, three thought patterns were considered for this simulation. These three thought patterns of visitors were dependent on the mode that takes the congestion related utility:

Thought pattern 1: believes that the parking area will be congested again next time, so this discourages a visitor from going to the Valley by car;

Thought pattern 2: believes that the parking area will be less congested next time, so this discourages a visitor from going to the Valley by bus;

Thought pattern 3: believes that the parking area will be less congested next time, so this discourages a visitor from cancelling the trip.

The explanation of these thought patterns above is just a possible explanation, but another explanation may be considered. In other words, since these utility functions are not as explicit as those of the IF-THEN rules in the conventional Minority Game, the transfer of the utility and utility functions themselves capture many rules of the Minority Game, i.e. benefit from the simplicity. Another benefit from implementing discrete choice analysis is adapting statistical and calibration techniques into the decision making of agents. The IF-THEN rule approach can be more realistic than discrete choice analysis to model human decision making since it is unlikely that we throw a die to make a decision all the time. Nevertheless, it is infeasible to use the rule-based approach in the real world. Rule-based behaviours are not observable in reality. Moreover, even if the rule-based behaviours are estimated, there is no established validation method to justify this approach. In contrast, the validation and calculation methods are robustly developed in discrete choice analysis. This issue is further discussed in Chapter 5.

3.6 Conclusion

This chapter showed how the situation in the Upper Derwent Valley could be explained as a game. The explanation of the game started from game theoretic concepts and transferred into the Minority-Game concept. This transformation enables one to analyse the game in the Upper Derwent Valley at a practical level. Moreover, implementing discrete choice analysis into the new stochastic Minority Game improves the practicality of the analysis due to its validation method and the benefit from the probabilistic approach.

Although the thought patterns of agents in the Upper Derwent Valley were approximated by deductive reasoning, the strategies of the agents have not been understood yet. It is because the choices of memories, which are the other key components of strategies[8], were not justified in this chapter. Therefore, the next chapter discusses the distribution of memories in reality and tests the practicality of the stochastic Minority Game with an online game.

[8]The strategies for the Minority Game in the Upper Derwent Valley were defined as the combination of thought patterns and the choices of memories on page 56.

Chapter 4

Online Minority Game and an conceptual stochastic agent-based modelling

4.1 Introduction

This chapter has two stages: 1) collecting real data about the distribution of memories through a web-based online game and 2) duplicating the data by using an agent-based simulation model. The main motivation for this chapter is to collect the distribution of strategies or more specifically the distribution of memories in the real world Minority Game. In the original Minority Game or El Farol Problem, allocating strategies is completely random (Arthur, 1994), as discussed in the last two chapters. Previous studies have modified some parts of the original game, but no study has verified the allocation of strategies in the real world (Challet and Zhang, 1997; Edmonds, 1999b; Bazzan et al., 2000). Therefore, it is essential to find out the real allocation of agent strategies before this concept is implemented in real world problems. Verifying simulation by comparing data and outcome is a common practice (Parunak et al., 1998;

Lofton, 2000) and it is an especially important process in agent-based modelling (Moss, 2002, p.7273).

4.2 Online game as data collection methodology

The data collection was conducted through my web site (`http://users.ox.ac.uk/cgi-bin/safeperl/scat1898/mg.cgi`) (Figure 4.1). The web site was written in a restricted subset of the Perl language called Safe Perl (Oxford University Computing Services, 2005). Since many useful functions were restricted in the subset language, the web site had to be simple. However, the web site dynamically changes according to the responses of game players, so that the web site is like an online role-playing game. It was not the aim of this thesis, but the web site is now used as an example in the interactive web design course at the Oxford University Computing Services.

On the web page mentioned above, the game is described as below:

> To win a game you have to choose a less congested side, i.e. either Side A or B, up to 10 times. You have to pay one point to play each game against other 99 artificial computer players. These artificial players are assumed to optimise the result of the last five games to estimate the less congested side. You can also see the result of the last five games to help your decision making. Then, the summed points on one side are shared by the players on the other side. Therefore, if you choose the less congested side you will gain otherwise you lose some points. Also, the score you gain or lose is determined by the level of congestion: if more players are on the other side and fewer players are on your side, you will get a bigger

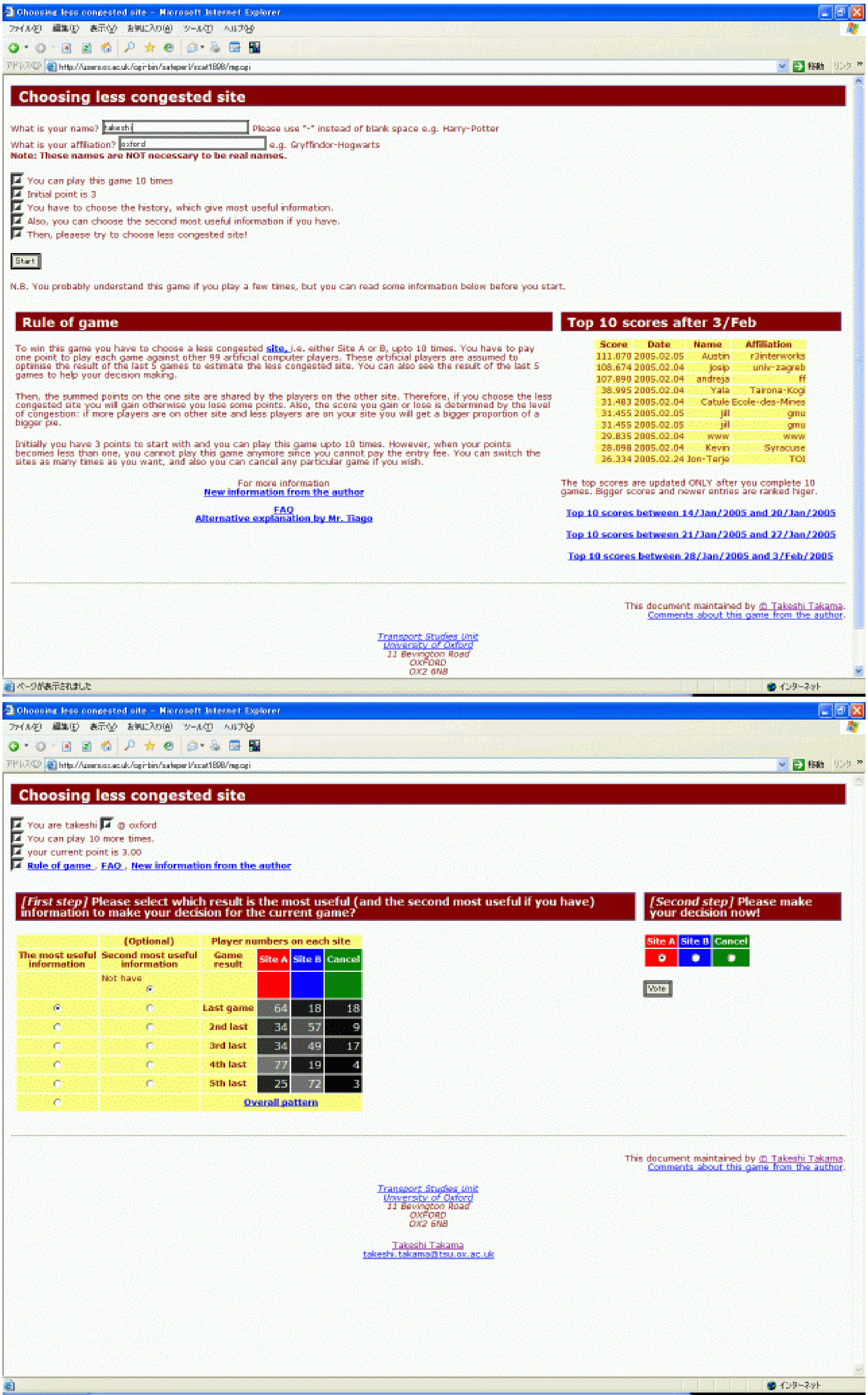


Figure 4.1: Online Minority Game

proportion of a bigger pie.

Besides the explanation on the web site, the dynamic feedbacks were given by me through email and my other web site (`http://www.geog.ox.ac.uk/~ttakama/blog/`).

As stated above, the main purpose of this online Minority Game is to find the distribution of strategies in the Minority Game, so that this is a web-based online questionnaire about human behaviour rather than a real online game. Therefore, the number of artificial agents in each side was generated by a pre-determined distribution before this online game was launched. First, the number of Cancel was generated from lognormal distribution with the mean and a standard deviation of 1.5 and 0.8 on the log scale, respectively. Then, the numbers of Side A and Side B were generated from a truncated normal distribution with the mean of (100 – Cancel) / 2 and a standard deviation of 20. If neither total size is 100 nor individual size is positive, the calculation was rejected.

The web site was launched on the 14th of January 2005. The web game may still be playable on the web site, but this analysis is based on the data collected between the 14th of January and 25th of February 2005. To publicise this game, the email lists of four research and academic organisations were used. They were the School of Geography and the Environment at the University of Oxford on the 14th January 2005, St Catherine's College of the University of Oxford on the 21st January 2005, the University Transport Studies Group on the 28th January 2005, and the Society of Computational Economics on the 4th February 2005. There is possibly sampling bias, which corresponds to these email lists. However, according to the personal response and domain name, people played this games from all over the world and these players possibly came via search engines with relevant keywords and my personal weblog (diary) web site.

Players were asked to select which previous result was the most useful (and the second most useful if applicable) information to make their decisions for a game they were facing at that time. Initially, the choice of past results (memories) is from the most recent game to the fifth recent game. Then, an overall pattern, which was defined as "a trend, average, or oscillation of player numbers etc., but not a specific memory", was added as an extra choice from the 4th February 2005[1].

During the specified period, 3,886 games were recorded. One set of games could be played as many as ten times, if the set was not terminated according to the rule. A total of 499 sets of games were recorded and 304 different individuals played the games according to the names of affiliations and players. This means that 40% of the games are played by returned participants. In addition, 314 out of 499 sets are complete sets of games (i.e. a set contains 10 games).

4.3 Justification of web-based questionnaire

A major concern about web-based online questionnaires is the quality of data. Some researchers may think that people (in this case, players of this game) do not think, but just select the choice randomly. The response time can be a good measurement as thinking time between questions (or games in this case). The response time is equivalent to lognormal distribution with the mean of 2.55 seconds and a standard deviation of 0.685 seconds on the log scale, respectively (Figure 4.2). The lognormal distribution is typical for such data (VanBreukelen, 1995; Ciuhandu and Murphy, 2002). More importantly, the response time shortens as players play more games probably because players get used to this game with time (Figure 4.3). Therefore, there are some signs

[1] `http://www.geog.ox.ac.uk/~ttakama/blog/index.php?entry=entry050204-112028`

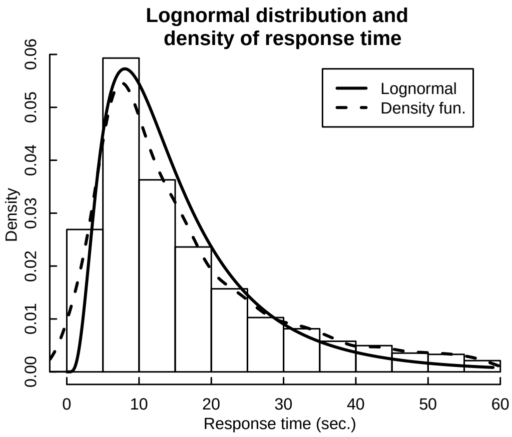


Figure 4.2: Distribution of response time in seconds with lognormal distribution function

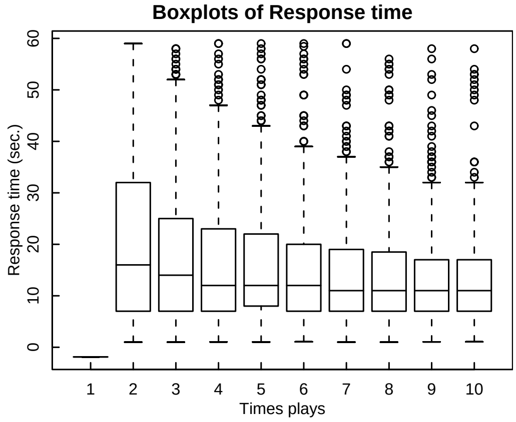


Figure 4.3: Boxplots of response times by the number of plays. The response times are seconds that players spent from the previous games to the games indicated on the x-axis.

suggesting that players are using their brains rather than randomly selecting their choices to play this game. In addition, web questionnaires are becoming acceptable as a research tool today (Rourke and Anderson, 2002), and people are more amenable to web questionnaires as they have experienced many web technologies already (Anderson and Kanuka, 2002).

4.4 Distribution of strategies

In this section, the distribution of strategies is examined. Only the past results were given to players and they were asked to use the information to win the games. Therefore, there were two components in the players' strategies, 1) choosing a referring memory and 2) choosing a thought pattern, which is either following the information of the referring memory or not. That is, 'Following the information' means that a player chooses the minority side of the selected memory. 'Not following the information' means that a player does not choose the minority side of the selected memory. The idea is the same as the thought patterns introduced in the last chapter on page 60. For example, a player chooses the memory of the most recent game as a reference and the minority side of the most recent game is 'Side A', so then the player decides to choose 'Side A'. In this case, the thought pattern is following a reference, i.e. choosing the minority side of the chosen memory[2]. The other thought pattern is the same process except that your decision is 'Side B'.

First, the equal distribution of memories among the different round of 10 games between 1 and 10 is tested with Chi Square (χ^2) statistics. The p-value of 0.521 does not reject the null hypothesis: distributed memories are the same for every round of 10 games. Therefore, players do not change the way of using memories with time.

Second, the equal distribution of memories with and without the overall pattern as an extra memory option is tested. The p-value below 0.01 from Chi Square statistics rejects the null hypothesis: the distribution of memories is the same for the existence of the extra choice. This means that at least one of the memory distributions is different. Figure 4.4 shows the relative frequencies of memories with and without the overall pattern[3]. This figure shows that the most recent memory is selected relatively more with the choice of the overall pattern. The trends are opposite to the rest of the memories used. That is, the change in the memory distribution is not proportionally equal. The difference may cause a significant problem to allocate the distribution of memories and consequently to allocate that of strategies. However, it is difficult to identify the real strategies of 'the overall pattern'. The pattern can be considered as moving average, minority times between two sides, weighted

[2]Unless someone has an emotional attachment to the letter 'A' or 'B'.

[3]The relative frequencies with the overall pattern were estimated as follows. The choices of the overall pattern, which were added from 4th February 2005 (See page 65), were excluded first. Then, the number of each memory chosen was divided by the sum of the un-excluded memories. Therefore, the relative frequencies with the overall pattern were comparable to the relative frequencies without the overall pattern.

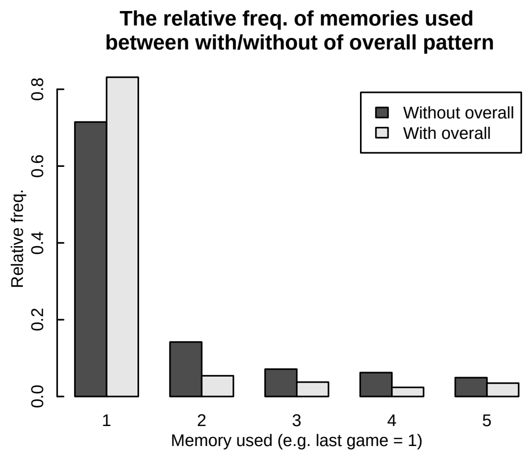


Figure 4.4: Relative frequencies of memories used between with and without the overall patterns

average, oscillating numbers, etc.. Therefore, sub-categorising the choice of the overall pattern may change the distribution of memories again. Eventually, this is the trade off between the quality and quantity of survey results. As mentioned on page 5, this whole project is regarded as experimental research for future research. Therefore, this problem should be seriously considered, but this thesis leaves this issue to future research.

Third, the equal distribution of memories and the thought patterns[4] are tested. The p-value of 0.2831 from Chi Square statistics does not reject the null hypothesis: the distribution of memories is the same if choosing the minority side of a selected memory or not. Therefore, players do not change the choice of memories regardless of the thought patterns. Also, the figure shows that the distribution of the thought patterns is 50/50. Figure 4.5 shows that the difference between the two thought patterns is nominal and statistically insignificant.

In conclusion, the distributions of memories selected were not equally distributed. Most players selected the most recent game result as a reference to play the current game (Table 4.1). Also, older memories are less frequently

[4] the choice of the minority side

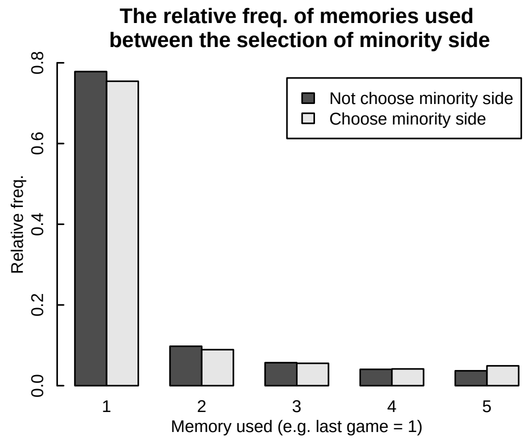


Figure 4.5: Relative frequencies of memories used between choosing or not choosing the minority side of a selected memory

selected by the players and this is consistent with the general time discounting[5] in economic theory (Samuelson, 1937; Frederick et al., 2002) and the Bayesian approach of statistical modelling[6] (Pole et al., 1999). Moreover, the distribution did not change regardless of the thought patterns throughout the time. Therefore, the one-dimensional distribution of memories, which is presented in Table 4.1, is discussed in the rest of the sections.

	recent	2nd recent	3rd recent	4th recent	5th recent
Freq	0.77	0.09	0.06	0.04	0.04

Table 4.1: Distribution of memories

[5]The older the information, the less the information is relevant.

[6]The Bayesian approach is dynamic modelling with predictor parameters (τ_t). The parameters are distinct but stochastically related through a system equation such as $\tau_t = G_t \tau_{t-1} + \omega_t$, where G_t is a matrix of known coefficients and ω_t is an unobservable stochastic term.

4.5 Theoretical discussion for the indifference between two thought patterns

The distributions between the two thought patterns, 'choose' or 'not choose' the minority side, in Figure 4.5 are indifferent and this section discusses the reason of the indifference further from game theoretic viewpoints. This indifference between two thought patterns can be explained by a famous example in game theory such as 'Meeting in New York game' (Schelling, 1960, pp.54–56)(Roy, 2002, 64–65). Mr. Schelling and Mr. Thomas know that they must

Table 4.2: Payoff matrix of meeting in New York game

		Mr. Schelling	
		Empire State	Grand Central
Mr. Thomas	Empire State	10, 10	–1, –1
	Grand Central	–1, –1	10, 10

meet each other in New York City on a specific day at noon, but do not know if they can meet at the top of the Empire State Building or at the clock in Grand Central Station (Table 4.2). Then, each of them has to choose which location they decide to meet the other at but they have to show up at the location at the same time (noon). If their selected locations are the same, they can meet, so they are happy, i.e. both win the game. If the locations are different, they cannot meet so they are disappointed, i.e. both lose the game. In this game, there is no pure strategic equilibrium or strict dominant strategy. However, each player chooses Empire State or Grand Central at the equal probability with a mixed strategy. When one of the players randomises his decision, it makes the opponent indifferent between playing with Empire State and Grand Central, or *vice versa*.

This concept is perfectly applicable to the situation of this online Minority Game. For example, Mr. Thomas is any player of the web game and Mr. Schelling is the result shown of the current game. The utility for 'Current result' does not make sense; therefore, it is omitted from Table 4.3. If a

Table 4.3: Payoff matrix of the online Minority Game. 'Choose' means choosing the minority side of a selected memory. Similarly, 'Not choose' means not choosing the minority side of a selected memory. Since the results of the games are pre-determined, the strategies of 'Current result' can be viewed as that in which the current game uses the thought pattern, as either choosing or not choosing the minority side of a selected memory, to determine the minority side of its result.

		Current result: Not choose Minority side	Current result: Choose Minority side
A player	Not choose	10	–1
	Choose	–1	10

player uses the thought pattern, which chooses the minority side of a selected memory, and the minority side of the current game is also the minority side of the selected memory[7], the player gains some points and wins the game, i.e. the top left corner of Table 4.3. The situation is the same for 'not choosing the minority side'. If a player uses the thought pattern, which does not choose the minority side of a selected memory, and the minority side of the current game is not the minority side of the selected memory, the player also gains some points and wins the game, i.e. the bottom right corner of Table 4.3.

On the other hand, if the selected thought pattern of a player is different from the thought pattern of current results, the player fails to choose the minority side of the current result; therefore, the player loses some point, i.e. the top right and bottom left corners of Table 4.3. In fact, all past results, which players see, were pre-determined in a randomised manner, as explained

[7]This can be viewed as that in which the current game uses the thought pattern, which chooses the minority side of a selected memory, to determine the minority side of its result.

on page 65, but the randomised mechanism was not shown to the players. So, the current result seems to randomise its decision and it makes the player of this game indifferent to choosing or not choosing the minority side of a selected memory.

Moreover, this is even true in the real Minority Game, introduced in Chapter 2. The movement of the Minority Game is indistinguishable from the random walk (Hughes, 1995) and the strategies of agents are also indistinguishable (Edmonds, 1999b) at a macro level, although different players have different experiences and strategies. In other words, the decisions of agents seem to be random moves, but it is a chain of thoughtful decisions at a micro level. The decision of an agent seems random to other agents and so the decision of these agents also seems random to some other agents since the decision of the second group of agents is based on the 'random' decision of the first agent. Moreover, the decision of the first agent seems to be a randomised decision, but in fact, it is not a randomised decision in the Minority Game (Mas-Colell et al., 1995, p.252).

For example, in the 1994 Football World Cup final Italy against Brazil, the Brazilian goal keeper, Cláudio André Taffarel, had to consider diving either the left or right when the penalty taker of Italian team, Roberto Baggio, aimed at the net[8]. Taffarel must have tried to read Baggio's mind and thought well, but the Baggio's decision to kick to the left or right might have seemed to be a randomised decision to Taffarel. Then, Taffarel dived to the left while Baggio kicked the ball into the sky. Unfortunately, the game was not made this time. In any case, the whole process does not seem to be different from the series of random decision making from the outside (at least for amateur observers like me). However, the randomness of Baggio's decision was just a false front and,

[8]There is actually another choice such as 'not dive'. However, this does not seem to be a choice for Taffarel according to Baggio (2002) (Read the next footnote).

in fact, he had some ideas about the kick[9] (Baggio, 2002) so that Taffarel's left dive was not a randomised choice, either.

This is the reason why the Minority Game seems like a random movement from a macro level, although players or agents play the game while they are using some algorithms to try to win the game. The next section shows that the process of agent decision making and random 'like' movements is simultaneously at a system level in a simulation.

4.6 Comparison between simulation results

In this section, the result of the web Minority Game is duplicated to verify the mechanism of the Minority Game. As a verification process, the situation of this simulation was set exactly the same. Only one agent plays this game at a time against the predetermined results like the online Minority Game. A simulation runs 10 games, and 499 sets of games are played, i.e. a total of 4990 games.

First, the strategies of real players have to be confirmed before modelling agent behaviour. Out of 3,886 games, Side A and Side B are selected 1,859 and 1,883 times, respectively. This means that only 144 games or 3.7% of total games are cancelled. In this web game, you can gain up to 99 points, but you may lose only up to 0.99. One maximum win is equivalent to 100 maximum losses, so that the skewed distribution to the big maximum win may motivate players not to cancel a game. Additionally, players are most unlikely to lose anything in their actual life by losing a game in the web game. Therefore, there can be non-commitment bias in this online game (Hovland and Sherif, 1952; Polydoropoulou et al., 1997). This web game seems to be a binary

[9] From his experience, Baggio knew (believed) that Taffarel always dived. So, Baggio kicked the ball into the middle, but he missed the kick. Taffarel possibly had a similar decision making process from his experience.

choice of Side A and Side B rather than a multinomial choice. Therefore, the simulation does not consider Cancel as a choice of the agent, but treats it as a choice randomly assigned with the probability of 0.037. This means that the set of thought patterns remains the same, i.e. choosing or not choosing the minority side of the selected memory. Then, there are five choices of memories so that there are 10 sets of strategies.

The results in the web game show that each individual has multiple strategies to play with, so that the agent in this simulation also has multiple strategies. It is difficult to determine how many strategies a real human has, so the threshold of five is used and this is based on the traditional human categorical mechanism – 'people are unlikely to remember more than several categories' (Miller, 1956). Therefore, up to five strategies out of 10 are assigned to the agent at the beginning of a simulation according to the calibrated distribution of memories (Table 4.4). The calibration was necessary since the distribution of the most recent memory was under- distributed and the rest of the memories are over-distributed with the original distribution. The distribution of the strategies is the same as the original distribution of memories when the calibrated distribution of memories is used to assign strategies to the agent. This means that the agent may decide its choice based on the most recent game result, but may use the oldest result in another occasion.

	recent	2nd recent	3rd recent	4th recent	5th recent
Freq	0.83	0.065	0.045	0.03	0.03

Table 4.4: Distribution of memories after a calibration process

We start looking at the mechanism of the strategy more closely in the following sections.

4.6.1 Strategies

The agent-based simulation model of the web game starts from creating strategies since strategies rather than the agent make a choice as explained in Chapter 2, i.e. the agent chooses the best strategy and follows the choice the strategy made.

The relationship of the two thought patterns can be expressed more formally in mathematical formulae. In this Minority Game, each agent has three choices, namely Side **A**, Side **B**, and **C**ancel. Side A and Side B have exactly the same mechanism to calculate their utilities. The study uses decision theory (i.e. probability theory and utility theory) as a decision making mechanism. The utility functions of both choices are based on the rule of gaining points, which is mentioned on page 63 in this chapter and on the web page, i.e. after you pay 1 point, the summed points on one site are shared by the players on the other site. The utility functions for the two choices are:

$$U_t^A = \mathrm{Fee}(N_{t-1}^B/N_{t-1}^A) - \mathrm{Fee} \tag{4.1}$$

$$U_t^B = \mathrm{Fee}(N_{t-1}^A/N_{t-1}^B) - \mathrm{Fee} \tag{4.2}$$

N^x is the number of players on Side x and t is the current game. Equations (4.1) and (4.2) mean that the entry fees (point) paid by an agent on one side are equally shared by the agents on the other side, 'benefit'. Then, these agents pay the fee on every game, 'cost', so the overall utilities are the benefit minus the cost. Having said that, there is only one agent in this simulation and the rest of the agents are just predetermined numbers. The predetermined results are set to make sure that neither side gets zero; therefore, the problem of 'division by zero' will not occur. Hence, these equations express the rule of game.

The decision is made based on the relative proportion of logarithm transformation of these utilities in a strategy, which is the most popular formula in discrete choice analysis (Ben-Akiva and Lerman, 1985; Greene, 2003). The two thought patterns affect the probability of each choice. For example, the probability of an agent choosing Side A with the thought pattern of choosing the minority side is:

$$P(\text{Side A}) = \frac{\exp(U_t^A)}{\exp(U_t^A) + \exp(U_t^B)} \tag{4.3}$$

Since the probability to choose 'Side B' is a mirror image of this equation, this thought pattern is likely to choose the minority side of the chosen memory. However, it is important to explain that this thought pattern is likely to choose the choice with the highest utility from the viewpoint of thought pattern of choosing the minority side. Similarly, the thought pattern of not choosing the minority side selects a choice by swapping the higher utility and the lower utility since the minority side has the higher utility in the utility function. In other words, when the choice with the higher utility is selected, an agent with the thought pattern of not choosing the minority side, in reality, chooses the choice of the lower utility. Therefore, the explanation of thought patterns can be re-written as below from the viewpoint of Equation (4.3): thought pattern 1 is likely to select the choice with the higher utility and thought pattern 2 is likely to select the choice with the lower utility. The thought pattern 2 is the second thought from the result of the binary discrete choice. Put differently, the thought pattern 2 is skeptical about the result of the binary discrete choice like Hume's evaluative skepticism[10] (Clark, 1998).

Introducing the idea of strategies into the decision making process has less

[10] "*Hume does not believe that it is possible to define evaluative terms. They are indefinable, primitive terms. Hume emphasizes that 'certain qualities in objects' are the occasions for our sentiments of approbation and disapprobation*" (Gracyk, 2003, §3.2). This skeptism can be viewed as a sequence of testings in every day language, i.e testing a value, re-testing the result, and so on.

assumptions than the decision making process only with discrete choice analysis. For example, with discrete choice analysis, a modeller has to decide which thought pattern the agent has without the idea of strategies. The conventional discrete choice models have only one set of probabilistic equations, which is similar to Equation (4.3). In contrast, the agent in this simulation model has as many sets of probabilistic equations as the number of thought patterns. Then, the simulation model lets the agent choose the best set of equations as a thought pattern and consequently the best strategy.

4.6.2 Choosing the best strategy

Each strategy calculates its own successfulness and the best strategy with the maximum success score is chosen before each game. Also, this Minority Game uses the horizon of strategy successfulness. The horizon is related with the adaptability of agents, since a long horizon makes the agent consider too much historical information, which may not be relevant to the current situation (Liu et al., 2004, pp.347-351). The length of the horizon is a parameter H, which represents the horizon for which each strategy records its score. Therefore, the success score of each strategy is only the virtual points in the last H steps an agent experienced:

$$\theta_t^s = \sum_{i=t-1-H}^{t-1} R_i^{x_i^s}/H \tag{4.4}$$

where:

$$\begin{aligned} x &= \text{Selected choice by strategy } s \text{ at } i \\ R^x &= \text{Return from the selected choice at } i \end{aligned}$$

R^x is calculated from equations (4.1) and (4.2), so that its unit is the point in the web game. As shown in Equation (4.4), the success score θ of any given strategy s at a time step t is the moving average of the return from a selected choice by the strategy within the scope of horizon H. The choice made by a strategy is not relevant with the choice used by an agent, which possesses the strategy. All strategy-scores θ were calculated whether or not the strategies were chosen by the agent. Similarly, although H was set to five in this model, the length of horizon was irrelevant with the length of experience remembered.

4.6.3 Example of the decision making process

This is an example of the decision making process. At a given week, each strategy calculates the probability of each choice according to the logit model including equation (4.3). However, the thought pattern 2 swaps the utilities or probabilities according to its rule. Then, there are five memories, so that a set of 10 possible strategies in an agent can be like the one in Table 4.5. These 10 strategies are possible strategies, but there are only a maximum five strategies for each agent in reality, according to the calibrated memory distribution in Table 4.4, i.e. some agents may have only one strategy. For example, a subset of five strategies can be like the one in Table 4.6.

Next, this agent needs to find the best strategy to make a mode choice. The set of strategies in Table 4.7 is the same set of strategies as in Table 4.6, and they have five horizon values. The choice in the table is the choice each strategy made in each experienced time step. R is the return from the predicted choice. Then, θ is the moving average of the five returns. In this example, the thought pattern 1 with memory 1 has the highest success score in Table 4.7, so this is the current best strategy. Nevertheless, this best strategy may change since it is a moving average. In the thought pattern 1 with memory 1, the side A has the probability of 0.6 (Table 4.6), so this choice is likely to be passed to this agent, but this is still determined according to the probability.

Table 4.5: Example set of 10 possible strategies in an agent

Prob. of choice	**Memory1**		**Memory2**		**Memory3**		**Memory4**		**Memory5**	
	A	B	A	B	A	B	A	B	A	B
Thought pattern1	0.6	0.4	0.5	0.5	0.7	0.3	0.4	0.6	0.8	0.2
Thought pattern2	0.4	0.6	0.5	0.5	0.3	0.7	0.6	0.4	0.2	0.8

Table 4.6: Example set of five assigned strategies in an agent

Prob. of choice	**Memory1**		**Memory2**		**Memory3**		**Memory4**		**Memory5**	
	A	B	A	B	A	B	A	B	A	B
Thought pattern1	0.6	0.4	0.5	0.5	0.7	0.3				
Thought pattern2	0.4	0.6	0.5	0.5						

Table 4.7: Finding the best strategy in an example strategy set. 'TP' stands for thought pattern and 'M' stands for memory.

Success	**Horizon1**		**Horizon2**		**Horizon3**		**Horizon4**		**Horizon5**		θ
score	Choice	R	Choice	R	Choice	R	Choice	R	Choice	R	
TP1M1	A	5.2	B	0.5	A	2.2	A	2.3	B	−0.4	**9.8***
TP2M1	B	−0.3	A	−5.3	B	−0.2	B	−0.3	A	1.3	**-4.8**
TP1M2	A	5.2	A	−5.3	B	−0.2	A	2.3	A	1.3	**3.3**
TP2M2	B	−0.3	B	0.5	A	2.2	B	−0.3	B	−0.4	**1.7**
TP1M3	B	−0.3	A	−5.3	A	2.2	B	−0.3	B	−0.4	**-4.1**

Moreover, from the viewpoint of an agent, the mechanisms of strategies and multinomial discrete choice models do not matter. For an agent, this is merely a trial-and-error decision making process.

4.6.4 Point gain and loss

The point changes according to the utility functions, but the attendance numbers are from the current game, but not a past game. The point does not change for Cancel since the agent with this choice does not play the Minority Game. The equations of point gain/loss for the three choices are:

$$\triangle \mathrm{Point}_t^A = \mathrm{Fee}(N_t^B/N_t^A) - \mathrm{Fee} \tag{4.5}$$

$$\triangle \mathrm{Point}_t^B = \mathrm{Fee}(N_t^A/N_t^B) - \mathrm{Fee} \tag{4.6}$$

$$\triangle \mathrm{Point}_t^C = 0 \tag{4.7}$$

4.6.5 Results

The distribution of the strategies used by the computer agent is the same as that of the original distribution shown in Table 4.1. The movements of the points are indistinguishable between the real web game played by human beings (Figure 4.6) and the simulation played by an artificial agent (Figure 4.7). Both results are heavily skewed, so that the two results are numerically compared with the five points statistics (Table 4.8). The five points statistics do not include the sets of games terminated before the 10th game. All five points statistics and central values are reasonably similar to each other. Additionally, the number of Side A and Side B were reasonably similar (i.e. 2250 and 2391) in the simulation just as that of online games. However, this does not mean a choice was a randomised decision making process as explained in Section 4.5. As the utility functions (4.1), (4.2), and a probability choice function (4.3)

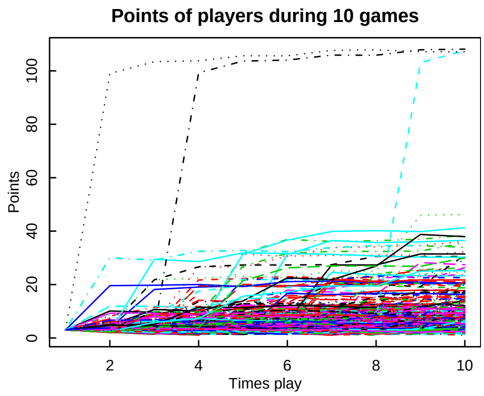


Figure 4.6: Point each player gets during 10 games (with 499 sets of games), i.e. online game

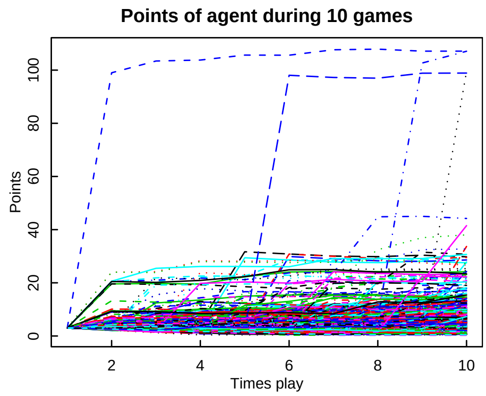


Figure 4.7: Point the agent gets during 10 games (with 499 times iteration), i.e. simulation

Table 4.8: Five points statistics (plus mean) of the distribution of points at the end of 10 games

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Online game	1.038	4.041	7.240	10.630	11.810	108.100
Simulation	1.051	4.500	6.735	9.620	10.920	107.100

show, the agent has an algorithm to make a choice so that it just looks at a randomised decision making process from a macro level.

4.7 Conclusion

The web online game was conducted to collect the data for the distribution of strategies, or more specifically the distribution of memories. Most players used the most recent information of the game and older memories were less likely to be used. In contrast, the distribution of thought patterns did not have any obvious tendency and this was consistent with the game theory.

The result of the web game was duplicated by agent-based modelling. The distribution of strategies was the same as the distribution of strategies in the real world after a small calibration. The movements of points that real players gained are indistinguishable from those that an artificial agent gained. These results may not be surprising since the rational behaviour of human players seems pseud-rational as mentioned earlier. However, this does not mean that thought patterns are a randomised decision making process as the utility functions (4.1), (4.2), and a probability choice function (4.3) show. Moreover, these results at least proved that simulating human behaviour through an agent-based model does not produce outcomes which are far from reality and conflict with theoretical arguments. It is concerned that the memory distribution determined in this chapter is used in subsequent chapters on the Upper Derwent Valley case study. Personally, this concern is considered the weakest point in this thesis; however, a feasible alternative approach is not available at this moment. The distribution of memories used by the visitors to the Upper Derwent Valley can be obtained by the combined approach of this chapter and stated preference analysis in future research. Moreover, although this is not the main focus of this thesis, this chapter shows the potential of online questionnaires to collect the large sets of data quickly without huge efforts and budgets.

4.8 Conclusion for the theoretical part of this thesis

This is the end of the theoretical part in this thesis. Chapter 2 explained the advantages and disadvantages of the Minority Game to analyse real world problems such as congestion in the Upper Derwent Valley. Chapter 3 discussed how the problem could be explained into the economic game theory and the Minority Game. The two approaches mutually help to analyse the situation in the Upper Derwent Valley leading to the development of a stochastic Minority

Game. In addition, that chapter defined the strategies of the stochastic Minority Game in the Valley as the combination of thought patterns and memories. The thought patterns were deductively discussed in that chapter. The distribution of memories was figured out while the mechanism of the stochastic Minority Game was briefly tested in this chapter.

From the next chapter, the components of the stochastic Minority Game including utilities, searching time as a congestion indicator, etc., are examined with the data collected around the Upper Derwent Valley. Although the theoretical part ends here, theoretical discussions come repeatedly in subsequent chapters due to the generative approach of this thesis. Some theoretical arguments are left in the successive chapters because the arguments are meaningless without practical examples and justification in the generative approach. The theoretical arguments will be clarified more by the end of the practical part of this thesis.

Part III

Practical part

Chapter 5

Background for practical study

5.1 Introduction

The practical side of this thesis has not been referred to much in the previous three chapters. Therefore, the practical background information specific to the practical study is now presented. These are generic concepts presented throughout this part of this thesis and more specific information is shown in each subsequent chapter.

First, the types and the methodology of data used in this practical study are explained. Next, the generic issues about practical modelling work are discussed. Following this, more specific topics namely agent-based modelling and the stochastic approach, are discussed. Finally, this chapter concludes the structure of this practical part and a stochastic agent-based modelling. Overall, this chapter explains reasons to use the stochastic agent-base modelling approach.

5.2 Data collection

The data in this project is described in detail in each chapter, so only the location, method, and the quality of data are explained here. The behavioural survey on mode and parking location choices was a destination survey, i.e. the survey was conducted at the parking areas around the Upper Derwent Valley (Figure 5.1).

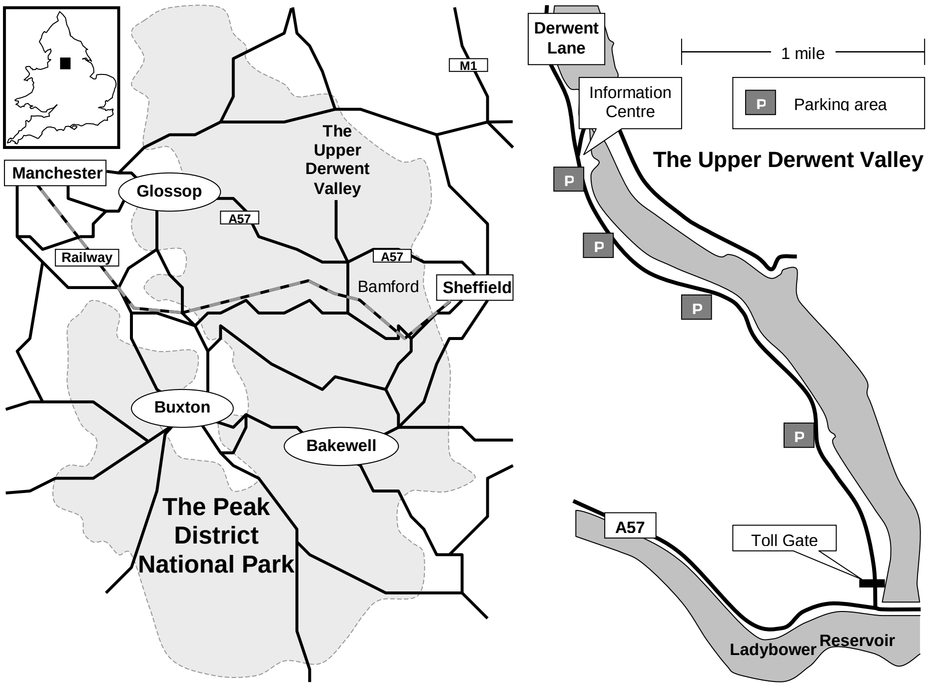


Figure 5.1: Map of Derwent Valley

After the pilot survey, 700 questionnaires were distributed and 323 were returned (46.1 %) during the summer of 2003. The same survey was attempted in the origins of visitors such as from Sheffield and Manchester in the summer of 2004 as an M.Sc. dissertation I supervised (Thomopoulos, 2004). However, only 10.65% or 69 questionnaires were returned out of 648 questionnaires. This low return rate showed that the origin survey was not practical in this case study. For example, proper surveying to collect 300 samples out of the entire Manchester population, i.e. 2,482,352 people (National Statistics, 2001), seems infeasible. The reasons are that it is more difficult to intercept potential visitors to the Upper Derwent Valley and dangerous to generalise the attributes of the potential willingness to travel for such a small proportion of the population. The low return rate of the origin survey can be explained as follows. Most of the people who received the questionnaire did not know the issues raised in the Upper Derwent Valley or they were not interested in the issues. Thus, this D.Phil. project excluded the origin survey data because of the inconsistent quality compared to the destination survey. The results from the origin survey are only briefly discussed in the concluding chapter as one of potential future research.

The traffic related data used in this project were car movements, the arrival rate of private cars, parking hours, and annual traffic flow on the A57. The arrival rate was collected in front of a parking area during the survey period mentioned earlier. The parking hours at each parking area were from the parking beat surveys, which were undertaken over three days during August 2001, i.e. the 23rd, 26th, and 27th of August 2001 by the Transport Office of Derbyshire County Council. The 23rd of August 2001 was a normal summer weekday, in contrast, the 26th and 27th of August 2001 was a Sunday and the summer Bank Holiday, which were usually the busiest days in the Upper Derwent Valley. All parked cars were recorded, so the data set acted as a population. In total, 1961 cars were recorded. When the data was given by the local authority, the fraction (minutes) of the time was rounded off. Therefore, the time showed only parking hours. In addition, the annual traffic flow on the A57 was collected by an automated system during 2003. The flow was westbound and eastbound from 7:00 to 19:00, and the average flow was used for this project.

5.3 Generic information of modelling

In this section, basic concepts and assumptions are explained. These assumptions are made not only for conventional approaches, which conduct analysis solely with discrete choice models, but also for the approaches with agent-based models. One of the advantages in the agent-based model is that fewer assumptions are made. However, unfortunately, this modelling is still constrained by time, budget, and other common costs associated with social surveys and consequently it is necessary to make a number of assumptions.

5.3.1 Validation and verification processes

Verification and validation of modelling are important issues throughout any modelling work. Generally, the verification checks whether a programme does what it is planned to do and the validation pertains to whether the simulation is a good model of the target (Gilbert and Troitzsch, 1999a). However, validation and verification processes are different in different types of models (Balci, 1994; Giunchiglia et al., 1998). Multinomial discrete choice models have clearer validation guidelines, such as Independent and Identical Distribution (IID)[1], Independence of Irrelevant Alternatives (IIA)[2] and Log-likelihood ratio test (McFadden, 1974)[3]. The variation of multinomial discrete choice models was not the problem in this thesis for economic analysis because the analysis was

[1]Each distribution has the same probability distribution as the others and all are mutually independent.

[2]"Where any two alternatives have a non-zero probability of being chosen, the ratio of one probability over the other is unaffected by the presence or absence of any additional alternative in the choice set" (Luce and Suppes, 1965)

[3]The log-likelihood ratio test is used to find out (Ortúzar and Willumsen, 2001, p.263): 1) Attribute generality – There are two main types of explanatory variables, generic and specific; the former have the same weight or meaning in all alternatives, whereas the latter have a different specific meaning in each of the choice options and therefore can take a zero value for certain elements of the choice set, and 2) Sample homogeneity – It is possible to test whether or not the same model coefficients are appropriate for two subpopulations (say living north and south of a river).

executed with an off-the-shelf software package, BIOGEME[4] (Bierlaire, 2003). Having said that, two algorithms, `BIO` and `DONLP2`, were used to check if the results from the two algorithms were consistent.

In contrast, although some guidelines for agent-based simulation models have been suggested in some studies (Moss et al., 1997; Gilbert and Troitzsch, 1999a; Raney et al., 2003; Peeta et al., 2005), there is no clear checklist of the validation for agent-based simulation models. However, the inconsistent guideline is understandable since agent-based modelling aims *"to simulate reality more realistically"* so that the modelling process itself is the validation process. Hence, validation processes could be as many as agent-based models. Thus, this project followed the guidelines of previous studies where it was appropriate and other validation were deductively explained in the theoretical part. Verification was more important since an agent-based simulation model, as software, was programmed in this project (Appendix B). An agent-based simulation model might be considered as an implementation of object-oriented programming and the programming paradigm has its own established verification method such as the Unit Test. This testing method was established by Beck (1994) and has been used as a standardised testing method in a lot of object-oriented programming. The approach is to write testing programmes for every non-trivial method to isolate each method from the problems of other methods. This project used the Java version of Unit Test, JUnit, and over 300 methods were tested at least every 10 minutes while the model was developed (Andou, 2003). It is common knowledge that "*there is at least one bug in every non-trivial program*", so the Unit Test does not guarantee to solve all verification problems. However, significant bugs were possibly eliminated in this project.

[4] `http://roso.epfl.ch/biogeme`

5.3.2 Policy as an exogenous factor

Traffic regulations and policies involve the property of authority, and the models in this project do not possess this authority. Therefore, the regulations are designed as exogenous or offline factors (Wooldridge, 2002, p.213). Thereby, the models do not take into account the possibility of policy emergence in current agent-based simulation models (Schlesinger, 2001). Observing the emergence of policy is not a part of the question in this research since the traffic regulations in this model are made by the policy makers in the real world.

In addition, the new regulations to control user demands were integrated with the model as scenario-based analysis. The policies such as the toll fee and bus fare were decided based on the interviews with local authorities. This issue is further discussed in Chapter 9 as potential future research.

5.3.3 Unit of agent

The lowest level of society in this agent-based simulation model is a visitor group, such as a family or hiking group, and the decision is made by a trip leader. Modelling collective decision making based on each individual would be a more realistic approach. However, collecting such data needs several rounds of surveys among each visitor group. This process is time consuming and difficult even within one corporation (Rose and Hensher, 2004). Therefore, the process is likely to fail with 300 households scattered around the Manchester and Sheffield regions. Moreover, from the view of agent-based modelling, it is common to model not from the individual, but from higher social levels such as departments, firms (Edmonds, 1999b), and visitor groups (Raney et al., 2003; Balmer et al., 2004). Thus, this study uses a visitor group by a unit of decision making as it has been applied in previous studies on the discrete choice model of road user charging scheme (Steiner and Bristow, 2000).

This section showed the importance of validation and verification in modelling work. In this project, these are handled with standardised testing methods in multinomial discrete choice models and object-oriented programming. Particularly, the validation methods of the choice models are useful since the decision making process is complex in the real world. Moreover, policies are treated as exogenous factors and the lowest level of society is set as a visitor group due to the complexity in the real world. These features are continuously discussed in the next section since the testing facility and simplification for speed are the key factors in stochastic agent-based modelling.

5.4 Concepts of the stochastic agent-based model

5.4.1 Differences between microsimulation and multi agent-based simulation

Microsimulation and multi agent-based simulation are sometimes confused with each other since the two simulation models describe macro phenomena bottom up. These two simulation models are used interchangeably in many articles. However, these simulation models are formally dissimilar. According to Gilbert and Troitzsch (1999a), the historical origins are different. Microsimulation was started as a direct descendent of the stochastic process in the 1950s. In contrast, the multi agent simulation was started in the late 1980s as a child of cellular automata[5] and artificial intelligence[6]. As their parents are different, the characteristics of the two simulation models are also different.

[5]Cellular automata models a world in which space is represented as a uniform grid, time advances by steps, and the 'laws' of the world are represented by a uniform set of rules which compute each cell's state from its own previous state and those of its close neighbours. (Gilbert and Troitzsch, 1999a, p.122)

[6]It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable (McCarthy, 2004, A.1).

Raney et al. (2003) explain the difference clearly. Microsimulation is based on the stochastic process of molecular dynamics, which means that if two molecules are in the same situation the expected movement is the same. In contrast, the multi agent simulation model treats the agent more intelligently and heterogeneously; therefore, the expected movement of two agents may not be the same even in the same situation. Intelligence can be translated into a rule-based code of artificial intelligence, in contrast to the continuous equation of stochastic process in the microsimulation. More importantly, microsimulation generally does not consider heterogeneity between molecules, or characteristics in agents. Therefore, although microsimulation is based on the individual particles, it is still based on aggregated information in some sense.

5.4.2 Modelling with uncertainty – stochastic approach

The first entry in the definition of 'Uncertainty' in the Oxford dictionary[7] is:

> The quality of being uncertain in respect of duration, continuance, occurrence, etc.; liability to chance or accident. Also, the quality of being indeterminate as to magnitude or value; the amount of variation in a numerical result that is consistent with observation.

These factors should be familiar to researchers dealing with modelling work. Especially, travel behaviours which are not always clear because the decision making process is often involving categorical choices (Bus or Car), which are considered more complicated than continuous choices (100 or 50 miles).

There are two possible ways to handle uncertainty in the discrete choice modelling with agent-based simulation, namely the fuzzy logic and the stochastic (statistic) approach. Fuzzy logic is a rather computational approach and all decisions have to be expressed by IF-THEN rules, and uncertainty in the fuzzy

[7] The Oxford dictionary online, `http://dictionary.oed.com/`

logic is ambiguous, which is different from randomness and chance. For example, a person can act exactly the same way if a situation is exactly the same with the conventional IF-THEN rule-based codes (Raney et al., 2003; Schleiffer, 2005). However, this may not be true with fuzzy logic if a situation is somewhere between the one of ambiguity (e.g. the fee is fairly expensive so I may or may not get on the bus). With fuzzy logic, all choices are still described in tree like IF-THEN rules and ambiguity is handled, so that any complex decision can be expressed in this fashion. However, this is a double-edged sword. Since all choices are expressed by IF-THEN rules, the amount of calculations is enormous. This approach has been used in some studies (Wu et al., 2000; Peeta et al., 2005) and will be used more often as computational power increases in larger projects. Furthermore, these numerous IF-THEN rules are unobservable in reality; therefore, it is likely to be impossible to justify IF-THEN rules without validation and verification tests. In other words, IF-THEN rules can be closer to the real decision making process at the theoretical level, but its application can be difficult without any justification mechanism. These performance and testing problems are resolved in the stochastic approach.

The stochastic approach to uncertainty has a longer history. In 1944, Neumann and Morgenstern (1953) developed the formal theory of risk and uncertainty as an economic approach, which has been applied to many decision making problems including discrete choice models as mentioned earlier. The stochastic approach in decision making has been explained in many ways in the past such as *relative frequency* with probability (Ortúzar and Willumsen, 2001, p.50) and as uncertainty associated with opponents' decision making in mixed strategy games (Rasmusen, 2001). Whatever the reason for the stochastic process, the important things are that this approach speeds up simulation and it enables the validation and verification tests of the decision making process to be undertaken.

Stochastic agent-based modelling

A physicist may support the idea of collective decision making since as everything can be a mass of particles in a theoretical viewpoint. For example, if the movements of all molecule particles in a human brain are found, the decision making of a human can be perfectly predicted. This is an extreme idea, but this is similar to the essential concept in the agent-based modelling, i.e. *modelling from bottom up*. According to the physicist's idea, combined with agent-based modelling, we can even find out the macroscopic social phenomena from the study of molecule movements in human brains.

The problem is that the more complex a model is the slower the speed of modelling is, with respect to the three types of speeds. First, there is the relative running speed of modelling against reality. The model of social simulation should run at least faster than reality (Raney et al., 2003). Otherwise, reality will catch up with a prediction before the simulation finishes. Then, the prediction is no longer a prediction in this sense and there is no point in carrying out the research. Second, there is data collection speed in reality. Disaggregated input generally needs more time to be collected. Then, if the data collection cannot be finished before the end period of its prediction, the data collection for the prediction is meaningless because the collected data will be the result of the prediction. The third type of speed is the speed of an agent's reaction time against the change in the real environment. This does not affect the main research because the environment that artificially exists in a computer system has the same speed as an agent's reaction time. However, if a simulation model has an interface with the real world, like the online web Minority Game in Chapter 4, it is an important matter. If the reaction time of an agent takes a year, the reaction will be irrelevant to the environment since the real environment changes at a much faster pace. So, the reaction of the agent does not reflect reality. More practically, people in the real world will get bored and forget about the agent and the simulation.

There is a good example about the third speed in computational intelligence of games as speed is the most important factor in this field (Baba and Jain, 2001). IBM's Deep Blue chess computer won a game against the world chess champion, Gary Kasparov, in 1997 for the first time at a tournament level (IBM, 2005). From the defeat of Deep Blue in previous year, the IBM researchers stopped an exhaustive search into every possible position because speed is the most important factor in chess. The faster a player moves, the stronger he is since he does not give a chance to change the environment (i.e. Gary Kasparov in this case). Thus, the situation is more predictable for the player.

This Deep Blue example shows that it is important not to find every detail in decision making, but to focus on the research target and speed even in the faster computer and bigger project. The main concern of this agent-based modelling is to analyse and forecast the effect of transportation policies on the mode choice of visitors and on congestion. Duplicating the Upper Derwent Valley and visitors' movement in the PC is also an interesting topic, but this should not be more important than the former concern when a real world policy is analysed. Additionally, agents in the current project are the simplified representatives of visitors to the Upper Derwent Valley. It is doubtful if real visitors seek the perfect decision in the real world. If it takes too long to obtain a perfect decision, the decision is no longer useful since the environment must be changed by that time (Schleiffer, 2002).

A stochastic approach is the study to see the particular information researchers are interested in from complex and noisy data. Therefore, the process cuts down factors which 'seem' not to be important for the outcome and this can be seen as the opposite process to the agent-based modelling approach.

However, at least in the current research situation, it is important to consider the speed of the agent-based model to reflect reality. Moreover, the stochastic process is used in the decision making of agents as the multinomial discrete choice model and this analysis has various validation and verification methods as mentioned earlier. Therefore, the stochastic agent-based modelling realistically reflects reality more than the conventional agent-based model at a practical level.

5.5 Conclusion and outline of successive chapters

The previous section compared microsimulation and agent-based simulation. The core of the simulation model in this thesis is the Minority Game, which is categorised into an agent-based modelling. However, stochastic microsimulation is partially used for traffic movement in the Upper Derwent Valley. Also, multinomial discrete choice models as a decision making process of agents are also a stochastic process. Therefore, the simulation model in this thesis is termed as stochastic agent-based modelling.

The reasons to choose the stochastic approach over the cellular automata and fizzy logic approaches are the better facility of testing methods and simulation speed. It is important to mention that these two reasons are consistent with the 'Aims at the theoretical level' stated in Chapter 1 on page 4, i.e. the facility of testing methods is for '*reality of model*' and the simulation speed for '*efficiency of model*'. This one sentence shows how the generative approach[8], which is the iterative process of the inductive and deductive approaches, worked in this thesis.

[8]This was explained in 'The conceptual framework and the approach of research' of Chapter 1 on page 7.

Since the stochastic approach has a longer history, the approach is facilitated with standardised testing methods at a practical level. In addition, the stochastic approach makes the simulation run significantly faster than the cellular automata and fizzy logic approaches. The issues on the simulation speed will be solved in the future research of well-equipped projects. However, for this thesis, this was the one important issue, since the best computer system, I could use for this project, was my laptop computer with a Pentium M 1.6GHz CPU and 750MB RAM. The same simulation on a Linux workstation with a Xeon 2.4GHz CPU and 1GB RAM did not run significantly faster than the laptop did. It is possibly because the computational resources had to be shared with other researchers in the workstation, but not in the laptop. A normal simulation process of the stochastic agent-based model in this thesis took several hours and long ones took from 24 hours to 8 days. Simulations have to run every time when the input data and scenarios are changed, so that the slower approach was impractical for this thesis. Hence, the stochastic approach was chosen over fizzy logic and cellular automata approaches for the agent-based simulation model and microsimulation model, respectively.

Figure 5.2 was shown in the first chapter and this figure represents the overall structure of the stochastic agent-based model of this project and the practical part of this thesis. The agent-based model has four sub modules: 1) Multinomial mixed logit model for mode choice, 2) Binary logit model for parking location choice, 3) Markov queue model for parking network, and 4) the Minority Game for parking congestion and learning. Chapter 6 analyses the visitor characteristics and travel behaviour with econometric models by using the first two modules. Chapter 7 develops the third module to simulate the movement of cars with the Markov queue network theory, which is considered as a microsimulation. Chapter 8 applies the Minority Game, which is discussed in the theoretical part, while all other sub modules are combined together with the agent-based modelling approach. Then, the results lead to the final chapter of this thesis.

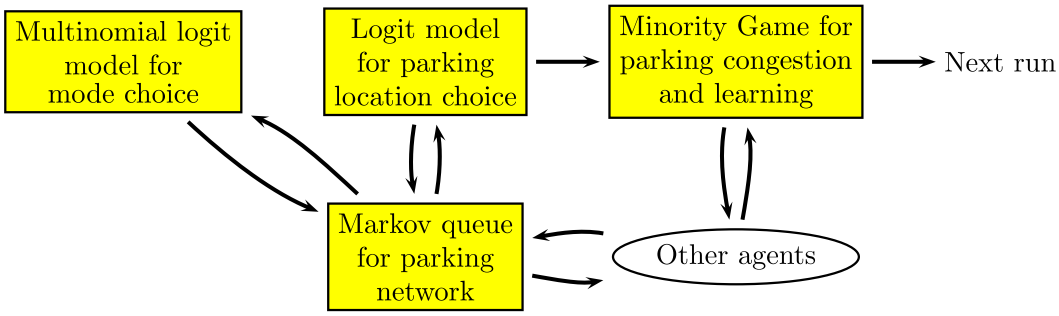


Figure 5.2: Structure of agent-based simulation

Chapter 6

Econometric analysis of road user charging at the Upper Derwent Valley, the Peak District National Park

6.1 Introduction

Private car use is a key component in the areas of outstanding natural beauty because these locations attract many visitors from local and urban areas by car. Transport policies such as private car access regulation or road user charging scheme potentially improve traffic congestion and the values of natural beauty by cutting visual intrusion and traffic noise. Therefore, today, policy makers consider implementing road user charging scheme in the areas of outstanding natural beauty. The Upper Derwent Valley in the Peak District National Park is one of the proposed areas for this new policy tool.

This chapter analyses the scheme with the econometric tools namely multinomial discrete choice models. The choice models are also used as a decision making mechanism in the agent-based simulation model in Chapter 8. In this research, the demand by the one-day visitor is focused on because it is the major component of traffic in the Valley. First, this chapter analyses the characteristics of the visitor. Then, the effect of a parking fee on a parking location choice is discussed with a binary logit model and regression analysis. Following this, future transport policies, road user charging and park & ride schemes, are analysed by a multinomial mixed logit model with stated preference data.

6.2 Methodology of data collection

The road user charging scheme in the Upper Derwent Valley is still under consideration, and there is no implemented road user charging scheme in a similar situation yet (Steiner and Bristow, 2000, p.96)(Eckton, 2003, p.310), so a revealed preference survey was impossible. Therefore, a stated preference survey was used for the question about the mode choice among Auto (Toll & drive), Bus (Park & ride scheme), and Cancel (Do not visit) options. Additionally, visitor characteristics and past trip experience were collected. The full questionnaire for the main survey is shown in Appendix A. Although the main concepts and methods used in this project are applicable to forecast entire travel demand (Anabel, 2002), this study focuses on day trip travel to the Valley due to the results from a pilot survey. The pilot survey showed that the main components of travel demand during the busy period were the day trip visitors from local towns and neighbouring cities, namely Manchester and Sheffield.

6.2.1 Pilot and main surveys

The pilot survey took place between the 1st and 3rd of August 2003 in good weather at around the first and second parking areas from the Information

Centre. Out of 130 distributed questionnaires, 41 of them were returned. The questionnaire was modified after careful examination of the responses. Only the mode choice within Derwent Lane was focused on in the stated preference questions of the main survey. Also, the questions about the actual modes and routes to the Peak District National Park were deleted in the main survey because the selections were complex so that it would have affected its return rate.

The main survey was carried out for 9 days from the 23rd to 31st of August 2003 including the bank holiday Monday on the 27th of August. The air temperature was cool and a short period of drizzle appeared during the survey period; however, it was generally fine weather overall. Survey locations were extended to the third parking area when the first two parking areas were extremely busy. Additionally, a small survey was conducted in public buses and at bus stops in the Valley. Overall, 700 questionnaires were distributed and 323 of them were returned (i.e. a return rate of 46.1%).

6.2.2 Stated preference questionnaire design

During weekends and holiday periods, Derwent Lane beyond the Information Centre is closed to private cars because of potential severe congestion. Visitors' destinations are usually beyond the Information Centre, otherwise, visitors relax around the Information Centre. Therefore, respondents were asked how they would travel to the Information Centre if the road user charging and park & ride schemes were put into effect in the Valley in *ceteris paribus* conditions (e.g. with same trip members). In addition, the visitors to the Upper Derwent Valley were expected to respond to the schemes in one of three ways:

'Auto' option: Pay a toll for road use and drive into Derwent Lane to get to the Information Centre.

'Bus' option: Come near the Valley with whichever travel mode, and then use the complementary park & ride service to get to the Upper Derwent Information Centre.

'Cancel' option: Cancel the trip to the Valley and instead go somewhere else or stay at home

Terms used in the questions and the brief explanations of the road user charging and the park & ride schemes were given before the hypothetical questions relating to mode choices. After reviewing previous research (Steiner and Bristow, 2000; Fowkes, 2000; Ortúzar and Willumsen, 2001, p.283), four attributes of the mode choices on travel time and costs were chosen, and four different levels were selected for each attribute:

Road user charging (£): a toll to enter Derwent Lane from the A57

Park & Ride fare (£): a fare for bus service, which links local parking areas, Bamford train station, and the Upper Derwent Information Centre.

Frequency of bus service (minutes): the period between departure times of the shuttle buses.

Searching & walking time (minutes): the combination of searching time for a parking space and walking time from the parking area to the Information Centre.

Parking fee difference (£): the difference between parking fees for the Auto and the Bus. The parking fee for the park & ride service is the fee visitors pay when they park their car before getting on a bus. The parking fee for toll & ride is the fee visitors pay when they park their car at one of the four parking areas along Derwent Lane.

The four levels were determined by using the boundary value evaluation technique (Fowkes, 2000). For the question about "parking fee difference", two sub-attributes were used – i.e. parking fees for the Bus and the Auto. The four values of all attributes were equally distributed in the 16 fractional factorial experiment and the design of 16 questions is known as a lattice square[1] (Table 6.1). This is explained in Section 6.5.1, but these four alternative specific variables were converted into generic variables comprising travel time and costs due to the insignificant log-likelihood test[2] (McFadden, 1974). Combined

Table 6.1: 16 hypothetical questions and attributes

Q.	Toll fee (£)	Bus fare (£)	Headway (min)	Seach & Walk (min)	Parking fee (£) Auto	Bus	Difference
1	20p	£1.00	5	1	50p	10p	40p
2	20p	£2.00	15	30	£2.50	50p	£2.00
3	20p	£3.00	30	50	£1.00	50p	50p
4	20p	£5.00	45	15	£2.00	50p	£1.50
5	50p	£1.00	15	15	£1.00	50p	50p
6	50p	£2.00	5	50	£2.00	50p	£1.50
7	50p	£3.00	45	30	50p	10p	40p
8	50p	£5.00	30	1	£2.50	50p	£2.00
9	80p	£1.00	30	30	£2.00	50p	£1.50
10	80p	£2.00	45	1	£1.00	50p	50p
11	80p	£3.00	5	15	£2.50	50p	£2.00
12	80p	£5.00	15	50	50p	10p	40p
13	£1.00	£1.00	45	50	£2.50	50p	£2.00
14	£1.00	£2.00	30	15	50p	10p	40p
15	£1.00	£3.00	15	1	£2.00	50p	£1.50
16	£1.00	£5.00	5	30	£1.00	50p	50p

attributes were presented in Figure 6.1, and respondents were asked to rank their preferences from the three stated options 1) Auto, 2) Bus, and 3) Cancel.

[1]For example, each variable in an $n \times n$ lattice square has $k = n$ levels. No orthogonal (row in this case) contains the same number twice (Lindner and Rodger, 1997). This is also known as latin square.

[2]See the footnote on page 89

Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**20p**	per person	Toll	£**2.00**	per car
A bus every	**15**	minutes	Searching for a parking space and walking to the centre takes	**30**	minutes
Parking fee	**50p**	per car	Parking fee	£**2.50**	per car

Under these circumstances, I would:

[ ] Park & ride

[ ] Pay toll and drive

[ ] None of them (don't visit the valley)

Figure 6.1: Example of stated preference question

6.2.3 Questions about characteristics and travel behaviour of visitors

Besides the hypothetical stated preference questions, revealed data was collected in the same questionnaire. For example, respondents were asked about their profession and incomes in the earlier section, and asked about the departure and arrival times and costs of their trips in the later section. Travel behaviour questions, similar to Anabel's (2002) work on memories about past trips, were briefly asked. Some revealed preference data were utilised with the agent-based modelling in Chapter 8. The final question was the maximum willingness to pay (WTP) to drive into Derwent Lane (i.e. road user charging). Initially, this question was asked to check if there was any bias in the stated-preference questionnaire since this was a known problem (Sugden, 1998, 316–317) (Section 6.3.2); however, this question was also utilised in the discrete choice model of parking location (Section 6.4.2).

6.3 Characteristics of visitors

6.3.1 Distributions of age, income, origin, and travel frequency of visitors

Only 16% of visitors came from the local council area, Derbyshire, probably because of the survey period. This survey was carried out during the busiest time of the year, and local visitors might have avoided visiting the Valley during the busiest period ('Visitors' origin' in Figure 6.2). However, most visitors (60%) to the Valley came from local towns and the neighbouring cities comprising of Manchester and Sheffield. The Upper Derwent Valley is easily accessible for the visitors from these two cities via the A57. For example, a visitor from Sheffield drives 20 minutes to the Valley about three times a week just to take a walk with his dog. Therefore, for these visitors, the Upper Derwent Valley is like a large backyard where they relax and take a walk.

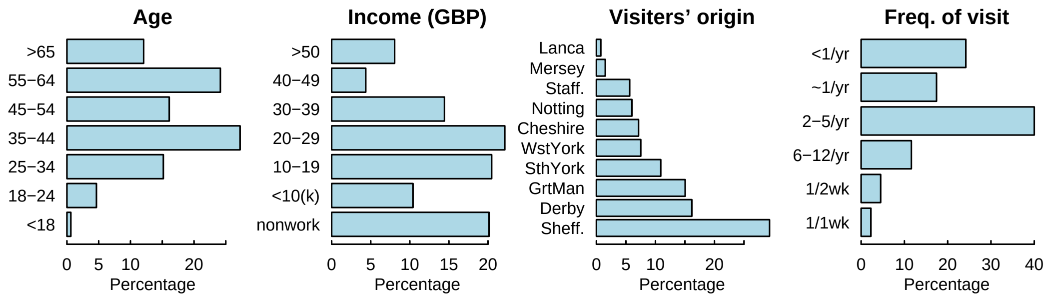


Figure 6.2: Proportions of visitors' characteristics

The age distribution of visitors is highly skewed, and two modes at '35-44' and '55-64' are present in the distribution ('Age' in Figure 6.2). This age distribution matches the observations made during the survey. Most visitors I met during the survey were either families or elderly visitors, and they could represent the two modes mentioned earlier, i.e. families for '35-44', and elderly visitors for '55-64'. In addition, income distribution ('Income' in Figure 6.2) also supports this trend. Some 20% of visitors to the Valley are non-workers, and most of these visitors are elderly people, since the proportion of students is nominal (i.e. 5%).

6.3.2 Distribution of willingness to pay to the road user charging

The average willingness to pay (WTP) to the road user charging scheme is £2.373 with a standard deviation of £1.75. However, the median of £2 seems to be a better representation of the central value. The distribution of the WTP is fairly normal with m=2 and σ=1 except for the small peak at £5 marked with an arrow in Figure 6.3. Also, the distribution is even closer to the normal distribution after omitting the observations of the WTP = £5 (i.e. WTP[!=5] in Figure 6.3). No clear reason for the small peak was found except for possible

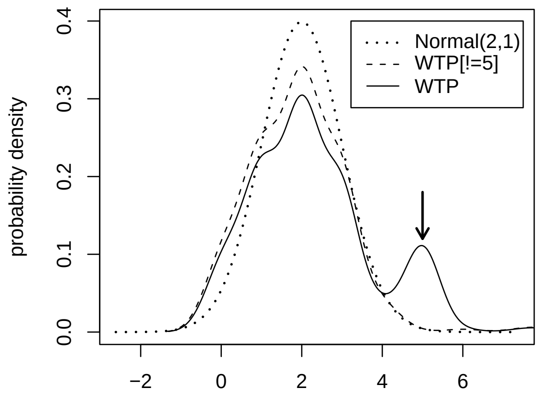


Figure 6.3: Density of the WTP

questionnaire bias. The previous section before this open-ended WTP question was the stated preference questions, and the highest value of the road user charging stated in the questions was £5. Thus, although no upper boundary was set for the question of the WTP, some respondents probably had assumed the upper boundary of £5 from the previous questions. The same questionnaire bias of assumed upper boundary in the WTP from previous questions was also observed in previous studies (Eckton, 2003, p.312). As a result, the upper tail of the WTP distribution would be elongated if respondents did not assume that the upper boundary was £5, and this meant that the sampled distribution of the WTP was possibly under estimated.

Therefore, from the observations above, the major components of visitors to the Upper Derwent Valley are families and elderly people from the local areas, Manchester, and Sheffield. Moreover, the central value of the WTP to the road user charging scheme is approximately £2.

6.4 Effects of the parking fee at the Information Centre as a current policy tool

6.4.1 Differences in the behaviours of parking costs by age, travel frequency, and region

In this section, the parking behaviours are analysed with respect to travel frequencies and parking costs. The only policy tool to suppress the private car use in the Upper Derwent Valley is currently a parking fee charged at the Information Centre. It is wise to buy a day ticket rather than pay per hour if parked at the Information Centre for more than five hours due to the charging system. The average hours visitors spend at the Valley is 4.1 hours[3]. Consequently, most visitors either buy a day ticket or go to the other the parking areas where they are not charged for parking. So, the distribution of parking costs is bimodal of '0 pounds' and '2.5 pounds' (Table 6.2). Furthermore, the distributions of parking cost are different among age categories (Figure 6.4), and the categories of visiting frequencies to the Valley (Figure 6.5). Between the ages of 18 and 24 and under the age of 18, approximately half of visitors parked at the Information Centre with a day ticket and the other half

[3]A standard deviation is 2.11 hours and the Shapiro-Wilk statistic (p-value = 0.98) shows samples significantly come from a normal distribution.

Table 6.2: Parking costs at the Information Centre

Park fee (£)	0	0.5	1	1.5	2	2.5
Parking time up to	–	1hour	2hours	3hours	4hours	1day
Park at Information Centre?	No	Yes	Yes	Yes	Yes	Yes
% of visitor	26.8	5.1	2.9	1.0	6.4	58.0

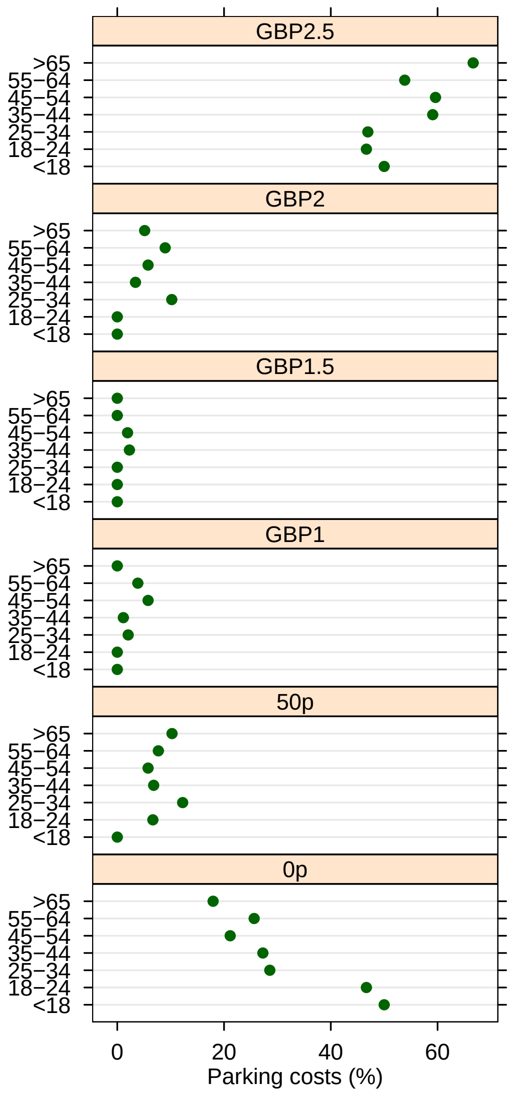


Figure 6.4: Parking cost by visitors' age

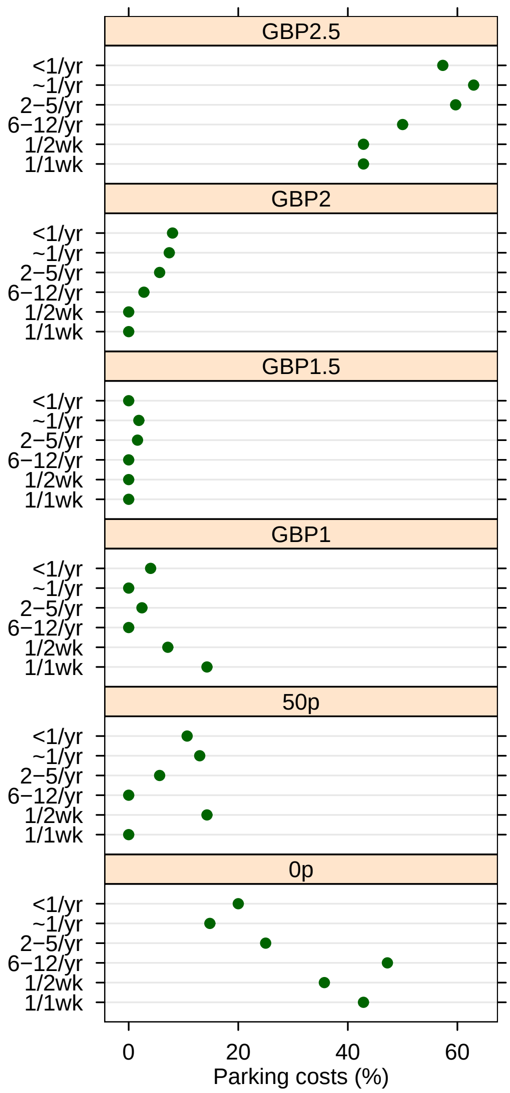


Figure 6.5: Parking cost by frequencies of visits

parked at the other parking areas. In contrast, no more than the 30% of visitors who are older than 25 and only 18% of visitors who are older than 65 did not park at the Information Centre. This difference among age categories may be partially due to an income effect. However, the distribution of parking costs by income categories does not show a similar differentiated pattern as clearly as that of age. Therefore, not only income but also other effects, such as distance from the Information Centre, contribute to the difference in the age categories. Even though the distance from the nearest free parking area (Derwent Overlook) is only 557 yards (ca. 510 metres) from the Information Centre, the distance could be too long for babies in young families and elderly people. Moreover, many families and elderly people brought chairs and other large equipment with their cars to relax around the Information Centre. It is difficult to carry the equipment by hand even for a few hundred yards.

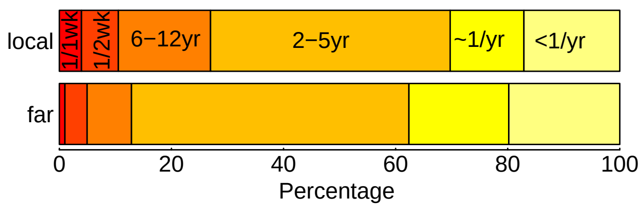


Figure 6.6: Visiting frequences between two regions

The distributions of the parking cost by the frequencies of visits are different (Figure 6.5). As shown in the figure, frequent visitors are less likely to park their car at the Upper Derwent Information Centre and thus avoid the parking fee. In contrast, the infrequent visitors do not mind paying £2.50 for a day ticket as much as frequent visitors do. This trend is easily predictable, e.g. you may not go to the Valley every week if you pay parking fees every time you visit, so you park somewhere else to visit the Valley more frequently.

Furthermore, the visiting frequencies are different between local areas including Manchester and Sheffield and other regions. Cochran-Mantel-Haenszel

Chi-Square statistic[4] was used to test the equal distribution of visiting frequencies. The result shows a p-value below 0.01, which means the distribution of visiting frequencies to the Valley is significantly different between the visitors from local area and other locations. This difference is predominant at the frequency levels below 'every other month', i.e. 1/1wk, 1/2wk, and 6-12yr in Figure 6.6. Some 27.0% of visitors from local areas come to the Upper Derwent Valley at least every other month. In contrast, only the 12.8% of visitors from other areas come to the Valley at least every other month.

6.4.2 Discrete choice analysis of parking location

As explained above, the distribution of parking costs is rather categorical, 'Park' or 'Not park' at the Information centre, and a binary logistic model was used with BIOGEME[5] (Bierlaire, 2003). If drivers park at the Information Centre, they do not need to walk to the centre, but they have to pay parking fees. In contrast, if the drivers do not park at the Information Centre, they have to walk to the Information Centre, but they do not have to pay the parking fees. The results of logit model shows a similar trend as the parking cost and visitors' characteristics in the previous section. The three significant factors in the logit model are visitors' age, visiting frequency per year, and the WTP to the road user charging (Table 6.3). The observed utility functions for the two choices are:

$$\text{1st parking area: } V_i^1 = \beta^{\text{often}}(\text{No. of visit}) + \beta^{\text{age}}(\text{Age}) + \beta^{\text{WTP}}(\text{WTP}) \tag{6.1}$$

$$\text{Other parking area: } V_i^O = \alpha^{\text{Other}} \tag{6.2}$$

[4]This is an appropriate test when both variables, such as the visiting frequency and locations (travel time or distance) in this analysis, lie in an ordinal scale (SAS Institute Inc., 1999)

[5]`http://roso.epfl.ch/biogeme`

The age of a trip leader is a categorical variable and its coding is: if age $< 18 =$ 1, else if age $< 24 = 2$, else if age $< 35 = 3$, else if age $< 45 = 4$, else if age < 55 $= 5$, else if age $< 65 = 6$, else if age $> 65 = 7$. The categorical age variable is commonly used in transport modelling (Bierlaire, 2001). Robust t-statistics for the zero coefficients are significant for all three factors and alternative specific constant (ASC), α^{Other}. Income level, travel hours from home to the Valley, and the hours respondents spent at the Valley were not significant factors in this model. No interaction terms are also significant. Income is unexpectedly insignificant because the relationship between income and car usage is commonly used (Redmond and Mokhtarian, 2001). However, the insignificance of the relationship has been observed in some studies (Tsamboulas, 2001). The

Table 6.3: Binary logit model for parking location choice

Coefficient	Estimate	Robust Std. Error	Robust t-value
α^{Other}	1.240	0.518	2.395
β^{often}	-0.033	0.015	-2.251
β^{age}	0.214	0.104	2.051
β^{WTP}	0.229	0.084	2.737

Number of observations = 268
$L(0)$ = -185.763
$L(\hat{\beta})$ = -144.046
$\bar{\rho} = 0.203038$

logistic form of the fitted model used for the results is below:

$$P(\text{1st parking area}) = \frac{\exp(U_i^1)}{\exp(U_i^1) + \exp(U_i^O)}$$

U_i^* is unobserved utility including an error term, i.e. $U_i = V_i + \varepsilon_i$. The positive ASC means that visitors prefer to park at the other three parking areas if all the three variables are the same, e.g. the visitors possibly want to avoid the parking fees. The cost effect is stronger than comfort effects by parking close to the Information Centre. Walking time and parking fees could not be integrated into the model in this study, but these effects can be captured with the stated preference questionnaire in future research. According to this model, the probability of parking at the Information Centre rises as age increases at a given WTP and frequency of visit level. In contrast, the probability declines as a visitor travels to the Valley more often. The positive relationship of age and the negative relationship of visiting frequency are consistent with the result displayed in Figure 6.4 and 6.5 in the preceding section.

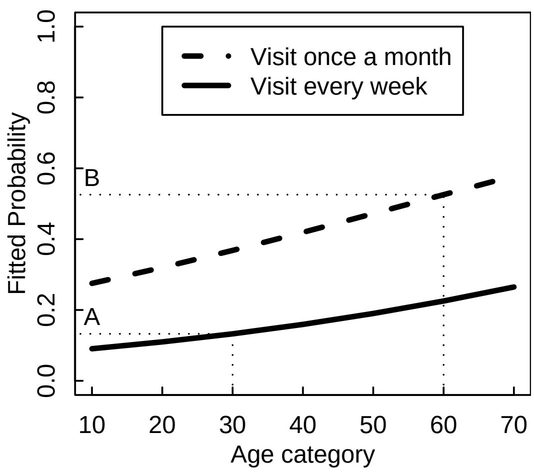


Figure 6.7: Probability of parking at the centre with the WTP=£2

For example, if a visitor is 60 years old, the WTP is £2, and the visiting frequency is once a month, the probability the visitor parks at the Information Centre is 52.6% ('B' in Figure 6.7). In contrast, if a visitor is 30 years old, the WTP is £2, the visiting frequency is every week, the probability is 13.3% ('A' in Figure 6.7). Therefore, elderly visitors want to park their car right in front of the Information Centre even though they have to pay a parking fee of up to £2.50. This behaviour is probably due to the physical disadvantages of the elderly to younger visitors because an income effect is not observed in this model. This trend is more emphasised as a visitor travels to the Valley less frequently.

Another important point is the positive relationship of parking location with the WTP to road user charging. The positive relationship means that if a visitor's WTP is more than those of others, they are likely to park at the Information Centre and buy a parking ticket. This binary logit model does not show causation, so the relationship between the WTP and parking location could be the other way around. This is discussed further in the next section.

From the modelling estimation, an equity problem is clearly presented with the current monetary policy tool, a parking fee. Elderly visitors are more willing to pay the parking fees to park at the Information Centre. In other words, elderly visitors are more disadvantaged when required to pay parking fees. Also, this policy tool may affect visitors differently by their travelling origins since visitors from local areas travel more frequently to the Valley than other visitors do. This example shows how a monetary policy tool causes an uneven effect to visitors.

6.4.3 Relationships of parking locations with willingness to pay road user charging

The decision on parking or not parking at the Information Centre shows the positive relationship with the WTP to the road user charging. Additionally, income level has a positive relationship with the WTP. The factors of age and the visiting numbers, which show a significant relationship with parking location, are not found significant with the WTP. No interaction effects are found significant. Therefore, the fitted linear functions from the regression

Table 6.4: Regression analysis for the WTP

Coefficients:	Estimate	Std. Error	t-value	Pr($<|t|$)
(Intercept)	1.269	0.340	3.730	0.0002
Income ($\times$£k/yr)	0.030	0.009	3.187	0.0017
Park location	0.642	0.293	2.190	0.0297

Overall R^2: 0.0747

model are below:

$$\begin{aligned}\text{Not park at the Centre: WTP} &= 1.269 + 0.03(\text{Income})\\ \text{Park at the Centre: WTP} &= 1.911 + 0.03(\text{Income})\end{aligned}$$

This means that visitors who parked their car at the Information Centre pay 64.2 pence (£1.911 minus £1.269) more than the others do for road user charging. This looks like a small amount; however this is not negligible since planned road user charging (toll) will be no more than £3 (Derbyshire County Council, 2003, per. com.), i.e. the effect of parking location is at least a quarter of the toll. The positive coefficient of income is a standard effect for any WTP analysis. If visitors earn more income, they do not mind paying a few extra pounds to visit where they want. The difference between the two groups is

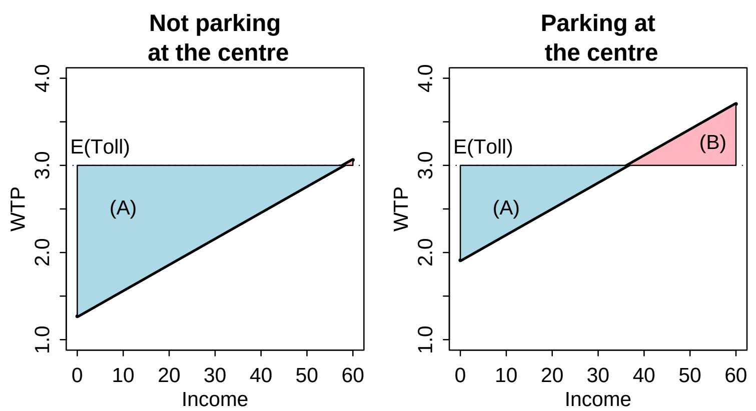


Figure 6.8: Difference between the WTP and maximum toll fee

more remarkable in Figure 6.8. For example, if the local authority charges toll fare at the maximum planned amount (i.e. £3), the user deficit[6] of visitors will be generated as shown in the triangular zones (A) in Figure 6.8. The WTP to the road user charging scheme is lower than the toll fare at any given income level in the triangular zones. Similarly, the other triangular zone marked (B) is user surplus[7] since the user pays less than what they are actually willing to pay. Then, the user deficit is much larger in the group of "Not parking their car at the Information Centre". The user surplus for this group (i.e. left side of Figure 6.8) is almost invisible in this scenario, i.e. suggesting that the £3 fee is too expensive for this group.

Consequently, the effects of the road user charging scheme will not be the same for all visitors to the Upper Derwent Valley. Some visitors avoid paying a parking fee and so they will be more reluctant to return to the Valley after the implementation of the scheme. This is a problem of the road user charging scheme, i.e. everyone pays the fixed fee no matter what is your WTP is (Eckton, 2003, p.309). This will be discussed in more detail in Section 6.5.2.

In conclusion, the parking fee scheme at the Information Centre had an uneven effect on visitors by age, visiting frequency, and the origin of travel. Similarly, the WTP effect was different between visitors parking at the Information Centre and at the other parking areas. The income effect of the potentially new policy, namely the road user charging scheme, was not observed. Therefore, this policy has a potential equity problem. Moreover, the results of this section showed the strong positive relationship between the WTP and parking locations that were also related with the parking costs. These two policy tools have the problem of double charging, i.e. the visitors, who are willing to pay a toll

[6]User deficit is the sum of price paid minus WTP in a population where WTP is lower than the price.

[7]User surplus is the sum of WTP minus price paid in a population where WTP is higher than the price.

fee, are likely to pay a parking fee. In the next section, the travel behaviours with the potential road user charging scheme and the complementary park & ride scheme are analysed with stated preference data.

6.5 Potential effects of the road user charging and the park & ride schemes

6.5.1 Results from the multinomial mixed logit model of mode choices

This section analyses potential mode choice of visitors after implementing the road user charging and park & ride schemes. The results in this section is a key output to compare and integrate with that of agent-based modelling in Chapter 8.

Overall, 48 respondents did not answer the section of stated preference questions properly; so 275 questionnaires were used for this analysis. The respondents were asked to answer their preferences in the questions by ranking them (i.e. the most preferred option is 1 and the least one is 3); however, the data set was transformed to binary data (i.e. the most preferred option is 1 and the rest of the options are 0). All possible combinations of models with several input variables and alternative-specific constants (ASCs) were tested.

The correlations among three alternatives were insignificant, so that nested logit[8] and the error component model of mixed logit model[9] were also insignifi-

[8]The nested logit model is derived when the random components have identical and non-independent Gumbel distribution, i.e IIA is invalid. The nested logit model solve this problem with decision-tree like structure. The assumption of independence of choices is retained only at each single node of the tree (McFadden, 1981). The structure is similar to the decision tree, but the decision may not necessarily to follow the sequences of a tree.

[9]The error component model tries to capture the correlation between alternatives, which share unrevealed attributes. Therefore, the idea is similar to that of the nested logit model explained above.

cant. The reason for the insignificant correlations among alternatives could be due to a simultaneous decision making process since destination (Trip | Cancel) and mode (Bus | Auto) are likely to affect the processes simultaneously in this situation (Steiner and Bristow, 2000, p.99).

The heteroscedastic taste of time and cost with multinomial mixed logit model are significant, but no socio-economic factors are significant. Possibly, socio-economic factors are efficiently captured by the taste variation of the mixed logit model[10].

The insignificant group size can be explained by the discussion on the marginal or average road pricing principle (Nash, 2003; Rothengatter, 2003). In this case, road user charging seems to be as effective as the marginal pricing principle, so that the additional trip members is not as important as the first member to calculate the travel cost, i.e. the toll is not simply divided by the number of trip members.

Recent studies show Johnson's S_B distribution[11] for the taste variation is more consistent with the microeconomic theory due to "the non-zero probability of positive coefficient" for travel related variables (Train, 2003; Hess et al., 2005). Although the normal distribution is most commonly used for the mixed logit model, the taste variation reaches the unexpected range from the microeconomic theory since the distribution is unbounded. In other words, with the normal distribution, the travel time and cost can positively affect the choice of Auto for some individuals. However, the normal distribution is unbiased in a mathematical view and the accuracy to estimate an expected value could be better than the Johnson's S_B distribution (Hess et al., 2005, p.233). The mixed logit model with S_B distribution was not significant in this analysis, so the multinomial mixed logit model with the normal distribution

[10]See the explanation on page 21

[11]Johnson's S_B distribution is similar to the logit transformation of the Normal distribution, but the distribution is bounded by the upper and lower limits.

was selected for analysis. Moreover, the lagged dependency from former to successive questions was inevitable in this situation since the data were collected by a stated preference survey, so that a panel data structure was also applied (Honore and Kyriazidou, 2000). The best-fitted utility functions with the multinomial mixed logit model are:

$$\text{Auto: } V_i^A = \alpha^{\text{Auto}} + \beta^{\text{cost}}(\text{Toll} + \text{Parking fee}) + \beta^{\text{time}}(\text{Search \& walk}) \tag{6.3}$$

$$\text{Bus: } V_i^B = \beta^{\text{cost}}(\text{Bus fare} + \text{Parking fee}) + \beta^{\text{time}}(\text{Headway}) \tag{6.4}$$

$$\text{Cancel: } V_i^C = \alpha^{\text{Cancel}} \tag{6.5}$$

The log-likelihood ratio test (McFadden, 1974), which compares the model fitness of a generic model with that of a specific model, showed that the parameters for costs and time are generic. Thus, no alternative specific coefficient is present in these utility functions. Also, the test showed the multinomial mixed logit model significantly improves the model fitness compared with the conventional multinomial logit model. These functions do not show a mean and standard error for the coefficients of cost, time, and lagged dependent variable, but these are expressed in the summary statistics for the estimates of the utility functions (Table 6.5). The cost and time attributes are presented in pounds and minutes, respectively.

For example, the logistic form of the fitted model for choosing the Auto option is:

$$P(\text{Auto}) = \frac{\exp(U_i^A)}{\exp(U_i^A) + \exp(U_i^B) + \exp(U_i^C)}$$

The other logistic forms for the other two options are similar to the one above. As mentioned before, U_i^* is unobserved utility including an error term, i.e. $U_i = V_i + \varepsilon_i$. The base unit of this model is a visitor group with an average

Table 6.5: Multinomial mixed logit model with normal distributed taste and panel data structure for mode choice. m and σ represent the mean and a standard error of a coefficient, respectively.

Coefficient		Estimate	Robust Std. Error	Robust t-value
α^{Cancel}		-4.627	0.299	-15.497
α^{Auto}		1.873	0.141	13.277
β^{cost}	m	-0.704	0.040	-17.463
	σ	0.089	0.043	2.058
β^{time}	m	-0.051	0.003	-15.019
	σ	0.025	0.004	7.120
lagged	m	0	-	-
	σ	3.070	0.222	13.819

Number of observations = 3840
$L(0)$ = -4218.67
$L(\hat{\beta})$ = -2730.05
$\bar{\rho} = 0.351$

membership of 2.99 people, but not an individual. Modelling based on individuals could be preferable. However, the same visitor group travelled together by the same mode, and the purpose of road user charging scheme is to reduce the number of cars going to the Valley. So, the model based on visitor groups is sensible for the purpose of this road user charging scheme. All six coefficients have no significant correlation with one another as calculated by the robust t-test.

6.5.2 Explanation and discussion on outcomes

The positive α^{Auto} shows a preference of Auto when the rest of the all remaining variables are constant. Similarly, the strong negative α^{Cancel} shows that the trip to the Valley is of great value to visitors. α^{Auto} is different from the average WTP for toll fee in Section 6.4.3. This means that the decision making process of visitors is rather complex and involves many factors including travel time and cost. As expected, the trip related time β^{time} and cost β^{cost} coefficient have negative signs. The higher the cost or time of an options is, the lower the utility is. Consequently, an option with strong negative coefficients is less likely to be chosen. The standard errors express that the probabilities of negative coefficients are $> 99.99\%$ for β^{time} and 97.93% for β^{cost}. Therefore, the problem of positive coefficient is negligible. The positive coefficients are irrational, but the probability of irrationality is small enough to be expressed by the mis-perception and mis-calculation. In addition, the positive time coefficient can be explained by the pleasure of walking and driving (Mokhtarian and Salomon, 2001; Redmond and Mokhtarian, 2001). For example, some visitor may enjoy walking time from the third parking area to the Information Centre. A lagged dependent variable is fixed to zero so only the standard error is estimated[12]. In this case, the value of time was 7.24 pence per minutes[13]. This is close to the non-commuting values of time in the report from the Department for Transport, i.e. 7.55 pence per minute[14] (Department for Transport, 2004).

Figure 6.9 visualises the results from the multinomial mixed logit model. Toll and searching & walking time are the concerns of this project, so the figure shows the probability of each mode that is chosen according to the change in these two variables. The parking fee for the Bus option is 50 pence, bus fare is 50 pence, and headway is 30 minutes according to the interview from the local authority. The Auto option has a negative trend, and the Bus and Cancel options have a positive trend against the toll fee. As the toll rises, visitors are likely to stop using their private cars and start using public buses to get to the Valley. Simultaneously, some visitors decide not to go to the Valley and go the

[12]The purpose of a lagged dependent variable is to control the variation from one question to another within an individual. Therefore, this should not affect the result between individuals. Hence, the mean value is fixed to zero

[13]Value of time = $\beta^{\text{time}}/\beta^{\text{cost}}$; so, -0.051/-0.704 = 7.24 pence per minute.

[14]£4.46 (non-working hour in 2002 price) / 60 minutes $\times$ 1.0158 (non-work value of time growth from 2002 to 2003) = 7.55 pence per minute

somewhere else or stay at home. However, this trend is not as strong as that of a mode-shift from Auto to Bus. Additionally, when the left side of the graphs (i.e. A, C) and the right side of the graphs (i.e. B, D) are compared, the strong effect of parking fee are recognised, which favours the Bus option. The effect of "searching time for a parking space and walking time to the Information Centre" works in a similar way as the effect of parking cost. Therefore, visitors in the top right graph (B in Figure 6.9) use the Bus option relatively more than the visitors in the bottom left graph (C in Figure 6.9) at any given toll level. All three effects seem to show sensible results in the situation of the Upper Derwent Valley.

The road user charging and the supplemental park & ride schemes are not yet put into effect, so strictly we cannot assume "Travel by car = Auto option". Therefore, the result may not be able to describe the current situation of travel behaviour around the Valley. Nevertheless, the travel behaviour at the Information Centre should be similar to the bottom right graph of Figure 6.9 (D). The parking behaviour of avoiding any parking fee at other parking area should be similar to the top left graph of Figure 6.9 (A). A visitor, who parks at the Information Centre, pays a parking fee and spends nominal time for searching and walking. Therefore, the visitor is more likely to change a travel mode from Auto to Bus compared to the other visitors arriving directly at the other parking areas.

After the implementation of schemes with the £3 toll fee, the probabilities of travel mode are the ones shown in Table 6.6. More than half of visitors, who used to park at the Information Centre, are likely to keep using their own cars to visit the Valley. In contrast, more than two third of the visitors, who used to park at the other parking areas, are likely to keep using their own cars to visit the Valley. This result shows that the effect from the road user charging scheme is not equal for all types of visitors. Elderly and infrequent visitors

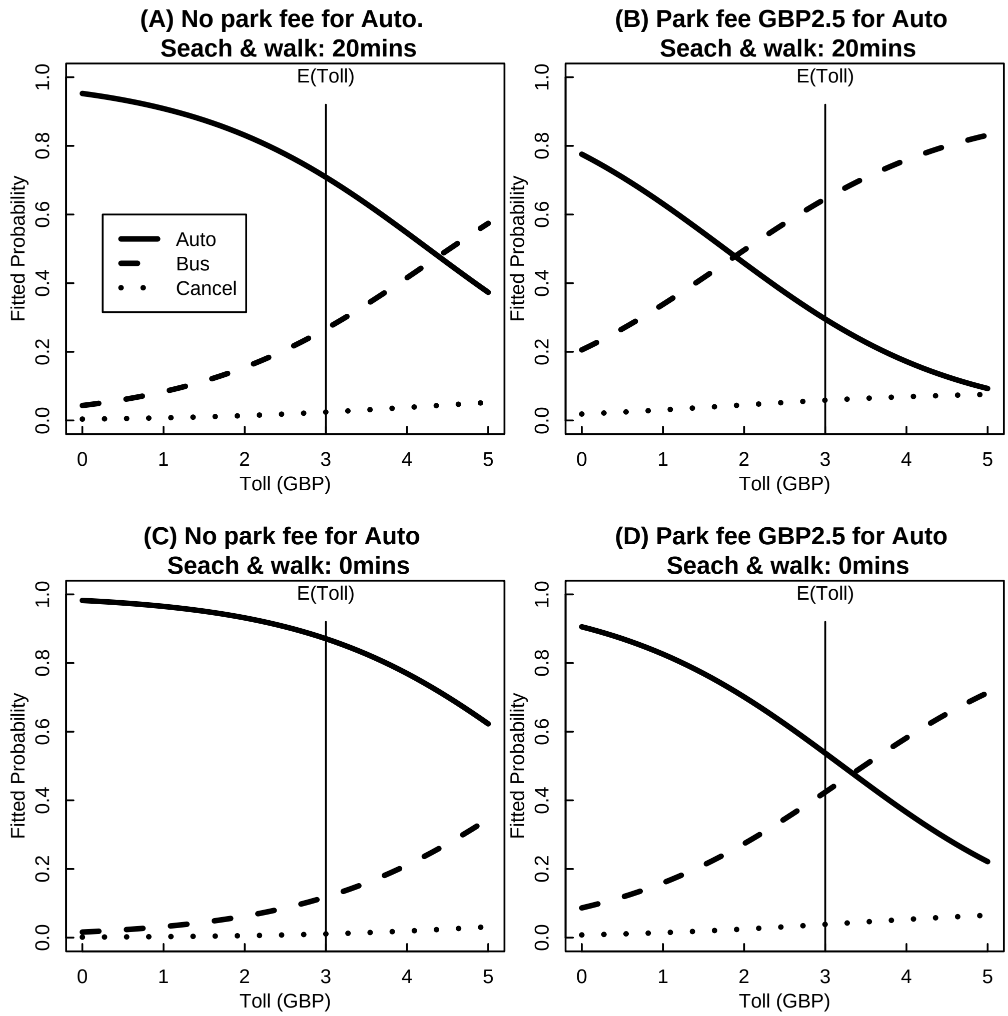


Figure 6.9: Results of logit model

are more likely to be affected by the road user charging scheme. They are most likely to park at the Information Centre, so that they are more likely to change their travel mode from private car to public bus than the other types of visitors are. Therefore, the road user charging scheme has a possible equity problem as Eckton (2003) suggested. On the other hand, the purpose of the road user charging scheme is to maintain the congestion level around Derwent Lane, and it is, consequently, effective to achieve this policy aim.

Table 6.6: Expected probabilities of each mode choice between parking locations. W + S stands for search and walking minutes and Parking means parking fee for the Auto option.

Park at	**Toll**	**W + S**	**Parking**	**Probability**		
	(pounds)	(mins)	(pounds)	Auto	Bus	Cancel
Centre	3	0	2.5	0.54	0.42	0.04
Other	3	20	0.0	0.71	0.27	0.02

6.6 Conclusion

This economic analysis of road user charging, first, identified the characteristics of one-day trip visitors to the Upper Derwent Valley. Most visitors come from local areas and two neighbouring cities and a large number of visitors are families and elderly people. Second, the characteristics of visitors were different among different parking areas. The parking locations were strongly correlated with the current policy tool, parking fee, so the policy tool affects visitors differently. Finally, the analysis with the stated preference data and the WTP to road user charging showed that the proposed road user charging and the park & ride schemes had equity problems. The equity problem is discussed again in the agent-based model analysis in Chapter 8 with the scenarios of elderly exemptions for the toll fee.

On the other hand, this econometric approaches, which conduct analysis solely with discrete choice models, has significant shortcomings. The multinomial mixed logit model for mode choice could not utilise the socio-economics characteristics of visitors in this case study although these characteristics were found significant in the parking location choice model.

Furthermore, the analysis solely with discrete choice models lacks the mechanism to observe the dynamic process of travel behaviour. This mathematical model does not show how visitors change the trip modes after the implementation of the schemes. Moreover, since any model about the parking network in the Upper Derwent Valley is absent in this analysis, the congestion level at each parking area is unclear. For example, we assume, by using the parking location model, that a visitor can definitely park at the Information Centre if the visitor decides to park there in the parking location model. This is because these models cannot formulate the concept of congestion, which requires dynamic interaction among the visitors – an example of "*Oversimplification*" (Stopher, 2004). These problems are overcome by the agent-based model, which dynamically models the traffic situation in the Upper Derwent Valley and considers visitor-interactions with the discrete choice models. The parking network is modelled in the next chapter.

Chapter 7

Markov-queue-like microsimulation of dynamic parking network

7.1 Introduction

This chapter explains the parking network modelling of the Upper Derwent Valley. The two main objectives of this chapter are: 1) to model a dynamic networked parking model, which represents the situation of the Upper Derwent Valley, and 2) to determine the searching time of a parking space and walking time between the parking space and the primary destination assumed to be the Information Centre. The parking microsimulation model formulates congestion levels, searching time, and walking time for agents and is used as a module in the agent-based simulation model in the next chapter.

The reason for the first objective is that the previous chapter, which conducted analysis solely with discrete choice models, did not explain the mechanics of a parking network system in the Valley, which is the necessary concept for assuming parking congestion and the interaction of agents. The reason for the second objective is that it is infeasible to collect dynamic searching time and walking time from surveys. Although some studies have used survey techniques to identify searching time (Axhausen and Polak, 1991), this approach cannot be applicable to the current study. From pilot survey, car drivers were found not to remember the exact searching time and walking time, or only answered the approximate time, such as 5 or 10 minutes. Known as a complex problem (Fortin and Rousseau, 1998; Casini and Macar, 1999; Bugmann and Coventry, 2004), the abstract cognition on travelling and walking time could be a D.Phil. topic. Additionally, conventional econometric search models with survey-based data are known to have problems due to the assumption[1] made in the econometric model (Thompson and Richardson, 1998, p.163). Moreover, it is not feasible to follow cars, which enter the parking network system, until they found a parking space to measure the searching time. The parking network system is defined as the series of four parking areas in the Upper Derwent Valley, which work like a system at a macro level view. While following a car, a researcher loses track of other cars, so that this approach needs an unnecessary large number of researchers. The survey with GPS tracking devices will be the way for future research, but it is also infeasible for the current research.

In contrast, arrival rates into the parking network system are easy to measure at the input point of the system. Parking time is also obtained easily by questionnaires since car drivers have a good memory about the arriving time at a parking area and departure time from a parking area. With the Markov queue theory (Chernick, 1999; Hinkley, 1988), departure rates are calculated

[1] This approach assumes that decision makers are risk neutral, have an unlimited time budget, face constant search costs, have full recall and possess a perfect knowledge of the utility distribution. However, since parking areas may be only temporarily available, the rejection of a parking area at any point in time means that it may not be available when and if the visitor decides to return to it at a later stage in the current trip. This general 'lack of recall' relating to the availability of previously inspected car parks, results in the current alternative being the most appropriate basis of comparison in the calculation of the expected gain in utility when deciding whether or not to continue the search.

from parking hours, i.e. the departure rate of the parking network system is the inverse of parking time. Simultaneously, the congestion level of a network system can be estimated with the Markov queue model. The major factor, which determines searching time, is a congestion level in a parking area, so searching time can also be estimated with the Markov queue model. A parking location is determined after a car driver finds a parking space and consequently walking distance and time are approximated though this process.

The mathematical Markov queue model of a parking network was considered first. However, previous studies showed that it was difficult to implement this approach to solve real world problems, so that simulation approaches had been recommended (Arnott and Rowse, 1999). The simulation model, which describes the parking network system of the Upper Derwent Valley, is based on a Markov queue model. This is one of the most popular stochastic models (Ripley, 1987, p.104) and one of the two most popular microsimulation transport models (Nagel, 2004). A Markov queue model is influenced only by a defined process and a current state, but not by past states to forecast the next state. This is called *Memoryless property* (Norris, 1997). In most cases, the process is based on an arrival rate to a system and a departure rate from a system. Additionally, time-driven and event-driven concepts are added to the Markov queue simulation model and these concepts are explained from Section 7.5.2. First, this chapter estimates the arrival rate into the parking network system and departure rate from the system. Then, the statistical simulation of a Markov queue model is used to duplicate the dynamic parking situation of the Upper Derwent Valley. Following this, the simulation model is used to find the searching time and walking time of car drivers. As discussed in the previous chapters, this parking model also focuses on the holiday periods.

Base unit of microsimulation

This chapter uses 'car' instead of 'visitor' as the base unit of the simulation model to distinguish the entities between them. Visitors in the last and next chapters have characteristics and are more intelligent than the cars in this chapter. Therefore, the visitors are called agents in the two chapters. In contrast, the base unit in this chapter is a homogeneous particle, but not an intelligent agent. Therefore, this simulation model is microsimulation, but not agent-based simulation model (See Section 5.4.1). Having said that, some terms like 'walking time of a car driver' are used in this chapter since 'walking time of a car' does not make sense.

7.2 Background of Markov queue model

The networked queue model (Hopcroft and Ullman, 1979) and cellular automata[2] are the two most popular microsimulation approaches (Nagel, 2004). The advantage of the queue model is simplicity and consequently the simulation speed (Raney et al., 2003). This was the reason this study implemented the queue approach. In addition, this is the same rational for DynaMIT[3] (Ben-Akiva, 2005), which is one of the most well-known software tools for microsimulation in transportation research. The disadvantage of the networked queue approaches is the limitation in modelling intersections and inhomogeneous cars (Febbraro and Sacco, 2004). Therefore, the cellular automata approach is more appropriate for detailed study of car movements (Lee et al., 2001). However, the main problem is speed. For example, the simulation of this study had to be run much faster than that of Lee et al. (2001) since a time-step in the model developed here was equivalent to an entire simulation period in the work of Lee

[2]See the footnote on page 92

[3]`http://mit.edu/its/dynamit.html`

et al.. This means that the current model needs to run the other simulation process a thousand times to complete one simulation process.

Moreover, from another perspective, the queue model is more realistic than the cellular automata when the approach is combined with the event-driven simulation, which is explained in detail in Section 7.5.2. The event-driven networked queue model uses a continuous time-line. In contrast, the cellular automata model has to use a discrete time-line, e.g. moving every 10 seconds. This is the reason for the cellular automata being significantly slower than the queue model (Nagel, 2004; Febbraro and Sacco, 2004). Furthermore, the event-driven approach is extremely fast in the single CPU environment, which is the situation of this study.

7.3 Distribution of data on arrival and departure rates

As stated above, the research started with identifying the arrival and departure rates for the Markov queue model. In Markov processes, the distribution of the arrival and departure rates are usually considered as coming from the Poisson distribution. In other words, the main purpose of this section is to validate Markov queue modelling. The distributions of rates are checked if they come from the underlying distributions by the bootstrap method, which treats the collected data as an hypothetical population and re-samples from it (Efron, 1979; Davison and Hinkley, 1997). As well as other statistical analysis, the bootstrap simulation method was conducted with R[4] (Ihaka and Gentleman, 1996).

[4] `http://www.r-project.org/`

7.3.1 Data collection and description of field

The departure rates at each parking areas were collected from parking beat surveys, which were undertaken over three days, namely the 23rd, 26th, and 27th of August 2001, by the Transport Office of Derbyshire County Council. The 23rd of August 2001 was a normal summer weekday, in contrast, the 26th and 27th of August 2001 were a Sunday and the summer bank holiday, which were usually the busiest days in the Valley. All parked cars were recorded, so the data set acted as a population. In total, 1961 cars were recorded. When the data was tabulated by the local authority, the fraction (minute) of the time was rounded, so the time showed only parking hours, but not minutes.

The rest of the data was collected during August, 2003 for 10 days. From the observations, the steps of the parking network are determined as below: 1) if a visitor decides to go to the Valley by car, the car enters the system from either the Information Centre or the second parking area (Derwent Overlook), 2) if the parking area is full, the car moves to the next parking area, and so on, and then 3) when the visitor decides to go home, the car goes home without entering any other parking areas (Figure 7.1). Therefore, there are two input

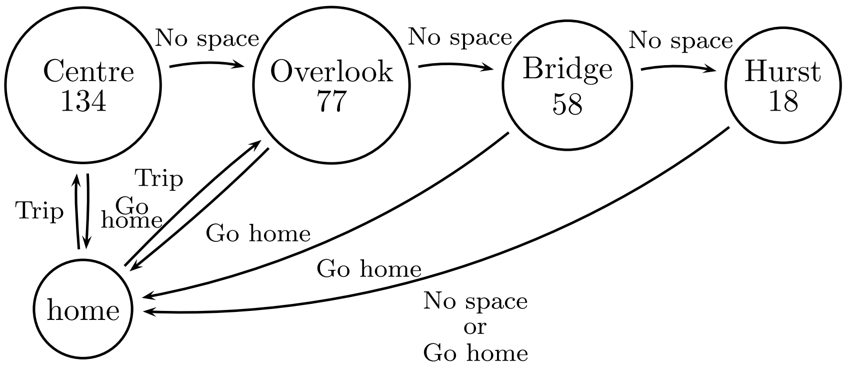


Figure 7.1: Markov queue network. The numbers in circles are parking capacities.

points (i.e. Information Centre and Derwent Overlook) and four output points (i.e. all four parking areas) in Figure 7.1. The state transition in the system at a macro level is represented as Figure 7.2. The arrival rate and the departure

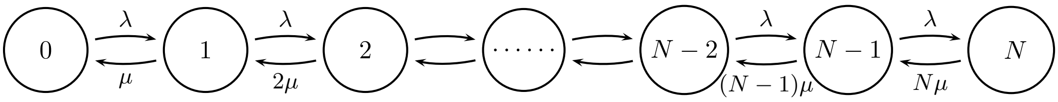


Figure 7.2: State transition diagram of the system

rate of car per minute are symbolised as λ and μ, respectively in this study. The number in the circle is the number of cars in the parking network system and the N is the overall parking capacity, i.e. $134 + 77 + 58 + 18 = 287$.

Arrival rate per minute was directly counted at the Information Centre with 30-minute intervals (30λ) between 10:00 and 15:00. As explained above, there were two input points in the parking network system and the arrival rate of Derwent Overlook was not large from the direct observations. It was therefore assumed it was the fraction of the arrival rate at the Information Centre and there were some concerns about the reliability of true observation. The simultaneous surveys on the arrival rate at the each parking area might be considered for future research, but they were infeasible for the current research. The data contains some mission values, so the total number of observed periods was 84. The departure rate per minute was calculated as the inverse of parking hour, which was collected by the parking beat surveys.

7.3.2 Arrival rate to the Upper Derwent Valley parking areas

Differences in the distributions of arrival rates between holiday periods and normal weekdays were examined (Figure 7.3 & 7.4). Non-parametric Wilcoxon rank sum test showed that the distributions were significantly different with 99% confidence. The arrival rate during holidays has a trend with the time of day. The arrival rate peaks around noon and gradually decreases afterward (Figure 7.3). On the other hand, there is no obvious trend against the time of day on Figure 7.4. This research focuses on the parking network system during holiday periods, so that only arrival rates during holidays are discussed after this point.

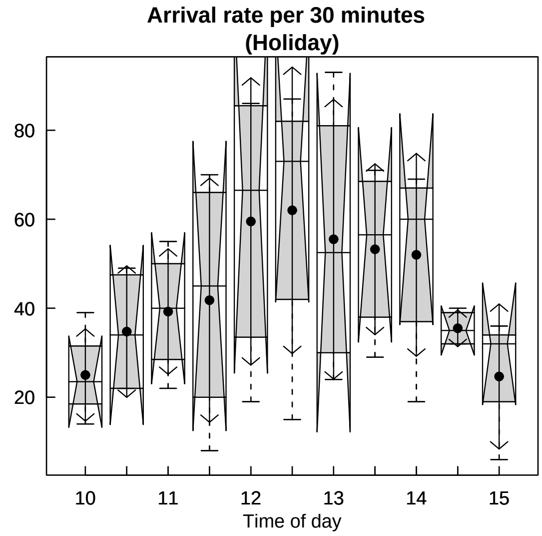


Figure 7.3: Boxplot and associated means and standard deviations of arrival rate during holidays

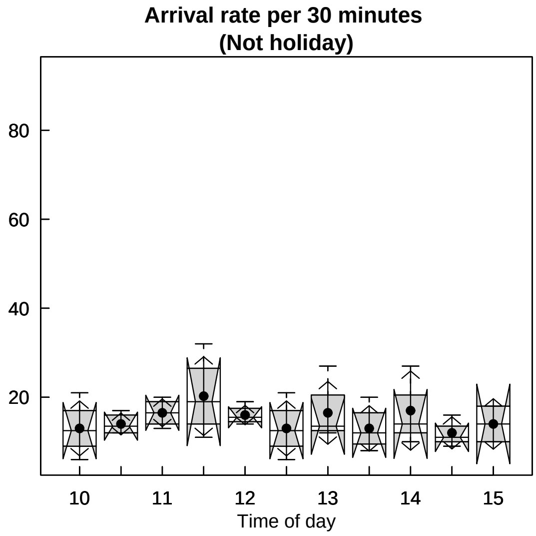


Figure 7.4: Boxplot and associated means and standard deviations of arrival rate during non-holidays

A commonly assumed distribution for counted data is the Poisson distribution (Pfeiffer and Schum, 1973, p.200), and if this assumption is valid, the Fano factor should be one, i.e. $\phi = \sigma^2/m = 1$ (Stevens and Zador, 1998, p.213). The bootstrap simulation method (Hinkley, 1988; Chernick, 1999) was used to estimate the Fano factors. Eight out of eleven bootstrapped ϕ for each observed time of day have confidence intervals containing '1' by the percentile method[5] (Table 7.1). Therefore, the evidence shows that the distributions of arrival rates during holidays come from the Poisson distribution.

A triangular function fits with the time dependency of arrival rates, $30\lambda_t$.

[5]Here the 100(1 - α)% confidence interval is simply given by the $\alpha/2$ and 1 - $\alpha/2$ quantiles of $T(X^*)$, which will be denoted by $t^*_{\alpha/2}$ and $t^*_{1-\alpha/2}$ respectively. That is, the 100(1 - α)% percentile bootstrap CI is given by $(t^*_{\alpha/2}, t^*_{1-\alpha/2})$. The bootstrap provides estimates of $t^*_{\alpha/2}$ and $t^*_{1-\alpha/2}$, i.e. the observed $\alpha/2$ and $1-\alpha/2$ quantiles of $T(X^*)$ from the bootstrap simulations.

Table 7.1: Fano factor for the arrival rate during holidays

Time of day	Boot. ϕ	Std. error	Quartiles (0.975)	(0.025)	Contain '1'
10:00	3.17	2.05	7.72	0.01	Yes
10:30	4.80	2.03	6.85	0.00	Yes
11:00	3.93	2.35	9.00	0.48	Yes
11:30	15.45	7.46	29.35	1.73	No
12:00	14.25	8.50	30.68	0.00	Yes
12:30	14.72	11.13	33.88	0.21	Yes
13:00	13.47	7.07	27.13	1.09	No
13:30	5.40	3.16	11.16	0.09	Yes
14:00	8.50	6.12	18.94	0.06	Yes
14:30	0.36	0.16	0.59	0.00	No
15:00	8.63	6.47	18.75	0.00	Yes

The function for $30\lambda_t$ is:

$$30\lambda_{\mathsf{hour}} = \begin{cases} \alpha^1 + \beta^1 \times \mathsf{hour} & 10{:}00 \leq \mathsf{hour} \leq 12{:}30 \\ \alpha^2 + \beta^2 \times \mathsf{hour} & \text{otherwise} \end{cases} \tag{7.1}$$

Where: hour is time of day and its interval is [10:00, 15:00]

Since the distribution is obviously heteroscedastic (Figure 7.3), the weighted least squares method (i.e. dividing each observation by the variance of the error term for that observation) is used to fit the linear models of the equation (7.1). The estimation of all parameters in equation (7.1) is significant with more than 99% of confidence (Table 7.2). The p-value for the model fitness is also significant at the 99% confidence level and the model fitness of R^2 is high, i.e. 0.92 for $10{:}00 \leq t \leq 12{:}30$ and 0.86 for $12{:}00 < t$. In addition, the large standard deviation around noon is explained by the property of the Poisson distribution (Figure 7.3). The standard deviation of the Poisson distribution is positively related with mean, so the larger the mean a distribution has, the larger the standard deviation that distribution has.

From the results, we can reasonably assume that the distribution of arrival

Table 7.2: Significance of triangular function of arrival rates, $30\lambda_t$

	α		β		Overall	
t	Est.	p-val.	Est.	p-val.	p-val.	R^2
10:00 $\leq t \leq$ 12:30	-129.18	0.008	15.36	0.002	0.002	0.92
otherwise	225.95	0.003	-12.98	0.008	0.008	0.86

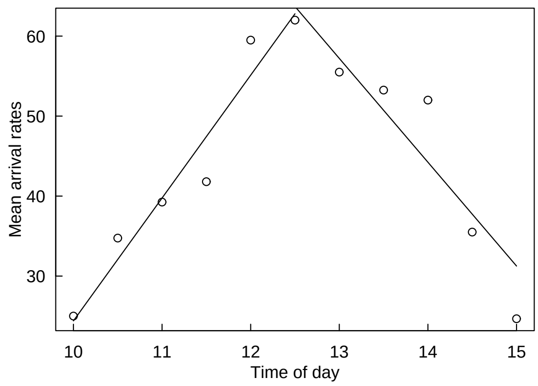


Figure 7.5: Means of time dependent upon arrival rate and a triangular function with the peak at 12:30

rates come from a time dependent Poisson distribution. Therefore, the equation (7.1) was used to produce the arrival rate in the Markov queue model of the parking network system.

7.3.3 Departure rate from the Upper Derwent Valley parking areas

The main objective of this section is to estimate the departure rate from the Upper Derwent Valley, which is defined as the expected number of cars leaving the Valley per hour. Moreover, departure rate is defined as the inverse of parking hour, so that this section starts with the analysis of the parking hour distribution.

The distributions of the parking hours between dates and parking areas were checked to see if they were different. Overall, the distribution of the parking hour is skewed to the right (Figure 7.7). The number after `D` means a sampled date and the number after `S` means a parking area from the Information Centre, e.g. `D23S1` means the parking hour at the Information Centre on the 23rd of August 2001. The boxplots show the difference in the parking hour among the categories more clearly. If the notches of two boxes are overlapped, the expected parking hours are not significantly different between any two categories. Eight medians out of 12 categories are the same, i.e. vertical lines in the middle of boxes overlap. The medians of the third and fourth parking areas on the 23rd are lower and those of the second and fourth parking areas on the 26th are higher than the other median. These differences could be a true difference. However, the differences are inconsistent within the same site and date, so that only the overall parking hour distribution was used to estimate the departure rate.

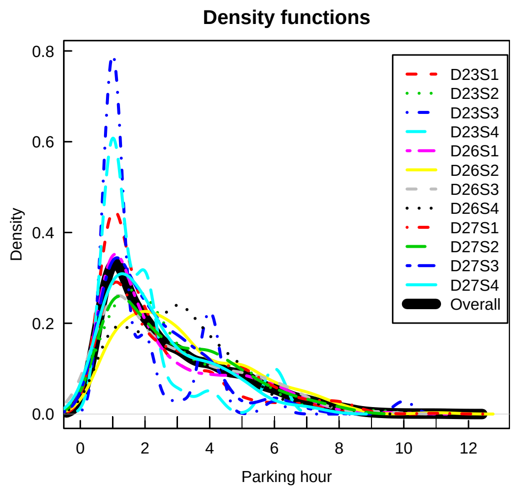


Figure 7.6: Density of the parking hour in each site and date

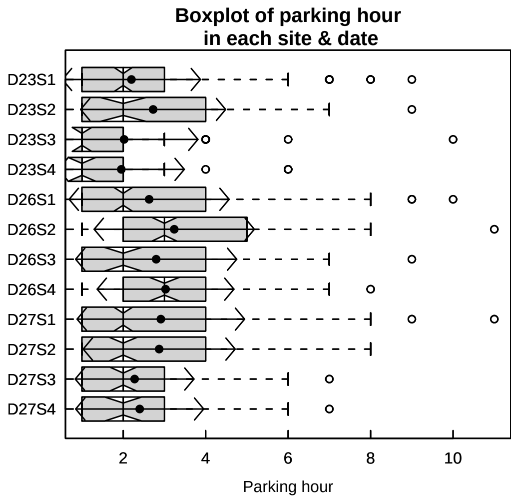


Figure 7.7: Boxplot of the parking hour in each site and date

Since the expected departure rate per hour (60μ) is the inverse of the parking hours, it is $1/2.736 = 0.367$ in this case. The bias from the non-linear transformation was found very small (< 0.00001) from a bootstrap simulation, so a correction for the parameter estimation is not necessary. Therefore, the departure rate per hour from the Upper Derwent Valley was determined as

0.367.

Moreover, in the Markov queue process, the service time (in this case, parking hour) is distributed exponentially (Norris, 1997, p.182). The exponential distribution with a rate of 0.367, is fairly close to the density function of the parking hour (Figure 7.8). Therefore, the departure rate is also assumed as Markov and consequently it comes from the Poisson distribution. In conclusion, the departure rate was estimated as 0.367 cars per hour and it was assumed as the Poisson distribution since the service time (parking hour) in the system came from the exponential distribution.

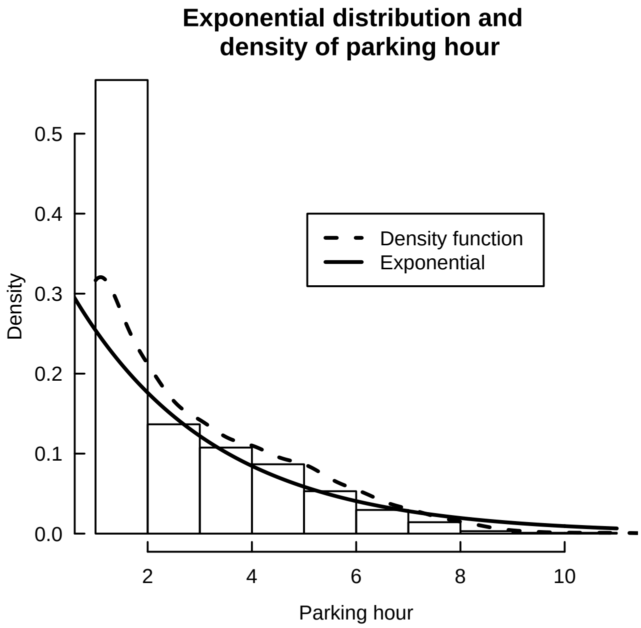


Figure 7.8: Exponential distribution of rate 0.367 and density of parking hour

7.4 Time dependent Markov Queue network simulation model

The distributions of arrival and departure rates satisfied the requirements for the Markov queue network model (Vose, 2000, p.235) and so the model is validated. Therefore, the Markov queue network model was developed to simulate the parking situation of the Upper Derwent Valley with RePast toolkit[6] (Collier et al., 2003). The main purpose of this section is to verify the simulation mechanism.

The simulation period of a day was bounded between 8:30 and 17:00 because the arrival rate was negative before and after the period. Also, the period reflected the active hours of the real parking situation in the Upper Derwent Valley. At the beginning of a simulated day (i.e. 8:30), 30 cars are assumed to be in the Information Centre. From the results in the previous sections, the arrival rate to the Information Centre is defined as:

$$\lambda^*_{\mathsf{hour}} = \begin{cases} -4.306 \quad + \quad 0.512 \times \mathsf{hour} & 10{:}00 \leq \mathsf{hour} \leq 12{:}30 \\ 7.532 \quad - \quad 0.433 \times \mathsf{hour} & \text{otherwise} \end{cases} \tag{7.2}$$

The expected cars to the Valley is approximately 595 from an analytical solution[7]. In contrast, the observed car numbers are 842 and 769 on the 26th and 27th of August 2001, i.e. the expected car number in the Valley on the busiest days is 805.5 from this data. If the model estimation is perfect, 210.5 cars do not go to the Information Centre and this number is equivalent to the 35% cars approaching the Information Centre. However, this model does not capture some cars. For example, cars, which approach neither the Information Centre nor Derwent Overlook, and cars, which arrive at the Upper Derwent Valley after 17:00, are not captured in this model. Therefore, these cars, which were equivalent to 5% of cars approaching the Information Centre, were excluded from this model. Hence, this study assumes that the distribution of arrival

[6]http://repast.sourceforge.net/

[7]These are two solutions, 561.73 or 569.17, on the 26th and 27th of August. The full calculation for the analytical solution of the car number is as follows. The large triangle is $539.0858 * 2.092/2$ or $539.0858 * 2.1196/2$. The small triangles outside of [8:30, 17:00] are $5.39064 * 0.046/2 + 23.6952 * 0.171/2$. So, the expected cars are the large triangle minus the small triangles plus pre-arrived cars of 30: 591.7338 or 599.1731. The average of the two is 595.4534.

rates approaching Derwent Overlook is 30% of the ones approaching the Information Centre. The arrival rate (λ) and departure rate (μ) per minute are defined as follows:

$$\lambda_{\mathsf{hour}} = \begin{cases} 1.3(-4.306 + 0.512 \times \mathsf{hour}) & 10{:}00 \leq \mathsf{hour} \leq 12{:}30 \\ 1.3(\ \ 7.532 - 0.433 \times \mathsf{hour}) & \text{otherwise} \end{cases} \tag{7.3}$$

$$\mu = 0.0061 \tag{7.4}$$

There is no reason to assume that a car, which comes to the Upper Derwent Valley, must leave after other cars, which have come earlier, i.e. First In First Out (FIFO). Therefore, the departure from the Valley in this simulation is System In Random Order (SIRO). The simulation ran 500 time steps and the algorithm for the Markov process was based on Ross (1997, pp.88-89).

The overall car number in the Valley was 774.818 and its confidence interval [777.2413, 772.3947] captured the expected car number in the Valley, i.e. $595 \times 1.3 = 773.5$. The number of cars in each parking area was similar to the actual data except for the Hurst Clough parking area (Figure 7.9). Therefore, the model is fairly verified. The reason for the different result in the Hurst Clough parking area could have resulted from excluding non-parked cars, i.e. on average, 21 cars could not find a parking space in the simulation model. Cars are not assumed to return to a parking area, which they have searched beforehand. However, in reality, cars may return to the Valley to find a parking space after they searched through all four parking areas. Moreover, parking area capacity in practice might be slightly higher than the proposed parking capacity in this thesis due to double parking and illegal on-street parking, which are actually observed during the survey period. This part could be the weakest point in this model, and this problem can be solved in future research with extra work.

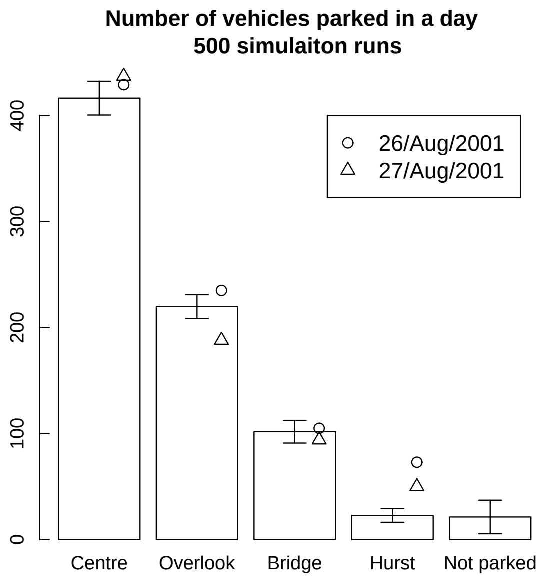


Figure 7.9: Number of cars parked in a day for 500 simulation runs. The bar graph represents the mean, and its error bar represents the standard deviation. The points are actual data from beats survey in August 2001.

Focusing on individual parking areas, the car numbers at the Information Centre and Derwent Overlook keep rising until around noon (Figure 7.10). Then, the car numbers at Bridge End Pasture and Hurst Clough shoot up due to the saturation of the previous two parking areas. Following this, these car numbers start declining after around 14:00. The mean peak time of this parking network system is 14:02 with a standard deviation of 30 minutes, although the peak arrival time is at 12:30. This is because of a buffer effect between arrival and departure. From this result, the first two parking areas are almost always saturated, approximately for two hours around 14:00 and the whole network is often saturated for one hour around 14:00.

7.5 Searching time and walking time

From this section onwards, this chapter focuses on the second aim, i.e. finding out the searching time and walking time of car drivers. An analytical approach is first carried out to show the difficulty of this problem. Then, the simula-

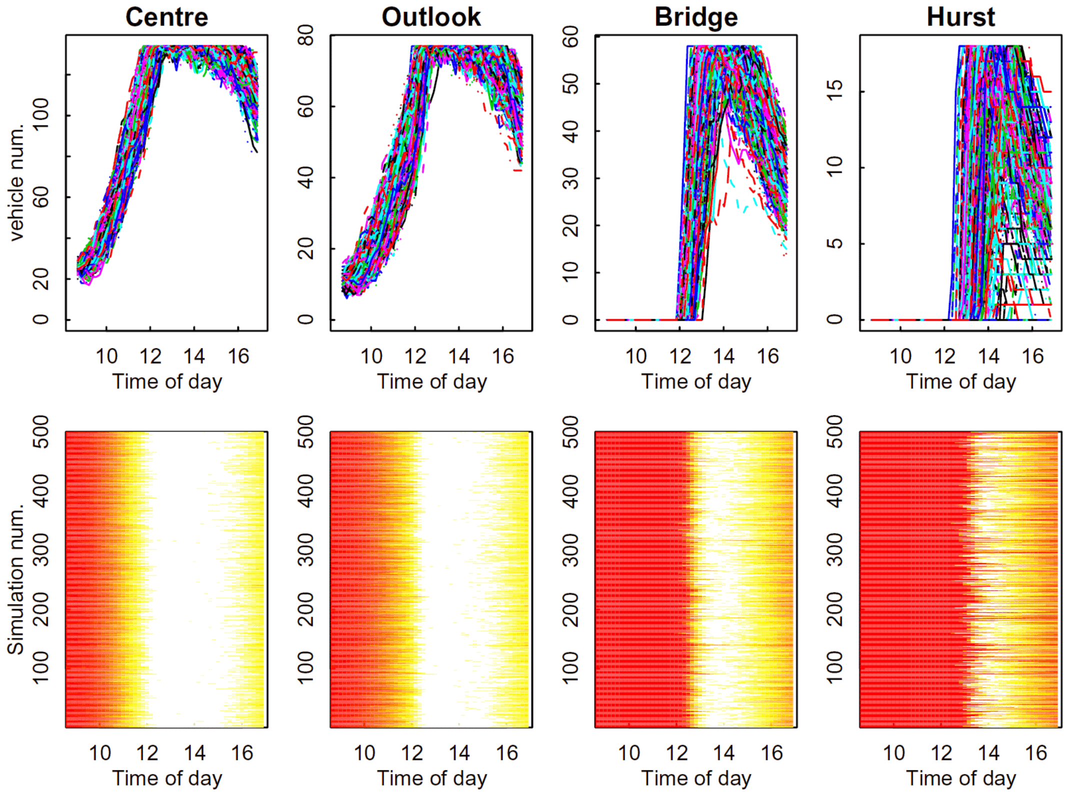


Figure 7.10: Car numbers in the four parking areas in the Upper Derwent Valley. The bottom four plots are expressed by a rule as follows: brighter (whiter) colour means more cars are present at a parking area at a given time

tion approach is implemented to solve the problem. At the same time, the mechanics of the Markov queue simulation model are explained in detail with the concepts of event-driven and time-driven approaches.

7.5.1 Analytical consideration

If the parking areas in the Upper Derwent Valley are linear, cars are likely to enter the first empty parking space they encounter. The minimum width of a parking areas space is set to 2.4 metres in the Peak District National Park (Peak District National Park Authority, 2001). Hence, the expected searching time is:

$$T_s = 2.4x/S_s + (\alpha_{1-2} \cdot D_{1-2} + \alpha_{2-3} \cdot D_{2-3} + \alpha_{3-4} \cdot D_{3-4})/S_d \tag{7.5}$$

The x is the number of parking spaces a car passed before finding an empty parking space. The S_s is the speed of a car in the parking area and the S_d is the driving speed between parking areas. The α is distance from one parking area to another parking area. The D is a corresponding dummy variables, which is '1' if a car passes the section, otherwise it is '0'. Thus, α_{1-2} applies only when a car approaches the Information Centre first and cannot find an empty space there.

The walking time function is similar to that of searching time:

$$T_w = (2.4x + \alpha_{1-2} \cdot D_{1-2} + \alpha_{2-3} \cdot D_{2-3} + \alpha_{3-4} \cdot D_{3-4})/S_w \tag{7.6}$$

The S_w is the walking speed. The only major difference between searching time and walking time is that a car driver will go to the Information Centre regardless of which parking area the car approaches first. This means α_{1-2} should be counted as a part of walking time even if cars approach from the

Derwent Overlook. Also, walking speed is assumed not to change within and between parking areas.

The equations (7.5) and (7.6) clearly show that the searching time and walking time are dependent on congestion level and the number of cars in the Valley. Moreover, the number of cars in the Valley is dependent on the time of day. For example, the expected number of cars in the Valley at the time t (minutes) in this simulation model is:

$$N_t = 30 + \int_{8.5\times60}^{t} \lambda t - \int_{8.5\times60}^{t} \mu(30 + \int_{8.5\times60}^{t} \lambda t) \tag{7.7}$$

This model uses SIRO[8], so the distribution of the empty space $x+1$ at t is uniform between zero and $N_t + 1$. Therefore, the expected empty space at t is $(N_t + 1)/2$ and consequently, the formula for expected x_t at t is:

$$x_t = (N_t + 1)/2 - 1 \tag{7.8}$$

Substituting equation (7.8) into equations (7.5) and (7.6) links the arrival and departure rates with searching time and walking time.

The following sections simulate the model based on the explanation above. One of the important facts skipped in this explanation is that there are two types of cars coming to the Valley, i.e. one approaches from the Information Centre and the other approaches from the Derwent Overlook. These two types of cars exist in the same system at the same time and affect the other type of cars. Thus, these two sub-systems should not be modelled separately. The simulation approach easily shows the results of searching time and walking time from two sub-systems, simultaneously.

Moreover, the result from the simulation can be deduced from the ana-

[8]See page 139.

lytical solution; however, the deductive approach cannot catch up with the simulation model if realistic concepts are introduced more into the simulation. For example, more complicated driving behaviour, such as following / anti-following previous car, will be coped with only by the simulation approach. Therefore, this study uses the simulation approach rather than seeking an optimal analytical approach.

7.5.2 Numeric solution from simulation

Besides the assumptions of the Markov queue model, there were six more assumptions for this simulation model:

1. Driving speed between parking areas is constant and assumed to be, on average, 20 miles or 32.187 KM per hour.

2. Driving speed while searching within parking areas and walking speed are assumed to be, on average, 3.5 feet or 1.067 metres per second.

3. Cars do not search again in the same parking area and move to the next parking area when they cannot find an empty space at the current parking area.

4. Cars take the first encountered empty space.

5. The arrival point of cars is at the entrance of Derwent Lane (the intersection from the A57).

6. The distances between parking areas are 509.90, 1,231.01, and 2,362.38 metres or 557.63, 1,346.25 and 2,583.53 yards from the Information Centre to the Hurst Clough. These numbers are estimated from map references[9].

[9]From the Information Centre to the Hurst Clough, SK 173894, SK174889, SK 180885, and SK188877

As mentioned before, the total parking capacity of the four parking areas is 287 and the width of a parking space is 2.4 metres. From these settings and equations (7.5) and (7.6), the range of searching time[10] is 0 and 18.41 minutes. The range of walking time[11] is 0 and 74.85 minutes.

Structure of combined event-driven and time-driven approach

This model is the combination of event-driven and time-driven approaches. Although the concept of the Markov simulation model in Section 7.4 was not described in detail, the concept was the same as that of the current model. The arrival and departure rates are determined according to the time of day as the Markov queue model in Section 7.4. A macroscopic timing determines the arrival and departure of cars, so these events are time-driven events for the system. Therefore, the arrival and departure are not controlled by each car in this simulation. At the same time, each car acts according to the micro level events they encounter between two macro level events. The micro level events are to enter and exit parking areas, and these events change the driving speed of cars, i.e. event-driven events for each car.

Time-driven simulation is suitable since arrival and departure are the key events in the situation of the Upper Derwent Valley. Also, by giving the authority of the timing events from each car to a macro system, the large computational overhead is reduced since the time-line of major events are clearly identified in this model (Cheng, 1998, p.235). This approach becomes more beneficial in a successive study with an agent-based model of the Upper Derwent Valley, in which the number of cars going to the Valley changes according to the number of private car visitors at every time step. In this case, only the arrival rate (λ) needs to change according to the change in car numbers.

[10]The max searching time is $2.4 \times 287/64.02 + (509.90 + 1,231.01 + 2362.38)/536.45$

[11]The max walking time is $(2.4 \times 287 + 509.90 + 1,231.01 + 2362.38)/64.02$

An event-driven approach is also suitable since the situation fits with the Petri Net[12], which is the key concept of the event-driven approach in transportation (Febbraro and Sacco, 2004, p.1226). The objective of a car in this model is to find a parking space. Events except for arrivals and departures can be concurrent and asynchronous in terms of Petri net to achieve the objective (Peterson, 1981), i.e. the event-driven events are to enter and exit a parking area. When each car enters a parking area, a car speeds down and starts searching for an empty space to park. It is too dangerous to drive fast in parking areas where there are many pedestrians as speed is strongly related with safety issues. Also, when a car cannot find a parking space in a parking area, the car exits the parking area and speeds up to get to the next parking area.

Figure 7.5.2 shows the pseudo-code for the main loop of the macroscopic timing in a simulation day (a representative day in a given week). First, *state* is a local variable about the current total number of cars in the parking network system. A variable with an under-bar-suffix is a field or global variable in the Class. Each car is created with an arrival event and the car enters the first array, which holds arrived cars (but not parked) at the Valley. There are two other arrays of cars: the second array which holds parked cars in a parking area and the third array which holds cars left in the Upper Derwent Valley. The lines numbered between 2 and 4 initialise the timings, the arrays of cars, and the state. Line 6 calculates the timing of the next event according to the merge theorem, i.e. minimum of exponential λ and μ (Jain and Neal, 2000). Within line 8, each car moves autonomously according to the event-driven events between the current time and the time the next event occurs. The next event-driven event is determined based on the relative proportion of λ and μ,

[12]Petri net is one of the discrete event system models. In this model, an event is in a discrete state (e.g. 'On', 'Off') and each event occurs at anytime (asynchronousnes) without the influence of other events (concurrency) (Peterson, 1981; Oota, 1995)

and then occurs in line 9. Line 11 finally updates the current time of day. This simulation runs until the end of a simulation day i.e. 17:00. This main loop

```
01 Start the Markov queue trip of cars
02     Initialise the current time and the next event time
03     Initialise the arrays of cars in the valley
04     Set initial state as the initial car number
05     while the current time is before the end time, repeat
06         Calculate the next event time based on state
07         if the next event time is before the end time
08             Move cars in the Valley and update state
09             Next event occurs and update state
10         End 'if' condition
11         Set the current time to the next event time
12     End 'while' loop
13 End the Markov queue trip of cars
```

Figure 7.11: Pseudo-code for Main loop of car movements.

shows that the model is time-driven at a macro view. However, between the programming codes, each car searches for a parking space and drives according to the events, so that the simulation is also event-driven.

7.6 Results and consideration

The simulation ran 500 times or 500 simulation days. Overall, both the searching time and walking time are dependent on the time of day since the parking areas are severely congested around early afternoon and so some cars have to search longer and park far from the Information Centre, as expected in the analytical solution in Section 7.5.1.

These times clearly depend on the parking location as well. The searching time (Figure 7.12) and walking time (Figure 7.13) are presented against arrival time of day. The distributions in these two figures have similar trends and are consistent with the explanation in the previous sections. There is a clear difference between the distributions of searching time and walking time. Searching time has two clusters in each parking area except for the Information Centre, but this does not exist in walking time. As explained in Section 7.5.1, searching time is different between a car approaching the Information Centre and a car approaching Derwent Overlook, and then, this produces two clusters within a parking area. In contrast, the entrance points do not matter for walking time since it is determined purely by the parking location.

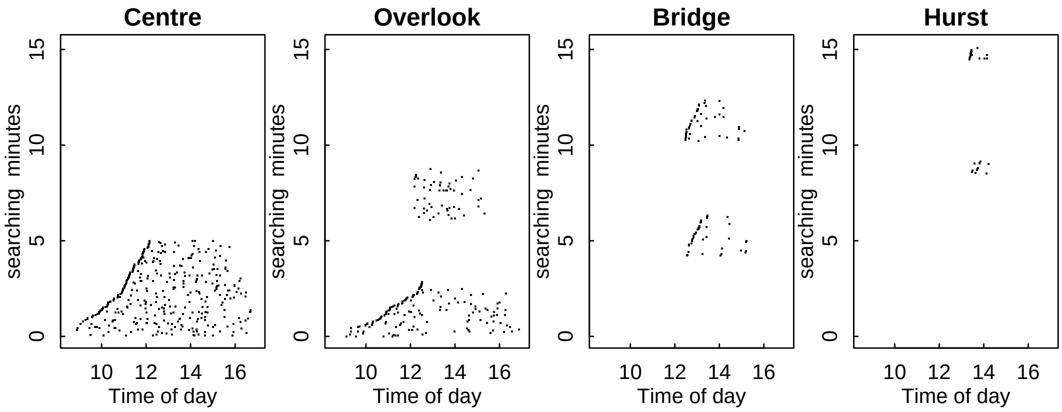


Figure 7.12: Searching time to find a parking space against the arrival time of day

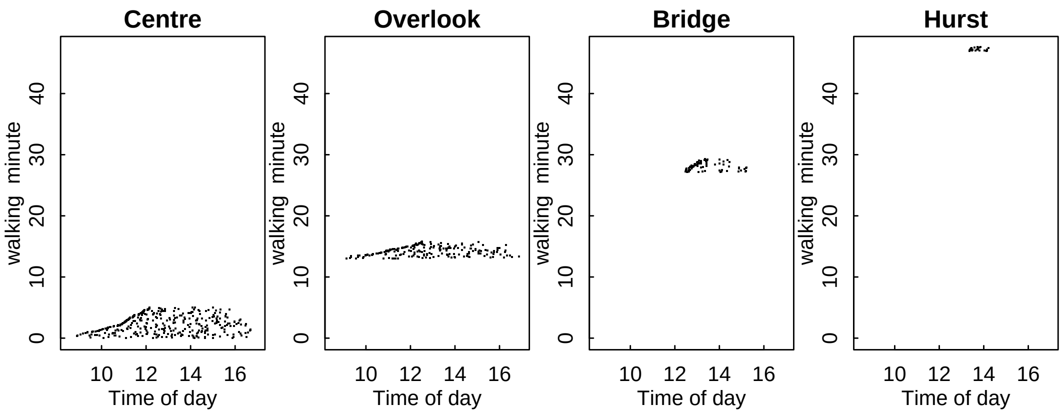


Figure 7.13: Walking time to the Information Centre against the arrival time of day

In addition, both searching time and walking time have either a trapezium or triangular shape within each cluster. This means that both searching time and walking time depend on the arrival time of day.

Overall, the mean of searching time and walking time was 3.582 minutes ($\sigma = 0.235$) and 10.675 minutes ($\sigma = 0.638$). Focusing on the relationship between these times and parking locations, the searching times are not different between the Information Centre and Derwent Overlook as long as a car driver parks at the initial target parking area, i.e. 2.37 and 2.32 minutes in Table 7.3.

Table 7.3: Means and standard deviations of search and walking minutes from 500 simulation runs. The name of the parking area in parentheses is the initial target parking area.

Parkings	**Search (Centre)**		**Search (Overlook)**		**Walking**	
	mean	sd	mean	sd	mean	sd
Centre	2.37	0.09			2.33	0.07
Overlook	3.03	0.41	2.32	0.14	14.29	0.05
Bridge	8.25	0.46	2.77	0.47	28.14	0.05
Hurst	11.63	1.01	8.20	0.68	47.26	0.07

In fact, the ways in which searching time increases are similar regardless of the initial parking areas. Therefore, the searching time are largely affected by the arrival time to the Valley and the number of parking areas a car passed, but not the initial target parking area and parked parking area. On the other hand, the walking time is determined by the arrival time of day and parking location. The expected walking time is 11 minutes shorter if a car can park at the Information Centre. In addition, the walking time is much longer than the searching time.

The standard deviations become larger as parked parking areas are at a distance from the Information Centre because the numbers of cars in the third and fourth parking area are not stable (Check Figure 7.10). On the other hand, a standard deviation of the walking time does not change much (Table 7.3) because the majority of the walking time is contributed by the walking time between parking areas, which is irrelevant to the number of cars in parking areas.

7.7 Conclusion

This chapter successfully showed the dynamic model of the parking network in the Upper Derwent Valley. Since this model used macroscopic timing from the Markov queue model, this was a statistical time-driven simulation. However, cars behaved according to their events independently, so this model was also a disaggregated event-driven model.

This chapter described how to estimate the searching time of a parking space and walking time with analytical and heuristical approaches. The results of the simulation showed the trend of searching time and walking time with the arrival time of day and parking locations.

Although the objectives of this chapter were achieved, some important issues remained for future research. For example, car drivers tried to approach from the Derwent Overlook even though they had to spend much more time walking (Table 7.3). The reason why this model could not explain this phenomenon was that this model did not consider the intelligence and the characteristics of cars and visitors. The cars of the current model do not consider optimising departure times or parking locations to avoid congestion. Such intelligence can be added to the current model with extra behavioural data from another stated preference analysis. Thus, this problem should be considered in future research. Also, this model treats cars as homogeneous particles and so this model is microsimulation as explained on page 93. Chapter 6 showed that different visitors have different parking preferences, which is determined by age, frequency of visits, and the willingness to pay to the road user charging (Section 6.4). Therefore, in the next chapter, the characteristics of visitor and (consequently[13]) their cars are integrated into this dynamic parking model together with the discrete choice models by an agent-based simulation. The simulation model in the next chapter allows researchers to analyse the more complicated situation of the Upper Derwent Valley.

[13]If we could really interpret cars' behaviours from car drivers' behaviours, that would be another question. But, this is not discussed in this thesis.

Chapter 8

Agent-based modelling with the Minority Game in the Upper Derwent Valley

8.1 Introduction

All four sub-modules developed so far are combined by agent-based modelling in this chapter. This chapter starts with the brief reviews on the difference between the structure of conventional approach solely with discrete choice models in Chapter 6 and the agent-based modelling in this chapter. Then, the stochastic Minority Game, which was discussed and validated in the theoretical part of this thesis, is rephrased in the context of the Upper Derwent Valley. After simulation settings are confirmed, three types of results are discussed, namely 1) results without seasonality, 2) results with seasonality, and 3) results with elderly exemption, in terms of mode choices and congestion levels. The last results further discuss equity and user utility while they are compared with the results from the conventional approaches, which conduct analysis solely with discrete choice models.

8.2 Approach of this study

8.2.1 Structure of conventional approaches solely with discrete choice models

Before explaining the current simulation model, the advantages and disadvantages in the conventional approaches, which conduct analysis solely with discrete choice models in Chapter 6 are briefly examined. The two types of equations on travel behaviour, i.e. mode choice and parking location choice, are estimated in that study. The multinomial mixed logit choice model on trip modes was based on travel information composing toll fee, bus fare, parking fees, searching time for parking space, and walking speed. The other discrete choice model on parking location was based on characteristics of visitors composing age, travel frequency, and willingness to pay the road user charging scheme. The two models clearly showed that the relationship between the inputs and the outputs were expressed in probabilistic form.

The major advantage of these discrete choice models is that the equations of probabilities simplify the problems of the real world (Parunak et al., 1998), so the approach, which conduct analysis solely with discrete choice models, requires relatively less detailed input data and decision making processes. Also, the equations clearly represent the behaviour of the problems individually. However, when the problems are interrelated to one another, these advantages cause oversimplification. It is important to underline that the characteristics of visitors were not directly modelled in the equation of mode choice. For example, Section 6.4 described that visitors who used to park at the Information Centre were likely to pay a parking fee and spend nominal time on searching and walking. Therefore, such visitors were more likely to change travel mode from Auto ('car') to Bus than other visitors, who arrived directly at the second parking area (Table 6.3). Hence, we can still link up the mode choice and the characteristics of visitors, but it is difficult to think how mode choices affect the decision making of the parking location. In other words, this approach connects the two multinomial discrete choice models unidirectionally only by the imagination of the researchers.

Moreover, in the case of a road user charging scheme, this D.Phil. research is concerned about how the scheme reduces congestion in the area with other relevant factors. From the conventional approach, there was no clear sign about the effect of a road user charging scheme on the congestion level at the parking areas. For example, a possible scenario about the effect of a toll fee on congestion is that "*if the probability of a visitor going by car is reduced by 10% due to a toll fee, the congestion level at parking areas in the Valley would be reduced by 10%*". However, this scenario is likely to overestimate the reduction in the congestion since an unblocked parking area will attract other potential car visitors. The problem came from the ignorance of a parking network and the concept of congestion.

Any model about the parking network in the Valley was absent it is modelled solely with discrete choice models. For example, a researcher solely using the multinomial discrete choice model assumes that a visitor can definitely park at the Information Centre if the visitor decides to park there. This is because these models cannot formulate the concept of congestion, which requires dynamic links among the visitors – an example of "*Oversimplification*".

In conclusion, although conventional approaches, which conduct analysis solely with discrete choice models, clearly presented the probabilities of visitors' choices independently, it could be dangerous to infer congestion levels at parking areas and visitors' mode choices based on these equations since this analysis depended on the only one ambiguous linkage and the process was non-dynamic (Figure 8.1).

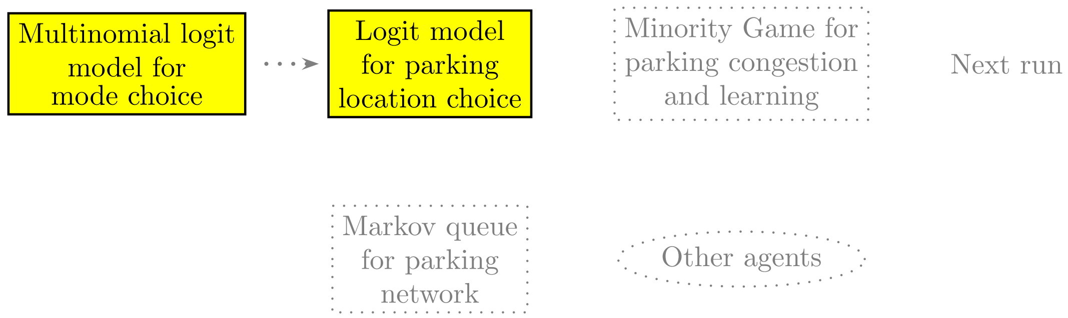


Figure 8.1: Structure of a conventional approach which conducts analysis solely with discrete choice models

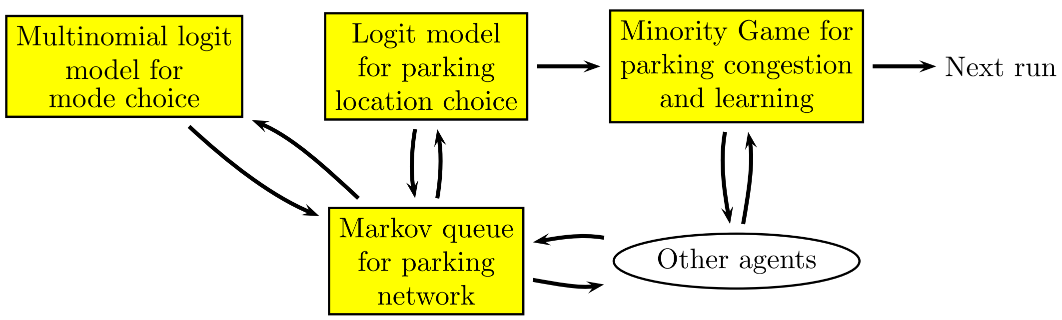


Figure 8.2: Structure of the agent-based simulation integrated with discrete choice models

The problems of conventional approaches which conduct analysis solely with discrete choice models are overcome by using agent-based modelling. The agent-based simulation model was developed in Java programming language with RePast toolkit[1] (Collier et al., 2003). Simultaneously, the advantages of discrete choice analysis are integrated into the simulation model. The alteration from the conventional approach into the stochastic agent-based simulation with discrete choice analysis is discussed throughout this chapter.

8.2.2 Structure of agent-based modelling

Finally, the product of the four modules that have been explained so far comes together (Figure 8.2). The two multinomial discrete choice models stayed as main modules of the agents' decision making at a micro level in this simulation model. However, a learning process was added to the multinomial discrete choice model with the strategies' success scores in the Minority Game. In addition, this simulation model included the Markov queue model of the parking network. The Markov queue model connected the two discrete choice models, bidirectionally. The outputs of the Markov queue model were inter-linked with the Minority Game through the interaction of various agents and the whole process was carried forward to the next time step. Therefore, this model was dynamic and had the concept of learning and congestion. To be consistent with the approach solely with discrete choice models in Chapter 6, an agent in the agent-based model was defined as a trip leader of a visitor group.

One more clear difference in the structures of the two types of analysis was that the calculation in Figure 8.1 occurred once at a system level, so there was only one set of calculations in discrete choice analysis. In contrast, in this agent-based model, the whole set of calculations in Figure 8.2 was carried out for each agent and at each time step, so that the system level results were

[1] http://repast.sourceforge.net/

the integration of these many small calculations. This was a useful concept since the user utility and behaviour of each individual agent were analysed by tracking down individual calculations.

8.3 The Minority Game in the Upper Derwent Valley

In this project, the Minority Game (Arthur, 1994; Challet and Zhang, 1997) was used to assess the congestion level of the four parking areas. Currently, there are only two mode choices in the Valley, 'Auto' (going to the Valley by car) and 'Cancel' (not going), since there is no other practical alternative. Then, after the road user charging scheme is implemented, one more choice 'Bus'[2] is added. Thus, there are three possible choices in the Upper Derwent Valley after the implementation of the road user charging scheme. In general, visitors should not come to the Valley by Auto when they cannot park at their target parking area. Also, visitors arriving by bus will be glad that they chose the Bus option if they find that there are no empty spaces in the parking areas where the buses pass and stop. These two situations indicate that the visitors are indirectly playing the Minority Game.

The Minority Game in the transportation sectors has been considered in the theoretical models (Lee et al., 2001; Bazzan et al., 2000); however these models have not been implemented in real world applications (Dia, 2002; Peeta et al., 2005). Some agent-based traffic models have used the concepts of congestion without being aware of the Minority Game (Klügl and Bazzan, 2004), although these studies are similar to the Minority Game. Nonetheless, these studies tend to be conceptual models. Therefore, this study contributes to fulfil the gap

[2]Parking a car before approaching a toll gate (Figure 5.1 on page 87), and going to the Upper Derwent Valley by bus.

between the theory and practice of the Minority Game in the transportation sector.

Utilities and actual returns were not made up from a theory or some imaginary threshold, but calculated by the utility functions of a multinomial mixed logit model, which were based on the survey at the Upper Derwent Valley (Chapter 6). Unlike conventional Minority Games, the winners and losers were not determined by a fixed proportion of agents in this model (e.g. 51%). This was the same as the online Minority Game in Chapter 4, i.e. if the return from a choice was less than expected, this visitor made a wrong choice. However, the winning side was determined more personally and either side was not necessarily a global winning side. For example, if too many visitors choose Auto, more people are likely to underestimate searching time to find a parking space and walking time to the Information Centre since these factors are on average longer in the congested condition. However, some of these visitors are still able to park where they want to if they are lucky. Therefore, in the same choice, there are both winners and losers in this Minority Game. Put differently, agents in this model used personal experiences rather than centralised information (i.e. perfect information) to make their decisions, so that the agents also had the personalised results of the Minority Game (de Cara et al., 2000).

Furthermore, one thing that the conventional Minority Game has not considered about in the transport sector is 'rigidity' in the decision making process. The rigidity or cognitive conservatism is observed in previous experiments and surveys (Fujii and Kitamura, 2003). The rigidity could be due to characteristics, which are specific to one travel mode such as comfort and convenience, and this study takes into account the rigidity as alternative specific constants in the multinomial discrete choice model (Fowkes, 2000; Ortúzar and Willumsen, 2001, pp.219–220). The coefficients in the mixed logit are not fixed, but oscillate with the normal distribution. Thus, the effect of travel times and costs are different among agents due to the taste variations of visitors.

8.3.1 Strategies

Most discussion about strategies are deductively validated in the theoretical part of this thesis, but some discussions are rephrased in the context. Although the former sections use the phrase: 'visitors make a choice', properly speaking, strategies make a choice instead of agents. Agents choose the best strategy and follow the choice that the strategy made. The reason for this two step decision making was due to imperfect information. It is impossible to obtain perfect information to win in the Minority Game since you need to know the decision making of many other agents. So, at least, the assumption of perfect information fails in this situation. Therefore, even if the mechanism of the decision making process was correct, the output may be wrong. It might be said that this was just a problem of imperfect information; therefore, the model does not need to be modified. Nonetheless, perfect information would never be achieved, so the output was likely to be inefficient. Therefore, it was sensible to give agents an ability to guess the results from the multinomial discrete choice model. According to the Minority Game, the two thought patterns were easily considered as explained on page 60 in Chapter 3. For example, if parking areas are severely congested at the time of travel, searching time and walking time tend to be longer for visitors with the Auto option. These visitors may think: 1) the parking area will be congested so I will not go to the Valley by car next time, or 2) many visitors will be discouraged to go to the Valley by car so parking areas will be empty; therefore, they think, I will go to the Valley by car next time. Thus, searching time and walking time can have a negative affect as well as a positive on the derived utility of Auto. From the description above, three thought patterns were considered for this simulation. The three thought patterns of visitors depend on which mode takes the congestion related utility:

Thought pattern 1: believes that the parking area will be congested again next time, so discourages a visitor from going to the Valley by car, i.e. add $\beta^{\text{time}}(\text{Search \& walk})$ into U^A

Thought pattern 2: believes that the parking area will be less congested next time, so discourages a visitor from going to the Valley by bus, i.e. add $\beta^{\text{time}}(\text{Search \& walk})$ into U^B

Thought pattern 3: believes that the parking area will be less congested next time, so discourages a visitor from cancelling the trip, i.e. add $\beta^{\text{time}}(\text{Search \& walk})$ into U^C

Thought pattern 1 is the same as the result from the multinomial mixed logit model mechanism. Thought pattern 2 and 3 try to cut the ground from under the feet of other agents. In other words, thought patterns 2 and 3 are the second thoughts from the result of the multinomial discrete choice models. Put differently, these thought patterns are skeptical about the result of the multinomial discrete choice like Hume's evaluative skepticism[3] (Clark, 1998).

Also, each thought pattern was sub-categorised into five strategies according to the experience agents had since real human beings were unlikely to remember more than several categories (Miller, 1956). Namely, agents could use any of the last five experiences to calculate the choice. From the survey carried out in the summer of 2003, it was unlikely that visitors remembered any detailed trip information older than the last five trips. Then, up to five strategies out of 15 strategies were distributed according to the memory distribution estimated in Chapter 4 (Table 4.1) at the beginning of a simulation

[3]See footnote on page 77

run. This meant that some agents decided on a travel mode based on the most recent trip experience while other agents used the oldest trip experience.

Best strategy

The best strategy with the maximum success score was chosen before each trip. This Minority Game used the horizon of strategy successfulness. The horizon is related to the adaptability of agents, since a long horizon makes agents consider too much historical information that may not be relevant to the current situation (Liu et al., 2004, pp.347-351). The length of the horizon is a parameter H, which represents the horizon for each strategy scores. Therefore, the success score of each strategy is only a virtual point in the last H steps an agent experienced:

$$\theta_t^s = \sum_{i=t-1-H}^{t-1} R_i^{x_i^s}/H \tag{8.1}$$

where:

$$\begin{aligned} x &= \text{The selected choice by strategy } s \text{ at } i \\ R^x &= \text{Return from the selected choice at } i \end{aligned}$$

R^x is calculated from the utility functions from the mixed logit model on page 119, so that its unit is utility. As shown in Equation (8.1), the success score θ of any given strategy s at a time step t is the moving average of the return from a selected choice by the strategy within the scope of horizon H. The choice made by a strategy is not relevant with the choice used by an agent, which possesses the strategy. All strategy-scores θ were calculated whether or not the strategies were chosen by the agent. Similarly, although H was set to five in this model, the length of horizon was irrelevant with the length of experience remembered.

Example of decision making process

This is an example of the decision making process. At a given week, each strategy calculates the probability of each choice according to the multinomial mixed logit model. Nevertheless, the thought patterns 2 and 3 swap the utility of searching & walking time from U^A according to their rules. Then, there are five memories, so that a set of 15 possible strategies in an agent can be like the one in Table 8.1. These 15 strategies are possible strategies, but in reality, there are only a maximum of five strategies for each agent according to the calibrated memory distribution in Table 4.4, i.e. some agents may have only one strategy. For example, a subset of five strategies can be like the one in Table 8.2.

Next, this agent needs to find the best strategy to make a mode choice. The set of strategies in Table 8.3 is the same set of strategies as in Table 8.2, and they have five horizon values. The choice in the table is the choice each strategy made in each experienced time step. R is the return from the predicted choice and actual parking condition in the Valley. Then, θ is the moving average of the five returns. It is important to mention that the returns R are not necessary to be the same among the strategies, even if the choice is the same at any given horizon since the game is based on localised experience, but not the centralised information[4]. Also R is expected to be negative according to microeconomic theory, i.e. cost and travel time expectedly affect the utility of visitors negatively (Hess et al., 2005).

In this example, thought pattern 2 with memory 1 has the highest success score, so this is the current best strategy. However, this best strategy may change in the future since it is a moving average. In the strategy of thought

[4]This is clear when Table 8.3 for this simulation and Table 4.7 for the web online Minority Game (page 80) are compared. The simulation for the web online Minority Game uses the centralised information so the returns R are the same if choices are the same.

Table 8.1: Example set of 15 possible strategies in an agent

Prob. of choice	**Memory1**			**Memory2**			**Memory3**			**Memory4**			**Memory5**		
	Auto	Bus	Cancel	Auto	Bus	Cancel	Auto	Bus	Cancel	Auto	Bus	Cancel	Auto	Bus	Cancel
Thought pattern1	0.5	0.4	0.1	0.4	0.4	0.2	0.65	0.35	0.1	0.4	0.45	0.15	0.45	0.3	0.25
Thought pattern2	0.6	0.2	0.2	0.45	0.3	0.25	0.7	0.25	0.15	0.45	0.35	0.2	0.55	0.1	0.35
Thought pattern3	0.55	0.45	0.0	0.45	0.45	0.1	0.7	0.3	0.0	0.45	0.5	0.05	0.55	0.4	0.05

Table 8.2: Example set of five assigned strategies in an agent

Prob. of choice	**Memory1**			**Memory2**			**Memory3**			**Memory4**			**Memory5**		
	Auto	Bus	Cancel	Auto	Bus	Cancel	Auto	Bus	Cancel	Auto	Bus	Cancel	Auto	Bus	Cancel
Thought pattern1	0.5	0.4	0.1				0.65	0.35	0.1						
Thought pattern2	0.6	0.2	0.2	0.45	0.3	0.25									
Thought pattern3	0.55	0.45	0.0												

Table 8.3: Finding the best strategy in an example strategy set.
'TP' stands for thought pattern and 'M' stands for memory.

Success score	**Horizon1**		**Horizon2**		**Horizon3**		**Horizon4**		**Horizon5**		θ
	Choice	R	Choice	R	Choice	R	Choice	R	Choice	R	
TP1M1	Auto	-2.4	Auto	-0.5	Auto	-2.2	Auto	-2.3	Bus	0.0	**-7.4**
TP1M3	Bus	0.0	Auto	-5.3	Cancel	-0.9	Bus	0.0	Bus	0.0	**-6.2**
TP2M1	Auto	0.4	Bus	0.0	Auto	-0.7	Auto	-2.3	Auto	-0.2	**-2.8***
TP2M2	Auto	-2.4	Auto	-0.4	Auto	-0.4	Bus	0.0	Auto	-3.2	**-6.4**
TP3M1	Cancel	-0.9	Auto	-1.0	Auto	-1.5	Cancel	-0.9	Cancel	-0.9	**-5.2**

pattern 2 with memory 1, Auto has the probability of 0.6 (Table 8.2), so this option is likely to be chosen by this agent, but this is still determined by the probabilities. Although this decision making process involves guessing and baffling other agents' calculations, the basis is still the multinomial discrete choice model. Thus, this process is not just throwing dice, but the result is still connected with the situation of the Upper Derwent Valley.

8.4 Population

The structure of society and the expected number of cars on the busiest days were estimated in the previous two chapters. The distributions of visitors in terms of origin and visiting frequency were already shown in Figure 6.6, but the table is displayed here once more (Table 8.4). The number of cars coming to the Upper Derwent Valley on the busiest days was estimated as 805 from observations[5] and simulation[6].

Table 8.4: Distribution table of origin and frequency of visit (percentage)

	1/week	1/2 week	6-12/year	2-5/year	1/year	< 1/year	Overall
Far	0.99	3.96	7.92	49.50	17.82	19.80	39.92
Local	3.95	6.58	16.45	42.76	13.16	17.11	60.08
Total	2.77	5.53	13.04	45.45	15.02	18.18	

From these results and a simple calculation[7], the visitors' population was also estimated. The visiting frequency below 'once a year' was assumed 0.5

[5]See page 138.

[6]This includes the cars before simulation (Section 7.4).

[7]For example, assume 10 people visit the Valley at a given week. Of those, five people visit weekly, three people visit monthly, and two people visit yearly. The five people really come every week, so there are only five people of this type, i.e. $10 \times 0.5 \times 1 = 5$. Another three monthly visitors come the next week, and this continues until the next month, so this means that four sets of three people of this type. So, there are 12 people of this type, $10 \times 0.3 \times 4 = 12$. Similarly, there are 52 sets of 2 yearly people who visit the Valley; therefore, 104 people of this type, $10 \times 0.2 \times 52 = 104$. In conclusion, there are 121 people who visit the Valley annually.

per year. The number of annual week was set at 52. The estimated population of the travel groups from the local area and neighbouring cities was 13,435.23, and that for other areas was 10,410.91. Therefore, the overall population with private car was 23,846.14 by calculation. The difference in population between two regions does not differ as much as that of the overall distribution shown in Table 8.4. This is because the visitors from the local area tend to go more frequently, and this calculation does not double count these frequent visitors. The problem is that this population is comprised of visitors by car and it excludes other visitors who currently use a public bus and potential visitors who cancel a trip on second thought. From the observations, the number of bus users was nominal, but the number of the potential visitors was unknown. There was no easy way to estimate the number of potential visitors since the origin survey was largely unsuccessful as mentioned on page 87 (Chapter 5). Therefore, this issue is left for future research.

As mentioned in Section 7.5.2 (from page 145), the number of cars to the Upper Derwent Valley was determined by a time-dependent arrival rate λ, so that alteration in the rate changed the car numbers in the Upper Derwent Valley. For example, if the number of agents is 3,577 and potential visitors are 50% of car visitors, the ratio of agent size and visitor population is 0.1 (i.e. $3577/(23,846.14 \times 1.5) = 0.1$). Then, on average, 121 agents are expected to go[8] to the Valley ($805 \times 1.5 \times 0.1$) per week in this scenario. The agent size, which is expected to choose the Auto option, is 0.1 of 805, namely 80 or 81. Then, if only 40 agents chose the Auto option, the arrival rate of cars and the number of cars in the Valley before simulation is also reduced by half. This means that Cancel applies only to the agents, which plan to go to the Valley and decide not to go on second thought.

[8]Including the ones who decide to cancel the trip later

8.4.1 Population as a collection of discrete individuals

One more issue about the population needs to be clarified. This study adapted the techniques used in the multinomial discrete choice model, which produces a result in probabilistic form instead of a discrete result. Traditionally, economists transform a probabilistic result as a proportion of a target population. For example, if the result of Auto is 0.5, 50% of a target population, on average, choose Auto. This transformation of probability to proportion produces the same output if input variables are the same. However, any population is not a continuous variable, but the sum of discrete individuals in the agent-based modelling. Therefore, each agent has to throw a die to decide a discrete answer from a probability. This is the first reason why agent numbers fluctuate before considering any other reason in the agent-based modelling. Additionally, this transformation of probability to the proportion of a population may cause biased results if interaction effects are significant as mentioned in the introduction on page 23.

8.5 Simulation setting

This agent-based model used numerous parameters and was rich in local rules[9], so only important settings are explained here.. Unless specified otherwise, the values in parameters were the same. The agent size or travel groups were 3,000. The agent size did not affect the results of simulation, since this was a sample from a larger agent population and car numbers were automatically reflected by the ratio between the population and the agent size.

Real bus fare was 50 pence per person, the parking fee for the Bus option was 50 pence per car, headway was 30 minutes, and the toll was £2 or £3.

[9] As explained on page 90, the total rules are over 300. Please see appendix B to see all rules.

Table 8.5: Range of believed values

Variable	Range	Explanation
Toll fee	[0, 5] (£)	Up to the R.U.C. in London.
Bus fare	{0, 0.1, 0.2, 0.5, 1} (£)	One-coin value from local authority
Searching times	[0, 18.41] (minute)	See Section 7.5.2.
Walk distance	[0, 4792.09] (metre)	See Section 7.5.2.
Parking fee for Auto	{0, 0.5, 1.5, 2.0, 2.5} (£)	The real range in the Valley.
Parking fee for Bus	{0, 0.1, 0.2, 0.5, 1} (£)	One-coin value.
Headway	{15, 30, 45, 60}(minute)	From interview and current situation

These costs come from interviews with the local authority (Derbyshire County Council, 2003, per. com.). Agents did not know this travel related information before they experienced it; therefore each agent picked up believed values randomly from possible ranges at the beginning of each simulation. The range of believed values are in Table 8.5. The first 520 time steps, which were 10 years in simulation time, were treated as an initial transient period (i.e. the warm-up period), so the outputs in the period were discarded from any analysis. The one time step was for a representative weekend day. Most agents were expected to go to the Upper Derwent Valley five times and fill all memory spaces by the end of this period, so that they were most likely to gather enough real experiences. The rest of the agents, which were a small part of infrequent visitors, may have used believed values even after the initial transient period.

When agents could not park: When a car could not be parked at any parking area in the Upper Derwent Valley, the associated agent was assumed to give up its current visit. The agent may have enjoyed the rest of the day somewhere else, but the maximum possible walking distance, 4792.09 metres, was added into the agents' experiences as a penalty, beside the searching time the agent spent in the Valley.

Walking speed: Walking speed was set, according to the agents' age, between 4.2 and 3.0 feet per second, i.e. the older, the slower. The difference in walking speeds between the older (the top three older categories) and the younger (the bottom four youngest categories) is 0.7 feet per second. Previous research on walking speed is in urban areas and not in recreational areas (TranSafety, Inc., 1997; Knoblauch et al., 1996), so these findings were used only as a rough standard.

Frequency of visit: Visiting frequency is probabilistic; therefore, "visit once every other week" did not guarantee that an agent visits the Upper Derwent Valley this time if it does not visit the last time step. Instead, this concept says that this agent is likely to visit the Valley, on average, 36 times a year. The frequency of visit was based on the real data (Table 8.4) and it depended on the travel origin. The mid point values were used for the categories '6-12/year' and '2-5/year', and also '$<$ 1/year' was set as 0.5.

8.6 Results: without seasonality

8.6.1 After the road user charging scheme becomes common knowledge

The results, after the road user charging scheme became common knowledge without a seasonal traffic pattern, were examined. Simulation in this section was steady state simulation. Policy tools such as a toll fee, bus fare, and parking fees were exogenous variables and these values were fixed throughout this study. The parking network model was based on the data on the busiest days and controlled by macroscopic timing. Therefore, by discounting the arrival rate λ, the model was able to control the level of travel demand. There were three traffic demand levels (i.e. the busiest, 80% of the busiest, and 50% of the busiest) and two levels of a toll fee (i.e. £2 and £3). It is important to mention that the traffic demand levels control the arrival rate of cars, but not the visiting frequencies of agents in this section[10]. As mentioned in Section 8.4, the agents, which are shown in this model, are sampled agents in the Upper Derwent Valley and the agent size is fixed to 3,000 in this model. Then, as the traffic demand level is higher, more cars are running in the Valley; therefore, there are more un-shown agents in the cars (i.e. agents in the population minus sampled agents). That is, the proportion of sampled agents in the Valley is lower with higher traffic demand levels. The bottom line is that the doubling traffic demand level does not mean doubling the agent size. The traffic demand levels can be purely considered as congestion levels in this section.

As expected, agents choose the Auto option more often as the parking areas are less congested or the toll fee is cheaper (Figure 8.3). In contrast, the other two alternatives are relatively less elastic with the changes in the two factors.

The lines in bold were smoothed by LOWESS method with value 0.1 (Cleveland, 1981), and raw data are shown in grey lines. The grey lines show that the mode choices are varied throughout any simulation and never get an equilibrium point although the system seems to achieve a steady state at a larger viewpoint. The variances are due to the probabilistic travel frequency and the discrete population as explained on page 166. More interestingly, the variance is due to the nature of the Minority Game. Each agent tries to forecast the most optimal choice at each time step with different travel experiences, but there is no such optimal choice in this Minority Game situation. The optimal decision making in the Minority Game situation is dependent on the other agents' decision making, which cannot be predicted perfectly since

[10]The traffic demand levels as seasonal demands that affect the visiting frequencies of agents in Section 8.7 and the mechanism is explained on page 172.

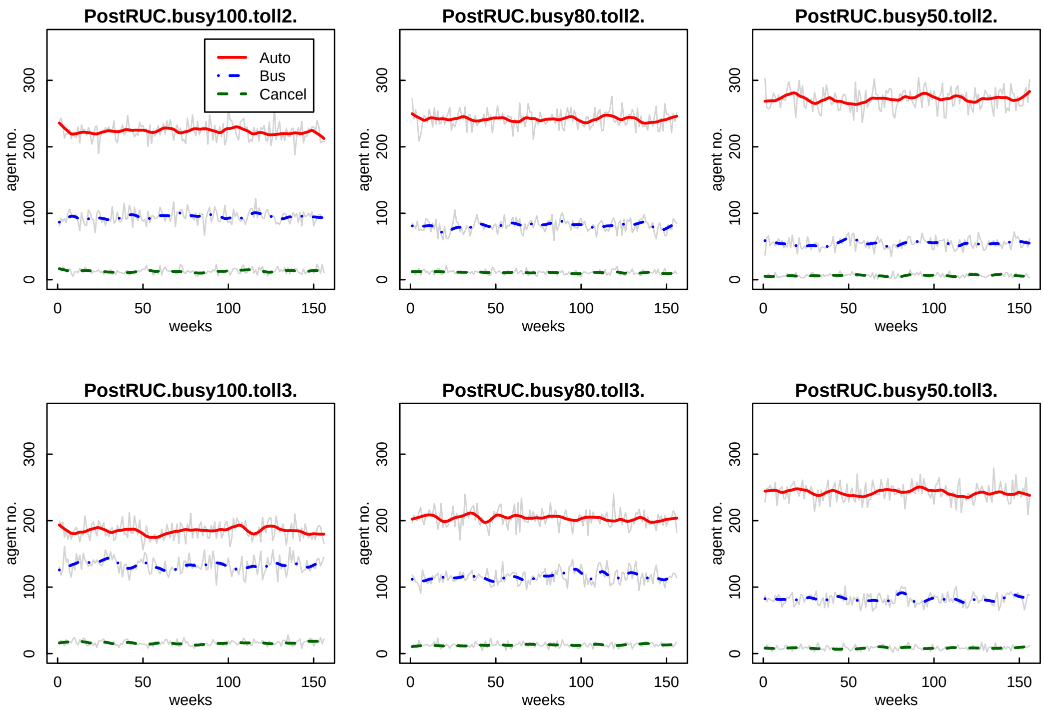


Figure 8.3: Mode choice after the road user charging becomes common knowledge. The traffic condition is 100%, 80%, and 50% busiest from left to right. And, the toll fee is £2 on the top figures and £3 on the bottom.

their decisions are also dependent on other agents and so on (Unger, 1978). This is an important issue so that the cause of variation by the Minority Game is discussed further in successive sections.

Although the decisions were based on different experiences on the different time steps, the experiences did not radically change at a system level when exogenous variables were fixed. Hence, the mode choices were varied within a limited boundary, and consequently, the system reached some degrees of steady state.

8.6.2 Transition period of the road user charging

In this section, the road user charging and bus service were implemented in the middle of the simulation, so that an exogenous variable, toll fee, was not constant. The road user charging and bus service were implemented at the

53rd week (i.e. the first week of the second year). The number of Auto and Cancel reduced while the number of Bus increased after the implementation (Figure 8.4).

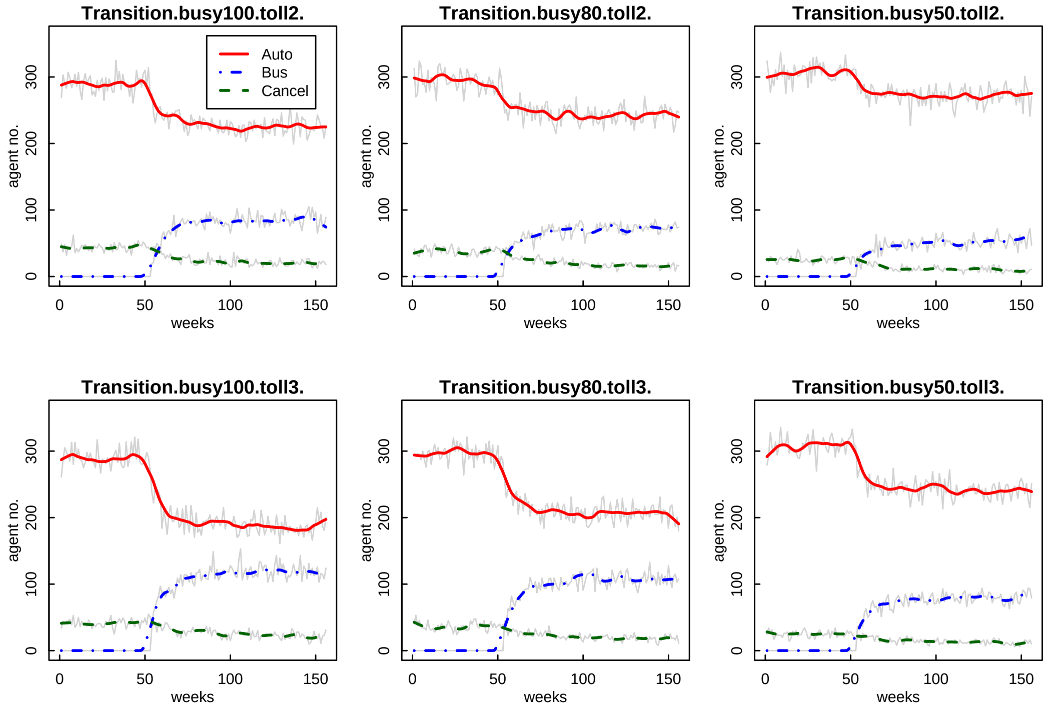


Figure 8.4: Mode choice during implementation of the road user charging. The traffic condition is 100%, 80%, and 50% busiest from left to right. And, the toll fee is £2 on the top figures and £3 on the bottom.

The change started rapidly soon after the implementation of the road user charging and the bus service. However, it took more than one year to become a new steady state, and the slight transition process continued for another year at least. At the end of the first road user charging year, 18.18% of agents were expected not to know the change (Table 8.4), so that the transition took sometime. Moreover, the transition period is likely to be longer as more agents choose Cancel, i.e. there are always a positive number of agents cancelling the trip on second thought every time they plan to go to the Upper Derwent Valley on the initial thought.

8.7 Results: with seasonality

8.7.1 Traffic pattern

From the traffic flow data of 2003 on the A57 explained in Section 5.2, the seasonal demands on the Valley were estimated. The final week of August and the first week of September were set as the busiest weeks (the vertical dotted lines in Figure 8.5). June to September were set as 90% of the busiest weeks, i.e. high season (the darkest background colour in Figure 8.5), April, May, and October were set as 80% of the busiest weeks, i.e. intermediate season (the intermediate background colour), and the rest of the periods were set as 60% of the busiest weeks, i.e. low season (the white background colour). Unlike the last section, the seasonality affected the frequency of agents' visits if their visiting frequency was less than 'once every other week'. For example, an agent, which is supposed to come once a year, still comes once a year, but it is more likely to come during high season than low season. The toll fee was fixed at £3, which was the highest possible value according to the local authority, from this section onward.

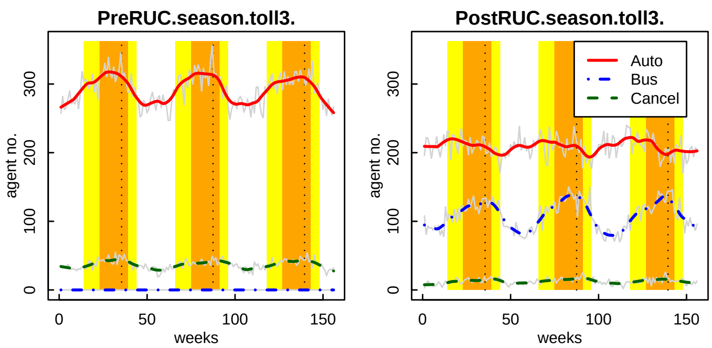


Figure 8.5: The left figure shows mode choices before the road user charging, and the right figure shows mode choices after the road user charging. The toll fee is set to £3. The background colours show the traffic seasonality in the Valley.

As expected from the previous results, the demand for Auto is reduced after implementing the road user charging as the demand shift to Bus and Cancel, and the shift is greater in the high seasons since the trend of Auto is relatively more flattened after the implementation. This is attributable to the fact that the preference of agents (or strategies strictly) was logarithmic and not linear with given parameters. This means that an extra 10 cars in the parking areas in an extremely congested situation discourages agents from choosing Auto more than the same extra 10 cars in a less congested situation. This phenomenon was consistent even from localised viewpoints. Figure 8.6 shows the proportion of time each parking area is congested. Generally, congestion levels are reduced after the road user charging scheme is implemented. However, the proportions of medium congested periods increase after the implementation, i.e. the black coloured areas (100% full) shrink while the red (medium) coloured areas (75% full) stretch in Figure 8.6.

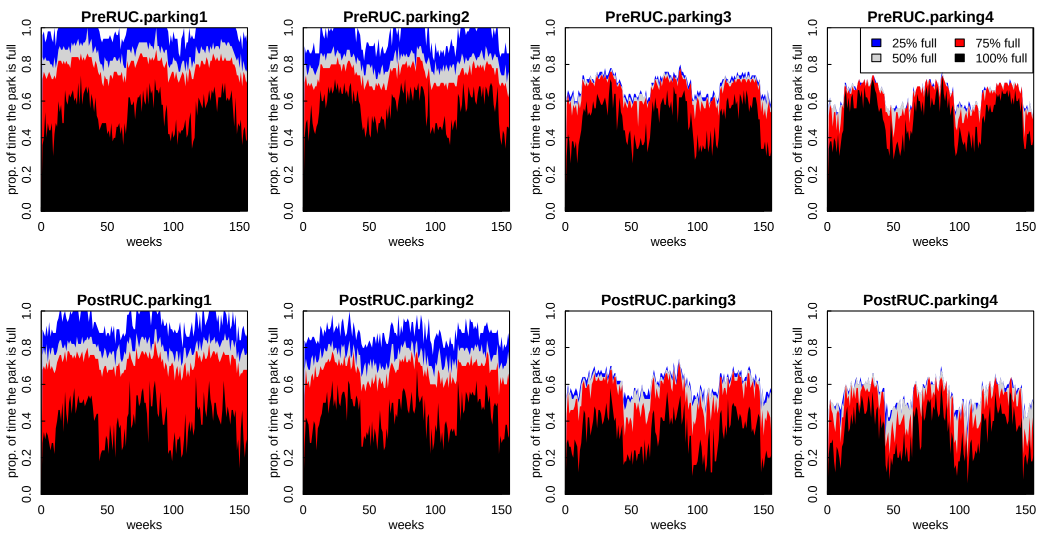


Figure 8.6: Congestion levels in four parking areas (Information Centre, Derwent Overlook, Bridge End Pasture, and Hurst Clough from left to right) with seasonality. The top graphs show congestion levels before the road user charging, and the bottom graphs show after the road user charging. The toll fee is set to £3.

Therefore, the model in this section showed that the road user charging scheme reduced the demand of Auto effectively in more realistic conditions and the reduction in the demand was corresponding with the reduction in the congestion level in the parking areas. Since the scheme reduces demand and congestion more efficiently at extreme conditions, it solves the severe congestion problem at parking areas while still attracting visitors in less congested conditions.

8.7.2 Timing of implementation

The time the policy scheme becomes common knowledge is different with implementation schedules. The three graphs in Figure 8.7 show the mode choices with different implementation timings. When the road user charging scheme is implemented at an intermediate season, the scheme becomes common knowledge by the end of this high season and so the majority of the process finishes within the first six months (left graph in Figure 8.7). In contrast, the process is slower when the implementation is in the middle of a high season or the end of an intermediate season (middle and right graphs in Figure 8.7, respectively).

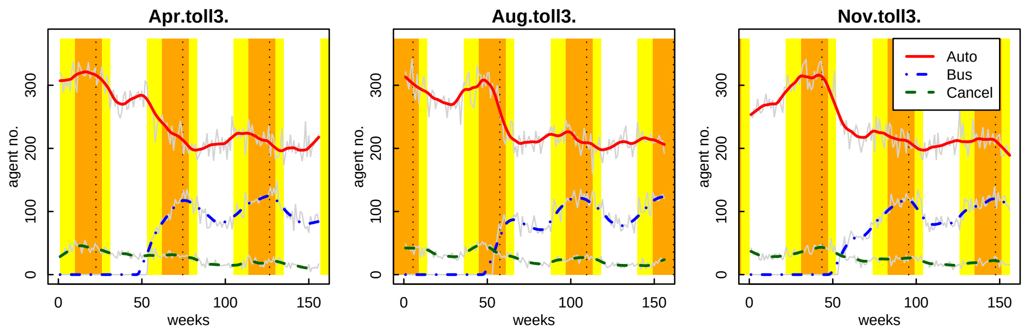


Figure 8.7: The road user charging scheme is implemented at the beginning of a intermediate season, the middle of a high season, and the end of a intermediate season, respectively from left to right. The toll fee is set to £3.

In the second case, the prevailing process was kick-started rapidly, but this fast process stopped before reaching an adequate level. In addition, the process was even more sluggish, in comparison with the other two cases, when the implementation was at the end of the intermediate season. In both cases, the process continues in the next high season and so it takes longer than six months. Therefore, the implementation timing at the beginning of an intermediate season is, time-wise, most efficient.

8.8 Results: with elderly exemption

This section focuses more on the elderly visitors and the toll fee was fixed at £3. From observations, a large proportion of visitors to the Valley are elderly visitors. The result from the questionnaire showed 24% of visitors were aged between 55 and 64 and 12% of visitors were aged over 65, so overall more than a third of visitors were elderly visitors. Many visitors to the Valley carry large equipment to have a picnic and it will be difficult for elderly visitors to carry this equipment without car. Although this model possibly takes this difficulty into account, within an existing factor ‘walking speed’, this internalisation is likely to be underestimated so that special care should be considered for this fact. There are five to six disability parking spaces in front of the Information Centre, which is the primary destination for the many visitors, but this space should be used by true disabilities, strictly speaking. Also, this space is insufficient during the high season and will never satisfy most elderly visitors even if the space is doubled.

One possible solution is to give the elderly a discount on the toll fee. Elderly visitors, who are eligible to receive the exemption, were defined by age, from 55 or 65, in this section. When the discount was only £1 off, overall demand for Auto was not increased significantly for both cases (Middle graphs in Figure 8.8). In contrast, when all visitors older than 55 received a full discount on the toll fee, the trend of Auto vertically rose and that of Bus fell (Top right graph in Figure 8.8). With 100 sets of this situation[11], the percentage rose in

[11] Each set has exactly the same setting including the seed of random number generator

overall Auto demand by 12.71% with a standard deviation of 1.07%. This case is rather remarkable, so this study concentrates on this case and looks into the inside of its result for the rest of this chapter.

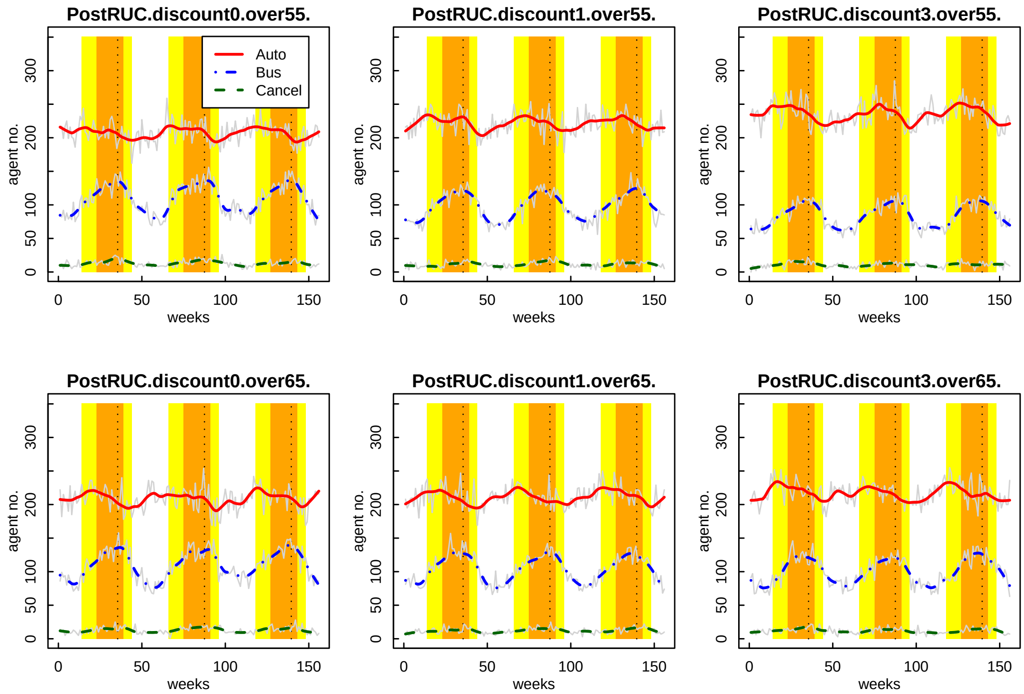


Figure 8.8: Mode choices with elderly exemption. Elderly visitors are charged discounted toll fee between £0 and £3 off, from left to right graphs. The visitor eligible to receive the discount is older than age from 55 on the top figures and 65 on the bottom figures. The toll fee is £3

8.8.1 Comparison with conventional econometric analysis solely using discrete choice models

When a researcher conducts the same analysis solely with the multinomial discrete choice model, the result could be similar but the contents of the result have to be examined carefully. The multinomial discrete choice model in Chapter 6 cannot figure out searching time, walking time, and a parking fee for Auto, so this approaches, which conduct analysis solely with discrete choice models, is not even on an equal footing with the agent-based simulation model.

except for the level of elderly exemption.

Leaving this question aside, assuming that these values are the central point of possible values and the other parameters are the same as those of the current agent simulation, the overall percentage increase in Auto by the discrete choice model is 15.09%. Thus, both approaches produce similar outputs, but the difference is obvious when the break downs by age categories are examined.

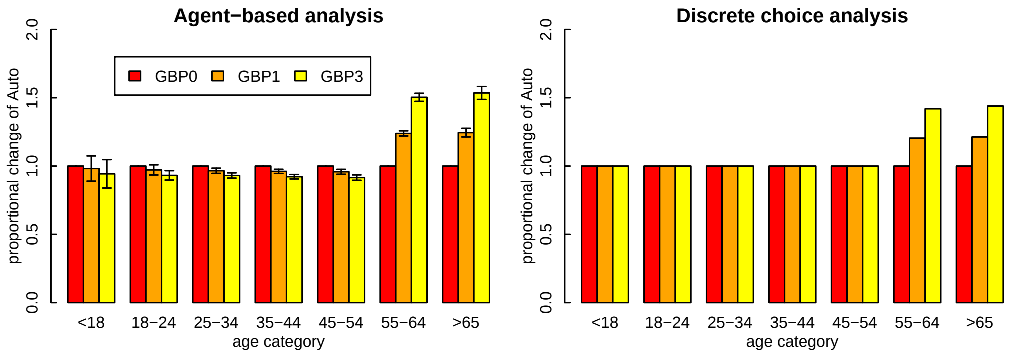


Figure 8.9: Proportional change in Auto by age categories with elderly exemption from the age 55. The left graph is result from agent-based modelling. The right graph is from the analysis solely with discrete choice models.

Auto demand rises only in the two elderly age categories and those in the other age categories stay the same in the conventional approach solely with discrete choice models (Right graph in Figure 8.9). In contrast, the demand of Auto is higher for the two elderly age categories in the agent-based analysis (Left graph). At the same time, the demand declines for younger age categories due to the side effect of the congestion in parking areas. As more elderly visitors come to the Valley by car, the parking areas are more congested, and consequently this situation discourages other visitors from coming to the Valley by car. Moreover, this discouragement reduces the parking congestion more than expected, so that this possibly encourages other visitors, namely the elderly visitors, to come to the Valley by car, simultaneously. Hence, the proportional rises of elderly visitors are more prominent and the demand for Auto by some agents is reduced in an agent-based analysis. There are also variations in the trend of Auto number in agent-based model. This is an important issue and discussed in the next section.

In the meantime, as shown in the Figure 8.1, the conventional approach solely with discrete choice models does not have the direct feedback mechanism with parking congestion or in other words, there is no concept of congestion in the analysis. Therefore, the type of visitors, who come by car due to less congested parking areas after the implementation, is ignored in the conventional approach solely with discrete choice models (Stopher, 2004). The conventional approach underestimates the Auto demand by elderly visitors and overestimates Auto demand by younger visitors because the analysis underrates the sub effect from parking congestion.

8.9 Variation in the model and unpredictability in the Minority Game

This section starts with the unanswered issues in the previous section. The change in the Auto number was varied in the agent-based model. Especially, error bars are spread widely in the age category younger than 18 in Figure 8.9.

This is the evidence of the variation amongst agents' decision making. As mentioned in the earlier section of this chapter, each agent behaves individually, so that it is not necessary that two agents make exactly the same result when inputs and characteristics are exactly the same. A macro level result as the sum of individual decisions could be the same as the result of a system level analysis when individual numbers are large. However, since the macro level result from an agent-based model is a collected result, deviations exist in the result and the stability of the result is strongly affected when the sample size is small. The proportion of age category below 18 is only 0.6% or only 18 or 19 out of 3,000 agents, and the variances are very large in this model. With the small agent size and wide variation, the result from the youngest age category is very sensitive.

It should be emphasised again that, the reason for the wide variation is partly due to the frequency of visits, but more importantly, it is partly because of unpredictability in the Minority Game. It is impossible to achieve perfect rationality in the Minority Game since this kind of rationality requires that an agent is aware of decisions from many other agents. Due to the cognitive limitations of individuals, this type of information is usually inaccessible in real life (Klügl and Bazzan, 2004), especially in the case of parking congestion (Thompson and Richardson, 1998, p.162). The current study uses thought patterns and strategies to formulate agents' decision making, and similarly, no strategy can be a globally best strategy in this dynamic situation. When many agents find the same best strategy, the strategy is no longer the best strategy since these agents move in the same direction and this direction is no longer the minority side.

In the end, these agents are making a decision, but at the same time, partly throwing a die to select a choice at every time step. This causes a wide variety in the decision making process of agents. For example, Section 8.8 shows that the Auto demand by elderly visitors increases when the exemption is given to them, but the change is not uniform even within the same age category.

Figure 8.10 is the cumulative number of Auto chosen by each elderly agent aged over 65 in the same visiting frequency between 2 and 5 times a year. These graphs show a cumulative sum with time, so a horizontal line at the bottom means the agent never chooses Auto. The lines are widely spread in both graphs, although these agents are in the same age and frequency category. The number of times Auto was chosen is distributed at a lower level without a toll exemption. In contrast, more lines are distributed at a higher level with

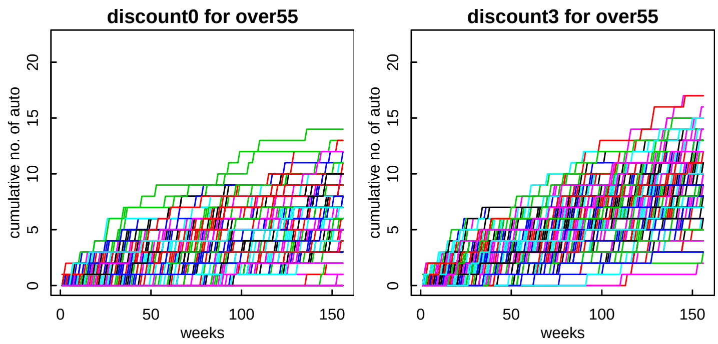


Figure 8.10: Cumulative numbers of Auto chosen by each elderly agents age over 65 without and with full elderly exemption from over the age of 55, respectively from the left to the right at the graphs.

the exemption of £3. These movements indicate that many elderly agents, who cannot afford to visit the Upper Derwent Valley by car, are supported by the exemption. However, some other elderly agents still prefer other choices such as Bus or Cancel. This is partially because extremely bad experiences discourage the agents from choosing Auto, but on the other hand, this could be because an internal preference assigned with other characteristics chooses alternatives. Therefore, analysis on the agent user utility is necessary to justify the comfort of elderly visitors.

8.10 Distribution of user utility amongst agents

User utilities were calculated from the utility functions of the mixed logit model (i.e. equations (6.3), (6.4), and (6.5) on page 119) in Figure 8.11; therefore, their units are utility. Therefore, these values are meaningful only in comparison, but not in absolute terms. In other words, the negative trends of the user utility do not mean that the agents in this category are worse off, but only that the relative difference between two plots has some explanations. The plots show the improvement of the agents' user utility after the implementation of elderly exemption in the same age and frequency category (Figure 8.10). Without the exemption, the user utility was ranged between -59.01 and 4.68 with the mean of -25.45. With the full exemption, the user utility ranged between -47.87 and 32.870 with the mean of 9.202. There is still equity problems within the same agent category since some agents keep increasing their user utility while some other agents are struggled to raise their user utility. This is the nature of the Minority Game, and could be the real phenomenon in many real competitive societies.

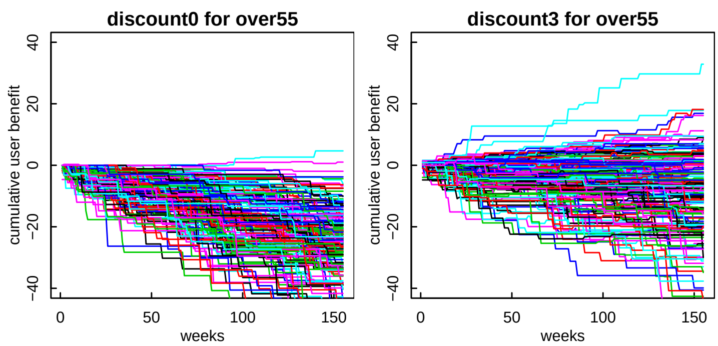


Figure 8.11: Cumulative user utility of each elderly agent aged over 65 without and with full elderly exemption from over the age of 55, respectively from the left to the right at the graphs.

Lastly, the cumulative utility is examined in the case of elderly exemption for those aged over 55. As discussed before, the variances of user utility are very large at the end of the simulation, also the user utility of agents are improved as the exemption rises in the top two elderly age categories (Figure 8.12). These are the expected phenomenon and the trend is similar to the bar graph about the relationship between the number of Auto uses and the exemption (left graph in Figure 8.9) except one important issue.

In the bar graph of the Auto number, the number chosen by the younger agents decreases as the exemption increases due to the congestion in the parking areas caused by the increase in elderly drivers. In contrast, the user utility of younger agents does not decrease much if not decreasing at all. The reason simply comes from the adaptation of agents. As parking areas are congested,

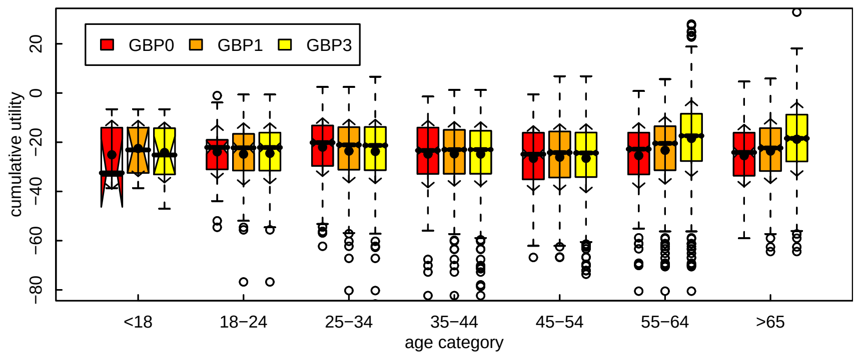


Figure 8.12: Boxplots of agent user utility at the end of simulation in the situations of elderly exemption from age over 55. The level of discount is none, £1, and £3 off. All agents have the visiting frequency category between 2 and 5 per year.

it is more likely that you get a bad experience and so younger agents learn that there is not much point to go by car in this situation. Then, younger agents stop using Auto and switch mode to Bus or even Cancel, but older agents still try to go to the Valley by car since the bad experience of the congestion is substituted by the exemption. For younger agents, alternatives are better options even though their initial motivation is Auto since the alternatives are less congested and they may feel more conformable with the choices. Because of this adaptation, the user utility of younger agents does not decrease much. Therefore, the exemption on the elderly visitors is a possible and suitable scheme to reduce the difficulty specific to the elderly visitors.

Nonetheless, the budget for the exemption comes from outside of this model and so this model cannot formulate whether or not the exemption is affordable. In addition, the overall Auto demand increases with the exemption in some cases. This means that the exemption partially ruins the chance to reduce the congestion level in the parking areas, which is one of the main purposes to implement the road user charging scheme. Therefore, this fact should also be considered when the exemption is implemented. Moreover, this study uses the exemption on elderly visitors, but the idea could be applicable to other social groups as well.

8.11 Conclusion

This simulation model produced comprehensive outputs including mode choices, congestion levels, and the user utility of visitors simultaneously. The study focuses more on concept and process. Having said that, the results showed that the road user charging scheme would reduce the Auto demand in the Upper Derwent Valley and proved that the reduction eased a concern about the congestion at the parking areas. The reduction in Auto demand and parking congestion was effective especially when overcrowding occurred, for example during the August Bank Holiday. Also, this study suggested that the scheme should be implemented sometime before the high season to make the scheme become common knowledge quickly. Although further study should be conducted to finalise the possibility of exemption to the elderly visitors, the model showed that the exemption improves the comfort of elderly visitors without sacrificing those of younger visitors significantly.

Looking back at the concept of the model, this model showed that the oversimplification in the conventional approach solely with discrete choice models gave significant biases when real world problems were analysed. In the case of the Upper Derwent Valley, the oversimplification was the ignorance of the parking network and consequently the concept of congestion, which required dynamic modelling of the linkages amongst visitors. Agent-based models had the advantages of the dynamic modelling and connecting between components in the model. Therefore, the agent-based model simulated the situation of the Upper Derwent Valley more realistically. Although the output from the model was still far from the real world situation due to the lack of some input data, the process of simulation demonstrated enough evidence to alert that the output of the conventional approach solely using discrete choice models underestimated the sub-effect of the parking congestion in the Upper Derwent Valley. In conclusion, the practical part of this thesis established an agent-based model to examine the advantages of the two different models and to overcome the disadvantage of the conventional approach solely with discrete choice models.

Part IV

Conclusion

Chapter 9

Overall discussion and thoughts on future research

9.1 Introduction

This is the final chapter of this thesis and concludes overall discussion in the previous chapters. The section of theoretical discussion answers the questions raised in the first chapter instead of summarising the theoretical arguments. The section of practical discussion recommends the implications from results in this thesis to potential policy makers. Then, these discussions are followed by thoughts about future research and some closing comments.

9.2 Theory: Looking back at the questions raised in the introduction

The key underlying question throughout this thesis was stated on page 11 in the first chapter.

"How do the results of conventional analysis, which solely

use discrete choice models, change when dynamic interaction and learning is integrated by agent-based modelling?".

This thesis integrated the discrete choice analysis with agent-base modelling. The efficiency of discrete choice analysis and stochastic approach was confirmed in this thesis in terms of computational power, simulation time, data collection, etc.. Therefore, the real question was how the integrated approach differed from the conventional approach solely with discrete choice models aside from the efficiency issues. This question is discussed by answering two sub-questions raised in the first chapter.

In the first chapter, one major assumption in economic models was stated on page 6:

> *"the summation of local optima equals a general optimum"*.

Also, this was more precisely explained with a mathematical formula on page 23;

$$n^x = \sum_N P^x$$

Detailed explanation is shown on the page mentioned above. In brief, this formula assumes that the summed probability of a mode x is the same as the proportion of tourists who choose a mode x in a given population N. Also, this formula does not take into account the interaction effects between travellers. In other words, the outcomes from the mixed logit presented in Chapter 6 can be the same as those of agent-based modelling in Chapter 8 if and only if the interaction effects between visitors are insignificant. Chapter 8 after Section 8.8.1 answered this question. As Figure 8.9 on page 177 showed, the proportions of visitors, who chose Auto, were not the same with (i.e. discrete choice model plus agent-based model) and without (i.e. solely by discrete choice model) the interaction and learning effects from the congestion at the parking areas. Therefore, this result disagreed with the formula and the assumption of *"the summation of local optima equals a general optimum"*.

Moreover, the comparative results in Chapter 8 also indirectly answer the hypothetical question on page 16:

H0: $\text{Model output}|\ \max\{U : e \xrightarrow{Ac}, \text{imperfect information}\} \rightarrow \mathbb{R} \quad =$
$\text{Model output}|\ \max\{U : r, \text{perfect information}\} \rightarrow \mathbb{R}$

H1: $\text{Model output}|\ \max\{U : e \xrightarrow{Ac}, \text{imperfect information}\} \rightarrow \mathbb{R} \quad \neq$
$\text{Model output}|\ \max\{U : r, \text{perfect information}\} \rightarrow \mathbb{R}$

In brief, this hypothetical question tests if a model based on an iterative process with imperfect information and another model based on a single calculation with perfect information produce the same output. As discussed in Section 1.2.2, perfect information includes the communication among agents and the communication between agents and the environment. Moreover, these communications as the pieces of information are usually ignored in traditional analysis and the information is assumed freely available to agents in the modelling. However, such information is not free in reality. Free and perfect information also means the ignorance of learning since if every agent has the perfect information there is nothing more to learn.

Moreover, the communication itself does not exist if there is no mechanism of interactions. For example, the restriction in the communication between agents and the environments causes shortsightedness in agents so that visitors in the agent-based simulation model in Chapter 8 use localised experience rather than the centralised perfect information. So, the utilities are calculated step by step, i.e. iterative process of $U : e \xrightarrow{Ac}$ in the simulation. In contrast, the conventional analysis solely based on discrete choice models in Chapter 6 does not have the concepts of interaction, communication, and learning so that the model was bounded to take one form of centralised information to analyse. So, the utilities are calculated in one run (step), i.e. $U : r$.

It is not discussed in this thesis if the centralised information is the perfect information. Leaving this question aside, Chapter 8 after Section 8.8.1 showed that the results based on localised or imperfect information were different from the results based on the non-localised information. In the case of the Upper Derwent Valley, the dynamic condition of imperfect information was inevitably associated with the Minority Game of the parking congestion. Therefore, the ignorance of this situation in conventional approaches which conduct analysis solely with discrete choice models also inevitably generated the biased results. Hence, this thesis rejects the null hypothesis of equal outputs with static perfect information and with dynamic imperfect information. The discrete choice analysis was less biased when dynamic interaction was integrated by agent-based modelling.

In conclusion, the theoretical achievements of this thesis conclude with the statement of a key problem and proposed solution mentioned on page 2.

The statement of a key problem in this thesis is:

> ***"Conventional approaches which conduct analysis solely with discrete choice models have the advantage of simplicity, but severe biases exist due to the neglect of some interaction and learning effects, which might be seen as oversimplification"***

The proposed solution to the problem in this thesis is:

> ***"Innovative interaction and learning are added to the conventional approaches by agent-based modelling together with discrete choice analysis"***

The statement of a key problem was observed and confirmed. Then, the proposed solution was achieved to improve the conventional approach with solely discrete choice models in this thesis. Interdisciplinary approaches into economic theories have become important to reflect reality (Clark and Wrigley, 1995), and this thesis confirmed the importance in current and future research.

9.3 Practice: Suggestion to policy makers

This section gives some recommendations from the results of practical analysis to potential policy makers. Generally, the road user charging and park & ride schemes cut the demands of private cars coming to the Upper Derwent Valley from the results of discrete choice analysis in Chapter 6 and the agent-based simulation model in Chapter 8. Consequently, the congestion levels in parking areas will be reduced according to the results of Chapters 7 and 8.

There are some concerns about equity from the road user charging scheme. An equity problem is always observed when policy makers try to implement policies. For example, Eckton (2003) is not supportive of the road user charging scheme in the Lake District National Park due to the horizontal and vertical equity problems[1]. This could be the same for the road user charging scheme in the Upper Derwent Valley since the effects of the road user charging were unequal between different visitors in the current research. However, the unique attribute of the road user charging scheme in the Upper Derwent Valley is that the road is a dead-end and only a few people live in the charging zone, so the other people visit the Valley only for leisure. Therefore, equity issues raised in this thesis may not be as critical as the ones in the other road user charging research.

[1]This could be defined as the right to mobility, and provision of identical conditions for citizens living in all parts of a certain country (Territorial equity) and associated with the protection of those in worst conditions, respectively (Viegas, 2001, p.291).

Having said that, from observation, the elderly visitors suffered most from the schemes since they carry large equipment, which is difficult to carry by public transport. Also, Section 6.3 revealed that a large proportion of visitors are comprised of the elderly people. The possible remedy for the elderly visitors is to give a discount on the toll fee. Section 8.8.1 showed that the elderly exemption encouraged elderly visitors to come to the Valley by cars, while Section 8.10 showed that the user utilities of younger visitors were not sacrificed significantly. Therefore, the elderly exemption is a recommended complementary policy tool alongside the park & ride scheme. Further research is needed to figure out how to finance the elderly exemption in this situation. Therefore, the level of the elderly exemption is not mentioned here.

Overall, the road user charging and park & ride schemes are recommended in the Upper Derwent Valley. Nonetheless, serious congestion was observed only around the Upper Derwent Information Centre, and road congestion was not observed on the A57 and Derwent Lane even during the August bank holiday weekend, which is the busiest time in the Valley. Therefore, policy makers have to think carefully about the severity of congestion in this area before they put into effect the schemes.

The new schemes were found effective to control the traffic in the Upper Derwent Valley. At this stage, the society is likely to obtain utility from the road user charging scheme if the congestion becomes a severe problem. The level of toll fee is effective at any level between £1 and £3 as planned by the local authorities, but £3 is the most effective. This research did not clarify the accounting of these schemes, so that I avoid stipulating the recommended level of toll fare. This will be the next research question.

9.4 Thoughts about future research

Some research questions are left and emerged from this thesis. First, the analysis on the Upper Derwent Valley was only with the destination data since the origin data was unsuccessfully collected in this thesis as mentioned in Section 5.2. The origin data was infeasible to collect in a D.Phil. research project, but it can be feasible to collect in a larger continuous project.

Second, it is still possible to improve the decision making mechanism of agents. For example, the utility functions of agents were assumed not to change throughout the simulation period. The assumption may not affect the results significantly since the simulation period is equivalent to three years in the real world. However We may have to wait for a while before we can use these new concepts at a practical level.

Third, this analysis focused only on the demand side of tourism at the Upper Derwent Valley and the analysis on the supply side was not included in this thesis. Therefore, toll fee, the frequency on bus schedule, bus fare, and parking fee had to be modelled as exogenous factors. In addition, the costs of these services are not known in this thesis. Consequently, this thesis could not suggest the level of the elderly exemption and the toll fee. Personally, it is interesting to see how the interaction between the demand and supply sides changes the results of this thesis. Moreover, if the supply side is integrated into the agent-based modelling, the producer surplus, which is the utility from the road user charging, the park & ride service, and the parking services, as well as the consumer surplus, which is the user utilities shown in Section 8.10, is calculated. Therefore, the improvement in the overall social welfare after the implementation of the schemes can be calculated.

Furthermore, if the supply side is integrated, we can compare the road user charging scheme with other policy options such as the expansion of parking areas. This comparison will be important to justify the usage of the road user charging scheme. Also, analysing the expansion of parking areas will introduce some interesting research since this policy option is likely to generate an induced travel demand[2], which impacts on the environment more significantly than the road user charging scheme (Fulton et al., 2000). The environmental cost is the factor the current model could not formulate in this thesis and it would be a key factor in calculating the overall social welfare more realistically.

Of course, finding out the supply side mechanism requires extra data from companies and governments. Unlike the behavioural data of visitors, such supply side data are usually more confidential, so that modelling the supply side behaviour may require more workload than this thesis. Nonetheless, it is worth thinking about, including the supply side into this agent-based simulation model in the future.

9.5 Closing comments: the shoulders of giants

Initially, I tried to criticise all conventional approaches in economics. My attitude can be described as 'fighting against the shoulders of giants' in contrast to a famous quote from Isaac Newton in a letter to Robert Hooke on the 5th of February 1676, i.e. "If I have seen further it is by standing on the shoulders of giants." (Newton, 1676). The phrase 'the shoulders of giants' means ancestral works, so that Newton meant his great works did not exist without the previous research. In the viewpoint of Newton's quote, my attitude was similar to treating the previous works as the enemies of my research.

However, as my research progressed, I realised how the previous forms of works in discrete choice analysis and stochastic simulation approach were advanced and represented the many real world phenomenon, realistically and

[2]Induced travel demand is an increase in traffic volume after a new road or new parking area is opened.

efficiently. This experience started telling me 'if you know your enemies, you may not need to fight with them', so I gradually changed my mind. This alteration broke the wall of my stupidity. What I needed for this thesis was the attitude before the level of Newton's quote, i.e. recognising the previous forms of work. Therefore, I say "if I have successfully achieved the aim of this thesis it is by looking at the shoulders of giants."

Having said that, the agent-based modelling of my thesis produced significant improvements from a conventional mathematical approach. Also, agent-based modelling as a new comer in science has not been penetrated in or interacted with the conventional research field as Leombruni and Richiardi (2005) mention; '*Despite many years of active research in the field and a number of fruitful applications, agent-based modelling has not yet made it through to the top ranking economic journals*'. As a researcher in the agent-based modelling, my next step toward the conventional economic approach as a giant is not only 'standing on the shoulders of giants', but also 'interacting with the shoulders of giants'. Interaction effects are always important.

Bibliography

Anabel, J. (2002). Segmenting the leisure travel market using attitudinal theory, *Proceedings of the Universities Transport Study Group, 34th Annual Conference 3-5 January 2002*, Edinburgh, pp. 67.1–67.14.

Anderson, T. and Kanuka, H. (2002). *E-Research: Methods, Strategies, and Issues*, Allyn & Bacon, Upper Saddle River, chapter 1 Introduction.

Andou, T. (2003). *Java/Eclipse Software test tutorial book*, Shuwa system, Tokyo.

Arentze, T. and Timmermans, H. (2003). Modeling learning and adaptation processes in activity-travel choice: A framework and numerical experiment, *Transportation* **30**(1): 37–62.

Arnott, R. and Rowse, J. (1999). Modeling parking, *Journal of Urban Economics* **45**(1): 97–124.

Arrow, K. J. (1951). *Social Choice and Individual Values*, New York: John Wiley.

Arthur, W. B. (1994). Inductive reasoning and bounded rationality, *American Economic Review* **84**: 406–411. `http://www.santafe.edu/arthur/Papers/El_Farol.html`.

Aumann, R. J. (1976). Agreeing to disagree, *Annals of Statistics* **4**(6): 1236–1239.

Aumann, R. J. (1995). Backward induction and common knowledge of rationality, *Games and Economic Behavior* **8**(1): 6–19.

Axelrod, R. (1997). Advancing the art of simulation in the social sciences, *in* R. Conte, R. Hegselmann and P. Terna (eds), *Simulating Social Phenomena*, Springer-Verlag, Berlin, pp. 21–40.

Axhausen, K. W. and Polak, J. W. (1991). Choice of parking: Stated preference approach, *Transportation* **18**(1): 59 – 81.

Baba, N. and Jain, L. C. (2001). *Computational Intelligence in Games*, Vol. 62 of *Studies in Fuzziness and Soft Computing*, Physica Verlag, Heidelberg.

Baggio, R. (2002). My penalty miss cost italy the world cup? Guardian Unlimited accessible from `http://observer.guardian.co.uk/osm/story/0,6903,716542,00.html`, an edited extract from *'Una Porta Nel Cielo' (A Goal In The Sky), Roberto Baggio's autobiography.*

Balci, O. (1994). Validation, verification, and testing techniques throughout the life cycle of a simulation study, *Proceedings of the 26th conference on Winter simulation*, Society for Computer Simulation International, pp. 215–220.

Balmer, M., Nagel, K. and Raney, B. (2004). Large-scale multi-agent simulations for transportation applications, *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations.*

Bates, J., Brewer, M., Hanson, P., McDonald, D. and Simmonds, D. C. (1991). Building a strategic model for edinburgh, *Proceedings of Seminar D, PTRC 19th Summer Annual Meeting*, PTRC, London.

Baumol, W. J. and Vinod, H. D. (1970). An inventory theoretic model of freight transport demand, *Management Science* **6**: 413–421.

Bazzan, A. L., Bordini, R. H., Andrioti, G. K. and Vicari, R. M. (2000). Wayward agents in a commuting scenario - personalities in the minority game. citeseer.nj.nec.com/bazzan00wayward.html.

Beck, K. (1994). Simple smalltalk testing: With patterns. http://www.xprogramming.com/testfram.htm.

Ben-Akiva and Bierlaire (2003). Discrete choice models with applications to departure time and route choice, *in* 2nd (ed.), *Handbook of Transportation Science*, Kluwer, pp. 7–37.

Ben-Akiva, M. and Bonsall, P. (2004). Increasing the relevance and utility to practice of transportation research, *Transport Policy* **11**(2): 101–103.

Ben-Akiva, M. and Lerman, S. R. (1985). *Discrete choice analysis : theory and application to travel demand*, MIT Press, Cambridge.

Ben-Akiva, M. E. (2005). *DynaMIT: Real Time Traffic Estimation and Prediction System*, Massachusetts Institute of Technology, Cambridge. http://web.mit.edu/its/papers/DynaMIT.pdf, Accessed on February 22, 2005.

Ben-Akiva, M. E., Ramming, M. S. and Walker, J. L. (1999). Improving behavioral realism of urban transportation models through explicit treatment of individuals' spatial ability, *the conference on Social Change and Sustainable Transport*, Berkeley, California.

Ben-Akiva, M., Walker, J., Bernardino, A. T., Gopinath, D. A., Morikawa, T. and Polydoropoulou, A. (2002). Integration of choice and latent variable models, *in* H. S. Mahmassani (ed.), *In Perpetual Motion: Travel Behaviour Research Opportunities and Application Challenges*, Elsevier Science, chapter 21, pp. 431–470.

Berger, T. (2001). Agent-based spacial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis, *Agricultural economics* **25**: 245–260.

Bierlaire, M. (2001). The acceptance of modal innovation: The case of swissmetro, *1st Swiss Transport Research Conference*, Monte Verità.

Bierlaire, M. (2003). Biogeme: a free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transport Research Conference*, Monte Verita, Ascona.

Bolduc, D. and Ben-Akiva, M. (1991). A multinomial probit formulation for large choice sets, *Proceedings of the 6th International Conference on Travel Behaviour*, Vol. 2, pp. 243–258.

Bratman, M., Israel, D. J. and Pollack, M. (1988). Plans and resource-bounded practical reasoning, *Computational intelligence* **4**: 349–355.

Bugmann, D. V. and Coventry, K. (2004). A method for studying representation of action and cognitive distance, *26th Annual Meeting of the Cognitive Science Society*, Chicago.

Casini, L. and Macar, F. (1999). Multiple approaches to investigate the existence of an internal clock using attentional resources, *Behavioural Processes* **45**: 73–85.

Challet, D. and Zhang, Y.-C. (1997). Emergence of cooperation and organization in an evolutionary game, *Physica A* **246**: 407–415.

Challet, D., Marsili, M. and Zhang, Y.-C. (2004). *Minority Games Interacting agents in financial markets*, Oxford University Press, Oxford.

Cheng, Y. (1998). Hybrid simulation for resolving resource conflicts in train traffic rescheduling, *Computers in Industry* **35**(3): 233–246.

Chernick, M. R. (1999). *Bootstrap Methods A Practitioner's Guide*, John Wiley & Sons Inc., New York.

Ciuhandu, O. and Murphy, J. (2002). Transaction distribution algorithms with user classes for distributed application performance optimisation, *Proceedings of 10th IEEE International Conference on Software, Telecommunications and Computer Networks*, Italy.

Clark, G. L. (1998). Stylized facts and close dialogue: Methodology in economic geography, *Annals of the Association of American Geographers* **88**(1): 73–87.

Clark, G. L. and Wrigley, N. (1995). Sunk costs: a framework for economic geography, *Transactions of the Institute of British Geographers* **20**(2): 204–223.

Cleveland, W. S. (1981). LOWESS: A program for smoothing scatterplots by robust locally weighted regression, *The American Statistician* **35**: 54.

Cohen, M. D., Riolo, R. L. and Axelrod, R. (1999). The emergence of social organization in the prisoner's dilemma: How context-preservation and other factors promote cooperation, Working paper at Santfe Institute, RePEc:wop:safiwp:99-01-002. `http://ideas.repec.org/p/wop/safiwp/99-01-002.html`.

Collier, N., Howe, T. and North, M. (2003). Onward and upward: The transition to repast 2.0, *Proceedings of the First Annual North American Association for Computational Social and Organizational Science Conference*, June, Electronic Proceedings, Pittsburgh, PA USA, p. 5.

Davison, A. and Hinkley, D. (1997). *Bootstrap Methods and their Application*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

de Cara, M. A. R., Pla, O. and Guinea, F. (2000). Learning, competition and cooperation in simple games, *The European Physical Journal B* **13**: 413–416.

Department for Culture, Media and Sport (1998). Tourism towards sustainability: A consultation paper on sustainable tourism in the UK. http://www.culture.gov.uk/pdf/sustour1.pdf.

Department for Transport (2003). Summary advice on modelling, *Technical report*, Transport Analysis Guidance Unit. http://www.webtag.org.uk/webdocuments/2_Project_Manager/4_Summary_Advice_on_Modelling/.

Department for Transport (2004). Values of time and operating costs, *Technical report*, Transport Analysis Guidance Unit. http://www.webtag.org.uk/webdocuments/3_Expert/5_Economy_Objective/3.5.6.htm#1_2b.

Department of Environment, Transport and Regions (1998). A new deal for transport: Better for everyone. the government's white paper on the future of transport. This department is currently known as Department for Transport. http://www.dft.gov.uk/itwp.

Derbyshire County Council (2000). Road charging scheme gets royal seal of approval, *News from Derbyshire County Council*. http://www.derbyshire.gov.uk/news/2000/dec/001211a.htm.

Derbyshire County Council (2002). Consultation on charging scheme delayed, *News from Derbyshire County Council*. http://www.derbyshire.gov.uk/news/2002/jul/020716a.htm.

Derbyshire County Council (2003). Re: toll and bus fare for road user charging at the Upper Derwent valley, *e-mail and interview through phone, 2nd – 6th, June*. Currently (Nov., 2003), road user charging scheme is sensitive issue so the interviewees' names are not printed here.

Dia, H. (2002). An agent-based approach to modelling driver route choice behaviour under the influence of real-time information, *Transportation Research Part C: Emerging Technologies* **10**(5-6): 331–349.

Dillon, J. (2002). Road tolls could cut congestion and pollution, *Independent*. 24 February.

Doran, J., Palmer, M., Gilbert, N., and Mellars, P. (1994). The EOS project: Modelling upper paleolithic social change, *in* N. Gilbert and J. Doran (eds), *Simulating Societies: the Computer Simulation of Social Phenomena*, UCL Press, London, pp. 195–221.

Downing, T. E., Moss, S. and Pahl-Wostl, C. (2001). Understanding climate policy using participatory agent-based social simulation, *in* S. Moss and P. Davidsson (eds), *Multi-Agent-Based Simulation: Second International Workshop, MABS 2000*, Vol. 1979 of *LNAI*, Springer, Berlin, pp. 198–213.

Eckton, G. D. C. (2003). Road-user charging and the Lake District National Park, *Journal of Transport Geography* **11**(4): 307–317.

Edmonds, B. (1999a). Gossip, sexual recombination and the El Farol bar: modelling the emergence of heterogeneity, *Journal of Artificial Societies and Social Simulation*. `http://jasss.soc.surrey.ac.uk/2/3/2.html`.

Edmonds, B. (1999b). Modelling socially intelligent agents, *Applied Artificial Intelligence* **12**: 677–699.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife, *The Annals of Statistics* **7**: 1–2.

Epstein, J. M. and Axtell, R. (1996). *Growing Artificial societies: Social science from the bottom up*, Brookings Institute, Washington D.C.

Febbraro, A. D. and Sacco, N. (2004). On modelling urban transportation networks via hybrid petri nets, *Control Engineering Practice* **12**(10): 1225–1239.

Fogel, D. B., Chellapilla, K. and Angeline, P. J. (1999). Inductive reasoning and bounded rationality reconsidered, *IEEE Trans. on Evolutionary Computation* **3**(2): 142–146.

Fortin, C. and Rousseau, R. (1998). Interference from short-term memory processing on encoding and reproducing brief durations, *Psychological Research* **61**(4): 269 – 276.

Fowkes, A. S. (2000). Recent developments in stated preference techniques in transport research, *in* J. Ortézar (ed.), *Stated Preference Modelling Techniques*, Vol. 4 of *PTRC Perspectives*, PTRC Education and Research Services Ltd, London, pp. 37–52.

Frederick, S., Loewenstein, G. and O'Donoghue, T. (2002). Time discounting and time preference: A critical review, *Journal of Economic Literature* **40**(2): 351–401.

Fudenberg, D. and Tirole, J. (1991). *Game theory*, MIT Press, Cambridge.

Fujii, S. and Kitamura, R. (2003). Drivers' cognition and dynamics of traffic condition induced by network modification, *in* M. Bell and Y. Iida (eds), *The Network Reliability of Transport*, Pergamon, Oxford, pp. 209–223.

Fulton, L. M., Noland, R. B., Meszler, D. J. and Thomas, J. V. (2000). A statistical analysis of induced travel effects in the U.S. mid-Atlantic region, *Journal of Transportation and Statistics* **3**(1): 1–14.

Gilbert, N. and Troitzsch, K. G. (1999a). *Simulation for the Social Scientist*, Open University Press, Buckingham.

Gilbert, N. and Troitzsch, K. G. (1999b). *Simulation for the Social Scientist*, Open University Press, Buckingham, chapter 9 Learing and evolutionary models, pp. 195–236.

Giunchiglia, F., Benerecetti, M. and Serafini, L. (1998). Multiagent systems verification via model checking, *in* F. V. Harmelen (ed.), *European workshop on validation and verification of knowledge-based systems*, Povo (Tn).

Gracyk, T. (2003). Hume's aesthetics, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, The Metaphysics Research Lab, Stanford University. `http://plato.stanford.edu/archives/win2003/entries/hume-aesthetics/`.

Greene, W. H. (2003). *Econometric analysis*, 5th edn, Prentice Hall International, London.

Haugen, K. K. and Hervik, A. (2004). A game theoretic 'mode-choice' model for freight transportation, *The Annals of Regional Science* **38**(38): 469.

Hess, S. and Polak, J. (2004). Mixed logit estimation of parking type choice, *the 83rd Annual Meeting of the Transportation Research Board*, Washington D.C.

Hess, S., Bierlaire, M., and Polak, J. W. (2005). Estimation of value of travel-time savings using mixed logit models, *Transportation Research Part A: Policy and Practice* **39**(2-3): 221–236.

Hinkley, D. V. (1988). Bootstrap methods, *Journal of the Royal Statistical Society. Series B* **50**(3): 321–337.

Honore, B. E. and Kyriazidou, E. (2000). Panel data discrete choice models with lagged dependent variables, *Econometrica* **68**(4): 839–874.

Hopcroft, J. E. and Ullman, J. D. (1979). *Introduction to Automata Theory, Languages, and Computation*, Addison Wesley, Reading, MA.

Hovland, C. I. and Sherif, M. (1952). Judgmental phenomena and scale of attitude measurement: Item displacement in thurstone scales, *Journal of Abnormal and Social Psychology* **47**: 822–832.

Hughes, B. D. (1995). *Random Walks*, Vol. 1 of *Random Walks and Random Environments*, Oxford University Press, Oxford.

Hurley, W. J. and Petersen, E. R. (1994). Nonlinear tariffs and freight network equilibrium, *Transportation Science* **28**: 236–245.

IBM (2005). Deep blue. `http://www.research.ibm.com/deepblue/`, Accessed on February 25, 2005.

Ihaka, R. and Gentleman, R. (1996). R: a language for data analysis and graphics, *Journal of Computational and Graphical Statistics* **5**: 299–314.

Jain, S. and Neal, R. (2000). A split-merge markov chain monte carlo procedure for the dirichlet process mixture model.

Jha, M., Madanat, S. and Peeta, S. (1998). Perception updating and day-to-day travel choice dynamics in traffic networks with information provision, *Transportation Research Part C: Emerging Technologies* **6**(3): 189–212.

Jotisankasa, A. and Polak, J. W. (2005). Experimental investigation of day-to-day traveller learning and adaptation in route and departure time choice behaviour, *Proceedings of 37th UTSG Conference*, 5C1.1, Bristol.

Karp, R. M. (1972). Reducibility among combinatorial problems, *in* R. E. Miller and J. W. Thatcher (eds), *Complexity of Computer Computations*, Plenum Press, pp. 85–103.

Klügl, F. and Bazzan, A. L. C. (2004). Route decision behaviour in a commuting scenario: Simple heuristics adaptation and effect of traffic forecast, *Journal of Artificial Societies and Social Simulation*. `http://jasss.soc.surrey.ac.uk/7/1/1.html`.

Knoblauch, R. L., Pietrucha, M. T. and Nitzburg, M. (1996). Field studies of pedestrian walking speed and start-up time, *Transportation Research Record*. No. 1538 Pedestrian and Bicycle Research.

Konolige, K. (1986). *A deduction models of belief*, Pitman, London.

Kripke, S. (1980). *Naming and Necessity*, Harvard University Press, Cambridge, MA.

Lee, K., Hui, P. M., Wang, B.-H. and Johnson, N. F. (2001). Effects of announcing global information in a two-route traffic flow model, *Journal of the Physical Society of Japan* **70**(12): 3507–3510.

Leombruni, R. and Richiardi, M. (2005). Why are economists sceptical about agent-based simulations?, *Physica A: Statistical and Theoretical Physics* **355**(1): 103–109.

Lindner, C. C. and Rodger, C. A. (1997). *Design Theory*, CRC Press, Boca Raton, FL.

Liu, X., Liang, X. and Tang, B. (2004). Minority game and anomalies in financial markets, *Physica A: Statistical and Theoretical Physics* **333**: 343–352.

Lofton, K. (2000). Bounded rationality and the El Farol bar, Working Paper POLS907. `http://lark.cc.ukans.edu/~lofton/swarm/`.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*, Wiley, New York.

Luce, R. D. and Suppes, P. (1965). Preference, utility, and subjective probability, *in* R. D. Luce, R. R. Bush and E. Galanter (eds), *Handbook of mathematical psychology*, Vol. 3, Wiley, New York, pp. 249–410.

Mackie, P. and Preston, J. (1998). Twenty-one sources of error and bias in transport project appraisal, *Transport Policy* **5**(1): 1–7.

Mackie, P., Fowkes, A., Wardman, M., Whelan, G., Nellthorp, J. and Bates, J. (2003). Value of travel time savings in the UK: summary report, *Report to department for transport*, Institute for Transport Studies, University of Leeds. `http://www.dft.gov.uk/stellent/groups/dft_econappr/documents/page/dft_econappr_022708.hcsp`.

Manski, C. (1977). The structure of random utility models, *Theory and decision* **8**: 229–254.

Marcus, G. (2004). *The Birth of the Mind: How a Tiny Number of Genes Creates the Complexities of Human Thought*, Basic Books.

Mas-Colell, A., Whinston, M. D. and Green, J. R. (1995). *Microeconomic theory*, Oxford University Press, Oxford, chapter 8 Simultaneous-move games.

McCarthy, J. (2004). What is artificial intelligence?, Computer Science Department, Stanford University, `http://www-formal.stanford.edu/jmc/whatisai/whatisai.html`, Revised November 24, 2004.

McCarthy, P. S. (2001). *Transportation Economics: Theory and practice a case study appoach*, Blackwell, Oxford.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour, *in* P. Zarembka (ed.), *Frontiers in Econometrics*, Academic Press, New York.

McFadden, D. (1981). Econometric models of probabilistic choice, *in* C. Manski and D. McFadden (eds), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, pp. 198–272.

McNally, M. G. (2000). The four step model, *in* D. A. Hensher and K. J. Button (eds), *Handbook of Transport Modelling*, Elsevier, Oxford, chapter 3, pp. 35–52.

Meyer, J.-J. and Lomuscio, A. (2003). Modal logic and multi-agent systems specifications, *Proceedings of 5th European Agent Systems Spring School*, Agent-link.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information, *The Psychological Review* **63**: 81–97.

Miller, H. J. and Lung Shaw, S. (2002). *Geographic information systems for transportation : principles and applications*, Spatial information systems, Oxford University Press, Oxford.

Mohring, H. (1972). Optimization and scale economies in urban bus transportation, *American Economic Review* **62**(4): 591–604.

Mokhtarian, P. L. and Salomon, I. (2001). How derived is the demand for travel? some conceptual and measurement considerations, *Transportation Research Part A: Policy and Practice* **35**(8): 695–719.

Moss, S. (2002). Policy analysis from first principles, *Proceedings of the National Academy of Sciences* **99**(3): 7267–7274.

Moss, S. and Edmonds, B. (2004). Towards good social science, *CABDyN (Complex Agent-based Dynamic Network) workshops 2004*, Oxford. `http://cfpm.org/papers/tgss/`.

Moss, S., Edmonds, B. and Wallis, S. (1997). Validation and verification of computational models with multiple cognitive agents, *Technical report*, Manchester Metropolitan University, Centre for Policy Modelling.

Nagel, K. (2004). Multi-agent transportation simulation, `http://www.sim.inf.ethz.ch/publications/book/` draft-in-progress for a book – February 2, 2004.

Nakayama, S. and Kitamura, R. (2000). Route choice model with inductive learning, *Transportation Research Record* **1725**: 63–70.

Nash, C. (2003). Marginal cost pricing and other pricing principles for user charging in transport: a comment, *Transport Policy* **10**: 345–348.

National Statistics (2001). Census 2001: Greater manchester, web-published. `www.statistics.gov.uk/census2001/pop2001/greater\_manchester\_met\_county.asp`, Accessed on April 18, 2003.

Neumann, J. V. and Morgenstern, O. (1953). *Theory of games and economic behavior*, 3rd edn, Princeton University Press, Princeton.

Newton, I. (1676). Letter to robert hooke, 5 feb 1676, *Bloomsbury Biographical Dictionary of Quotations (1997)*, Bloomsbury. http://www.xreferplus.com/entry.jsp?xrefid=169561&secid=-.

Norris, J. (1997). *Markov chains*, Cambridge University Press, Cambridge.

Oi, W. and Shuldiner, P. (1962). *An Analysis of Urban Travel Demands*, Northwestern University Press, Evanston, IL.

Oota, J. (1995). Introduction to petri net, *Technical report*, Faculty of Information Science and Technology, Aichi prefectural university. http://www.aichi-pu.ac.jp/ist/~qua/introPN.pdf.

Ortúzar, J. d. D. and Willumsen, L. G. (2001). *Modelling transport*, 3rd edn, John Wiley, Chichester.

Oxford University Computing Services (2005). Writing your own safe perl. http://www.oucs.ox.ac.uk/web/faq/index.xml.ID=safeperl, Accessed on February 28, 2005.

Page, S. J. (1999). *Tranport and Tourism*, Longman, Essex.

Parunak, V., Savit, R. and Riolo, R. (1998). Agent-based modeling vs. equation-based modeling: A case study and users guide, *Proceedings of Workshop on Multi-agent systems and Agent based Simulation (MABS'98)* pp. 10–25.

Peak District National Park Authority (2001). The Local Plan: Appendix 1 Car Parking Standards in Derbyshire November 1994. http://www.peakdistrict.org/pubs/lp/appendix1.htm.

Peeta, S., Zhang, P. and Zhou, W. (2005). Behavior-based analysis of freeway car-truck interactions and related mitigation strategies, *Transportation Research Part B: Methodological* **39**(5): 417–451.

Peterson, J. L. (1981). *Petri net theory and the modeling of systems*, Prentice-Hall, Englewood Cliffs, N.J.

Pfeiffer, P. E. and Schum, D. A. (1973). *Introduction to Applied Probability Theory*, Academic Press, New York.

Pole, A., West, M. and Harrison, J. (1999). *Applied Bayesian Forecasting and Time Series Analysis*, Chapman & Hall. p409.

Polydoropoulou, A., Gopinath, D. A. and Ben-Akiva, M. (1997). Willingness to pay for advanced traveler information systems : SmarTraveler case study, *Transportation Research Record*. paper no. 970926.

Raney, B., Cetin, N., Vollmy, A., Vrtic, M. and Axhausen, K. (2003). An agent-based microsimulation model of Swiss travel: First results, *Networks and Spatial Economics* **3**(1): 23–41.

Rao, A. S. and Georgeff, M. P. (1995). BDI-agents: from theory to practice, *Proceedings of the First Intl. Conference on Multiagent Systems*, San Francisco.

Rasmusen, E. (2001). *Games and Information: An Introduction to Game Theory*, 3rd edn, Blackwell, Oxford.

Redmond, L. S. and Mokhtarian, P. L. (2001). The positive utility of the commute: modeling ideal commute time and relative desired commute amount, *Transportation* **28**(2): 179–205.

Ripley, B. D. (1987). *Stochastic simulation*, John Wiley & Sons Inc., New York.

Rose, J. and Hensher, D. A. (2004). Modelling agent interdependency in group decision making, *Transportation Research Part E: Logistics and Transportation Review* **40**(1): 63–79.

Ross, S. M. (1997). *Simulation*, 2nd edn, Academic.

Rothengatter, W. (2003). How good is first best? marginal cost and other pricing principles for user charging in transport, *Transport Policy* **10**: 121–130.

Rourke, L. and Anderson, T. (2002). Using peer teams to lead online discussions, *Journal of Interactive Media in Education*.

Roy, J. A. (2002). Carolene products: a game-theoretic approach, *Brigham Young University Law Review* pp. 53–109.

Samuelson, P. A. (1937). A note on measurement of utility, *Review of Economic Studies* **4**: 155–61.

Sandholm, T. (1999). Distributed rational decision making, *in* G. Weiß (ed.), *Multiagent Systems*, MIT Press, Cambridge, pp. 201–258.

SAS Institute Inc. (1999). SAS OnlineDoc, Version 8 – Chi-Square Tests and Statistics. `www.id.unizh.ch/software/unix/statmath/sas/sasdoc/stat/chap28/sect19.htm`.

Savit, R., Manuca, R., Li, Y. and Riolo, R. (1998). The dynamics of minority competition, UM Center for Study of Complex Systems Working Paper. `http://cscs.umich.edu/research/techReports.html`.

Schelling, T. C. (1960). *The strategy of conflict*, Harvard University Press, Cambridge, Mass.

Schleiffer, R. (2002). Intelligent agents in traffic and transportation, *Transportation Research Part C: Emerging Technologies* **10**(5-6): 325–329.

Schleiffer, R. (2005). An intelligent agent model, *European Journal of Operational Research* **166**(3): 666–693.

Schlesinger, M. (2001). The agent-based approach: a new direction for computational models of development, *Development Review* **21**: 121–146.

Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games, *International journal of game theory* **4**: 25–55.

Siegfried, J. J. (1970). A first lesson in econometrics, *Journal of Political Economy* **78**(6): 1378–1379.

Simon, H. A. (1957). *Models of man*, Macmillian Press.

Spear, B. D. (1977). Applications of new travel demand forecasting techniques to transportation planning: A study of individual choice models, *Final report*, Urban Planning Division, Office of Highway Planning, Federal Highway Administration, U.S. Department of Transportation, Washington D.C. `http://ntl.bts.gov/DOCS/SICM.html`.

Steiner, T. J. and Bristow, A. L. (2000). Road pricing in national parks: a case study in the Yorkshire Dales National Parks, *Transport Policy* **7**: 93–103.

Stevens, C. F. and Zador, A. M. (1998). Input synchrony and the irregular firing of cortical neurons, *Nature Neuroscience* **1**: 210–217.

Stopher, P. R. (2004). Reducing road congestion: a reality check, *Transport Policy* **11**(2): 117–131.

Sugden, R. (1998). Anomalies and biases in the contingent valuation method, *in* D. Banister (ed.), *Transport policy and the environment*, E and FN Spon, London, chapter 14, pp. 298–319.

Sunitiyoso, Y. and Matsumoto, S. (2005). Modelling behavioural change in a social dilemma of commuters' mode choice, *Proceedings of 37th UTSG Conference*, 5C3.1, Bristol.

Sutton, R. and Barto, A. (1998). *Reinforcement learning: an introduction*, MIT Press, Cambridge, Mass. http://www.cs.ualberta.ca/%7Esutton/book/ebook/the-book.html.

Takama, T. (2004a). Agent-based modelling of a transport mode choice with the minority game, *CABDyN (Complex Agent-based Dynamic Network) workshops 2004*, Oxford.

Takama, T. (2004b). Methodology and a case study of statistical agent based modelling of discrete choices: Comparison and integration with econometrics models, *Proceedings of 8th World Multiconference on Systemics, Cybernetics and Informatics*, Vol. XIV, International Institute of Information and Systemics, Orlando, pp. 134–139.

Takama, T. (2005). Agent-based model on the road user charging scheme at the Upper Derwent Valley, the Peak District National Park, *Proceedings of 37th UTSG Conference*, 5C3.2, Bristol.

Thomopoulos, N. (2004). *Road user charging in Upper Derwent Valley: Commodification in a national park?*, Dissertation for M.Sc. in Nature, Society and Environmental Policy, University of Oxford, School of Geography and the Environment.

Thompson, R. G. and Richardson, A. J. (1998). A parking search model, *Transportation Research Part A: Policy and Practice* **32**(3): 159–170.

Tobias, R. and Hofmann, C. (2004). Evaluation of free java-libraries for social-scientific agent based simulation, *Journal of Artificial Societies and Social Simulation*. http://jasss.soc.surrey.ac.uk/7/1/6.html.

Tolley, R. and Turton, B. (1995). *Transport systems, policy and planning: a geographical approach*, Longman Scientific & Technical, Harlow.

Train, K. (2003). *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge.

TranSafety, Inc. (1997). Study compares older and younger pedestrian walking speeds, *Road Engineering Journal*.

Tsamboulas, D. A. (2001). Parking fare thresholds: a policy tool, *Transport Policy* **8**(2): 115–124.

Unger, P. (1978). *Ignorance: A Case for Scepticism*, University Press, Oxford, chapter 5 From Ignorance to Irrationality, pp. 197–211. Oxford Scholarship Online `http://www.oxfordscholarship.com/oso/public/content/philosophy/0198244177/toc.html`.

van Berkum, E. C. and van Der Mede, P. H. J. (1998). The impact of dynamic traffic information: Modelling approach and empirical results, *in* J. d. Ortuzar, D. A. Hensher and S. R. Jara-Diaz (eds), *Travel Behaviour Research: Updating the State of Play*, Pergamon, Oxford, pp. 401–421.

van Damme, E. (1983). *Refinements of the Nash Equilibrium Concept*, Springer, Berlin.

VanBreukelen, G. J. P. (1995). Parallel information processing models compatible with lognormally distributed response times, *Journal of Mathematical Psychology* **39**(4): 396–399.

Varian, H. R. (1993). *Intermediate microeconomics : a modern approach*, 3rd edn, W.W. Norton, London.

Viegas, J. M. (2001). Making urban road pricing acceptable and effective: searching for quality and equity in urban mobility, *Transport Policy* **8**(4): 289–294.

Vose, D. (2000). *Risk analysis : a quantitative guide*, 2nd edn, John Wiley, Chichester.

Weisstein, E. W. (2003). NP-complete problem, Eric Weisstein's World of Mathematics. `http://mathworld.wolfram.com/NP-CompleteProblem.html`, Accessed on January 17, 2003.

Winston, C. (1985). Conceptual developments in the economics of transportation, *Journal of Economic Literature* pp. 57–94.

Wooldridge, M. (2002). *An introduction to multiagent systems*, John Wiley and Sons, Chichester.

Wooldridge, M. and Jennings, N. L. (1995). Intelligent agents: Theory and practice, *Knowledge Engineering Review* **10**(2): 115–152.

Wu, J., Brackstone, M. and McDonald, M. (2000). Fuzzy sets and systems for a motorway microscopic simulation model, *Fuzzy Sets and Systems* **116**(1): 65–76.

Appendix A

Questionnaire distributed at the Upper Derwent Valley

Questionnaires were distributed during the summer of 2003 for this thesis. Pages from 217 to 224 present the master copy of the questionnaires. To fulfill the regulation of this thesis for the left margin (i.e the left margin should be between 1.5 and 1.25 inches), the questionnaire was reduced in size. The box surrounding the questionnaire was the A4 size when it was distributed in 2003.

D.Phil student:
(Mr.) Takeshi TAKAMA

Transport Studies Unit
University of Oxford
11, Bevington Road
OXFORD, OX2 6NB

Tel. : ++44 (1865) 274715
Tel. : ++44 (1865) 274709
Fax : ++44 (1865) 515194
Email: takeshi.takama@tsu.ox.ac.uk
Homepage: users.ox.ac.uk/∼scat1898

The impact of transport around the Upper Derwent Valley

Dear Sir / Madam,

I am a doctoral (D.Phil) student at the University of Oxford and I am conducting the above project for my doctoral thesis. The main objective of the project is to assess the extent to which transport impacts on the natural beauty of the Upper Derwent Valley. **IT WOULD HELP ME GREATLY IF YOU FILLED IN THIS QUESTIONNAIRE.**

This survey is purely for academic purpose only. Your information will be used only for this research and your privacy will be respected at all times.

If any of the questions are not clear please get in touch with me (Takeshi Takama), by e-mail `takeshi.takama@tsu.ox.ac.uk`, or call me at (01865) 274709. More contact details are listed at the top of this page. Also, you can find the latest information on this project on my homepage, "*http://users.ox.ac.uk/∼scat1898/*". Please return the questionnaire with the supplied FREEPOST envelope, as soon as possible. No additional stamp is required. Thank you for assisting me with this research project.

Yours faithfully,

Takeshi Takama
Transport Studies Unit
University of Oxford

PART 1: Household characteristics

Please tick the relevant boxes ☑, except where otherwise specified.

Section 1 - Personal Details

1. How old are you?
   - ☐**1:** Under18
   - ☐**2:** 18 – 24
   - ☐**3:** 25 – 34
   - ☐**4:** 35 – 44
   - ☐**5:** 45 – 54
   - ☐**6:** 55 – 64
   - ☐**7:** Over 65
2. Are you
   - ☐**1:** Male
   - ☐**2:** Female?
3. What is your occupation?
   - ☐**1:** Managerial
   - ☐**2:** Professional
   - ☐**3:** Skilled Non-Manual
   - ☐**4:** Skilled Manual
   - ☐**5:** Semi/Unskilled Manual
   - ☐**6:** Student/School child ⇒ **Go to question 5**
   - ☐**7:** Housewife ⇒ **Go to question 6**
   - ☐**8:** Unemployed ⇒ **Go to question 6**
   - ☐**9:** Retired ⇒ **Go to question 6**
   - ☐**10:** Others
4. What is your personal income before tax?
   - ☐**1:** None
   - ☐**2:** Below £10,000
   - ☐**3:** £10,000 – £19,999
   - ☐**4:** £20,000 – £29,999
   - ☐**5:** £30,000 – £39,999
   - ☐**6:** £40,000 – £49,999
   - ☐**7:** Over £50,000
5. What is your main method to get to work/school?
   - ☐**1:** Car driver
   - ☐**2:** Car passenger
   - ☐**3:** Bus/Coach
   - ☐**4:** Park and Ride
   - ☐**5:** Train
   - ☐**6:** Walking
   - ☐**7:** Bicycle
   - ☐**8:** Other
6. Do you have a full driving licence?
   - ☐**1:** Yes, Motor Vehicle
   - ☐**2:** Yes, Motorbike
   - ☐**3:** Yes, both
   - ☐**4:** No licence
7. Please read the description below and indicate which most closely reflects your view **(Tick one box only)**
   - ☐**1:** I would be willing to reduce my car use, but I do not see the need at present
   - ☐**2:** I would be willing to reduce my car use, but I do not have sufficient information about the environmental and social impact of reducing car use.
   - ☐**3:** I am strongly resistant to reducing car use.
   - ☐**4:** I could reduce my car use, but have problems using the alternatives.
   - ☐**5:** I could be willing to reduce car use due to awareness of the environmental impact of car use.
   - ☐**6:** I do not use a car because I have problems or difficulty using it. **(Please state your problem or difficultly below.)**

     ____

Section 2 - Household Details

1. How many people are in your household (including yourself)? **(Please state the number below)**

   ____

2. How many cars and vans do you have in your household? **(Please state the number below)**

   ____

3. Where is your home location?

   ☐**1:** Greater Manchester
   ☐**2:** Sheffield
   ☐**3:** Other locations in South Yorkshire
   ☐**4:** Buxton
   ☐**5:** Bakewell
   ☐**6:** Other locations in Derbyshire
   ☐**7:** West Yorkshire
   ☐**8:** Staffordshire
   ☐**9:** Lancashire
   ☐**10:** Nottinghamshire
   ☐**11:** Cheshire
   ☐**12:** Merseyside
   ☐**13:** Other (please state)

   ____

PART 2: The trip to Upper Derwent Valley

Please tick the relevant boxes ☑, except where otherwise specified.

Section 1 - Questions about today's trip to the Valley

1. Please give travel times.

   [From home to the Valley]

   What time did you leave **your home**? At ________________

   What time did you arrive at **a parking area in the Upper Derwent Valley**? At ________________

   [From the Valley to home]

   What time did you leave **the parking area in the Upper Derwent Valley**? At ________________

   What time did you arrive at **your home**? At ________________

2. How did you travel on Derwent Lane to the Information Centre, today?

   ☐**1:** By car / van
   ☐**2:** By Bus / coach
   ☐**3:** By cycling
   ☐**4:** By walking

3. Please state the costs of today's trip.

   (a) Parking costs **(Please state below)**

   £________________

   (b) Other transportation cost **(Please state below)**

   £________________ (e.g. petrol)

   (c) Any other costs **(Please state below)**

   £________________

4. Did you have a car available for the trip you made today?

   ☐**1:** Yes
   ☐**2:** No

5. How many people are travelling with you today (**including yourself**)? ___________

   (a) How many of these are children under 5? ___________

   (b) How many of these are children between 5 and 15? ___________

   (c) How many of these are full time students or between 15 and 26? ___________

Section 2 - General questions about your travel to the valley

1. How did you hear about the Upper Derwent Valley?

   ☐**1:** Common knowledge
   ☐**2:** Tourist leaflet
   ☐**3:** Tourist information centre
   ☐**4:** From a friend
   ☐**5:** Advertising on TV *etc.*
   ☐**6:** Other

2. Can you guess how many of your friends go to the Upper Derwent Valley more than once a year?

   ☐**1:** No idea ⇒ **Go to question 3**
   ☐**2:** Yes **(Please state the number below)**

   ____

   (**Please guess if you are not sure about this**)

   (a) How many of them go to the valley between 2 and 5 times a year? ___________
   (b) How many of them go to the valley between 6 and 12 times? ___________
   (c) How many of them go to the valley around once every two weeks?___________
   (d) How many of them go to the valley around or more than once a week?___________

3. How often do you come to the Upper Derwent Valley?

   ☐**1:** less than once a year
   ☐**2:** around once a year
   ☐**3:** between 2 and 5 times a year
   ☐**4:** between 6 and 12 times a year
   ☐**5:** around once every two weeks
   ☐**6:** around once a week or more

4. What do you think of the impact of other cars in the car parks of the Valley when it is busy (e.g. noise, visual impact and exhaust from cars rounding in a car park to find a space to park)

   ☐**1:** Significant
   ☐**2:** Fairly significant
   ☐**3:** Moderate
   ☐**4:** Not a big problem
   ☐**5:** No problem
   ☐**6:** Not experienced

5. Did you go to the Upper Derwent Valley before today?

   ☐**1:** Yes ⇒ **Go to question 6**
   ☐**2:** No ⇒ **Go to question 7**

6. Please tick **ALL** things you remember in **EACH** past trip to the Upper Derwent Valley.

	The last trip	The 2nd last trip	The 3rd last trip	The 4th last trip	The 5th last trip
1:Travel time & date	☐**1**	☐**1**	☐**1**	☐**1**	☐**1**
2:Travel costs	☐**2**	☐**2**	☐**2**	☐**2**	☐**2**
3:The location you park your car	☐**3**	☐**3**	☐**3**	☐**3**	☐**3**
4:The time you spent in the Valley	☐**4**	☐**4**	☐**4**	☐**4**	☐**4**
5:Traffic situation (including in car parks)	☐**5**	☐**5**	☐**5**	☐**5**	☐**5**

7. What problems, if any, have you encountered on your trip to the Upper Derwent Valley?

Please write comments here!

PART 3: Hypothetical Trip to the Valley

Explanation of Terms used - "Please see the map below"

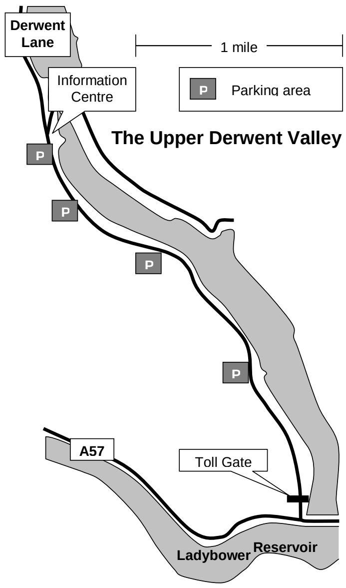
Road User Charge (£): It aims to limit traffic volumes by charging motorists to use Derwent Lane. A toll to enter Derwent Lane from the A57 will be charged.

Park & Ride (£): Those who do not want to pay the road user charge will be able to use a new bus service with a low fare. The service links to local car parks, Bamford train station, and the Upper Derwent Valley. The location for the parking site around the toll gate of the road user charging scheme has not been decided.

Frequency of bus service (minutes): This is the period between departure times of the shuttle buses for Park & Ride.

Searching & walking time (minutes): This is the combination of the period of searching for a space to park and walking time from a parking area to the Information Centre.

Parking fee (£): The parking fee for the Park & Ride service is the fee you pay when you park your car before getting on a bus. The parking fee for Toll & Ride is the fee you pay when you park your car at one of the four car parks along Derwent Lane.



In this part of questionnaire, I would like you to imagine travelling to the Upper Derwent Valley next weekend **WITH THE SAME TRIP MEMBERS AS TODAY**. Imagine you arrive at the entrance of Derwent Lane and you need to visit the Information Centre at Fairholmes, Upper Derwent Valley. Please **RANK** three options in each scenario. **The most preferred option is ranked "1" and the least preferred option is ranked "3".**

This is an example

1. Conditions to visit the Information Centre:

Park & Ride service		
Fare	**20p**	per person
A bus every	**15**	minutes
Parking fee	**50p**	per car

Toll & drive		
Toll	£**2.00**	per car
Searching for a parking space and walking to the centre takes	**30**	minutes
Parking fee	£**2.50**	per car

Under these circumstances, I would:

[2] **A:** Park & ride

[1] **B:** Pay toll and drive ⇐ **"Most preferred"**

[3] **C:** None of them (don't visit the valley)
⇑ **"The least preferred option"**

This is an example

PLEASE COMPLETE ALL 16 QUESTIONS, even though the stated conditions may be unrealistic to you. These are HYPOTHETICAL questions. Please answer the questions by **RANKING**, but **NOT TICKING** (i.e. CORRECT = "1, 2, 3" WRONG = "✓"). Please see the example on the previous page.

1. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**20p**	per person	Toll	**£1.00**	per car
A bus every	**5**	minutes	Searching for a parking space and walking to the centre takes	**1**	minutes
Parking fee	**10p**	per car	Parking fee	**50p**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

2. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**20p**	per person	Toll	**£2.00**	per car
A bus every	**15**	minutes	Searching for a parking space and walking to the centre takes	**30**	minutes
Parking fee	**50p**	per car	Parking fee	**£2.50**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

3. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**20p**	per person	Toll	**£3.00**	per car
A bus every	**30**	minutes	Searching for a parking space and walking to the centre takes	**50**	minutes
Parking fee	**50p**	per car	Parking fee	**£1.00**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

4. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**20p**	per person	Toll	**£5.00**	per car
A bus every	**45**	minutes	Searching for a parking space and walking to the centre takes	**15**	minutes
Parking fee	**50p**	per car	Parking fee	**£2.00**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

5. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**50p**	per person	Toll	**£1.00**	per car
A bus every	**15**	minutes	Searching for a parking space and walking to the centre takes	**15**	minutes
Parking fee	**50p**	per car	Parking fee	**£1.00**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

6. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**50p**	per person	Toll	**£2.00**	per car
A bus every	**5**	minutes	Searching for a parking space and walking to the centre takes	**50**	minutes
Parking fee	**50p**	per car	Parking fee	**£2.00**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

7. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**50p**	per person	Toll	£**3.00**	per car
A bus every	**45**	minutes	Searching for a parking space and walking to the centre takes	**30**	minutes
Parking fee	**10p**	per car	Parking fee	**50p**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

8. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**50p**	per person	Toll	£**5.00**	per car
A bus every	**30**	minutes	Searching for a parking space and walking to the centre takes	**1**	minutes
Parking fee	**50p**	per car	Parking fee	£**2.50**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

9. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**80p**	per person	Toll	£**1.00**	per car
A bus every	**30**	minutes	Searching for a parking space and walking to the centre takes	**30**	minutes
Parking fee	**50p**	per car	Parking fee	£**2.00**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

10. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**80p**	per person	Toll	£**2.00**	per car
A bus every	**45**	minutes	Searching for a parking space and walking to the centre takes	**1**	minutes
Parking fee	**50p**	per car	Parking fee	£**1.00**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

11. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**80p**	per person	Toll	£**3.00**	per car
A bus every	**5**	minutes	Searching for a parking space and walking to the centre takes	**15**	minutes
Parking fee	**50p**	per car	Parking fee	£**2.50**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

12. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	**80p**	per person	Toll	£**5.00**	per car
A bus every	**15**	minutes	Searching for a parking space and walking to the centre takes	**50**	minutes
Parking fee	**10p**	per car	Parking fee	**50p**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

13. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	£**1.00**	per person	Toll	£**1.00**	per car
A bus every	**45**	minutes	Searching for a parking space and walking to the centre takes	**50**	minutes
Parking fee	**50p**	per car	Parking fee	£**2.50**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

14. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	£**1.00**	per person	Toll	£**2.00**	per car
A bus every	**30**	minutes	Searching for a parking space and walking to the centre takes	**15**	minutes
Parking fee	**10p**	per car	Parking fee	**50p**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

15. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	£**1.00**	per person	Toll	£**3.00**	per car
A bus every	**15**	minutes	Searching for a parking space and walking to the centre takes	**1**	minutes
Parking fee	**50p**	per car	Parking fee	£**2.00**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

16. Conditions to visit the Information Centre:

Park & Ride service			Toll & drive		
Fare	£**1.00**	per person	Toll	£**5.00**	per car
A bus every	**5**	minutes	Searching for a parking space and walking to the centre takes	**30**	minutes
Parking fee	**50p**	per car	Parking fee	£**1.00**	per car

Under these circumstances, I would:

[ ] **A:** Park & ride

[ ] **B:** Pay toll and drive

[ ] **C:** None of them (don't visit the valley)

- How much would you be willing to pay to drive into the Derwent Lane if a road user charging is enforced?
  (Please state) __________

Please return the completed questionnaire to me in the envelope provided.

Thank you for taking part in the study.

If you are willing to participate further in this study by exchanging messages with me through email or telephone, could you please provide your contact details below?

Name: ______________________________

e-mail: ______________________________

Phone: ______________________________

Your information will be used only for this research and your privacy and e-mail address will be respected at all times.

You can find the latest information on this project in my homepage:

http://users.ox.ac.uk/∼scat1898/

Please feel free to visit.

Appendix B

Java source codes created in this thesis

Overall 221 classes were created during this thesis project. If they are printed out, the length of the source codes are much longer than this thesis (i.e. over 600 pages). Therefore, the hard copy of the source codes is not printed, but they are freely accessible within foreseeable future from the URL link below:

`http://www.geog.ox.ac.uk/~ttakama/files/`

When the URL link is opened, a popup window asks you for a user ID and password. Therefore, please type these correspondingly:

User ID: `thesis`

Password: `javasource`