Bringing ‘smarter choices’ into multi-modal modelling

Helen Bowkett

A thesis submitted in partial fulfilment of the requirements of the University of the West of England, Bristol for the degree of Doctor of Philosophy

Faculty of Environment and Technology
University of the West of England, Bristol

May 2015

Acknowledgements

I would like to express my gratitude to all members of the Centre of Transport Studies at the University of the West of England for their warm welcome and continued encouragement while I attempted to bring their insights into the transport related behaviour of people into my computer simulations of transport systems. In particular I would like to thank Dr. Kiron Chatterjee, my main supervisor, for his valued guidance during this process, Professor Graham Parkhurst for his supervisory support and Professor Glenn Lyons for facilitating my re-introduction into academic life.

I would also like to record my gratitude to my family for their patience while Mum was so often lost in her books or writing code on her laptop. My husband has been incredibly supportive and understanding, as have the children.

My attendance at Dr. Nathaniel Osgood's agent based modelling boot camp for health researchers in Saskatchewan was made possible by generous sponsorship from my former employer, Peter Brett Associates LLP. This intensive induction into the world of agent based models was the turning point in this research and provided me with a valuable new set of modelling skills. I am grateful for everyone who made it possible for me to attend the course and for Nate's enthusiasm and accessible teaching.

Finally I wish to thank friends in the transport modelling community for their continued interest in my work and for being patient listeners while I talked about my research. The 'Modelling on the Move' series of seminars from 2012 – 2014 organised by Dr. Rachel Aldred was particularly helpful as an opportunity to hear from and share with others. Hopefully together we can keep the focus of our models on the people and the decisions they make.

Abstract

Transport models are widely used in the preparation of advice to policy makers on the future performance of the transport system and the appraisal of transport schemes and policies. A recent focus of funding in the U.K. has been on ‘smarter choices’ measures which aim to increase the use of public transport, walking and cycling through information, marketing and low cost infrastructure interventions.

This study starts with the recent experience of UK transport modellers when assessing the impact of ‘smarter choices’ measures, using the widely applied ‘four stage’ model framework. It reports on interviews with the teams who built the latest generation of such transport models, developed with central government funding, and considers the reasons for their reported limited success in integrating the effects of ‘smarter choices’ measures into their multi-modal models.

Alternative modelling approaches which could be used to develop multi-modal transport models are reviewed including sketch-plan methods, system dynamics, micro-simulation and agent based modelling. The main advantage of the latter two methods is the detailed representation of individuals which complements the targeted and individualised nature of many ‘smarter choices’ measures.

The strengths and limitations of using an agent-based approach for modelling mode choice are investigated through the building of an agent based model for commuting trips using data collected from the 626 respondents to the 2010 Department for Transport funded stated preference survey for the ‘Climate Change and Transport Choices’ project. The agent based model is based on Triandis’ Theory of Interpersonal Behaviour, where both habits and intentions can influence a person’s observed behaviour. The incorporation of habitual behaviour into the model results in lagged responses to changes in transport costs as observed in the real world. Modelling at the level of the individual allows for more precise specification of the choices facing each person and any external or personal constraints on the modes available to them.

Finally the issues that would be encountered when applying the model to a particular area are considered, including the difficulties encountered in obtaining data on the attributes and constraints of each agent and their preferences. The use of latent class analysis is recommended as a method for grouping people together on the basis of their unobserved but shared preferences, rather than on directly observable characteristics such as their journey purpose or time of day of travel.

Word count: 73,400

Contents

Figures

Tables

1 Introduction

1.1 The use of transport models

'A model is a simplified representation of a more complex phenomena, process or system'. (Barnsley, 2007). Models are built for a variety of purposes and the actual process of building the model often leads to an increased understanding of the model's subject. Strategic transport models are frequently developed as a tool that can be used to estimate the state of the transport system at a future point in time if current trends continue, or how it may respond to proposed interventions. This enables policy makers to investigate the likely consequences of various policy options and provide evidence to assist in decisions that need to be taken, for instance over the allocation of the transport budget. Visualisation and the clear reporting of the output of model runs can be used to share the knowledge and understanding gained from creating and running the model with the wider audience of stakeholders involved in transport related decisions and those affected by the future state of the transport network.

Models of transport systems are often used to assist planners and policy makers in understanding the current and future performance of the transport system in an area and how this is likely to change in response to proposed changes in land use and the provision of transport infrastructure and services. Until recently the main changes to the transport system under consideration worked through the supply side, for example through the provision of new roads and bus services. Transport modelling techniques were developed to forecast the response of travellers to such changes. Now, however, attention is increasingly given to measures which directly influence the demand for transport which, in turn, brings the challenge of incorporating techniques capable of predicting the effects of these policies into models of transport systems. Various terms are used to describe these policies such as 'travel demand management', 'soft measures' and 'smarter choices'.

1.2 The emergence of ‘smarter choices’ in the UK

The publication of the Labour government’s white paper, ‘A New Deal for Transport’, in 1998 marked a new direction for transport policy in the UK (DfT, 1998). It stated that ‘people know we cannot build our way out of congestion with new roads’ and proposed improvements to public transport and ‘everyone doing their bit’ to ‘cut down on car use’. It encouraged organisations such as local authorities and employers ‘to produce their own green travel plans’ and for people to consider ‘their own travel habits’. The DfT co-funded pilot residential, school and workplace projects to encourage people to drive fewer miles through marketing campaigns, increased awareness of the public transport services in the area and the promotion of the benefits of walking and cycling.

The term ‘smarter choices’ was first used in a report by Cairns *et al.* (2004) which documented the potential reduction in car trips that these demand management measures could achieve. The types of interventions they considered to be ‘smarter’ measures covered:

• ‘Workplace and school travel plans;

• Personalised travel planning, travel awareness campaigns, and public transport information and marketing;

• Car clubs and car sharing schemes;

• Teleworking, teleconferencing and home shopping.’

These are all policy interventions aimed at directly affecting the demand for transport rather than influencing it indirectly through changing the quantity or quality of the supply of transport provision. The objective of a ‘smarter choices’ approach to transport planning is to reduce the amount of car travel by encouraging people to walk, cycle, use public transport or to travel less. A typical ‘smarter choices’ package includes a variety of ‘carrots’ and ‘sticks’ as well as information and marketing campaigns. For example, a workplace travel plan for an employment site on the edge of a town may include a minibus service from the nearest railway station to make public transport a more attractive option, together with the constraint of a restriction on the number of car parking spaces available on site and an incentive for car sharing by providing parking spaces for cars registered within such a scheme. Information events may be provided to inform staff of local bus services that run near the site and to promote the health benefits of walking and cycling.

The coalition government in the UK, shortly after its formation following the 2010 general election, announced its intention to continue the pursuit of policies that fit within the 'smarter choices' umbrella. The Department for Transport's vision (2011) was 'for a transport system that is an engine for economic growth but one that is also greener and safer and improves quality of life in our communities'. As part of their programme to encourage sustainable local travel they stated their intention to 'use insights from behavioural science to encourage lower carbon forms of travel' and their desire to make public transport, cycling and walking 'more attractive and effective'.

1.3 The need to include 'smarter choices' in transport models

The need to incorporate 'smarter choices' in transport models is a result of the noticeable difference that such measures can make on the level of demand in an area. The evaluation by Sloman *et al.* (2010) of the impact of the Smarter Choices Programme funded by the Department for Transport (DfT) in three towns in the UK, Darlington, Worcester and Peterborough (referred to as 'Sustainable Travel Towns'), reported that when a variety of smarter choices measures are implemented together the impacts can be quite substantial. Their evaluation work showed that 'in their four-year appraisal period, this produced a reduction of 5-7% in car driver distance travelled by residents for those journeys under 50km that were in-scope.' Given the magnitude of this change, the outputs of transport models that do not allow for the implementation of smarter choice measures, if they are to be undertaken in an area, could be misleading and lead to poor transport-related decisions being taken.

The scale of the reduction in car trips achieved in the Sustainable Travel Towns and the findings from other studies influenced the decision of the incoming UK Government in 2010 to allocate up to £560m towards funding similar programmes during the lifetime of that parliament through the Local Sustainable Transport Fund (LSTF). Since the initial announcement, the Government showed their continued support for these programmes through the allocation of additional funding. The successful bids in the first LSTF round all consisted of a mixture of infrastructure improvements, public transport changes, large scale personalised travel planning programmes and other smarter choices measures aimed at both the demand and supply side of the transport system. For example, the Plymouth bid (2011) included the building of a new cycle bridge over the river Plym, new cycle paths and improved lighting on existing cycle paths, complemented by a programme of home visits by travel advisors, cycle training schemes and marketing events.

Developers have also been seeking to use transport models to justify their statements as to the extent to which smarter choices measures would mitigate the impact of their developments on the surrounding road network by reducing the number of car trips generated. For example, the 2012 planning application for the Alconbury Weald development near Huntingdon for 5000 homes and employment space for 8,000 jobs, was supported by a detailed transport strategy that combined new infrastructure such as road junctions with a mix of land uses on-site to maximise the number of internal and therefore short distance trips and a set of smarter choices measures aimed at reducing car use. (Urban and Civic, 2012).

The challenge for the transport modelling is to find a method which can capture the impacts of such a transport strategy, which includes the provision of new infrastructure as well as a whole range of softer measures, on the forecasts of the impact of the development on the transport networks in the area. Ideally the method would be unified, so that the model can predict the impact of a change in any one element of the overall transport package on the overall forecast of the number and pattern of trips associated with the development. Such a model would be a valuable tool for use in the process of developing a cost-effective package of transport interventions acceptable to the planning authorities.

The motivation for this research comes from the need of the transport modelling community to extend the range of questions that can be answered with their current transport models, and in particular to meet the difficulties being experienced by practitioners when seeking to include the impact of smarter choices programmes into these models. An extension of current transport modelling practice is required to support policy makers who are deciding whether to invest in smarter choices programmes, and decision makers who need to understand the likely impact of such measures and their interaction with other proposed changes, such as the provision of new infrastructure and transport services, on the overall performance of the transport system in an area. This research therefore has a methodological focus and seeks to consider an approach to modelling that could bring 'smarter choices' into multi-modal transport models.

1.4 The standard framework for strategic transport models

The systematic analysis of the operation of the transport system in an area and the use of models to predict future conditions started in the United States of America in the 1950s (Bruce and Williams, 2003). A standard framework for these models was rapidly established, now known as the four stage transport model (McNally, 2007) and illustrated in Figure 1.1 overleaf. This remains the predominant modelling framework used in transport models today.

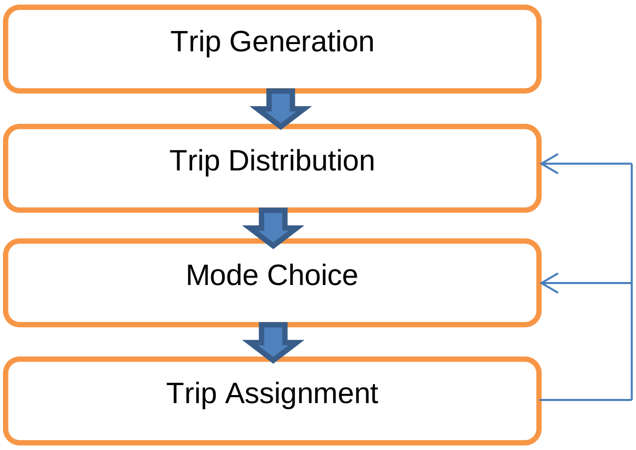


FIGURE 1.1 THE FOUR STAGE MODEL

In Stage 1 *trip generation*, the number of trips produced by or attracted to each zone is calculated as a function of the population and land use in that zone. In the second stage, *trip distribution*, the trips produced by each origin zone are allocated between possible destination zones on the basis of the travel time and / or cost to the possible destination zones and the land use in those destination zones. This gives the trip matrix, which is a two dimensional grid containing all the zones in the area on each axis. The number of trips in each cell represents the number of trips going from zone a to zone b. In Stage 3, *mode choice,* the trips in each cell in the matrix are allocated to each of the transport modes included in the model based on the relative time and cost of using each of the available modes. Finally in Stage 4, *trip assignment*, the trip matrix for each mode is assigned to the relevant transport network so predicting the number of vehicles using each road section and the number of passengers using each public transport service.

Many criticisms have been levelled at four stage transport models over the years. These include the high level of specialised technical knowledge needed to implement these models (Vigar, 2006), the lack of transparency in the processes involved (Evans *et al.*, 2007), the accuracy of the predictions made (Flyvberg, 2008) and the lack of behavioural realism in the assumptions and processes contained in the models (Givoni *et al.,* 2012, Sivakumar, 2007).

The development of this modelling framework was strongly influenced by the original purpose of the transport models being built in the 1950s, the constraints of the processing power and memory of the computers then available and the prevailing economic ideas of the main neo-classical school of economics taught in universities at that time.

The dominance of the need to evaluate highway schemes when the framework was developed resulted in design decisions that now affect the ability of the framework to assess 'smarter choices' measures, such as the use of zones as the basic building block of the model, the characteristics that are used to describe the links in the network and the use of the logit model for mode choice.

The four stage model 'was primarily designed for the analysis of urban highway investment.' (Hensher, 2007) in American cities in the 1950s which were facing an unprecedented growth in the ownership and use of private cars. Automobile production jumped from 70,000 vehicles per year in 1945 to 3.5 million in 1947 and highway travel was increasing by 6% per annum (Weiner, 1992). The Housing Act 1954 introduced the requirement for regional urban planning of which transport planning was a key component. In 1954 Mitchel and Rapkin developed trip generation models with trip rates based on the demographics in each zone. Alan Voorhees used a gravity model approach in 1955 to model the distribution of trips. Trips were divided between car and public transport using diversion curves (where the proportion of trips using public transport was a function of the public transport travel time). Shortest path routeing algorithms were developed in 1957 for highway trip assignment.

The Detroit and Chicago Area Transportation Studies in the 1950s brought these highway-oriented elements together into the full four stage modelling framework. The team in Chicago made use of computers to assist in their calculations but at that time the processing speed and memory of computers was considerably slower and smaller than today's machines. This led to the design decision to divide the study area into zones and to hold all trip information in matrices. The calculations during the modelling process are applied to the number of trips in each cell of a matrix.

The use of zones and matrices to hold data lies behind many of the difficulties encountered in incorporating smarter choices into the current generation of four stage models. Many walk and cycle trips are of a short distance and occur within zones and are therefore 'lost' or ignored by the models. The size of the zones results in aggregation bias and other errors, also encountered in Geographical Information Systems (Haining, 2003); for example the mean walk time from a zone to the nearest public transport stop is not representative of the actual walk time experienced by almost everyone in that zone. This leads to a discrepancy between the public transport journey times actually faced by the people in a zone which influences their mode choice decision and the mean value used in the modelling.

The software developed to run four stage models is often limited in the number of matrices it can handle during certain modelling processes. 'Smarter choices' often target particular groups of people, for example with fare discounts, and the impacts then get lost in the aggregate measures used in the modelling when the same weighted average fare is applied in the model to all people. The restrictions on the number of matrices that can be handled by the software limits the degree of segmentation of travellers that can be implemented. This reduces the ability of the model to capture the impact of highly-targeted policies and can lead to the exclusion of key influencing variables such as income, season ticket holding status and employment type. The design decision to use matrices has led to a modelling system that cannot easily accommodate the very detailed information needed to represent many of the features of 'smarter choices' measures.

The transport networks included in the original four stage models were designed primarily to represent the highway network and serve the zoning system. They often did not contain minor routes or walk and cycle-only links. Even when these are included today, the software is set up to handle those link and junction characteristics that influence highway capacity and route choice rather than attributes that are relevant for decisions affecting other modes such as, for cyclists and pedestrians, the degree of segregation from cars and the quality of the urban realm. If these aspects of the network links are not coded, the models are not sensitive to changes in these factors. Again, this means that the models do not capture interventions that are often major components of 'smarter choices' packages such as cycle lanes, lighting, signing and other improvements to the walking environment (Eash, 1999).

1.5 The influence of neo-classical economics

Paul Samuelson's influence on the people developing transport models in North America in the 1950s came through his work on utility theory (1938), his introduction of the use of mathematical optimisation models to explain economic behaviour (1947) and his widely adopted university textbook, Economics (1948), which popularised the neo-classical school of economics. The key working assumptions used in neo-classical economic analysis include rationality, perfect knowledge and the existence of a state of equilibrium in the market.

The logit model now used in stage 3 of the four stage transport model to forecast the mode used by travellers replaced the use of diversion curves in the 1970s. It was developed by McFadden (1974) who first used it to forecast the patronage for work trips on the proposed Bay Area Rapid Transit public transport system in San Francisco. The logit model approach is based on Samuelson's work on revealed preferences and utility theory (1938). Samuelson started from the assumption that individuals seek to maximise the utility they gain from the total bundle of goods they can purchase subject to their budget constraint. He argued that it was possible to derive the utility functions underlying a person's choices from observations of their revealed preferences. The logit model developed by McFadden then extended Samuelson's work to the mathematical modelling of the choices people make between discrete alternatives such as transport modes.

In the logit model as applied by McFadden to mode choice, the utility of each mode is calculated by combining together those attributes of a journey that can be quantified, usually the time and monetary cost elements of the journey by each mode. For a car journey this includes in-vehicle time, vehicle operating costs and any parking charges or tolls. For a public transport journey this could include in-vehicle time, walk time, waiting time, fares, boarding and interchange penalties. Revealed and stated preference surveys are often carried out to determine the weightings that people apply to these various factors, as for instance a minute spent walking has greater disutility than a minute spent in the bus. Values of time are used to convert monetary items into time equivalents.

The model is calibrated to match the observed choices of groups of respondents by incorporating a mode specific constant into the utility function which captures aspects of generalised costs that are not captured by the time and cost variables and a scaling parameter which controls the sensitivity of changes in the proportion of people using each mode as the relative utility of using each mode changes.

The logit model has become the engine of the four stage modelling process although the suitability of the logit model for this central role has been criticised. Beimborn (2006), looking at the application of the logit model in transport modelling, summarises these arguments as

- the logit model is only sensitive to changes in elements that are captured in the generalised cost functions and avoids items such as levels of comfort and reliability (unless they can somehow be quantified)
- the times used in generalised costs come from network models and are often simplified and inaccurate (acknowledged by McFadden (1984) and confirmed in Bhatta and Larsen (2011))
- people are not able to mentally combine time and cost elements to derive a single ‘value’ for each option they are facing

- it is not the case, although assumed in the application of the logit model, that each person applies the same weights to the different components of the generalised cost functions.

The logit model is based on the standard neo-classical economic assumptions, promulgated by Samuelson, that people have perfect knowledge of the options available to them, can calculate the total utility provided by each option and then consistently choose that option which maximises their utility. The realism of these assumptions was challenged in the 1950s by Herbert Simon (1955) who introduced the notion of 'bounded rationality', aiming to 'replace the global rationality of economic man with a kind of rational behaviour that is compatible with the access to information and the computational capabilities that are actually possessed by organisms, including man, in the kinds of environments in which such organisms exist'.

The work of behavioural economists, Ariely (2008), Lund (2008) and cognitive psychologists have continued this strong challenge to Samuelson's basic assumptions and his *'homo economicus'* and adds to the evidence that the logit model is based on an incomplete view of human decision making.

Behavioural theories, such as Ajzen's Theory of Planned Behaviour (1991) emphasise the role of other factors in the decision making process such as the influence of beliefs, attitudes, personal and social norms and intentions on the final choice observed in a person's behaviour. These theories suggest there is a need to include a wider range of objective and subjective factors in the modelling of the mode choice decision process.

1.6 Influence of habits and life events on mode choice decisions

The deliberative nature of travel mode decisions has itself been challenged by the view that, when considering trips made on a regular basis to the same place for the same reason, the use of a particular mode is a habitual action (Verplanken *et al.*, 1998, Aarts *et al.*, 1998, Goodwin, 1977). Habits are ‘strong associations between goals (e.g. going to the supermarket) and actions (e.g. using a car)’ (Aarts and Dijksterhuis, 2000). The association between the goal and the means used to achieve it occurs as a result of the consistency of the choice made, such as always choosing to use a car. After a while, the use of the car for such a trip becomes automatic and no deliberation over the choice of mode occurs; rather the behaviour becomes script-driven (Garling *et al.*, 2001).

The strength of the habit developed to use a particular mode increases the challenge to policy makers hoping to voluntarily achieve a change in the mode used by people (Gardner, 2009, Klockner *et al.,* 2004). For example, decreasing bus journey times may not result in an immediate transfer of trips from car if potential travellers are not actively considering their travel options at the time of the change. The enactment of habitual choices may be interrupted by external shocks (Adamowicz *et al.,* 2013) and can be forcefully curtailed, for example by temporary road closures, withdrawing parking spaces, (Gardner 2009, Brown, Werner and Kim, 2003; Fujii *et al.,* 2001) or by changes in the internal context of people’s lives, such as following a change in residential or work location for commuting trips. (Bamberg, 2006).

Studies into the influence of key life events on transport mode choice have highlighted the importance of changes in residential and workplace location in prompting changes in travel mode (Clark *et al.,* 2014). Van der Waerden *et al.,* (2003) suggest that key life events, such as passing the driving test and the acquisition of a car, can result in changes in the number of alternative modes available. Other changes may affect the characteristics of the modes available for a particular trip, for example a change of work location may affect the time and cost of the car journey to work. The main influential life events on travel mode choice for adults are moving house, changing work location, acquiring a car and becoming a parent (Chatterjee *et al.,* 2013, Prillwitz *et al.,* 2006; Klockner, 2004; Van der Wealden, 2003).

The interaction between these key life events and travel choices can be bi-directional, with considerations of travel options influencing the choice of residential location (Stanbridge *et al.,* 2004). The outcome is to provide a 'window of opportunity' (Franke, 2001) in which travel choice is once again a deliberative process and open to external influence. Smarter choices measures have been designed to recognise the opportunity of life events. Some measures directly affect the availability of modes or their characteristics, some aim to change attitudes towards modes other than car and yet others seek to trigger the re-evaluation of transport choices which have become habitual.

This study contributes to research on habitual behaviour by showing that the incorporation of habits in modal choice models can be achieved. The inclusion of habitual behaviour makes a significant difference to the predicted number of users of each transport mode in an area following the introduction of a change in the transport system. This work highlights the importance of relaxing the neo-classical economic assumptions of perfect knowledge and the constant re-assessment of utility maximising choices based on complete knowledge of the utility of each option. The replacement of these assumptions, made for the sake of simplicity and mathematical convenience, with more behaviourally realistic choice models results in a better representation of the gradual change observed in usage numbers of the various transport modes in an area following a change in the system.

This study also contributes to the understanding of the importance of feedbacks between the numerous elements of the transport system and its users. This is currently only acknowledged in the transport models used to assess transport policies as a feedback between journey times and costs, when trips are assigned to the network, and the utility of transport modes in the choice modelling component of a model. Transport models are customarily run with several iterations between the assignment modelling and the demand modelling components, until an equilibrium situation is reached where there are minimal changes between iterations in travel times and the number of trips predicted to use each mode. This study shows that it is possible to incorporate feedbacks between the demand for public transport and the level of supply into transport models.

This work also extends current understanding of the effect of aggregation bias on the results of transport models. Castilglione *et al.*, (2014) report that grouping people together and using an average cost change, for example an average parking charge increase of 10% for everyone in a zone, produces a different forecast change in the overall number of car users than if the actual change in costs for each person is modelled, i.e. a zero increase for those with free parking and a higher increase for those who pay to park. This study extends the treatment of aggregation bias by showing that it occurs, not only when average rather than individual input values are used with the logit models within transport models, but also when average rather than individual preference functions are used.

1.7 The four stage model and smarter choices

The UK Department for Transport issued a report in 2008 written by WSP (DfT, 2008) that considered possible methods for incorporating various 'smarter choices' measures that affected the journey to work into the standard four stage models used in the UK. A summary table of their findings regarding the modelling of workplace travel plans is presented in Appendix 1. They concluded that some measures could be modelled if:

- the zones were made small enough to pick up location specific impacts e.g. running buses to particular employment sites
- a sufficient number of matrices could be created to handle groups of people with different costs e.g. as a result of subsidised public transport fares or availability of free car parking.

However, the experience of practitioners is that the software packages currently in use become cumbersome if many matrices are needed, which also increases the possibility of user error when setting up the models. There is also a limit to the number of matrices that can be handled by some of the popular modelling software packages, such as SATURN, which is widely used in the UK because of the quality of its highway assignment algorithms. There are also practical and financial difficulties in obtaining enough data to calibrate the models for small groups of people.

WSP also reported that some measures cannot be directly incorporated into the current logit based mode choice models (e.g. offering personalised travel plans to staff, providing secure cycle facilities) and they suggested that these measures be reflected in the modelling by adjusting the mode-specific constant used in calculating the utility of each transport mode. They recognised, however, that there is no evidence on what level of adjustment should be made to this constant, which itself often varies between models as it is adjusted as part of the model calibration process.

The report also acknowledged that there are some ‘smarter choices’ measures that cannot be reflected in the four stage model, such as preferential car parking for car sharers, demand responsive bus services, the option of working at home and flexible working hours. This is due both to the inability of the four stage model to handle these measures and the lack of observed data about the likely impact of these measures.

The conclusion of the WSP report was that some measures can be incorporated into widely used four stage models although this requires much greater levels of segmentation than currently used. The impact of some measures could be approximated by adjusting the utility value of different modes but there still remain some ‘smarter choices’ measures that cannot be handled by the four stage modelling framework.

This study aims to investigate whether there have been advances in the four stage framework since the WSP report was written which now make it more suitable as a tool for modelling ‘smarter choices’. It reviews a new approach to transport modelling, known as activity based modelling, and then investigates modelling systems used in other disciplines such as systems dynamics and agent based modelling, which may be suitable for capturing the impacts of the wide variety of interventions that can be part of a ‘smarter choices’ package. The focus throughout this work is placed on commuting trips as many of these trips take place during the morning ‘peak’ period, which is generally the busiest time of day on the road network, and are the main target of many ‘smarter choices’ programmes.

1.8 Research questions

The primary question for this study is whether there are other modelling approaches, apart from the four stage framework required by the DfT in the UK by its WebTAG guidance, which would be better suited to tackle the task of modelling ‘smarter choices’ measures. This is addressed by answering three specific questions:

First, what has been the experience of transport modellers when using four stage models to include the impact of ‘smarter choices’ programmes on the mode used for commuting trips?

Second, what other modelling approaches could be used to model the impact of ‘smarter choices’ programmes on the mode chosen for commuting trips?

Third, what are the strengths and limitations of using an agent-based approach for modelling the impact of a ‘smarter choices’ programme on the mode chosen for commuting trips?

1.9 Thesis structure

The research process undertaken for this study is illustrated in Figure 1.1 below and is reflected in the organisation of this thesis. This first chapter has presented the background to the research. Chapter Two reports on a qualitative research exercise into the experience of UK transport modellers in incorporating smarter choices into the four stage transport modelling framework. This chapter addresses the first research question.

Chapter Three considers the work undertaken in the academic field that could assist in extending the existing four stage modelling framework to encompass smarter choices. It also reviews alternative modelling techniques which could be capable of modelling the impact of ‘smarter choices’ measures.

Chapter Four presents the philosophical perspective on modelling adopted in this study. It considers the purpose and nature of modelling and how the modelling process can assist in understanding the subject of the model. This chapter addresses the second research question and the research perspective informs the selection of the modelling technique used in the remainder of this study

Chapter Five describes the overall steps followed in the modelling exercise undertaken in this research. Chapter Six describes the structure of the agent based model built for this study and the data used in the model.

Chapter Seven presents the results from using the agent based mode choice model of commuting trips. The purpose of the modelling exercise was to test the potential of this approach for incorporating smarter choices into the modelling of the mode choice decision. The model results are presented so as to illustrate aspects of agent based modelling which are particularly relevant to the task of incorporating the impact of ‘smarter choices’ measures into multi-modal models.

Chapter Eight considers practical and theoretical issues that arise in applying agent based modelling to mode choice decisions and the impact of ‘smarter choices’ measures.

The final chapter (nine) is based on the experience of building an agent based model to forecast the mode choice for commuting trips. It addresses the third research question and considers the suitability of agent based simulation modelling as a tool for incorporating smarter choices into multi-modal transport models.

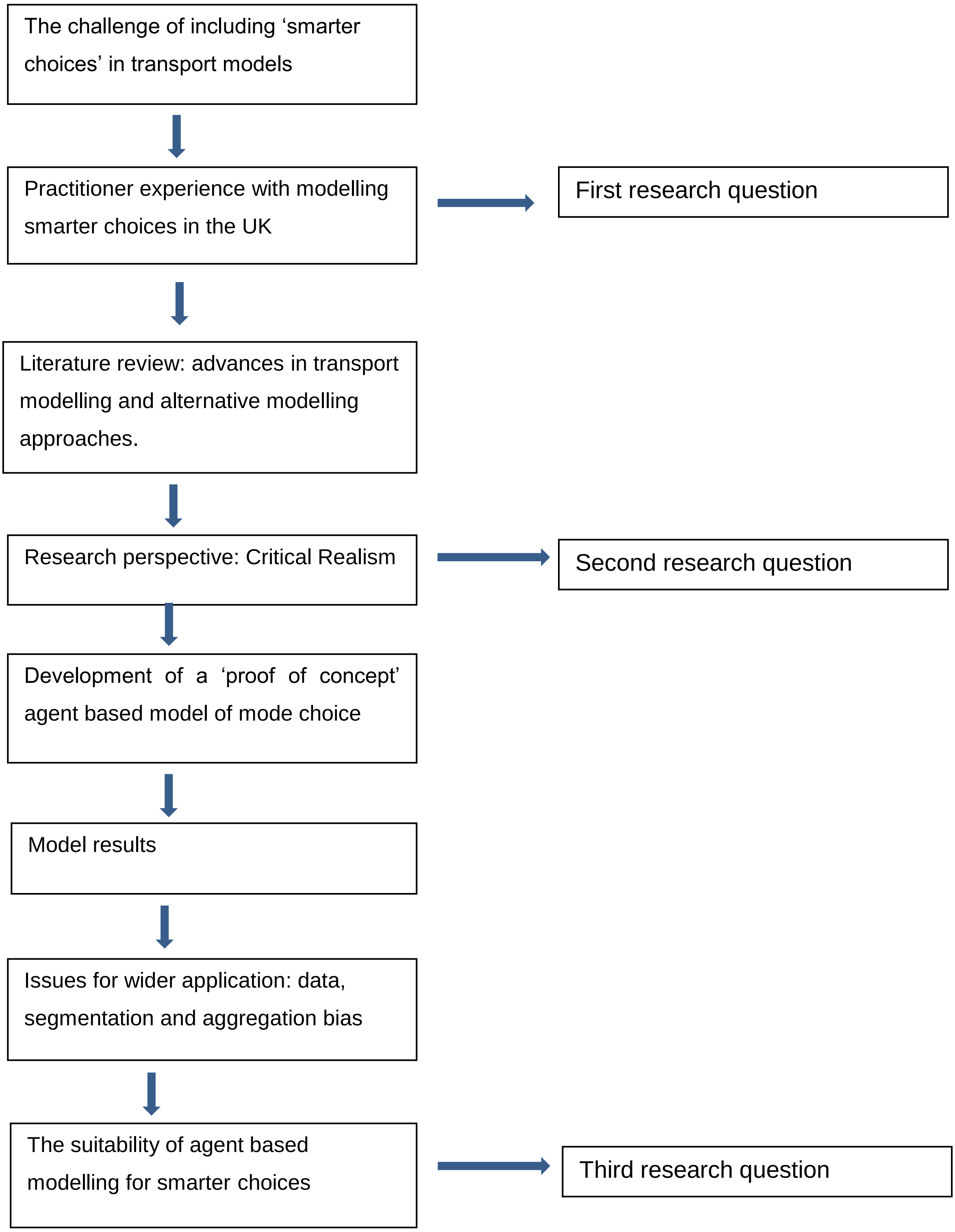


FIGURE 1.2 THESIS STRUCTURE

2 Practitioner experience with modelling smarter choices

2.1 Research methodology for practitioner study

The motivation for this research comes from the desire to incorporate smarter choices into transport models. The WSP report for the DfT (2008) identified ways in which four stage models could incorporate some ‘smarter choices’ measures by using more matrices to handle a greater level of segmentation of travellers, having a finer zoning system, modifying the generalised costs and adding additional detail into the network coding. The first research question for this study is to investigate the recent experience of transport modellers of including the impact of ‘smarter choices’ programmes in their four stage models.

Many of the most advanced transport models currently in use in the UK were built by local authorities when they were preparing bids to the Department for Transport’s Transport Innovation Fund (TIF). Substantial financial contributions were made by the Department for Transport (DfT) towards the cost of developing state-of-the-art transport models which were meant to be capable of testing the introduction of either a congestion charge or a workplace parking levy accompanied by complementary infrastructure improvements and ‘smarter choice’ measures to improve the attractiveness of public transport, walking and cycling. Ten local authorities or consortium of local authorities built transport models to support bids to the Transport Innovation Fund. These were Manchester, Birmingham, East Midlands, Bristol, Reading, Cambridge, Durham, Tyne and Wear, Shrewsbury and Norwich. They employed many of the most experienced transport modellers in the UK, which made this a suitable set of models to investigate for this study.

The research exercise undertaken for this study was designed to gain an understanding of current practice in the UK for modelling smarter choices. I had personal involvement in the TIF process as the author of the business case for the Reading bid. In this area the transport model was able to test the road pricing and bus rapid transit system elements of the proposed package of measures. As there were no forecasts available for the impact of the proposed personalised travel planning programme and investments in walking and cycling infrastructure, I had to draw on the limited evidence from elsewhere in the UK to assess the value for money of this element of the Reading bid. Through attendance at meetings held by the DfT I was aware that some of the TIF modelling teams had been considering how to incorporate 'softer' measures such as personalised travel planning into their models.

The aim of the first, exploratory, research exercise undertaken for this study was to investigate what methods practitioners in the UK had been using to incorporate 'smarter choices' into their multi-modal models and whether they regarded these as satisfactory. The research design was to undertake this task in two phases. First the model documentation would be reviewed, as the Local Model Validation Reports are intended to provide a detailed account of how the models are structured and validated. They should provide details of the way in which the practitioners had incorporated 'smarter choices' into their models. The second task was to conduct a semi-structured qualitative survey with each of the TIF modelling teams to investigate their views on whether they regarded the methods they had used for modelling 'smarter choices' as satisfactory and whether they felt further work on developing these or alternative methods was required.

The review of the available documentation for the multi-modal models built for the ten TIF congestion fund bidders showed that there was very limited documentation publicly available on the models that had used for modelling road pricing. As the preparation of the TIF bids progressed local politicians had become reluctant to introduce a road user charge in their area and the modelling work was either halted or diverted away from the specific requirements of the TIF projects. The models' documentation recorded their innovations in modelling variable demand matrices for car and public transport use but, with the exception of Bristol, did not discuss the modelling of 'smarter choices'. The Cambridge TIF bid explicitly said 'the modelling framework does not capture the mode shift attributable to a smarter choice strategy'.

The research design was to conduct a semi-structured interview in person with each of the modelling teams. The review of the model documentation informed the development of the survey questions used in the semi-structured interviews with TIF modellers. After securing permission from the University's ethics committee, each of the local authorities which owned the transport models produced by the TIF modelling teams, was approached with a request to interview their TIF modelling team. The purpose of the interview was to ensure that the modelling methodology adopted had been correctly understood from the reading of the documentation and to ask the respondents additional details about their approach to and experience of modelling smarter choices. The local authorities were contacted first with an email, outlining the purpose of the study and requesting an interview with their consultants. If they were willing for their modelling teams to participate in the study, these teams were then contacted by email. Many of the consultants preferred to be interviewed on the telephone or to reply by email rather than be interviewed in person. Where a personal or telephone interview was agreed a follow-up email was sent confirming the date of the interview, accompanied by a project information sheet, which detailed the purpose of the study. An advance copy of the questions was also provided. It was made clear that the consent of the participants was voluntary and that the participant could withdraw at any time, even retrospectively, until the project was completed and published.

Table 2.1 below summarises the response from each TIF bidder to the request to participate in the survey. In the West Midlands and Reading, the local authority officers responsible for the modelling work asked to be interviewed as well as their consultants, as they felt they had been closely involved in the technical modelling work supporting their bids.

Area	Response	Consultant	Preferred interview method
Greater Manchester	arranged interview but withdrew	MVA	None
West Midlands	personal interview	PRISM team	personal interview
Cambridgeshire	telephone interview	WSP and Atkins	telephone interview
Durham	telephone interview	Jacobs	telephone interview
Greater Bristol	supplied documents	Atkins	no interview
Tyne & Wear	no response	Jacobs	telephone interview
Reading	personal interview	PBA	personal interview
Shrewsbury	response by email	Mouchel Parkman	answers provided by email
Norfolk	response by email	Mott MacDonald	answers provided by email
Three Cities (Leicester)	no response	WSP	None

TABLE 2.1 CONTACT WITH TIF MODELLING TEAMS

2.2 Results from practitioner survey

The questions in the survey and the responses are presented here.

Did you have any smarter choices measures, including walking and cycling, in your future year packages?

The Transport Innovation Fund was politically contentious as a requirement of a bid to the fund was that the local authority had to introduce either a road user charge or a workplace parking levy. Both of these measures were unpopular with the general public and elected council members and in most cases the work was discontinued on bid modelling and preparation before the complete package of measures had been designed and tested. The bid packages most fully developed were those prepared by Manchester, Birmingham, Bristol, Cambridge and Reading. These all contained smarter choices measures and walk and cycle schemes. Work on the bids in Durham, Tyne and Wear, Shrewsbury and Norfolk did not reach the stage of defining the exact components of the proposed package of transport measures.

How were these incorporated into the modelling?

Durham, Tyne and Wear and Norfolk did not get as far as trying to model smarter measures. Norfolk used DIADEM, which is software produced by the DfT to conduct variable demand modelling for a highway scheme. It incorporates the facility to model a variety of responses to changes in the supply of transport including changes in the frequency with which trips are made, the destination of trips, the time of departure and the mode used. It uses a logit model for mode choice based on describing journey alternatives in terms of their mode, time and cost. The model iterates between the demand modelling and the highway assignment so that the time and cost of travelling by car changes in response to changes in the level of highway demand. The time and cost of using public transport remains fixed throughout the modelling process i.e. a change in the number of people using public transport does not affect the time or cost of making a public transport trip. The use of fixed times and costs for public transport and the exclusion of walk and cycling means that the software is not suitable for testing many 'smarter choices' measures.

Manchester, Birmingham, Tyne and Wear, Bristol, Cambridge, Reading and Shrewsbury used other software packages and did develop walk and cycle matrices but these were not fully used in the demand modelling in the same way as the highway and public transport matrices. The walk and cycle matrices were developed with the intention of being able to use them in the demand modelling but each modelling team reported that in the end they were not used as they each failed to establish a suitable methodology for modelling changes in the demand for walk and cycle. There was a remarked degree of reluctance to discuss why these efforts had failed; the general feeling was that the problem was too difficult to be tackled in the time available and that there was a lack of data against which to calibrate the models.

Birmingham mentioned the difficulties involved in constructing matrices of walk and cycle trips. As is usual practice, their modelling work was based on car park surveys, roadside interviews, public transport and household surveys. Walk and cycle trips were only covered in some additional surveys carried out later once modelling was already underway. Counting the number of pedestrians and cyclists and carrying our interviews with them proved to be very labour intensive and expensive. Given the limited budgets for modelling, even with the financial contribution from the DfT, only a few surveys and counts were carried out, which resulted in only a small dataset being available to the modelling teams.

There were concerns over using this data due to the small sample sizes. For example for vehicle counts the DfT advise using automatic traffic counters for at least 14 days at sites (DfT WebTAG Unit M1.2, 2014) in order to provide information on the daily variation in the number of vehicles. Some sites have permanent counters which provide additional information on the seasonal variation in vehicle numbers. As the pedestrian and cycle counts were conducted manually, which is expensive, they were only undertaken for a single day. As a result Birmingham council has since installed footfall cameras in the city centre, but these only supply count data, not origin and destination information, and only cover a very small proportion of all the walk and cycle trips carried out in the study area.

Birmingham and Reading both commented that the TIF models were designed to be strategic models and covered a wide area in order to capture the possibility of people changing their destination as a response to the introduction of road user charging, As a consequence of this, each zone was quite large, for example in Birmingham the city centre was covered by nine zones. This meant that many walk and cycle trips were intra-zonal and as such never appeared in the trip matrices or were assigned onto the network.

Birmingham, Bristol, Reading and Cambridge incorporated smarter choices into their modelling results by manually manipulating the vehicle matrices that came out of their transport model. They removed vehicles from the model output matrices in accordance with evidence from other studies as to the level of reduction in vehicles trips that could be expected if a smarter choices programme was implemented. They varied the percentage reduction in trips according to the distance of the trip, with a higher percentage reduction applied to shorter trips. The modified matrices were then reassigned to the highway network to provide results of the impact of these reductions on journey times and congestion levels.

No iterations with the transport model were then made, so there was no modelling of trips changing their mode or destination as a consequence of these changes.

Bristol applied different reductions in vehicle numbers for each of the components of their smarter choices package in turn; workplace travel plans, school travel plans, tele-conferencing, tele-working, car sharing, home shopping, car clubs, public transport marketing, travel awareness and personalised travel planning. Figure 2.1 below illustrates their handling of vehicle matrices, after the application of a logit model to predict mode and destination choices, in order to reflect the impact of smarter choices.

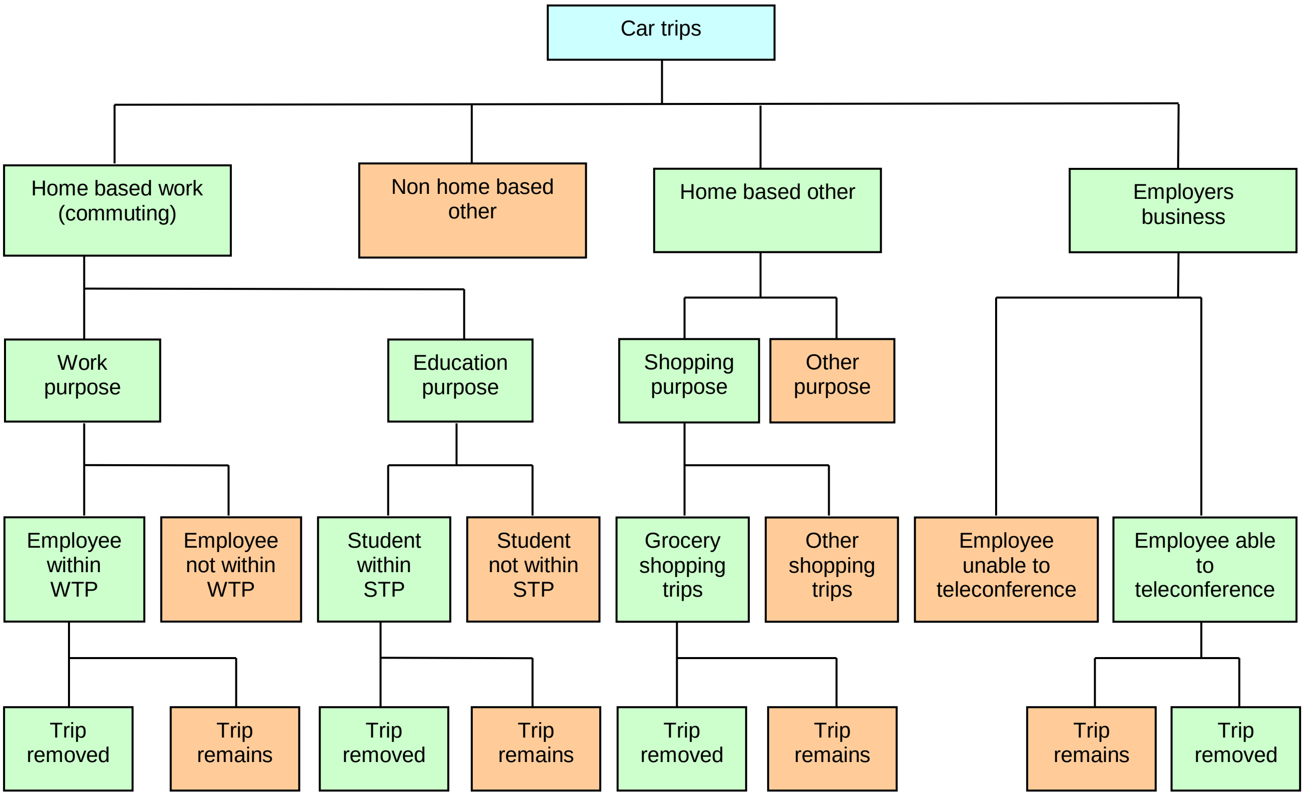


FIGURE 2.1 BRISTOL'S SUBDIVISION OF THE CAR TRIP MATRIX BY SOFT MEASURE IMPACT

Source: Greater Bristol Strategic Transport Study, Working Paper - Soft Measures' Unit Rates, 2005

WTP: workplace travel plan

STP: school travel plan

For workplace travel plans, for example, the modellers predicted the number of people who would be working at a place with a workplace travel plan, and reduced the number of commuter car trips to these destinations by 18%. All schools were assumed to have school travel plans which would reduce car travel by 12%. For tele-working, an assumption was made on the number of employees who could tele-work and how often they then worked at home. The number of car commute trips was then reduced accordingly.

A similar approach, that is deciding what percentage of people were 'in scope' for each 'smarter choices' measure and then reducing the relevant car trips based on evidence of achievements elsewhere was followed for the other 'smarter choice' measures with the exception of car clubs, public transport marketing and travel awareness where no evidence was found and so no reductions were made. Finally for personalised travel planning no specific reductions were made to avoid double counting with the other measures already considered.

How satisfied were you with this way of modelling smarter choices?'

None of the teams had modelled smarter choices using their preferred solution of building an integrated and fully multi-modal transport model that they could use to predict the changes in trip frequency, mode choice or destination that would result from measures such as road user pricing, changes in public transport provision, improvement to the walking and cycling infrastructure and 'smarter choices' measures.

The DfT had provided guidance to the modelling teams on methods to incorporate income segmentation and variable demand modelling in the TIF models but had provided no guidance on how to model smarter choices. The view was expressed by Birmingham, Reading and Cambridge that they were disappointed in not being able to work out a method themselves to incorporate smarter choices within their modelling framework. They had manually manipulated the matrices by mode output from their transport model and used these matrices in their appraisal work only because they could not devise a satisfactory alternative approach.

Did you conduct a separate appraisal of any of the smarter choices components of the TIF package?

Separate appraisals of their ‘smarter choices’ measures were carried out by Manchester, Birmingham, Bristol, Cambridge and Reading, They produced estimates of the number of people changing their behaviour as a result of these measures and then valued these impacts. The type of impacts which were valued and the valuations used varied between the bidders.

Did you consider alternative ways of modelling walking and cycling to the approach you adopted? Why were these not implemented?

The teams were generally unwilling to discuss the ideas they had considered for modelling ‘smarter choices’. Reading had considered changing the lambda value in the logit model they used for mode choice modelling so as to increase its sensitivity to changes in time and cost between using public transport and car. The intention was to increase the public transport mode share in response to the same level of reduction in the cost of travel by public transport compared to car as a way to reflect ‘smarter choices’ resulting in a greater decrease in car use than would otherwise be expected, but this was approach was rejected because of the lack of calibration data. The DfT also rejected this method on theoretical grounds as although it gives a greater proportion of people using public transport if public transport costs go down it also increases the number of people predicted to use car if highway times decrease, and a reduction in car journey times is an intended consequence of implementing road user charging.

The TIF modelling teams concentrated their initial efforts on developing modelling frameworks that encompassed variable demand modelling and the segmentation of matrices by income. When improvements to walking and cycling infrastructure and ‘smarter choices’ measures were added to the packages of proposed transport interventions, the modellers faced the challenge of adapting their models to face a task which was not envisaged when the models had been initially designed. Different amounts of effort were expended on dealing with this issue by the teams, including commissioning other consultants to work on the task, but none of the TIF bidders could devise a suitable way of incorporating 'smarter choices' into their models.

A major difficulty mentioned by the modelling teams was the production of the complete current year walk and cycle matrices needed for model calibration from very limited survey and count data. Birmingham raised the issue that the structure of the software used to implement four stage transport models is designed to operate over the whole of a matrix at the same time. Dealing with separate areas is possible but convoluted and requires the use of 'masking' matrices or code to apply actions to particular origin – destination zone combinations. 'Smarter choices' measures are often aimed at specific corridors and so the impacts need to be modelled only on trips that use this corridor or have origins and destinations in specific areas. This spatial detail makes the treatment of such policies extremely complicated using existing matrix based software.

The three teams that continued as far as producing a bid to the DfT for TIF funding, Reading, Bristol and Cambridge, all relied on manually manipulating the matrices that came out of their transport model to reflect the changes they expected to occur given the limited evidence available from experience elsewhere of implementing these measures. As the adjustment to the matrices was made outside of the automated modelling process there was no feedback within the model between the impacts of the smarter choices component of the TIF packages and the other policy measures tested.

2.3 Reflections on the practitioner experience

The TIF teams were already committed to a particular modelling framework before they started to consider how to model smarter choices. They had started work on their models before they were asked by the DfT to include ‘smarter choices’ in the package of complementary measures to be introduced at the same time as road user charging. This extended the type of interventions that needed to be tested in the model beyond those which were usually tested, such as changes to the supply of public transport and fares policy. The TIF teams may have been more successful in incorporating smarter choices into their models if they had been aware of this requirement at the start of the process so that it could have been considered during the initial model design and development of the data collection strategy.

The TIF modelling teams were required to build a model that was compliant with DfT guidance as contained in WebTAG. This forced them to use a particular modelling framework, the standard four-stage modelling approach, as this is the mandatory structure for a DfT WebTAG compliant variable demand transport model. This may however not be the most appropriate modelling framework to use for modelling smarter choices.

The TIF teams had already started the data collection for their modelling before they considered how to extend the modelling to incorporate smarter choices. This again imposed a restriction as they may have collected different or additional data if they knew from the start that they would be modelling smarter choices. For example, the consideration of ‘smarter choices’ places a greater emphasis on walking and cycling modes and so more attention may have been paid to the collection of data about the current number of walk and cycle trips and people’s preference for walking and cycling rather than using other modes.

The segmentation used in the model had already been decided on the basis of needing to model responses to the introduction of road user charging. The trips in the area were grouped into matrices based on:

- Car availability
- Time of day
- Journey purpose e.g. commuting, education, employer’s business, leisure

- Income

The segmentation by income was novel in the UK at the time and was introduced because it was felt that people might respond differently towards being charged to use certain roads or to park at work depending upon the relative size of the charge compared to their income. Once this segmentation was established changing it to also divide people up into groups that might be appropriate for modelling smarter choices, such as by attitude towards safety (which affects preference towards cycling) or by employment in a job suitable for tele-working was not feasible. This was because the data had not been collected with such segmentation in mind, the matrices had already been created and there was not time in the work programme to go back and re-do this work. Adding additional segments would also increase already long run times and the DfT's TUBA software used for the economic appraisal work was already struggling with the memory requirements of the increased number of matrices as a result of the income segmentation.

As the TIF teams were faced with adapting their current models to incorporate smarter choices, they moved directly to considering solutions that were technically possible with minimal change to their current modelling framework, such as extending their mode choice model to include the option of walking and cycling as well as public transport and cars. There was little or no thought as to **how** smarter choices affected the travel decisions and whether those mechanisms were present in the model. For example, personalised travel plans place a great emphasis on providing information on travel choices to people but the transport model assumes that people have perfect knowledge already of the options available to them and their time and cost.

It is also possible that people are making rational travel choices but that these are based on other considerations as well as time and cost, such as the health benefits from active travel and the greater certainty over travel time if the trip is made on foot. There was no mechanism by which these factors could influence the mode choice in the modelling because the original choice of variables for inclusion in the TIF models, which omitted these factors, was made before the need to include ‘smarter choices’ arose.

2.4 First research question

The first research question for this study is ‘what has been the experience of transport modellers when using four stage models to include the impact of ‘smarter choices’ programmes on the mode chosen for commuting trips?’

It was clear from the practitioner interviews that including ‘smarter choices’ programmes did not prove to be as straight forward as they anticipated before attempting it. All the modelling teams faced issues with collecting sufficient amounts of observed data on the current level of walking and cycling in their area to use when building base year matrices and calibrating their logit based mode choice models.

The TIF modellers also found that some of the interventions they planned such as improving safety could not be dealt with directly in the model except by adjusting the mode specific constant for cycling, but there was little data available to indicate by how much the mode specific constant should be adjusted. In the end, rather than adjusting the model parameters to get the changes they expected, the modelling teams chose to adjust the output matrices. This was a more transparent exercise but an admission that the models could not be extended to cover the proposed ‘smarter choices’ measures.

A key lesson learnt from the experience of the TIF modelling teams is that the requirement to include responses to ‘smarter choices’ measures should be part of the initial design specification of the modelling tool. The model developed should be able to handle a greater level of segmentation than is currently accommodated so that more differentiation is possible on the time and cost of travel by different modes for different groups of travellers. It should also use a fine zoning system in order to produce more precise time and cost information on trips and to reduce the loss of short distance trips to the intra-zonal category, as these are often walk and cycle trips.

In conclusion, the experience of practitioners reinforced the WSP report (2008). There is a lack of evidence on the impact of the various components of ‘smarter choices’ packages which hinders model calibration. Although some features of ‘smarter choices’ packages could be incorporated in a four stage transport model by, for example, increasing the segmentation of trips, using finer zoning systems and more detailed networks, difficulties were encountered in implementing these changes due to software constraints, increased run times and lack of data.

The practitioners could not find a satisfactory method for incorporating many ‘smarter choices’ initiatives within their modelling framework. Measures such as personalised travel planning and marketing campaigns do not affect the variables, time and cost, which determine mode choice in the current modelling approach but have been observed to affect the number of people using each mode in the real world. These issues became the subject of a literature review which considered the work of other modellers in extending the current four stage modelling framework and then examined alternative modelling methodologies that may be better suited to the task of modelling ‘smarter choices’ within a multi-modal transport model.

3 Literature review

3.1 Introduction

This literature review first considers recent work to enhance the four stage modelling framework. It then reviews alternative approaches to modelling that could be used to study the transport system and offer the potential of incorporating ‘smarter choices’ measures. The review considers sketch plan methods, system dynamics modelling and the individual based modelling approaches of microsimulation and agent based modelling.

3.2 Enhancements to the four stage model

3.2.1 DfT review of the four stage model

The recommendations of the WSP (2008) report for the DfT into ways of incorporating ‘smarter choices’ such as workplace travel plans, school travel plans and targeted marketing initiatives, into current four stage models were to:

- identify those aspects of a smarter choices package that affect the time and cost of travel and apply those to the people affected by using a finer zoning system
- increase the segmentation of travellers so as to be able to better match people with the costs they actually face, and
- to use a more accurate representation of those costs e.g. through parking charges and bus fare concessions.

3.2.2 Finer zoning systems

A more accurate representation of travel time can be gained by using a finer zoning system. The zoning system used in a model is a way of aggregating trips together spatially; combining trips in the same locality and giving them a common, representative travel time by loading them onto the network for trip assignment at the same point. Trips that start and finish within the same zone are never loaded onto the network and do not contribute towards flows on the network or travel times between zones. The smaller the zones, the more representative the travel time is for all people in that zone and a higher percentage of trips are loaded onto the network but the time taken to run the model increases.

Originally four stage models were used for modelling vehicle trips and as cars are seldom used for very short distance trips, the use of zones was a practical solution for reducing model run times while retaining virtually all trips in the model. When the need to consider public transport was added to the modelling requirement, extra care was needed when designing the zoning system to ensure that the travel times for a zone were truly representative. As the first part of a public transport trip is walk, zones tended to become smaller so as to distinguish between people with a short or a long walk to a bus stop or station.

Eash (1999) investigated the impact of the distribution of housing in a zone on the difference between the actual and modelled average access time from a zone to a bus stop on the eastern edge of the zone. He noted that with an uneven spatial distribution of housing within a zone the distance between the geographical centre of the zone and the bus stop was different than the average distance between the houses and the bus stop. He recommended the use of an ‘access mode’ centroid rather than a geographic centroid in order to gain a truer estimate of the actual access time to public transport from a zone. He also recommended designing the zoning system so that public transport stops were at the centre of a zone. However this may conflict with the ideal design of zones and location of centroids needed to produce an accurate representation of typical highway times.

When models are further required to incorporate walking and cycling modes, there is a need to have a yet finer zoning system in order to reduce the percentage of intra-zonal trips which are not assigned onto the network. A possible solution is to have a transport model that operates with several zoning systems and contains procedures to seamlessly convert matrices between the various systems during the modelling process. This is easier to achieve for trip matrices when the zones are designed with this in mind so that the process becomes one of simple aggregation and disaggregation. Great care is needed when using time and distance matrix skims prepared from a fine zoning system for use in other routines in the modelling system, such as demand modelling, which usually operates with a coarser zoning system to avoid the difficulties arising from calibrating mode choice models to matrices which have many cells with very few or zero trips.

3.2.3 Customised networks

When public transport modes are introduced into four stage models the coding of the network is extended to include public transport services. Key bus stops and railway stations are selected and coded into the model and services described either using actual timetables, which means that accurate interchange times are used, or the frequencies of services. In this case approximate waiting and interchange times are used based on applying a factor to the headway time between services. This enables the models to calculate public transport times between zones and assignment algorithms were used to allocate trips to routes based on time, cost and crowding conditions.

The extension of these models to include walking and cycling, which are important modes in many smarter choices packages, requires that the coding of the networks is extended to include walk and cycle only routes. The starting point for a walk and cycle network is often the highway network but additional links are required for walk / cycle only links and one-way links for cars need to be coded as two-way for pedestrians to ensure that the journey routes and times are accurate. There is an implicit assumption that people can cross the road where-ever they want.

Assigning pedestrians and cyclists to the shortest route does not always match their actual route choice and the trip assignment may benefit from the coding of additional information, such as the degree of segregation from traffic, the quality of the urban realm and the amount of street lighting, as these factors can influence route choice.

Castiglione (2014) notes that 'a number of studies have considered the factors that affect bicycle choices, but very little of this research has been incorporated into traditional four-step or activity-based travel demand model systems'. He reports on work in San Diego where the complete cycling networks were coded, including all cycle tracks. Additional variables which affected cycle route choice such as elevation changes along a link were coded and used in the trip assignment stage. Desyllas *et al.* (2003) built a strategic pedestrian model for London which had a pedestrian route assignment component. The walk network included a variable indicating the pedestrian capacity of each link which can affect the attractiveness of certain walkways for pedestrians.

Using these techniques it is possible for four stage models to incorporate the impact of some smarter choices measures, such as improvements to public realm and lighting. It is resource intensive to maintain an inventory of the network at this level of detail and to code the information into a network model (Connectics Transportation Group, 2008). Further work is also required to determine which network characteristics are the most influential on route choice and should be included in the network coding and then to calibrate the route choice algorithms that use these variables. There is also a need to develop methods of incorporating this information into the mode choice modelling so that improvements, such as better lighting, can influence the overall number of walk trips as well as the routes that such trips take.

3.2.4 Increased segmentation

The four stage model stores and handles trips in matrices which represent the number of trips going between each pair of zones in the model area. The trips in each cell of the matrix share a common travel time and cost, derived by skimming the routes chosen in the network assignment between the relevant zone pairs for that cell. All travellers within each matrix are assumed in the choice modelling to have identical weightings or preferences for time, cost and any other variables included in the modelling. The four stage model can segment travellers into different groups by creating separate matrices for the travellers in these groups.

The benefits of increased segmentation are that the matrix holds a more accurate representation of the actual times and costs for that group of travellers. For example, travellers may be segmented by whether they qualify for a concessionary bus pass which affects their cost of travel by bus or whether they have to pay for a parking space at their destination. A more accurate representation of travel times can be achieved by segmenting travellers by the time of day when they travel, which allows, for example, for longer travel times in peak rather than off-peak periods. The TIF modellers used segmentation by income as a method of segmenting travellers by their value of time, or willingness to pay for time savings achieved by road user pricing. Segmentation by journey purpose is also common practice. It assumes that people can be grouped together with common values of time according to the purpose of their trip with those travelling on employers business having a higher value of time than commuters or those travelling for leisure purposes.

The WSP report (2008) recommended greater segmentation as a method of incorporating some ‘smarter choices’ measures into four stage models. It could be used to separate out those travellers receiving free travel to a workplace for example, or those facing high parking charges. Simmonds *et al.* (2001) distinguish two separate tasks associated with segmentation; first the distribution of the segments, that is to determine how many trips there are in each cell of the matrix and secondly to determine the demand response for each segment. Even with moderate levels of complexity reflected in the segmentation through cross-classification, for example, time of day by income by availability of free workplace bus service, there is ‘the possibility of ending up with more categories than travellers. The question then arises as to whether it is either efficient or plausible to represent all the feasible combinations of traveller characteristics if the result is that each cell of the matrix contains only minute fractions of travellers’. The report suggests that microsimulation may be a more suitable tool for implementing models with a high degree of segmentation. Microsimulation is considered further in Section 3.6.

3.2.5 Variable demand modelling

A key finding of the DfT's Sustainable Travel Towns study (Sloman *et al.,* 2010) was that the impact of smarter choices was not just on mode choice but also on trip distribution; a car trip to a destination some distance away may become a trip to somewhere nearer and then may switch mode as well from car to walk or cycle. The concept of changing both mode and destination in response to changes in travel costs was introduced into four stage models in the UK by the DfT in 2006 as a response to the SACTRA committee work on induced traffic. The WebTAG Units 3.10.1 to 3.10.4 (2006) explained that variable demand modelling would be required when a proposed transport intervention would produce a significant change in time of travel that could in itself influence the number of trips people make (quicker travel times leading a higher frequency of travel), and their destination (faster average journey speeds or lower costs leading to people travelling to other destinations further away) as well as mode switch (as journeys by one mode become relatively quicker some people switch to that mode).

The modelling community developed procedures to implement variable demand modelling for car and public transport trips but were hindered by the lack of evidence on the strength of the responses and factors affecting them and their efforts have concentrated on the impact of variable demand on the number of highway and public transport trips. There is as yet little work on how changes on the supply side affect the number and destination of walk and cycle trips. The principle that these measures could affect more than just mode choice is accepted but the challenge of representing this mathematically in the model is still unmet, as illustrated by the experience of the TIF modellers.

The current approach in the UK for variable demand modelling is to use an elasticity approach for trip frequency. A logit model is used for destination choice, mode choice and time of day choice modelling. (DfT WebTAG Unit M2, 2014). This highlights the critical importance of the logit model in the modelling of the travel decisions that 'smarter choices' measures are intended to influence. The next section therefore examines recent work on enhancements to the logit model that may assist in the modelling of 'smarter choices'.

3.3 Enhancements to the logit model

3.3.1 The logit model

Discrete choice modelling has been widely used as the mechanism for determining the choices made by travellers in transport models. A logit model is a discrete choice model that predicts the probability that a person will choose a particular alternative when faced with a choice between a set of options, such as which mode of travel to use for a particular journey. (Gordon, 2009). When these probabilities are applied to a number of people this then gives the mode share. The binary mode choice is given by

$$P_i = \frac{\exp(U_i)}{(\exp U_i) + \exp(U_j)}$$

where:

P_i is the probability of choosing alternative i

U_i is the utility of alternative i

U_j is the utility of alternative j

The utility of each mode is calculated by combining together attributes of a journey that can be quantified, usually the time and monetary cost elements of the journey by each mode.

$$U_i = \lambda\, c_i$$

where

U_i is the utility of alternative i

c_i is the generalised cost of alternative i

λ is a negative scaling parameter

For a car journey this would include in-vehicle time, vehicle operating costs and any parking charges or tolls. For a public transport journey this could include in-vehicle time, walk time, waiting time, fares, boarding and interchange penalties. Stated preference surveys are often carried out to determine the weightings that people apply to these various factors, as for instance a minute spent walking has greater disutility than a minute spent in the bus. Values of time are used to convert monetary items into time equivalents. The utility function for each mode often includes a mode specific constant which captures aspects of generalised costs that are not otherwise captured in the utility function. Without it, if the generalised cost of travel by car and public transport is identical then 50% of travellers would use car and 50% would use public transport.

The scaling parameter is vital as it dictates the slope of the curve in the logit model and the sensitivity of mode share to the difference in the generalised cost of travel between each mode. This is illustrated in Figure 3.1 below. The calibration of a logit model involves adjusting the weightings applied to construct the generalised costs, adjusting the mode specific constants and adjusting the value of the scaling parameter (also known as lambda) so that the logit model replicates the observed mode shares in an area and changes in mode share as one of the variables in the generalised cost functions such as public transport fares are changed (DfT, 2009).

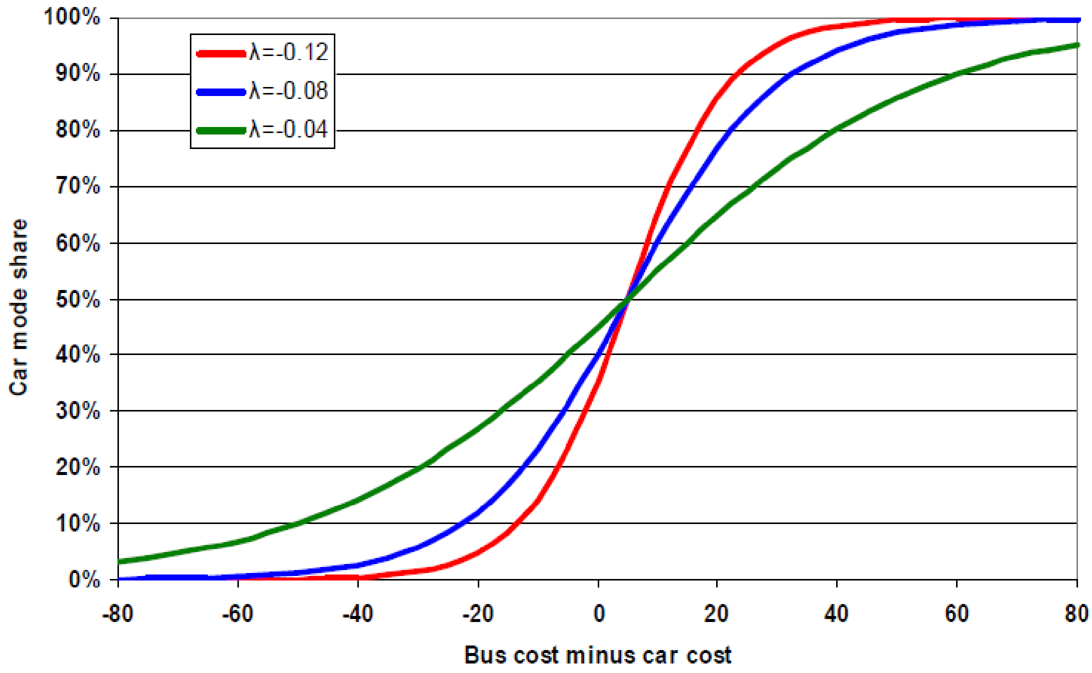


FIGURE 3.1 SENSITIVITY OF MODE SHARE TO VALUE OF LAMBDA IN A BINARY LOGIT MODEL

Source: Diadem manual, DfT 2011.

Originally in 1973, McFadden proposed the logit model as a behavioural model based on the assumption that people seek to maximise their utility. By 2002 he was a co-author of a paper which presented random utility logit models only as **predictive** choice models which 'emphasise the regularities of choice behaviour in quantitative models that can be used for prediction' and acknowledged that they only captured some of the issues involved in how people made decisions.

Recent development work on the logit model has considered two aspects of particular relevance for the modelling of 'smarter choices'. The first is to segment people into groups not based on external observable characteristics such as journey purpose, but on unobservable characteristics, such as attitudes of the travellers towards using public transport. The intention of segmentation by attitudes is to produce segments which have a more accurate response to changes in the transport system by the individuals within each segment. This work is described in section 3.3.2 below on latent classes.

The other development has been on the inclusion of unobserved but important variables that affect the mode choice decision in the model. This work is described in section 3.3.3 below on latent variables.

3.3.2 Latent classes

Walker and Ben-Akiva (2002) extended the logit model by introducing segmentation based on unobserved attitudes or preferences, with travellers divided into segments knows as latent classes. Latent Class Analysis (LCA) is described by McCutcheon (1987) as a methodology that 'is well suited for the analysis of typologies'.

The premise behind latent class analysis is that a group of people can be divided into distinct sub-groups based on their attitudes or behaviours with the members of each sub-group sharing the same attitudes or behaviours. The approach is based on the concept that there is an underlying unobserved or latent variable which accounts for the difference between groups and this latent variable can be used to allocate people into groups. The observed choices made by people within each group based on their preferences are similar, but those for each group are distinct from the other groups. The differences between the groups are due to the unobserved latent variable.

The result of a latent class analysis is a distinct set of clusters known as classes based on unobserved or latent characteristics of the respondents. LCA also provides for each respondent the probability of their belonging to each of the classes. They are then allocated to the class for which they have the highest probability of membership. When setting up the latent class analysis key variables such as time and cost preferences are selected as primary drivers for the segmentation. Secondary drivers can also be designated, such as demographic data and these can be treated as either active, in which case they influence the clustering, or inactive, in which case they are used only in the profiling of individuals within each class.

The latent class solution is not dissimilar to the cluster analysis approach often implemented using the K-means technique, but LCA has some distinct advantages over cluster analysis. It can readily use data of different types, so it is possible to determine classes on the basis of a person's preference weighting for time, cost, etc. and responses to other parts of a survey such as income, car ownership, location and even attitudinal responses to Likert scale type questions. It also produces output that gives the probability of a respondent belonging to a particular grouping.

Atasoy *et al.* (2011) used latent classes in a Swiss study of mode choice. A qualitative survey of 20 people was used to investigate people's mobility habits, residential choices and opinions on different transport modes. This informed the design of a revealed preference survey, completed by 1763 respondents, which contained a one-day travel diary and a set of five-point Likert scale questions on respondents' opinions on topics related to environment, mobility, residential choice, lifestyle and their perceptions of different transport modes. Data was also collected on their household composition and socio-economic characteristics.

Latent class analysis was carried out to discover what natural groupings there were amongst the respondents in terms of their preferences. A logit model was calibrated for mode choice for each of these groups. The work revealed two latent classes or groups of people who shared similar preferences. The first group consisted of individuals who were mostly middle aged, living with family and children, highly educated and with high incomes. The second class were young individuals, mostly students and old people who were mostly retired. The latter group were more sensitive to cost than the former.

The performance of the choice models was tested by calibrating the latent class logit model using 80% of the observations selected at random and comparing the predicted mode with the reported mode choice for the remaining 20% of the observations. The logit model with the two latent classes had better prediction power compared to a single class logit model, with choice probabilities higher than 0.5 for 75.00% of the withheld data when using latent classes compared to 72.87% for the base model. The percentage of choice probabilities higher than 0.9 was 27.93% with latent classes compared to 25.80% without.

The use of latent classes offers the possibility of improving the overall performance of mode choice modelling as it provides a way of using the data to decide on the relevant segments and allocation of individuals to those segments rather than the groupings being pre-determined before the data is collected and the models calibrated. For example, in the application of the logit model monetary costs are converted into a time equivalent by dividing the cost by the values of time. Applying logit models with a different value of time to market segments shown to have different values of time rather than a single model with one value of time applied to everyone, would be expected to improve the forecasts of those aspects of a smarter choices programme which sought to influence mode choice by affecting the cost of travel, for example, by providing subsidised bus services.

3.3.3 Latent variables

Latent variables are unobserved variables. With latent variables, attributes and values which cannot be directly quantified can be brought into the utility function. The use of latent variables has been developed to address the omission of factors which influence choice but cannot be directly observed, measured and incorporated in the utility function. These latent variables can be used to include unobservable psychological factors such as the desire for safety, convenience and attitudes towards environmental considerations. (Ben-Akiva et al., 2002, McFadden, 1986).

Work by Temme (2007) and Yanez *et al.* (2010) has shown that it is feasible using available software to use latent variables to introduce perceptions and attitudes into logit models and that this improves the ability of these models to replicate observed choices and improves the forecasts produced by logit models. Temme (2007) used latent variables in a logit model to look at the impacts of a change in the terms and conditions of a German railpass which reduced the cost of travel slightly but decreased the flexibility of its use. A traditional logit model forecast an increase in rail travel as a result of the decrease in costs, but the inclusion of a latent variable that captured the value of flexibility correctly predicted the actual decline in rail travel that was observed as rail users rejected the new pass because it imposed too onerous a set of travel time restraints.

Yanez *et al.* (2010) used panel surveys in Santiago, Chile, to record mode choice before and after a major change in the public transport network. Three latent variables were used in the modelling; reliability, comfort/safety and accessibility. The value that a person placed on each of these attributes is deduced from knowledge of the individual's income, education, gender and age as shown in Figure 3.2 below. The relationship between these observable attributes and the value the individual would accord to the latent variables was estimated using a Multiple Indicator Multiple Cause model. The latent variables, with their deduced values, were then used in the logit model to extend the generalised cost functions to contain reliability, comfort/safety and accessibility.

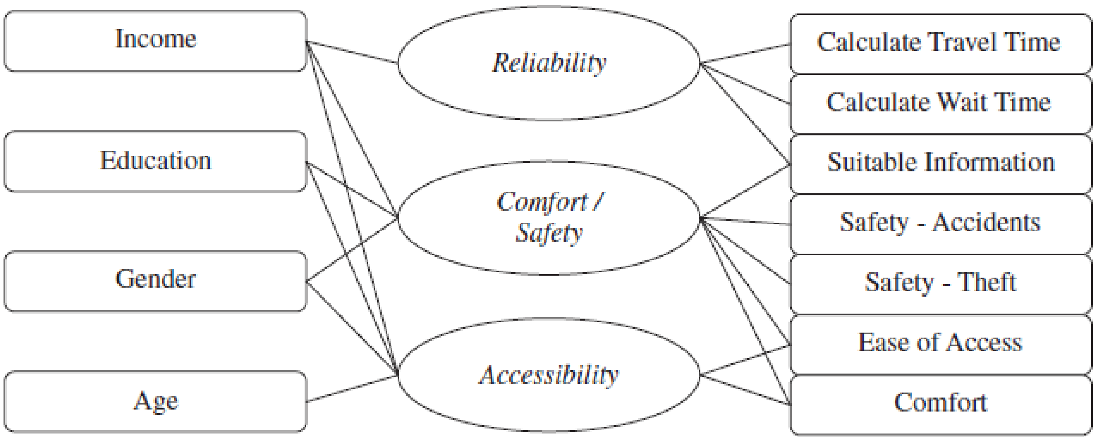


FIGURE 3.2 LATENT VARIABLES USED IN SANTIAGO

Source: Yanez *et al.*, 2010

Yanez *et al.* (2010) reported that this hybrid choice model produced better predictions of the observed changes in mode share than standard logit models. They conclude that 'hybrid models are clearly superior in fit to traditional discrete choice models that do not incorporate latent variables'.

The next section takes up the suggestion of Simmonds *et al.* (2001) that alternative modelling approaches may offer potential as ways of developing transport models with the greater degree of segmentation required for the modelling of 'smarter choices'. It reviews five modelling approaches; sketch plan methods, system dynamics, microsimulation, activity-based modelling and agent based modelling and considers their potential contributions towards the modelling of 'smarter choice' measures within an integrated multi-modal transport model.

3.4 Sketch plan methods

The popularity of sketch plan methods is due in part to the length of time it takes to build a four stage model for an area and even to undertake model runs from a model that is already built. Sketch plan methods aim to produce a similar result more quickly and offer the opportunity to make adjustments to the numbers based on professional judgement or factors not included in the full four stage model. They are often implemented in spreadsheets or GIS systems.

Marshall and Grady (2006) used a sketch plan model to model the impact of land use on public transport patronage. The reasons provided for not using the full four stage model for the region were concerns over the model's accuracy, its insensitivity to land use changes and institutional and cost barriers to using the full model. The accuracy concerns related to the validation of the four stage regional model concentrating on a correct representation of highway trips rather than public transport trips, and an insufficient level of detail in the validation, which considered only overall boarding numbers rather than validating individual stops and lines. The trip generation model in the full model was not very sensitive to land use changes in a zone. The institutional barriers they encountered included the high degree of training required of staff before they could use the model and the Washington

Metropolitan Planning Organisation, which owns the model, wishing to concentrate resources on modelling other issues such as air quality conformity.

A sketch plan model was developed to include additional relevant variables such as household density, employment density and the presence of a Metro service in the area that would influence the number of trips made by public transport. It still used a logit model for mode choice but by concentrating on just one aspect of the full model and implementing it outside the complete structure of the four stage model they were able to develop a tool that could be used with a finer zoning system and incorporate more explanatory variables than were included in the full model. The sketch plan model produced a better match between the observed and modelled number of public transport trips for each zone, with a correlation at the home end of 0.899 for the regional model and 0.974 for the sketch plan model. When smarter choices can often result in small changes in small numbers, the opportunity to improve the accuracy of a model is valuable.

As the model was implemented as a stand-alone module it was possible to run many tests in a short amount of time and it was feasible to test a greater range of measures that might influence the levels of public transport patronage than would have been possible using the full model.

Sketch plan models have been used by planning authorities in the US for modelling smarter choices both to supplement and replace the mode choice component of the four stage model. This approach uses manual manipulation of trips matrices using spreadsheets or bespoke computer programmes which contain a database of the likely responses to particular smarter choices measures to provide an indication of the response to the package of measures under consideration (Jotisankasa and Polak, 2008).

One of the more sophisticated of these bespoke programs is the COMMUTER model developed by the US Environmental Protection Agency (Carlson *et al.*, 2005) which combines the database approach with logit models. First the current baseline mode shares are adjusted to reflect the impact of support and incentive programmes, for example car share programmes and providing cycle facilities which do not affect the time or cost of travel and so are not captured in the utility functions used in the logit models. The extent of the impact on current mode shares is taken from an extensive database of the likely impact of such measure compiled from experience all over the United States. Then a logit model based mode choice model is used to estimate the impacts of strategies which are designed to change the time or cost of travel such as car park charges, high occupancy lanes and fare reductions. A pivot point logit model is used with either a set of locally calibrated model co-efficients or a default set of national values.

This approach has some value in that it makes the practitioner review each component of the package of measures and decide if their effects can be captured in the logit model or whether its influence on trips needs to be made explicit outside of the modelling framework. However, other than providing a database of the experiences of implementing these measures elsewhere it does not show in what ways the trips will be affected. Although the number of vehicle trips is reduced, is that because people have switched to other modes or chosen not to travel at all? It also requires the practitioner to be able to decide how much of the observed effect elsewhere came from elements that were captured in the logit model and how much came from other factors.

The difficulties of these methods that rely on professional judgement to alter the matrices to account for smarter choices is that even experienced professionals are unsure of the nature and scale of the effects that will be produced in their area. Results observed elsewhere will have been affected by their local context and may not have the same impact when applied in a different context,

Bonsall (2009), Stopher (2009) and Chatterjee (2009) have all noted the many failings of evaluation studies, such as the lack of sufficient sample size, inadequate description of the measures actually undertaken, only interviewing participants in a travel behaviour programme, self-reporting of behaviour and the lack of a control group in order to understand the counterfactual. Even if a practitioner did have access to a high quality evidence base on the efficacy of these interventions, the context in which they were applied may well differ to the context of the area in which they are proposed and the actual combination of measures is likely to be different from those other studies. It places an impossible cognitive burden on the practitioner to transfer and modify mentally the results of programmes elsewhere into his area, while also separating out those impacts which will be due to aspects of the programme that can be modelled with the logit model.

3.5 System dynamics

Systems dynamics modelling considers the way the system to be modelled operates. It is a way of 'systems thinking' (Sterman, 2000) based on the belief that the subject of the model is a complex system, full of inter-dependencies between its component parts.

The aim of systems dynamics is to gain an understanding of the system being modelled to assist in the prediction of the consequences of changing one part of it. Thomas (1974) warns that 'you cannot meddle with one part of a complex system from the outside without the almost certain risk of setting off disastrous events that you hadn't counted on, in another, remote, part. If you want to fix something you are first obliged to understand…the whole system…Intervening is a way of causing trouble'.

Forrester (1989) traces the development of the systems dynamics approach to modelling back to when he was approached by General Electrics to consider why they experienced periods of working three or four shifts a day in the factory and then periods when they had to make staff redundant. Forrester mapped out the dynamics of the operations in the factory, their employment policies, production processes and inventories with pen and paper, concentrating on the dynamics in the system. The methodology he devised to explain the system forms the core of system dynamics.

A systems dynamics model is based on the links between the elements included in the model. It represents the direction and strength of causality of these links, which often leads to the description of causal loops. These loops may be positive (reinforcing) which amplify whatever is happening or negative (balancing) as shown in Figure 3.3 below. An arrow indicates that there is a causal relationship between two elements in the systems dynamics model. The '+' sign means that if the cause increases, for example the adoption rate of a new product such as mobile phones rises, the effect increases, and there is an increase in the number of people with a mobile phone. The '-' sign means that if the cause increases the effect decreases; so as the adoption rate rises the number of people without a mobile phone and therefore potential customers falls.

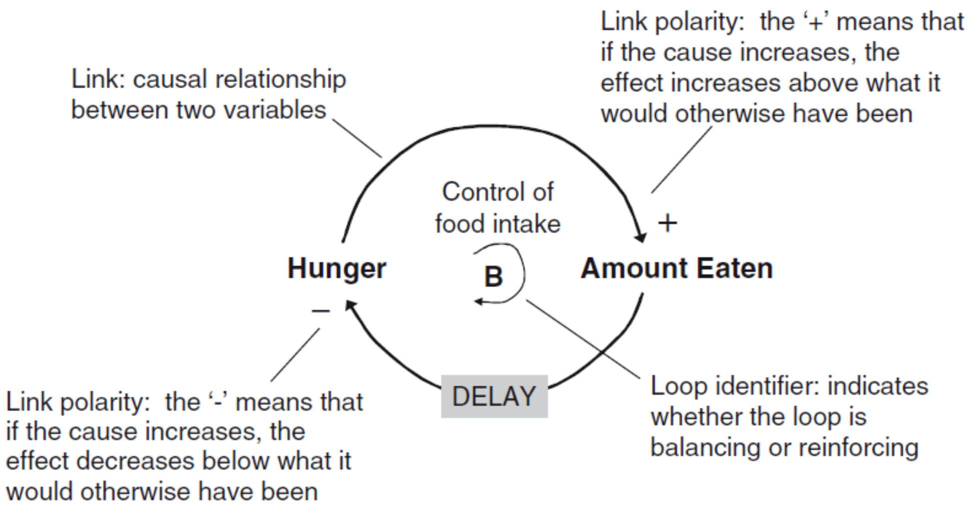


FIGURE 3.3 CAUSAL LOOPS IN A SYSTEMS DYNAMICS MODEL

Source: Blleininger, 2010

The feedbacks within the system are often not instantaneous and the system is modelled through time in order to capture the time lags within these loops. This results in dynamic models that can capture non-linear relationships caused by the varying strength of several inter-related causal loops. The systems do not always reach a steady state, or equilibrium, and may show instability and oscillations, such as the staffing level cycle observed by Forrester at General Electric. A common role of a systems dynamics model is to test possible interventions and to observe whether the consequences of these changes assist in the meeting of a policy objective, such as the maintenance of stable employment levels.

A further standard feature of systems dynamics models is the use of stocks and flows to illustrate the workings of the system. For example the number of people employed at the factory is a stock, and the rate of recruitment and dismissal are flows. Stocks contribute to the modelling of the dynamics of a system as they act as a kind of memory, storing the result of past actions. In a feedback process, past decisions and actions come to influence present decisions and actions through the amount of stock available.

The systems dynamics approach has been applied to several aspects of the transport system. Armah et al (2010) illustrated the issue of road congestion as a contributing factor in the air pollution problems in Accra in a causal loop diagram reproduced below in Figure 3.4. Complaints about road congestion leads to the building of additional road capacity which reduces travel times and results in more people using the car which leads to increased road congestion and worsening air pollution. This is the ‘induced traffic’ affect resulting from the feedback between the supply side of transport (build new roads) and the demand for transport. This extra traffic in turn leads to higher levels of air pollution which poses a health risk to the local population.

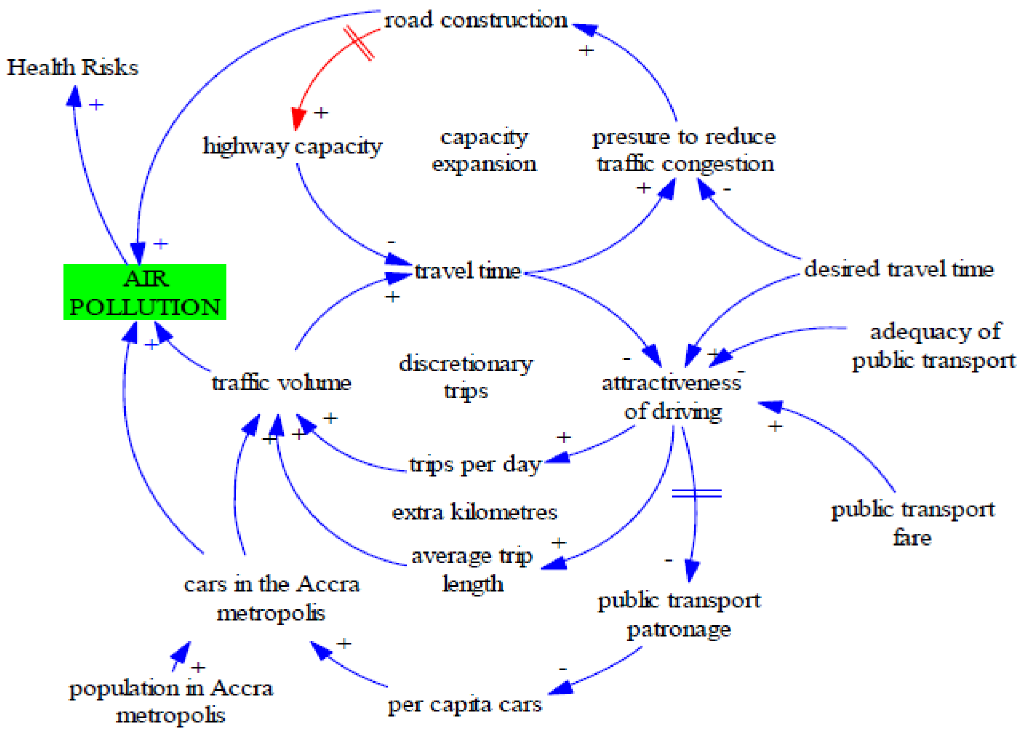


FIGURE 3.4 A CAUSAL LOOP DIAGRAM FOR ROAD CONGESTION

Source: Armah, 2010

Land use is a common subject for systems dynamics models, starting with Forrester's Urban Dynamics model (1969) which concluded that building social housing without also providing employment in the same area exacerbated urban poverty and eventually led to the boarding up of the new houses as people left to find work.

Swanson (2008) developed a model to capture the wider economic impacts of investment in transport infrastructure and used this model to support bids to DfT for investment in Leeds (Steer Davies Gleave, 2014). The model incorporates a zoning system which allows transport times and costs to vary across the modelled area. These transport costs are one of the variables that affect the attractiveness of an area for housing; the more attractive an area the rate at which people wish to move into an area rises and the rate at which they leave falls. The rate at which new houses are built in an area depends upon the demand for housing there, the level of existing housing stock, the amount of land available for building and the rate of return on new houses. For businesses, the attractiveness of an area depends on the transport times and costs which affects access to customers and suppliers, the ability to recruit staff and the availability of premises. The model was used in Leeds to trace the change in land use and GDP growth following the delivery of a new piece of transport infrastructure.

In relation to smarter choices, MacMillan (2014) developed a system dynamics model of the choice of cycling for the journey to work in Auckland, New Zealand. The model was built from evidence gained by interviewing 16 stakeholders to establish the relationships affecting the level of cycling use for commuting trips. The preliminary causal loop diagram for the model was then refined at two workshops with over 30 participants involved either directly or indirectly in designing transport policy in Auckland.

The causal loop diagram for the model, with two balancing and four reinforcing feedbacks, is shown in Figure 3.5 below. The dominant feedback loop is the balancing loop, B1. When there are more cyclists but no change in the bicycling infrastructure to accommodate them the number of cyclists involved in accidents rises, news of which deters people from cycling.

The main reinforcing feedback is R1, more investment in cycle-friendly infrastructure reduces the real and perceived risk of cycling, encouraging more people to cycle which in turn encourages more cyclists. The other two reinforcing loops are, R2, as more people cycle it becomes more normal and socially acceptable to cycle and, R3, there is safety in numbers with cycling and that as more people cycle the rate of cycle injuries per total miles cycled reduces.

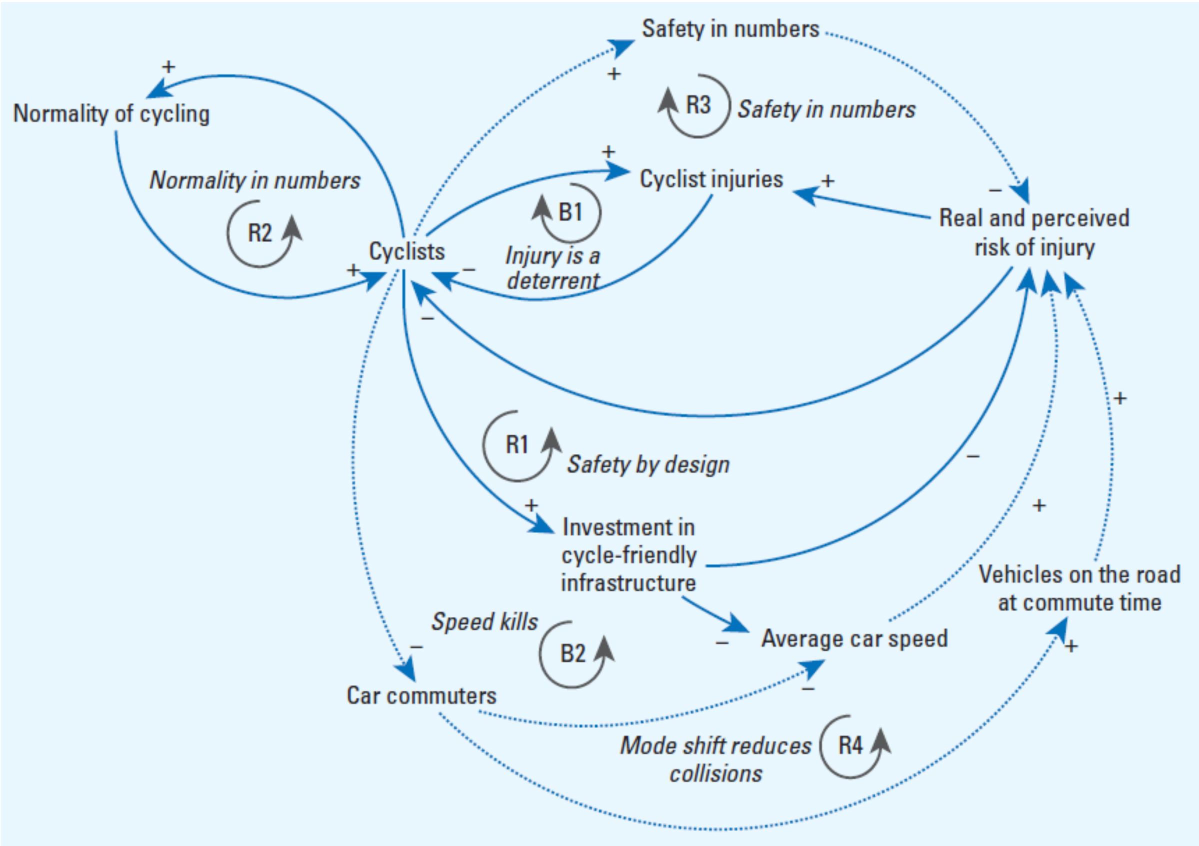


FIGURE 3.5 A CAUSAL LOOP DIAGRAM FOR BICYCLE COMMUTING TO WORK

Source: MacMillan, 2014

The strength and time delay in the feedback loops was quantified using available evidence from other studies. The model was validated by starting the model in 1991 and modelling forward to 2011, then checking against current cycle mode share in Auckland. The model was used to forecast annual cycling mode share forward to 2050. Cycling rates in Auckland between 1991 and 2011 oscillated around a low level of around 1% of all commuting trips during this period and this was replicated by the model. The point at which the strength of the reinforcing loops counteracts the balancing loops has not yet occurred in Auckland and is not shown in the model runs up to 2011 so, in my opinion, the conditions for this increase in cycle mode share, its timing and strength are not validated.

The model was used to test five policy scenarios;

- Baseline (no investment in cycling)

- Regional cycle network (RCN) investment - the current policy to invest in on-road cycle lanes and some new off-road shared footpaths
- Arterial segregated bicycle lanes (ASBL) - a major investment in segregated cycle lanes on every arterial route and provision for bicycles at intersections
- Self-explaining roads (SER) - the introduction of traffic calming measures and lower speed limits on all local roads
- Arterial segregated bicycle lanes and self-explaining roads.

The model was run for each of these scenarios and outputs reported for each year, such as the number of cyclists, the number of cycling injuries and savings in emissions level as trips switch from motorised modes. The cycling mode share predictions are reproduced in Figure 3.6 below. It shows that a modest increase in cycling rates can be achieved by the regional cycle network and traffic calming local roads but that a significant modal shift can only be achieved with the provision of segregated cycle lanes on arterial routes.

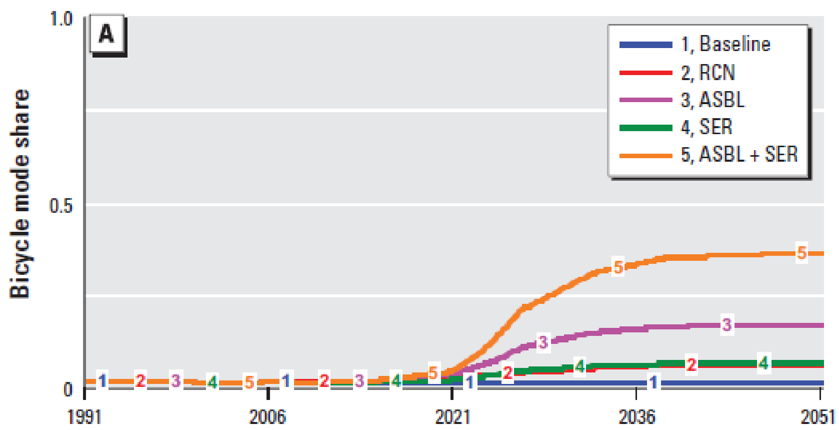


FIGURE 3.6 PREDICTED CYCLING MODE SHARE FOR COMMUTING TRIPS IN AUCKLAND UP TO 2051

Source: MacMillan, 2014

The strength of systems dynamics lies in its ability to capture the dynamic complexity of systems. This complexity arises because systems are, as Sterman (2000) outlines:

- constantly changing but at different rates

- tightly coupled, with elements interacting with other elements within the system
- governed by feedbacks
- non-linear, with an effect rarely proportional to its cause
- history dependent, where 'taking one road often precludes taking others and determines where you end up'
- self-organising, with the dynamics of systems arising 'spontaneously from their internal structure'. Small random perturbations can be amplified and modified by feedbacks which generate patterns and create path dependencies.
- adaptive, as a system evolves over time and participants learn from experience and change their actions
- affected by time delays in feedback channels which may have different strengths of responses, resulting in the long-run response in the system to an intervention being stronger or contradictory to the short-run response
- counterintuitive, the long term responses can be unexpected as unforeseen feedback loops come into force.

Systems dynamics models in the transport field have mainly been developed in order to develop an understanding of the relationships between various elements of the transport system and to explore the way these elements interact. It could be useful for exploring some elements of 'smarter choices' measures and has the potential to include a wider range of causal links affecting mode choice rather than just time and cost. It highlights that there are alternative ways of considering the transport system other than through the neo-classical framework of the four stage model. The transport system in an area can be considered as a dynamic complex system, with a whole myriad of causal links of differing strengths, direction and timing.

The main weakness for using systems dynamics to model 'smarter choices' is the very limited amount of segmentation that can be readily handled by these models. Given the heterogeneity of travellers in terms of their characteristics, preferences and the trips that they make, a modelling approach which recognises the complexity of the transport system but attempts to model it in a far more dis-aggregate way may be more successful in modelling ‘smarter choices’.

The ‘top down’ approach of system dynamics means that it concentrates on the processes at work in a system rather than the consequences of the changes in the system on individual players within it. When transport policies and schemes are appraised, a key consideration, as well as any impacts on mode share, time savings and vehicle operating costs, is the social impacts of the proposal. An assessment, however, of the distributional impacts of a scheme, which shows how people, including non-users of the scheme, are affected and in what way, is hard to achieve with a whole-system modelling approach. This suggests that a modelling approach that starts with the individual people using or affected by the transport system may be more appropriate for the appraisal of ’smarter choices’.

3.6 Microsimulation

Orcutt (1957) developed a ‘bottom-up’ modelling approach based on modelling each individual decision-making unit in a system. These units are ‘elemental decision making entities such as individuals, families, firms, labour unions and governmental units’. He was prompted to consider new approaches to modelling by his reflection ‘that current models of our socio-economic system only predict aggregates and fail to predict distributions of individuals, households, or firms in single or multi-variate classifications’. He felt that ‘there is an inherent difficulty, if not practical impossibility, in aggregating anything but absurdly simple relationships about elemental decision-making units’. If most relationships are non-linear then aggregation becomes problematic and he proposed modelling at the micro level. Predictions about the future state of the system would be obtained by aggregating the state of each of these individual units.

He proposed that this micro-level modelling was carried out in a series of discrete time steps. In each time step, the decision or action of each modelled unit depends upon its inputs. ‘An input into a unit is anything which enters into, acts upon, or is taken account of, by the unit such as its age, gender, employment status’. The output is often drawn from a probability distribution, for example, that shows the likelihood of a person getting married given their age or from the application of a set of rules. The model traces the state of these units over time and the overall state of the system is obtained by aggregating the state of each of the individual units.

Static microsimulation models which model only a single time step became popular in the 1970s, primarily in the modelling of the impact of changes in the tax and social security systems. The rules in these systems are often non-linear and the exact change experienced by a person depends on that individual’s circumstances, such as the age and income of other members of the same household. A microsimulation model can calculate the impact of the proposed change for each individual and it is by aggregating the individual outputs that the total absolute change is predicted, for example the change in the total level of social security benefits after a rule change in entitlement. When aggregating the individual changes it is possible to present the distribution and scale of the impacts and to identify the winners and losers from a proposed change.

Static models calculate the immediate effect of a change whereas dynamic models consider the effect over time. Static models consist of two parts (Martini and Trivellato, 1997); an initial data set which contains the relevant attributes such as age or income of each individual or household unit, and a set of accounting rules which are used to calculate the effect of a proposed change. The accounting rules are applied **once** to model the immediate effects of a proposed change and so are sometimes termed ‘morning after’ models. (Collins, 2006). The attributes of each individual remain constant in a static microsimulation model. If the model is used to predict the effect of a proposed change some time into the future then the individual units are re-weighted to reflect a forecast of the future composition of the population for the relevant model year (Spielauer, 2010)

In Europe the EUROMOD model provides a standard microsimulation framework for modelling European tax systems and has been applied in over a dozen European countries, using the relevant database of individuals and the tax rules for each country. (Harding,2007). In the UK, H.M. Treasury (2014) has a tax and benefit microsimulation model that is used to model the distribution of the impacts of budget changes. They regularly publish the results of their modelling work.

Orcutt's vision for microsimulation was to model a whole series of time steps, now known as dynamic microsimulation. In dynamic models the state of a unit in one time period can affect the outcome in the next time period, which means that the approach can model path dependency and collect information on the life-histories of the modelled units. Orcutt's attempts at building such models in the late 1950s though were hampered by 'the lack of sufficient computer power and data availability at that time' (Spielauer, 2010).

The main areas of application of this modelling approach in the transport field are the modelling of the movement of vehicles on the road network and activity based modelling. During the 1990's a number of commercial software packages became available, such as Vissim, Paramics and Aimsun, which model the movement of each vehicle individually through the network. The vehicles are assigned routes and car-following and lane-changing rules are applied to the movement of vehicles through the network, with vehicles responding to the actions of other vehicles or agents. The decisions of each agent depend on factors such as 'their age, gender, risk-taking behaviour, driving skill, vehicle size, and vehicle performance characteristics', the position and speed of other agents and the physical characteristics of the environment as specified in the description of the road network. (Panwai and Dai, 2005, Laufer, 2007). This approach to trip assignment is commonly known in transport modelling as microsimulation.

3.7 Activity based models

3.7.1 The structure of activity based models

The activity based approach to transport modelling has established itself in the past 10 years as an alternative approach to the four stage model. Most of the activity based models currently in use have been developed in the United States. As of 2012, there were 12 completed models in the USA and a further 10 under development, Activity based models are also used in Toronto, Jerusalem, Tel Aviv and Copenhagen (Gliebe and Picado, 2012).

Activity based modelling is based on the idea that travel is a derived demand, not undertaken for its own sake, but in order to undertake an activity at a specific location (Jones, 1977). The modeller focuses on predicting the schedule of each person's activities for the day. The trips made by that person then follow from this schedule, often known as a 'day plan'.

The activity based modelling approach is based on the work of Hägerstrand (1970) who proposed a theoretical framework of time geography where each individual could undertake their activities with a set of resources but facing a set of constraints. The resources are time and space and the constraints are:

- capability constraints on the physical and technological limitations of an individual.
- coupling constraints arising from the need for a person to undertake certain activities at the same time-space locations as other people
- authority constraints from institutionally imposed restrictions and regulations, such as shop opening hours.

The two most common modelling platforms for activity based models are CT-Ramp and DaySim. The developers of both these systems describe their models as microsimulations. 'The CT-RAMP system is implemented in a fully-disaggregate microsimulation framework' (Parsons Brinckerhoff, 2011) and DaySim 'uses a microsimulation structure' (Bowman *et al.*, 2006). Activity based models (ActBM) are micro-simulation models; the units in their models are individuals and a set of rules are followed to determine the activities they are likely to follow and the time, destination and mode of the associated trips. Most ActBMs use a set of nested logit models to determine the probability of a person undertaking each of the possible activity schedules and associated trips in their choice set and then compares these probabilities with a random seed to determine which activity schedule and transport choice is assigned to that individual. A minority of ActBMs, such as the ALBATROSS model, use a set of heuristics to model the activity plans and transport choices of each person.

The typical structure of an ActBM is shown in Figure 3.7 below. The modelling of a person's choices is unified by using of a series of linked logit models for each choice. The utilities are based on journey times and costs. These are passed up the modelling structure, while the components of each person's choice sets are passed down the model, to ensure consistency in the application of time/space constraints and in the modelled decisions for each individual. This consistency in mode choice, for example, means that if a person drives to work in the morning they are not allocated to public transport for their final trip home at the end of the day.

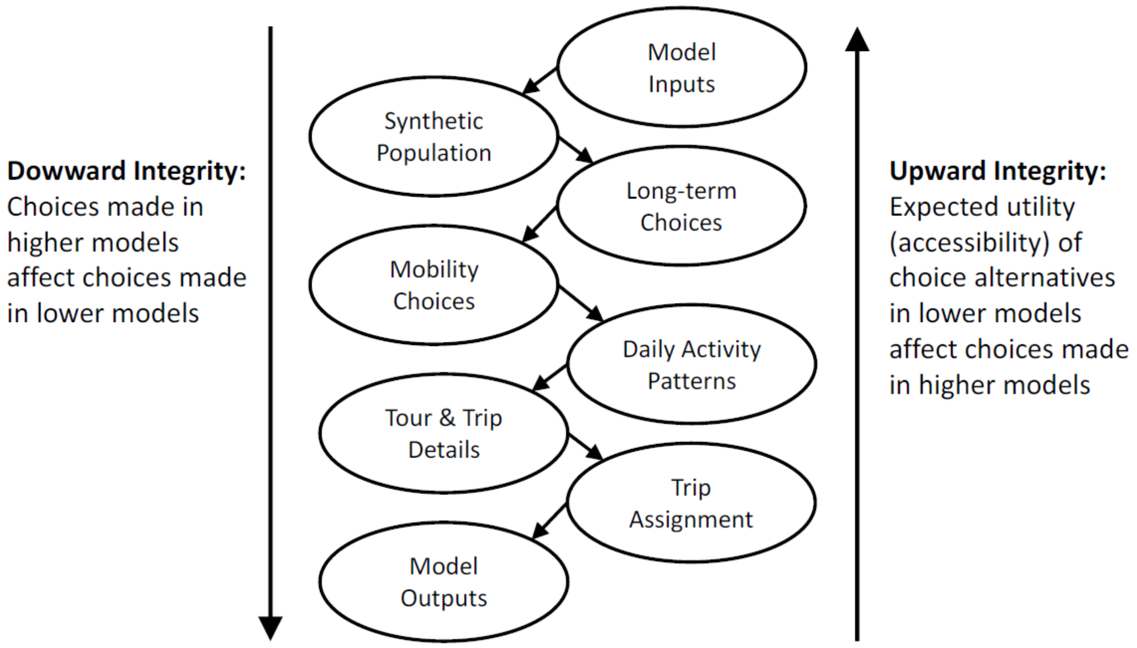


FIGURE 3.7 MAJOR STEPS AND INFORMATION FLOWS IN AN ACTIVITY BASED MODELLING SYSTEM

Source: Castiglione, Bradley and Gliebe, 2014

The basic components of an activity based model are as follows:

Synthetic Population: This is a database of all the individuals living in an area, containing information on the attributes for each person required for the model and the location of their home. Typical attributes, such as those used in the Chicago model (Parsons Brinckerhoff 2011) are, for persons, their age and employment status and for households, their size, number of workers, income, car ownership and housing type. The techniques used to create 'synthetic populations' are considered in more detail in Chapter 8.

When modelling future years, a synthetic population is needed which matches forecast numbers for key household and person variables such as the number of households, number of individuals in each age category and future employment levels. A current research area is the further development of microsimulation household evolution models to predict the changes in a population over time, for example through births, deaths and migration, in order to produce the future population for input into forecast year runs of an ActBM directly from the base year synthetic population. (Gliebe and Vovsha, 2012). This approach has already been adopted in the ILUTE model for Toronto (Miller 2009).

Long term choices: In recognition that some choices that affect a person's activity plans are not made on a daily basis, separate models are used to predict a person's usual work and school locations. (Outwater and Vovsha, 2012).

Mobility choices: Separate models are also often used to predict factors that affect a person's mobility decisions but do not change on a daily basis such as car ownership, possession of a driver's licence, bicycle ownership and season ticket ownership. The results of these models, for example, whether a person owns a car or possesses a season ticket affects the modes in an individual's choice set and the costs for that individual of using each mode.

Daily activity patterns: The creation of individual daily activity patterns 'is the cornerstone of activity based modelling and key differentiating feature from 4-step'. (Vovsha and Gliebe, 2012). The daily activity pattern lists the time and place for each of the activities scheduled for each person during the model's 24 hour time period. A typical classification system is:

- location: at home or out-of-home
- type: mandatory, maintenance, or discretionary
- purpose: work, education, shop
- priority: primary activity/destination vs. secondary activity/stop
- intra-household interaction: individual or joint.

For each person there are a large number of potential daily activity plans. For example in the Sacramento DaySim model, even after excluding unobserved and infrequent combinations, each person has a choice of one out of 2080 possible daily plans when considering the order and type of activity alone (Bowman el al., 2006). The choice of daily activity plan is modelled using a logit model with the utility of each plan based on the sum of a constant value, the utility of each activity which varies for each household/person type and the utility of the travel options.

Many ActBMs take account of household characteristics, such as income and car ownership, when constructing activity plans for the people in those households. Some also take account of joint travel arrangements between household members and the constraints imposed by family obligations such as escorting children to school.

Tour and trip details: A tour is the linked set of trips from the primary origin point, to a series of secondary destinations, and back to the origin point. Once an activity plan has been selected for an individual it is converted into a set of tours and their component trips. For example, for a person with a daily activity plan of leaving home, dropping the children at school, going to the office, going to a café at lunchtime and then shopping at a supermarket on the way home, their daily activity plan converts into two tours, with a total of six trips:

- Home - school - work - shop - home (4 trips)
- Work - café - work (2 trips)

A set of logit models is used to determine the time of day, destination and mode used for each of the trips, although the order in which the decisions are taken varies between different implementations of the ActBM framework. The software is designed to ensure that consistency is maintained between all trips and the other elements of the model, so that for instance if the person does not have access to a car they are not able to choose the car driver mode or if they take cycle to work in the morning they do not return home as a car driver.

Auxiliary demand: some trips in an area are not derived from the population’s daily activity plans. This includes goods vehicle movements and trips made by people living outside of the study area who are visiting or travelling through the study area. These trips are modelled separately.

Trip assignment: once the origin, destination, mode and time of day for each trip has been determined, the trips are combined into trip matrices and assigned to the transport network to derive travel times and costs. In most ActBMs the assignment is carried out with the static assignment algorithms used in the software that implements four stage models but some areas, such as San Francisco, are moving towards the use of traffic micro-simulation, (Erhardt, 2013).

3.7.2 Strengths of activity based modelling

One of the main advantages of activity based models compared to four stage models is that they can be sensitive to more policy variables. In addition to modelling responses to changes in the time and cost of trips, they can consider changes brought about by policies that affect a person’s activity schedule, such as home-working, flexible working hours and extended shop opening hours. They can also model the ‘knock-on’ impact of a person’s mode choice on their daily plan For example if a person chooses to walk to work it may well take them longer and that additional time may lead to them foregoing another activity such as a visit to the gym, or altering the time of other activities in their schedule for the day.

The ActBM approach achieves greater internal consistency within the model. Origin – destination (OD) based four stage models work solely on the basis of trips so a person could be assigned to public transport in the morning when congestion is high and car for the return trip in the evening when congestion is lower and/or public transport frequencies are lower. Greater consistency can be achieved by using Production – Attraction (PA) matrices with a four stage model but there are issues with converting these PA matrices into the appropriate time slice matrices for assignment which requires OD matrices.

The ActBN approach should produce more accurate forecasts of the number of people using each mode, as the use of specific costs for each individual reduces the aggregation bias found in the matrix based four stage models. By modelling each individual, the model uses a more accurate set of costs for that person than with the four stage approach, where the same set of costs is used for all the trips between the same origin – destination and the only level of differentiation in travel time/costs between travellers is provided by the limited degree of segmentation in the matrices. As Figure 3.8 below shows, the use of an average cost in the logit model when predicting mode choice, rather than an individual's actual costs leads to aggregation bias. The figure shows the typical S-shape of a logit curve and two travellers with different costs, a and b. The average of the probabilities of using a particular mode is different when the costs are used separately in the logit model, 'average P', than if the average cost faced by the travellers is used, which results in 'P at average cost'.

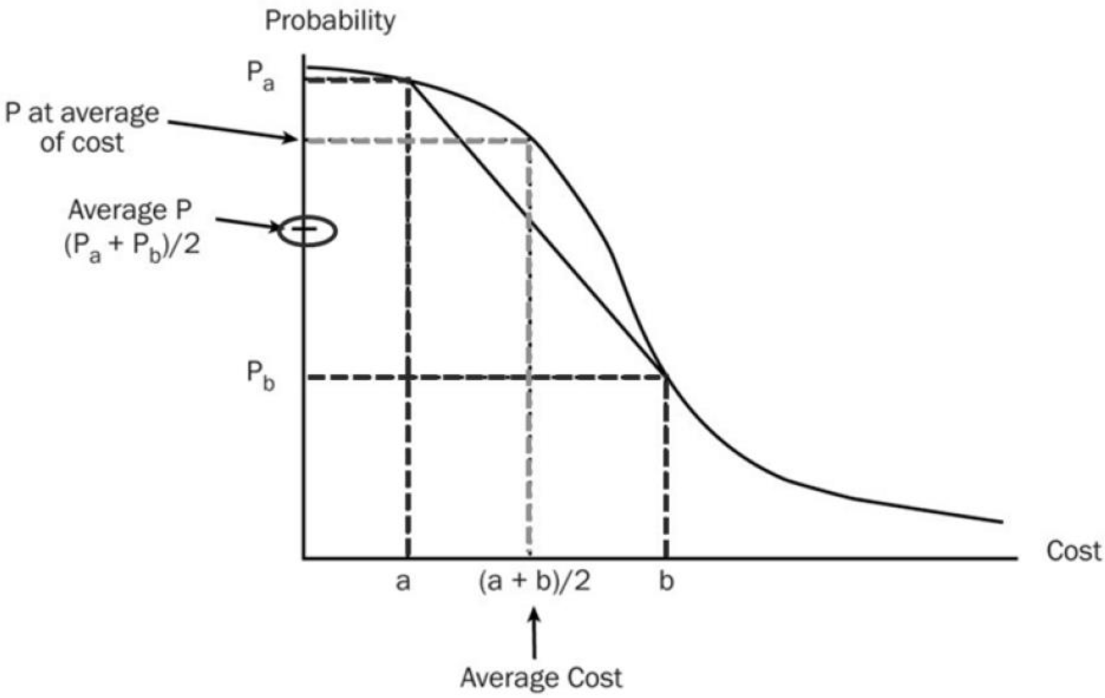


FIGURE 3.8 AGGREGATION BIAS THROUGH THE USE OF AVERAGE COSTS

Source: Castiglione, Bradley and Gliebe, 2014

ActBMs work at an individual level which also allows for greater spatial detail in the model outputs. Policy-makers are concerned with the details of who will be affected by proposed changes in the transport system and the more detailed output provided by ActBMs improves the information required for social and distributional impact assessments and equity analysis.

The benefit of the ActBM approach for the modelling of ‘smarter choices’ is that it provides a way of meeting some of the recommendations made by WSP (see section 3.2.1) for the enhancements needed to four stage models to enable them to better model the impact of ‘smarter choices’ measures. By modelling individuals, it enables the models to use a more accurate set of times and costs for each traveller. This enables it to model ‘smarter choices’ policies that affect the times and cost of travel of very specific groups, such as dedicated employer provided bus services and on-site parking restrictions or aim to remove constraints on mode choices such as providing bicycles and secure cycling facilities. It also offers the possibility of modelling some policies which WSP acknowledge could never be incorporated in four stage models, such as the impact of employer policies to permit some home - working which would reduce the need to travel for commuting purposes and flexible working hours which may enable staff to commute outside of peak times.

3.7.3 Weaknesses of activity based modelling

The main disadvantages of activity based models come from their ambition to model both a person's activity choices and their transport choices which are inter-related decisions. This means that activity models become large and complex and, in practice, simplifications are often made when calculating the journey times faced by individuals in order to reduce the run times of the models (Bekhor *et al.*, 2010). The use of ever more powerful computers, sometimes using cloud based resources, is reducing the impact of the complexity of the models on run times but there are often budget limitations which affect the use that can be made of these expensive computing resources.

The complexity of the nested logit models used in ActBMs makes the calibration of these models a skilled task and there are a limited number of practitioners who can undertake this very specialist work. This can mean that the practitioners using the final models are not fully aware of the compromises made during the calibration process and how these may be influencing the final model outputs.

There is also some difficulty in being confident in the final model calibration as the model incorporates a multitude of decisions both on the nature, place and length of activities as well as travel decisions. There is the danger that as Polak (2011) observed 'the mismatch between ambition and capability means we just create bigger messes'.

Recent work in the USA has considered the impact on model results of the way in which the random numbers are produced and applied which affects the conversion of probabilities produced by the logit modelling into the actual travel pattern for a person. Bowman *et al.*, (2006) reported that different runs of a model's base case, with absolutely identical inputs, would produce varying results due to the different random numbers used during the model run. When the model is then run to test a policy scenario ‘the difference in the predictions resulting from changing the random numbers…will be mixed with the differences resulting from changes in the policy variables, with no way of separating the two’. When the authors tested the Sacramento model with a variety of policy measures such as a congestion charge for the central business district and increasing the connectivity of the network they found ‘that the difference in results for the same scenario but with different random numbers is typically greater than the difference in results for different scenarios’.

They concluded that ‘because most policy differences at the regional level are quite small, it does not take a great deal of random error to outweigh the policy effect, and thus the “law of large numbers” is not sufficient if different random sequences are used’. In the Sacramento model they programmed the model so that the same random number was used for the same resident/tour/trip combination in both the base case and the scenario test in order to reduce the amount of variation in model results that was due to the use of random numbers. There remains though a concern that the effect of the use of random numbers is not well understood by model developers and how or whether it is handled in the modelling can affect the final results produced.

Another concern raised over ActBMs is the amount of data they require. They are very reliant on data from household travel diaries which contain information on the trips and activities undertaken by each person, as well as their personal characteristics, in order to develop and calibrate the scheduling and travel choice models. This data is expensive to obtain, especially if care is taken to ensure that diaries are completed by a representative sample of the population in an area and that all trips made by each person are recorded.

Activity based models are heavily reliant on the use of random utility logit models and there are concerns as to whether these models are able to capture all the behavioural responses to changes in a transport system. This has led some activity based modellers, such as the developers of the MATSIM software, developed at ETH University and the University of Berlin, to incorporate some aspects of agent based modelling into the ActBM framework in order to increase the behavioural realism of these models (Balmer, Nagel and Axhausen,2005). Consideration is now given to agent based modelling.

3.8 Agent based modelling

Agent based modelling (ABM) is a bottom up approach to modelling which produces dynamic models of the development of a system over time. The basic unit of these models is agents, 'autonomous units … capable of processing information and exchanging this information with other agents in order to make independent decisions' Castle and Crooks (2006). Each agent has a set of rules that determine its behaviour and a set of goals to achieve. The agents interact with each other and their environment to pursue these goals and the state of the overall system emerges from the individual actions of its agents.

Agent based modelling (ABM) is a recent but rapidly developing approach to simulation modelling that has its roots in the modelling of complex systems. A common theme from the investigation of complex systems is that the behaviour of the complex system emerges from the activities of its lower level components as they interact with each other (Miller and Page, 2007). The patterns detected at the more aggregate level come from the 'bottom up' in systems which can have both positive and negative feedbacks as the lower-level components carry out their own agendas (Arthur, 1995). It provides a way of examining the dynamics of a system as it develops over time and tackles issues such as the level of dependence on initial conditions, thresholds, tipping points (Gladwell, 2000), criticality and phase transitions.

3.8.1 Agents

An agent based model has three components:

- the agents
- the environment

- the interactions between agents and between each agent and the environment.

3.8.1.1 Agent characteristics

O'Sullivan and Haklay (2000) state that 'a precise definition of agent-based models is elusive', but the general idea is portrayed by envisaging an artificial world of heterogeneous, autonomous agents each following a set of rules governing their behaviour in an attempt to achieve their own goals.

Wooldridge and Jennings (1995) describe the characteristics of agents from the perspective of computer science and propose that agents are:

- Autonomous - agents are independent and separate, operating without the direct intervention of humans or others, and have some kind of control over their actions and internal state (Castelfranchi, 1995).
- Socially able - agents can interact with other agents (Genesereth and Ketchpel, 1994).
- Reactive - agents perceive their environment and are able to respond to changes in it.
- Pro-active - agents do not simply act in response to their environment, they take the initiative to achieve their goals.

Franklin and Graessner (1996), built on this definition by adding the following characteristics:

- Temporally continuous – an agent is continuously existing
- Goal - oriented – an agent has a goal or goals
- Learns – from its experience
- Adaptable – can change behaviour based on previous experience
- Mobile – can move around
- Flexible – it's actions are not scripted
- Has character – an agent has a believable "personality" and emotional state.

Not all agents in a particular implementation of an agent based model need to have all these characteristics. The aim of modelling is to distil the essence of a system and to capture it in the modelling, using just enough detail to make the model realistic and true to the original but preserving as much simplicity as possible to aid understanding of the processes contained in the model.

3.8.1.2 Agent behaviours

The ABM approach is well suited to the modelling of alternative approaches to decision making to those used in the neo-classical economic model. Standard methodologies used by economists involve solving sets of mathematical equations to deduce the equilibrium state, with the underlying assumption that all the entities in the model always make perfectly rational choices. Rational choice is an optimisation process where the preferred choice among alternatives is that which maximises utility given constraints (Gilboa, 2010). This approach incorporates the assumption that everyone has perfect knowledge of the alternatives available to them, is capable of trading-off the different attributes of the costs and benefits of each alternative e.g. the cost and time elements and can pick the option that will be best for them. Agent based modelling provides a tool that enables the development of models that relax these assumptions.

ABMs are able to control and monitor the information available to each agent at the time they take their actions. It can record what information they retain about the past and what predictions, if any, they may make about the future. This enables it to depart from the assumption that all agents have perfect knowledge of the past, present and future. It is also able to incorporate other constraints that may operate on rational choice.

ABMs can incorporate a variety of methods for handling constraints on optimisation. These include constraints on human cognitive and processing constraints and response constraints (Maynard 2010), such as the impact of habitual behaviour which leads to inertia and an undue influence of a past decision on a current choice. An ABM can also record an agent's history and so be aware of past decisions and incorporate a mechanism for allowing these to influence its current behaviour.

The work of behavioural economists and psychologists has reported a wide range of heuristics used by people to make decisions. The growing evidence on human decision making being garnered by behavioural economists can be used to inform the design of the decision making process in ABM models. An interesting opportunity presented by agent based models is that different individuals can use different choice processes within the same model.

Girengezer (2011) has conducted research for many years into heuristics used by individuals to make decisions and believes these better represent the way people make decisions than utility maximising models. These heuristics (Schwanen and Lucas, 2011) can include:

- maximum, where the highest score of a single attribute is used to make the choice,
- dominance, where an alternative is selected if it scores the highest on each of a key set of attributes
- conjunctive and disjunctive choices where the scores on several characteristics are used to select or eliminate some options, so narrowing the field of possible alternatives
- lexicographic, where options are ranked on a scale for a number of attributes valued by the decision maker and the option with the highest combined score is selected.

Schwanen and Lucas (2011) note that there has been very little research into these heuristics as applied to transport decisions. Foerster (1979) found evidence from a survey of mode choice decisions of 91 respondents that conjunctive and lexicographic models fitted observed mode choice data better than models with people making trade-offs between time, cost and other journey attributes as assumed in logit modelling.

Research in other fields, such as the choice of food in a canteen (Scheiberhenne *et al.*, 2007) found that 'everyday food decisions can be understood and predicted based on a surprisingly small amount of information and very simple rules of thumb'. When decisions were modelled based on utility functions, the model replicated current decisions better than rules of thumb but when the same models were used to forecast the food decisions of an alternative set of diners, the rule of thumb models performed better. They suggest that this is because the abundance of data for the calibration enabled the utility functions to be fitted well with the observed data. However these observed parameters did not match the choices of the unobserved diners and the rules of thumb performed better at predicting their choices.

Hess *et al.* (2008) reviewed mode choice data from stated preference surveys in Denmark, the United Kingdom and Australia. They found evidence for inconsistent responses, non-trading choices, lexicographic decision-making as well as the anticipated utility-maximising choices. In the Danish study 22% of respondents displayed non-trading decision making, with 16% always choosing the cheapest of option and 6% choosing the quickest. The Australian survey found that the degree of non-trading varied by journey purpose, with 12% of commuters and 22% of non-commuters always choosing the cheaper option when a toll was included in the choice set. The United Kingdom study found that the degree of non-trading behaviour varied according to the current mode used with 46% of car users never changing mode and 20% of bus users always choosing the bus. The general split of respondents between non-trading, lexicographic and compensatory decision making cannot be determined from these surveys, as this was not the intention of the original survey design. For example, non-trading could be lexicographic behaviour but the level of attributes needed to detect this for a particular individual were more extreme than those used in a survey intended to produce weightings for attributes in a utility maximising decision making model.

Hess *et al.* (2012) found that the removal of these non-trading individuals from the dataset affects the resulting weightings applied in the calibration of a logit model.

3.8.2 Environment

The environment is defined by Teahan (2010) as ‘being everything that surrounds the agents, but which is distinct from the agent and its behaviour’. The environment is the world in which the agent operates. It usually has a spatial dimension and each agent has a position in a space bounded in two or three dimensions. Sometimes this is a representation of a specific physical location and the agent based model is linked to a geographic information system.

Russell and Norvig (1995) suggest that an environment can be described in terms of its position along several axes, as shown in Figure 3.9 below. Their five key characteristics of an environment are:

Accessible vs inaccessible*.* In an accessible environment, an agent has ready access to all the information that is relevant to its choices i.e. it has perfect knowledge. This is seldom the case and most environments are inaccessible to some degree.

Deterministic vs non-deterministic. An environment is deterministic if the effect of a single action is certain, that is ‘the next state of the environment is completely determined by the current state and the actions selected by the agents’. Many ABMs contain stochastic processes which make the outcome non-deterministic.

Episodic vs non-episodic. In an episodic environment, the agent experiences a set of separate episodes in which it perceives and then acts. Each episode is self-contained; it is not affected by the previous episode and it does not affect future episodes.

Static vs dynamic. An environment is dynamic if it can change while the agent is deliberating on its course of action.

Discrete vs continuous. An environment is discrete if there are a limited number of possible actions.

Accessible	**---------------------**	**Inaccessible**
Deterministic	**---------------------**	**Non-deterministic**
Episodic	**---------------------**	**Non-episodic**
Static	**---------------------**	**Dynamic**
Discrete	**---------------------**	**Continuous**
Simple	**---------------------**	**Complex**

FIGURE 3.9 THE ENVIRONMENT

Source: Russell and Norvig, 1995

3.8.3 Interactions

The interactions in an agent based model consider with which other agents a particular agent interacts and how it interacts with its environment. One of the defining characteristics of an agent based model is that only local information is available to agents so each agent only interacts with a subset of all the agents in the model. The typology of a model describes which agents are connected with which others.

The role of networks is crucial in many agent based models and part of the model building process is to determine how links are established between agents and hence the extent, in each point in time, of the network for each agent. There could be a consistent rule which determines the links from each agent, for example, they are only linked to the closest two agents if the model has a spatial dimension. The number of links coming into, or out of a node (or agent) in a network is known as its degree of clustering. In a tight knit community, most of the agents would have many connections and the spread of activity such as messages regarding the quality of a newly introduced bus service between the agents would be more rapid than in a loosely connected community.

3.8.4 Agent based modelling and microsimulation

Agent based modelling and microsimulation are both individual-level modelling approaches, In pure agent based models the emphasis is on interactions between individuals and their environment and the behaviour of the individuals which may evolve over time. The emphasis is on incorporating a few simple behavioural rules into the model and then observing the emerging behaviour of the system as the individuals follow these over time.

In pure microsimulation models, a set of mathematical rules or transition probabilities are applied to each individual to determine their state in the next time period. (Williamson, 2007). These models are usually very rich in data on the attributes and circumstances of each individual. There is a debate within the ABM community as to whether these two approaches will combine over time as microsimulation models begin to add behavioural rules and ABMs seek to empirically ground their models (Rounsevell, 2012) by adding more detail into the models. Chattoe-Brown (2009) believes that it is not ‘the case that agent based modelling and microsimulation will naturally meet in the middle’ as the use of behavioural rules makes a model distinctively ABM. This allows ABMs models to seek to explain the observed behaviour of the system being modelled rather than just to predict its outcome as the result of the application of transition probabilities as in microsimulation models. Williamson (2009) believes that there is ‘a continuum from a ‘pure’ microsimulation model, totally fitted to empirical data, on the one hand, to a ‘pure’ agent based model, entirely based on theory driven rule sets on the other’. The software that supports agent based models can readily handle microsimulation models, including links with GIS and the spatial detail that is needed for spatial microsimulations. This allows for the development of models that draw upon both techniques. AnyLogic software takes this further and allows the creation of models that combine elements of system dynamics, agent based modelling and microsimulation within the same model.

3.8.5 Agent based modelling and transport applications

The dynamic microsimulation traffic assignment models developed for the final, trip assignment, stage in the four stage model are an example of the merging of agent based modelling and microsimulation. These packages, described in section 3.6 above are known as 'microsimulation packages' but the drivers could be considered as simple agents, displaying a few of the characteristics listed by Wooldridge and Jennings (1995). Macal and North (2007) use the term 'proto-agents' to describe drivers in these traffic simulation tools to acknowledge their limited number of agent based features.

Further additions to agent behaviours have been made in some of these models by adding interactions between agents and their environment. Dia *et al.*, (2002) added the influence of real time information on driver behaviour. A discrete choice survey was used to calibrate a logit model which was then used to allocate drivers to one of a set of choices when faced with real time information about delays such as whether to continue on the current route or to change route. The extension of microsimulation assignment models to include richer methods of modelling behaviour reflects the more general move towards the merger of the two techniques seen in other disciplines.

The SUSTAPARK parking model (Dieussaert *et al.*, 2009) is described as an ABM of parking search behaviour. It relaxes the assumption of perfect knowledge by the agents but applies a logit model to the choices available to the agent. The utility of the option of parking in an on or off street place near the driver is calculated and a logit model, based on work by Hess and Polak (2004), is used to determine the choice made by drivers. Benenson's (2008) PARKAGENT model was developed for residential parking in the evening. It uses a set of complex rules constructed by the author to model the decisions made by drivers. The rules are applied at various stages of the journey, such as the probability of a driver choosing to park or drive on depending upon the number of free spaces passed. Once the destination has been reached the driver will choose any free space as long as it is not too far from the final destination and as search time increases the driver considers paying to park.

McConnell and Zellner (2011) built a prototype ABM for a single route with the mode choice decision making using a simple rule based approach rather than a logit model. People have the choice between driving or using a bus rapid transit (BRT) service. In the model initialisation, a set proportion of users were assigned to bus on the first day. The model was run for 20 days and on each day a user changed mode if the time they experienced on the previous day was higher than their pre-set threshold. The model was used to assess the effectiveness of an exclusive lane for BRT rather than the BRT using the lanes for general traffic. The threshold values did not come from observed data. Sensitivity tests were carried out on the impact on the final mode share of BRT of varying the initial proportion of bus users and the threshold levels that trigger mode switching.

Han *et al.*, (2011) applied ABM techniques to the modelling of destination choice. In a prototype model they considered the choice of shopping destination in a synthetic city. Agents are only aware of a limited number of shopping destinations. They vary in their knowledge of the attributes of each destination in their choice set and in the weighting they give to those attributes. The model concentrates on the role of social networks in the choice of destination. Information from contacts in an agent's network can make him aware of new destinations which may enter his choice set and update his knowledge of the details of locations he has not visited recently. The agent may amend the weightings applied to the various attributes of a location to better reflect those of other members of his network. As long as a destination remains 'satisfactory', agents exhibit habitual behaviour. Otherwise they consider alternative locations and choose the destination which gives them the highest utility.

The MATSIM software (Balmer, Axhausen and Nagel, 2005) adds features of agent based modelling into an activity based modelling framework with its 'strategy' module. In MATSIM each agent starts with a day plan of activities and travel between them. These agents are all assigned to the network and the time taken to make the trips is recorded. The travel times come from a dynamic microsimulation traffic assignment model within the MATSIM package.

Each agent's day plan is then given an overall utility score, with the activities having a positive utility and travel a negative utility. In the strategy module an evolutionary algorithm is used to improve the value of an agent's plan. In each iteration a copy of the original plan is made for a proportion of the agents and random alterations made to their planned activities. The utility of the new plan is calculated and the plan adopted if it has a higher overall utility than the current plan. The model is run for many iterations. Over these iterations the average score of agents' plans rises as more highly valued plans replace those with lesser values. This approach allows for the current travel times on the network to feed back into the scoring and possible replacement of a person's day plan affecting their activity and travel behaviour.

3.8.6 Agent based modelling and smarter choices

The key characteristics of an agent based model are:

- each agent can carry detailed personal information which reflects their heterogeneity
- the model is dynamic and shows the development of the system over time
- the model has built in mechanisms for feedback between the environment and agents which leads to emergent behaviour
- it is possible to model different decision rules for different agents
- the history of each agent is recorded.

This provides the opportunity for agent based modelling to tackle many of the modelling refinements needed to incorporate smarter choices as suggested by WSP. The use of individual agent characteristics means that the model can be aware of very detailed data such as the exact location of the origin and destination of a trip. Combined with networks based on the fine level of detail available from GIS networks, accurate representations of the time and distance of even short distance trips is possible. The implementation of policies to restrain the availability of car parking at certain locations, or to levy differential prices for parking and public transport fares can be modelled more accurately as the actual individuals affected can be identified.

The dynamic nature of the model means that it can provide insights into how a transport system may develop over time and provide measurements of key indicators over time. The four stage model assumes that a stable equilibrium state is achieved by the transport system. In an agent based model this is not necessarily the case and the model may show either the progress toward such an equilibrium state, or constant flux as no such state is reached. Being a dynamic modelling tool the model can show how long the system takes to respond to changes, reflecting the level of change in response to a policy measure (such as increasing car parking charges) in both the short and long term.

The interactions between agents mean that the outcome of the model is not obtained by solving a set of mathematical equations but rather by recording the emergent patterns. This is particularly valuable when the relationship between two variables is non-linear. An ABM enables observation of processes as they pass through these critical points. It also allows the modelling of processes operating in the real world that may reinforce or counteract each other and suggests to policymakers how they could intervene, if desired, to affect these processes.

The bottom up approach to modelling enables the model to apply different decision making rules to different individuals which may be a better representation of reality. It is also possible to use the same decision making rule, for example utility maximisation, but to vary the preference given to different elements of the utility function. (Rounsevell *et al.*, 2012). A further characteristic of ABM, of particular relevance to the modelling of ‘smarter choices’, is its ability to record the history of each agent. This means that the decision making of an individual does not have to be based on the assumption of perfect knowledge. A person may only know the travel times and costs they have personally experienced or of which they have been informed by people in their social network or external agencies. ABMs can therefore be used to model responses to marketing initiatives.

The ability of ABM to model the networks between agents also provides a mechanism for representing the spread of information about travel modes and the influence of others on an individual's attitudes and choices.

In the WSP review on the capabilities of the four stage model to handle 'smarter choices' measures, they reviewed a variety of such measures and considered how they could be incorporated in current models. The table presented here in Appendix 1 comes from this report and considers workplace travel plans which are relevant for commuting trips. It is extended in this study by the addition of a final column, shaded grey, which shows the characteristics of agent based modelling that could be utilised to incorporate each measure in an ABM model of a transport system.

3.9 Conclusion

The second research question for this study asks 'what modelling approaches could be used to model the impact of 'smarter choices' programmes on the mode chosen for commuting trips?'. This chapter has reviewed the academic literature for work on both enhancements to the current four stage modelling framework and alternative modelling approaches which may be suitable. Before addressing the second research question consideration is given in the next chapter to the philosophical approach to transport modelling underlying this research project and which influenced the response to the second research question.

4 Research perspective

4.1 Introduction

This research focuses on the ability of different forms of transport models to assess the impacts of ‘smarter choices’ measures. The interviews with practitioners led to the understanding that ‘smarter choices’ pose a real challenge to the current four stage modelling framework used in transport modelling. The previous chapter reviewed recent work on extending the capabilities of four stage models and considered the potential of alternative modelling approaches. This chapter considers the essential characteristics needed in a model of a transport system. It presents critical realism as an appropriate philosophy of science to apply to the task of assessing the merits of different modelling approaches. The second research question is addressed at the conclusion of this chapter.

4.2 Purpose of modelling

In an explicit model the assumptions are laid out, the processes contained in the model specified and the outputs presented for inspection by others (Epstein, 2008). A key objective of building a transport model is to be able to use it to predict the future state of the modelled system in a way that the knowledge of the predicted outcome can be shared. Parties interested in the prediction include those who will be affected by the outcome and those involved in making plans to either accommodate that outcome or to attempt to influence the system such that an alternative outcome is realised.

As well as providing predictions a model can produce information on a range of possible outcomes, yield an understanding of how the system being modelled operates and assist in communication amongst interested parties. Exact knowledge of the future state of the inputs into a model are seldom known but an estimate can often be made of the likely range of values. For example, a key input into many strategic transport models is the level of GDP growth. In the UK a ‘most likely’ value is issued by the Bank of England within fan charts which also show the probability distribution of the likely future growth rates in GDP. An explicit model can be run a number of times using different input values taken from within this range to provide an indication of the range of likely outcomes. Models can also be used to test the range of outcomes when a number of input values are changed simultaneously.

The running of a model over a range of input values and the analysis of the subsequent outputs gives an insight into how the system being modelled operates and which variables have the greatest influence on the outcome achieved. This will assist in identifying those areas likely to be the most effective for policy makers to influence and which variables warrant careful monitoring. A close examination of the results may also provide an insight into the processes at work in the system, particularly which reinforce each other or produce counteracting influences.

A model can assist in communication amongst people as it requires making explicit the assumptions and the relationships incorporated in the model. This provides a basis for discussion as to whether the current state is accurately portrayed in the model and the values which should be assumed for the future. The model also provides a concrete portrayal of a future state so that discussions on future policies can be based on a shared understanding of future conditions. As Epstein (2008) comments, ‘models do *not* obviate the need for judgment. However, by revealing trade-offs, uncertainties, and sensitivities, models can *discipline the dialogue* about options and make unavoidable judgments more considered’.

4.3 A model and its subject

Having considered why a model can be useful for policy makers, this section considers the features required of a model. Miller and Page (2007) suggest that a model is like a map that allows ‘people to easily acquire and productively use information about a complex reality’. It portrays the area it covers, distilling the essential information and revealing insights otherwise obscured by the detail. It can be used not only to provide directions but also to assist in our understanding of the world. The simplicity of a map, which highlights selected aspects of reality, can make apparent patterns which are lost in the complexities of the real world. Snow's seminal mapping of the cholera epidemic (Brody *et al.*, 2000) is an early example of using a model to reveal the underlying cause of observed events.

Holland *et al.* (1986) expanded this metaphor of maps to illustrate the process of modelling a system that changes over time. It is not sufficient for a model to accurately represent the current state of the system it is modelling, it should also contain representations of the key processes at work in that system. Such a model can then be used to predict the future state of the modelled system. The model is validated by showing both that:

- it provides an accurate representation of the current state of the modelled system
- when the model is used for prediction, the future state in the model is still an accurate representation of the future state of the system it is modelling.

The main implications of Holland's work for transport modelling are that the model should not only contain an accurate representation of the current state of the transport system but also of the processes at work in its transformation. The scope of the model also needs to be well defined so that it captures all the key influences on these processes.

4.4 Philosophical influences on transport modelling

Timms (2008) observes that 'the main philosophy actually used in transport modelling has been positivism' which is unsurprising given the fields from which it has borrowed its tools and the engineering background of many transport modellers. This is manifest in the reverence that is accorded to observed data in model building and the great attention that is played to ensuring a transport model accurately replicates current conditions. Current transport modelling practice uses models which consist of a series of mathematical equations. The transitions coded into the models are 'laws' of travel behaviour based on key concepts taken from neo-classical economics such as the assumption that people have perfect knowledge of all the transport options available to them and seek to minimise their travel costs and time.

Timms (2008) calls for transport modellers to consider an alternative philosophy of science to underpin their modelling and to consider other approaches. He puts forward the idea of modelling as a process of communication and 'perceives transport modelling as a linguistic activity within the overall context of transport planning, which is in turn considered as a communication process'. Timms' perspective draws on the hermeneutic philosophical tradition and highlights the usefulness of using models as part of a communication process.

Following Timms' admonition, the research reported in this thesis is based on an alternative philosophy; one which places an emphasis on the processes at work in the real world. Models are trying to capture the essence of this world and for this, the uncovering of the processes at play within it (which are more than manifests of language) is vital. The perspective of this research is that the process of building a model and presenting its results can assist in the communication of ideas but that an understanding of the processes at play in the transport system is essential to producing models that can provide accurate predictions.

4.5 Critical Realism

Critical realism offers a comprehensive philosophical background in which to embed the task of transport modelling. This research adopts as its basis for determining the worth of a model the critical realism framework as set out by one of its leading originators, Roy Bhaskar, in his works 'A Realist Theory of Science' (1975) and 'The Possibility of Naturalism' (1979). The next section outlines the main ideas of critical realism. The implications of the insights gained from critical realism for the choice of a modelling approach for incorporating 'smarter choices' into transport modelling are drawn out in Section 4.6.

4.5.1 Key concepts in critical realism

Bhaskar started his thinking that led to critical realism when he tried in the late 1960s to apply the theories of neo-classical economics to the problems facing the economies of developing countries. He was confronted with the issue that the assumptions lying behind neo-classical economics, such as perfect knowledge and utility maximising behaviour, did not seem to describe the world he was studying.

Bhaskar asked himself what the world was actually like, rather than the world as simplified by economists in their modelling. He sought the answer to the transcendental question “what must the world be like for science to be possible?’’. He concluded that the world must consist of real objects independent of whether we observe them or not, and tendencies (or mechanisms or laws of nature) which may be ‘possessed unexercised, exercised unrealised, and realised unperceived (or undetected)’.

The **real** is whatever exists regardless of whether we understand it or not. It does not have to be a physical object; it could be a mechanism such as gravity. The **actual** world refers to what happens when these mechanisms are activated. The **empirical** world is defined as the domain of experience. It is the things we actually observe or experience. This is illustrated in Table 4.1 below. In a modelling exercise the aim is to model the world by capturing the essence of its characteristics and operations in the area of interest, while being aware that our observation and understanding of these mechanisms is hampered by the fact that many aspects of the world exist in open systems which contain many mechanisms, often with countervailing impacts at work simultaneously.

	Domain of Real	*Domain of Actual*	*Domain of Empirical*
Mechanisms	√		
Events	√	√	
Experiences	√	√	√

TABLE 4.1 THE STRUCTURED WORLD

Source: Bhaskar, 1975

4.5.2 Critical realism and transport modelling

The critical realist philosophy of science provides four insights which can usefully be applied when designing models to test ‘smarter choices’ policy interventions in transport systems.

First, there is a need to be explicit about which world is being modelled; the actual, the real or the empirical. Some models are simply mapping the world at the level of the empirical (observed) level. They may be accurate descriptions of what is observed but lack any explanatory powers as they do not attempt to capture the processes involved. For a model to act as a useful forecasting tool for policy makers, it needs to be modelling the real world and aiming to replicate not only observed events but the underlying observed and unobserved processes that give rise to the observed events.

Second, the transport modeller needs to be aware of the danger of attributing causality to the mere conjunction of events. ‘Explanation depends instead on identifying causal mechanisms and how they work, and discovering if they have been activated and under what conditions. Events arise from the working of mechanisms which derive from the structures of objects and they take place within geo-historical contexts (Sayer, 2000). The mathematical relationships contained in many transport models describe a conjunction of events e.g. the number of cars as a function of the number of houses rather than explaining why a certain number of houses generates a particular number of car trips. This means that they are of limited value in forecasting the number of car trips in the future and how this number could be changed. They can only vary the number of car trips as the number of houses changes but the underlying conjunction between the number of houses and the number of car trips may not be constant over time. The transport modeller needs to consider whether he is modelling conjunctions or causality.

Third, the modeller needs to consider whether the model contains sufficient depth in order to capture all the relevant mechanisms at work. The processes operating in the system can lead to the emergence of other structures. Danermark *et al.* (2001) writes that ‘the outcome of the mechanisms – the events we can observe – is a complex combination of the influences from other mechanisms reinforcing each other while others counteract each other's manifestations'. The transport modeller needs to consider whether his model has sufficient depth and complexity to capture all the relevant influences and emergent behaviour.

Finally, the observation that the social world operates mainly as a series of open systems is a reminder that transport models are delineating part of an open system to create a closed system. Care needs to be exercised in the setting of the boundaries of transport models and a clear description provided of the chosen boundaries and which factors are considered to be exogenous to the model. When making predictions using the model, it must be remembered that the assumed values of these exogenous factors can change and the influence of the exogenous processes may change over time.

4.5.3 Critical realism and choice of modelling approach

A critical realist philosophy has three ground rules, ontological realism (our world is real), epistemological relativism (our knowledge is fallible) and judgemental rationality (we can choose between theories). It allows for the recognition that it is possible to influence the outcomes in the transport arena, and that although our knowledge of the processes at work may be fallible, these should be captured in the transport model. Exploring a transport system in the artificial laboratory of a transport model may provide insights into the inter-relationship between factors, improve our understanding of them and produce clues as to which variables could be targeted in order to move the performance of the transport system from one state to another.

For a critical realist, the minimum needed to appraise an intervention is:

- the context
- the action
- the intervention

- the mechanism
- the outcome.

This provides a set of criteria which can be used to judge the appropriateness of a transport modelling methodology for capturing the effects of 'smarter choices' measures. It can be assessed with regard to its ability to describe the context of the transport system, its success in replicating the current activity observed in the system, the way in which it formalises the processes at work in the system and the decisions made by the individuals within it, and finally how well it predicts the outcome of an intervention against the observed outcome when it is implemented.

The Critical Realist philosophy suggests that the choice of modelling methodology should be guided by the needs of the study and its ability to illuminate the mechanisms at work in the area under study. Fleetwood (2005) considered critical realism as applied to organisational and management studies, and noted that critical realism states that there is 'one' reality but 'advocates selecting research methods and techniques according to the nature of the phenomena under investigation'. Outhwaite (1987) adds that critical realism does not 'exclude any method *a priori*, but the choice of method should be governed, on the one hand by what we want to know, and on the other by what we can learn with the help of different methods'.

Critical realism also emphasises that the events we observe are the outcome of the interactions between individuals and groups within society and that investigations should not be conducted solely at a single level, i.e. individuals, groups or society level. 'One cannot concentrate solely on a single level of investigation of the society, group or individual: critical realism argues for a relational perspective' (Dobson, 2002).

These insights from critical realism are of relevance to the selection of a modelling approach that can accommodate the whole range of 'smarter choices' measures. They led to the selection of two criteria for use when assessing possible modelling approaches. The first is that the resulting model needs to be able to contain behaviourally realistic representations of the processes involved in a person's travel decisions. The second is that the modelling technique should operate at the level of the individual. It should be capable of capturing the impact of a person's context on their travel behaviour choices, the constraints and capabilities of each person and the interactions both between individuals and between individuals and their environment. This insight was combined with the findings of the literature review recorded in the previous chapter to address the second research question.

4.6 Second research question

The second research question for this study is 'what modelling approaches from other fields could be used to model the impact of 'smarter choices' programmes on the mode chosen for commuting trips?. The previous chapter reviewed extensions to the current four stage model that could assist in modelling 'smarter choices' such as using finer zones and more detailed networks to achieve a better representation of travel times and costs, the inclusion of latent variables such as attitudes in the choice modelling and implementing greater segmentation of travellers so as to better group them according to their preferences. Latent class analysis provides a useful clustering technique for identifying such groupings and allocating individuals to the appropriate segment. Even if these enhancements are made though, there will always be difficulties arising from the holding of information in the form of matrices rather than having the trip information associated with particular individuals. The framework also does not lend itself to the modelling of some 'smarter choices' such as car sharing which require very detailed knowledge on the potential car sharers and their working arrangements.

Critical Realism points to the importance of modelling the individual and therefore the selection of a modelling approach which uses the person as the basic unit, building the pattern of travel in an area from the bottom up based on the decisions of these individuals. This led to the rejection of the four stage modelling approach as it is based on the opposite approach. It starts at the aggregate level, with total person trips, and subdivides them until it reaches the number of trips made by mode and time of day between zones used in the assignment of trips to the network.

The use of sketch plan methods was rejected as this is not really a distinct modelling framework but rather a method developed for the swift application of selected parts of a four stage model. It is often adopted to compensate for deficiencies in a complete four stage model such as the long length of time taken to build a complete model, the lack of fine levels of detail about the actual origin and destination of trips and local influences on mode choice decisions. The spreadsheet and GIS tools used to implement these models restrict the ability to expand them to cover the full range of choices affecting trips in an area such as destination and time of day choice. The approach also does not provide a coherent framework in which to model the actual processes involved in transport decisions and the way in which the full range of 'smarter choices' measures may influence them.

Systems dynamics is a useful tool for considering the processes involved in the system as a whole and how they may re-inforce or counteract each other over time. It could be usefully employed to consider which processes need to feature in a more detailed model. It was rejected for use in this study, as it cannot handle information on the different constraints operating on each person and the fine level of detail required on the time, cost and other characteristics of each of their journey options.

Microsimulation meets the 'individual' criteria adopted for testing possible modelling approaches as it adopts a 'bottom up ' rather than 'top down' approach. By modelling at the level of the component units of a system, the aggregate state of the system emerges from the state and decisions of each part of the system. Microsimulation operates, however, in a very formulaic manner, applying rules to determine the state of each agent in each time period or drawing a state at random from a probability distribution. Activity based models have shown that a modelling approach that is based on the person as the basic modelling unit has the capability of modelling more components of a 'smarter choices' package of policies than a four stage model. It does however lack the ability to offer a richer capture of the processes involved in the transport system which is offered by agent based modelling.

Agent based modelling appears to offer all the benefits of microsimulation, with its emphasis on modelling individual people, the basic component units of the system, but combines this with a greater emphasis on modelling behaviour processes together with the interaction between agents and between agents and their environment. Bonabeau (2002) proposes that ABM is suitable when one wishes to describe a system from the perspective of its constituent units and their activities. This is appropriate when:

- the behaviour of individuals cannot be clearly defined through aggregate transition rates
- individual behaviour is complex and so describing it with equations becomes an intractable mathematical problem
- activities are a more natural way of describing the system than processes
- stochasticity applies to agents' behaviour. ABM can be used to apply randomness in the right places rather than as a more general noise term as used in aggregate equations
- when validation and calibration of the model will need to be assisted by the use of expert judgement.

Agent based modelling was selected for use in the third stage of this research. It offers great potential for the modelling of 'smarter choices' but there is a gap in our knowledge as to how it can be applied to this task. The remainder of this study seeks to address this by building a proof-of-concept agent based model of commuter mode choice and using this model to consider the insights that the ABM approach could provide to the assessment of 'smarter choices' packages under consideration by policy makers.

5 Overview of the modelling process

5.1 Introduction

This chapter presents the methodology used to build the agent based model used in this study to assess, through practical application, the potential for modelling mode choice decisions and incorporating 'smarter choices' into multi-modal models. The research design adopted for the third stage of the research project was to develop a model using the stages shown in the modelling cycle shown in Figure 5.1 overleaf. The model would then be used to test techniques made possible by the agent based modelling approach which could be usefully employed in modelling 'smarter choices'. Finally the insights gained from the experience of building and using the proof of concept model were used to present guidance on how the technique could be used in applications which modelled the transport system for a specific area.

The stages used in the development of the agent based model for commuting trips developed in this research project are based on those presented by Lay-Yee and Cotterell (2012) supplemented with a prior stage proposed by Thalheim (2011) and a final stage advocated by Rounsevell (2012) and Railsback and Grimm (2011). The seven stages in the modelling process are:

- Statement of purpose
- Conceptualisation
- Computing platform
- Data integration
- Implementation
- Application
- Presentation of results

Apart from the initial stage, the stages form a modelling cycle, shown in Figure 5.1 below, which may be followed through in its entirety several times and/or involve the repetition of several stages in smaller loops, such as between data issues, implementation and application (Railbacks and Grimm, 2012).

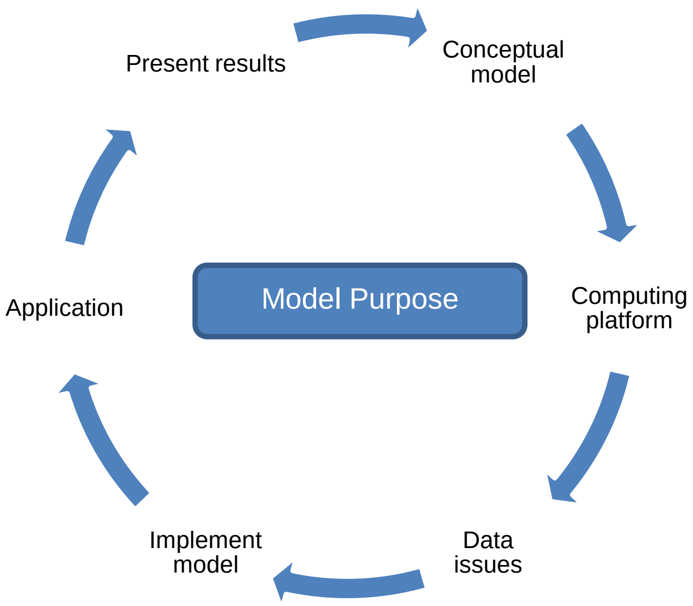


FIGURE 5.1 MODELLING CYCLE

5.2 Setting the model purpose

The purpose of a model is the 'intentions, goals, aims, and tasks that are going to be solved by the model' (Thalheim, 2011). Thalheim proposes that the whole modelling process should be governed by the purpose of the model and this purpose provides the overriding consideration whenever decisions are made. Once the purpose is settled it should be preserved by the model and be invariant during the modelling process. This means that the model will be particular and appropriate to its own objectives and purpose and should not then be used for other purposes without consideration as to whether it is transferable to these other proposed contexts and uses (Law, 2006).

The purpose of the model developed in this project is to assist in answering the third research question 'what are the strengths and limitations of using an agent-based approach for modelling the impact of a ‘smarter choices’ programme on the mode chosen for commuting trips?’

Shemli (2010) describes three main purposes of modelling; description, explanation and prediction. A descriptive model aims to present and summarise the main features of the subject being modelled, without attempting to examine any underlying causal mechanisms. The focus of these models is on the measurement and representation of the elements in the model’s subject domain, for example a regression model which shows the association between journey distance and the proportion of trips made on foot.

Explanatory modelling in contrast has the specific purpose of aiming to test and uncover causal explanations; the emphasis is on determining causality rather than just describing and recording correlation. Epstein (2008) provides the example of the plate tectonics model which can usefully explain why earthquakes occur but does not predict when and where the next one will happen.

Predictive modelling uses a model in order to predict the outcomes associated with a particular set of inputs. It can be used to predict values at a future point in time or in another place. This sort of modelling is of particular value to policy makers in predicting the results of policy interventions which will change the values of key inputs into the transport system (for example changing fuel prices), as the model will show the outcomes of those policies. The complexities of the real world and the strength of second order impacts can make it impossible for the likely results of a policy change to be realised by a ‘thought exercise’ alone (MacLeod *et al.*, 2013).

Heath *et al.* (2009), in their review of 279 simulation models published between 1998 and 2008 categorised each model’s purpose in terms of the level of understanding available to the modeller of the workings of the real / target system which was the subject of the model. When little is known about the real system, then only a Generator model can be built, capable of generating hypotheses and theories about the behaviour of target system. The purpose of such a model is to test whether a particular theory is capable of generating the behaviour observed in the real system.

When the working of the real system is moderately understood a Mediator model (Morrison and Morgan, 1999) can be built. A Mediator model does not contain a complete representation of how the target system behaves but can be used to test theories and provides insights into the characteristics and behaviours of the real system. As understanding of the real system grows and real data is collected a Mediator model has the potential to be developed into a Predictor Model.

A Predictor model can be built when the real system is well understood and can be used to give clear predictions of the future behaviour of the real system.

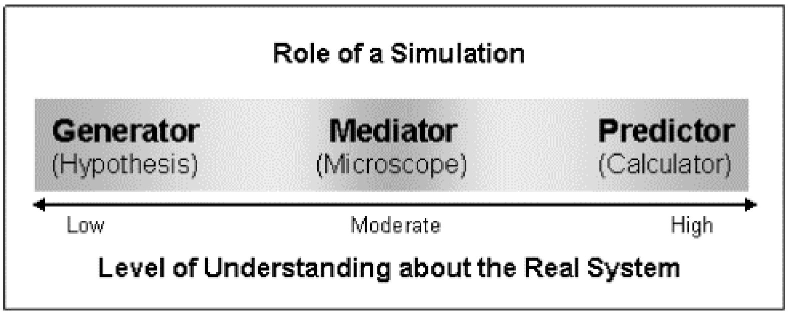


FIGURE 5.2 ROLE OF A SIMULATION

Source: Heath *et al,* 2009

A critical realist approach would aim at producing a Mediator model and the needs of policy makers to make decisions regarding which transport interventions to fund suggest that the agent based modelling approach should also be tested to investigate whether it is capable of producing Predictor models. This requirement means that the model developed for this project will have to be capable of incorporating real data relating to transport systems.

The availability of new software languages based on object oriented techniques means that data no longer needs to be stored in matrices and that processes can be coded to operate directly on individuals rather than to manipulate the numbers stored in matrix cells. It also allows for the relaxation of some of the basic assumptions of neo-classical economics such as perfect knowledge and consistent rational decision making.

The model purpose is therefore to illustrate that by using agent based modelling techniques it is possible to model mode choice decisions on an individual basis and incorporate added behavioural realism, such as the influence of habitual behaviour, on the number of people using each mode. The model is not intended to provide accurate predictions of the mode shares in a particular place but rather to be a more generic model that can be used to investigate the potential of this approach for modelling mode choice. The model will have the characteristics of a 'Mediator' model. It will not contain a complete representation of all the factors affecting commuting mode choice decisions but it should be capable of testing theories. The model will provide insights into the characteristics and behaviours of the real system and assist in the design of the collection of data required to develop a Predictor model.

5.3 The conceptual model

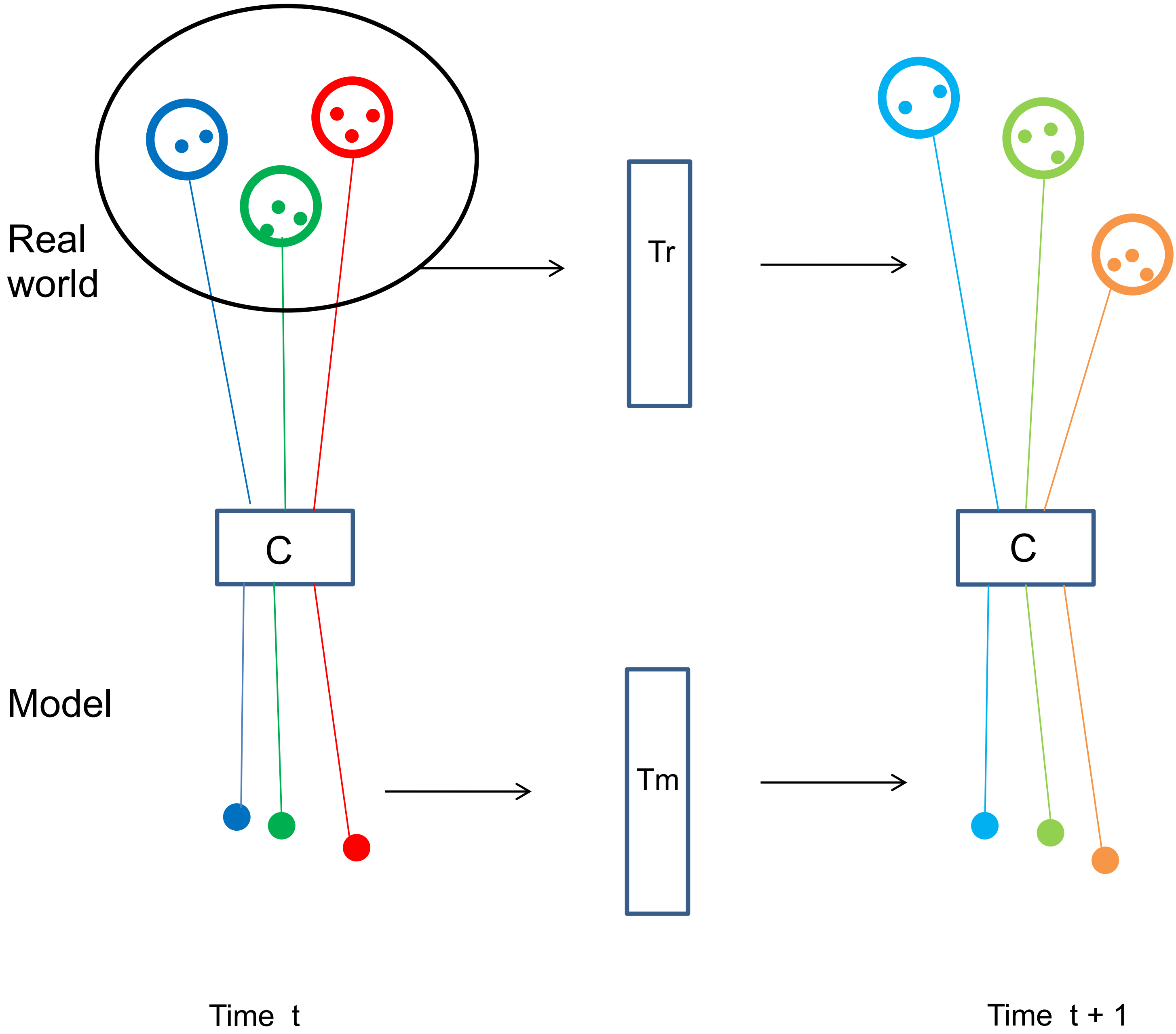
5.3.1 Mapping the world

After the model purpose is set the next step is to develop a conceptual model which will then be implemented in the computer model. Conceptual modelling is defined by Mylopoulos (1992) as the "activity of formally describing some aspects of the physical and social world around us for the purposes of understanding and communication". The conceptual model presents descriptions of the entities and processes included in the model. The static elements of the model are the real world **entities** which are described by their attributes and relationships with each other. The dynamics elements of the model are described by the **processes** included: their interfaces with the entities and their behaviours (Kung, 1989). In the language of critical realism, the processes are the mechanisms at work in the world and the entities are the objects in the world upon which they operate.

Mapping the real world to the model is crucial in the development of the conceptual model. It necessitates formalising decisions which establish both the scope and the scale of the model. The scope of the model sets the model boundaries: what will be internal to the model and what influences will be considered as exogenous. The scale of the model, whether it is a micro, meso or macro level model influences the units chosen for the objects in the model such as, for example, individual animals or species.

Figure 5.3 below adapted from Holland *et al.* (1986) illustrates the mapping of the real world to its model. The top section of the diagram reflects the real world and the bottom section depicts the modelled version of the real world. The boundaries for the domain to be captured in the model are shown by the black circle. This boundary needs to be chosen with care in order to ensure that it contains all the relevant objects and the influences on them that are required for the model to fulfil its purpose.

Once the boundary is established the objects or entities required for inclusion in the model and their attributes are recorded. A key principle of modelling is abstraction or simplification. One way of simplifying the model is to group similar objects together using a classifying rule. This means that many objects which are identical as far as the relevant attributes for the purpose of the model are concerned can be represented by a single object in the model. This abstraction is shown in the left hand side of the figure, where the classification of objects is recorded using colour. All the objects sharing the attribute red are represented by a single red object in the model. The classification rules are represented in Figure 5.3 below by the box marked C.



C: classification rules

Tr: transmission processes in real world

Tm: transmission processes in model

FIGURE 5.3 MAPPING THE REAL WORLD (ADAPTED FROM HOLLAND, 1986)

Abstraction also applies to processes and is achieved by isolating and often simplifying the mechanism at work in the real world. The model mapped in Figure 5.3 depicts a system that changes over time, from time t to time t+1. The transitions (in Holland's language or mechanisms in the language of critical realism) that affect the objects over time are represented in the upper central box marked Tr.

Consider Holland's (1986) example that the figure illustrates a model of how different areas of an oil painting picture fade over time. The different areas of a picture are classified according to the colour of the pigment used in the paint applied to that area. The conceptual modellers may decide that they consider the transmission mechanism in the real world relates the decrease in pigment colour to the strength of the sunlight to which the picture is exposed, for a given number of hours. This results over time in the picture having less bright colours. The real world is shown in the top half of Figure 5.3 above.

This transmission mechanism is encoded into the model so that the model can be used to predict the future state of the picture. The modelled world is shown in the lower half of Figure 5.3. The final state of the picture in the model, lower right hand corner, could be compared with the final state of the real world, shown in the upper right hand corner. Differences between the two can be investigated and could lead to further improvement of the model such as a refinement of the classification rules or the number and nature of the transmissions. For example it could be observed that some parts of the picture are covered with varnish. The entities in the real world (areas of the picture) could then be classified according to their pigment and whether they are covered in varnish. The transition rules could be expanded to have different rates of colour loss for varnished and unvarnished painted areas. If the modelled state in time t + 1 then better matches the actual state of the world in time t + 1 these changes can be considered to have improved the model.

The quality of the final model is dependent on the quality of the underlying conceptual model which in turn depends on the quality of understanding of the real world system and the skill with which the key entities, their attributes and processes that affect them are identified and maintained in the trade-off between model simplicity and its accuracy in replicating the real world.

5.3.2 Theories of behaviour

The conceptual model for this study is based on Triandis' Theory of Interpersonal Behaviour (1977) which builds on Fishbein and Ajzen's Theory of Reasoned Action (1975) shown in Figure 5.4 below. Their model proposed that a person's observed behaviour could be predicted by their intended behaviour, which in turn was influenced by their own attitude towards the proposed act and their beliefs about what other people think about the proposed act (social norms).

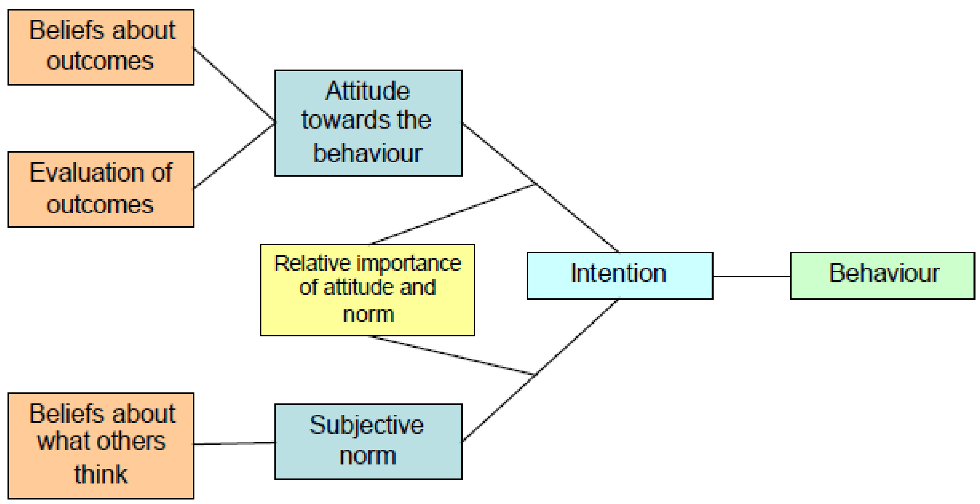


FIGURE 5.4 THEORY OF REASONED ACTION, FISHBEIN AND AJZEN 1975

Source: Jackson, 2005

Triandis expanded this theory in a number of ways, crucially for this research by proposing that observed behaviour is not solely influence by intentions but is a result of the relative strength of both habitual and intended behaviour. His theory is illustrated in Figure 5.5 below.

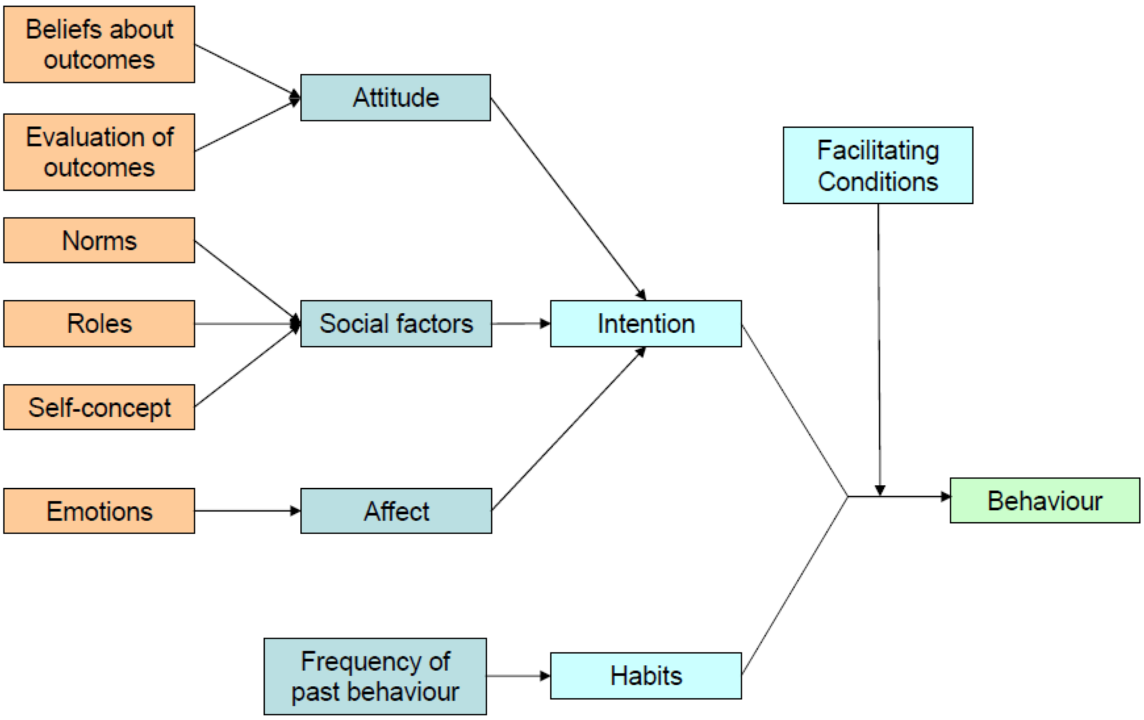


FIGURE 5.5 THEORY OF INTERPERSONAL BEHAVIOUR, TRIANDIS 1977

Source: Jackson, 2005

A person's intended behaviour is determined by their attitude, social factors, and affect. Attitude refers to the extent to which the individual has a positive or negative evaluation of performing the behaviour under consideration. This depends upon their beliefs about the outcome, for example, in the desirability of the outcome and their evaluation of the outcome, for example, how likely is it to succeed and whether the benefits exceed the costs of the undertaking the behaviour and any associated risks (Chatterton, 2011).

Social factors include norms, roles and a person's self-concept. Social norms are the usually expected behaviours in society and relate to how people behave in general in society. Roles are defined by Triandis as "sets of behaviours that are considered appropriate for persons holding particular positions in a group" and relates to how other people holding a similar position in society behave. Self-concept is a person's perceived identity, such as, for example if they identify with pro-environmental climate change believers and wish to avoid the use of fossil fuels in their travel arrangements.

Emotions (affect) are a person’s mood at the time of the behavioural decision, which combined with their attitude and social factors result in a person’s intended behaviour. The actual behaviour, though, may be different from this intended behaviour as a result of the influence of habits.

Habit is routine behaviour undertaken without any conscious thought. Habitual behaviour is developed as a result of frequent and consistent choices made to achieve specific goals (Aarts and Fijksterhuis, 2000). Applying this to the context of the commuting mode choice decision, a habit could be always to use the car to travel to work. Originally the choice of travel mode is the result of conscious choice, but over time the consistency and frequency of choosing to use the car increases the strength of this habitual behaviour. Triandis considered there to be a trade-off between intention and habit, so that a strong habitual behaviour can overwhelm any changed intention, for example to cycle to work instead.

The final behaviour undertaken is affected by facilitating conditions. These are personal or external factors which may help or hinder a person from carrying out their intended or habitual behaviour. Triandis’ description of facilitating conditions covers a person’s ‘ability to perform the act’, their ‘level of arousal in regard to the act, the difficulty of the act… possession of the knowledge required to perform the act and environmental factors that increase the probability of the act’. (Triandis, 1977). So for example, the absence of a bus service or lack of knowledge of the availability of a bus service thwarts an intention to travel by bus while the ownership of a bicycle enables the fulfilment of an intention to cycle to work.

Azjen extended his and Fishbein’s Theory of Reasoned Action in 1985 to cover its application to situations where the possibility of a person carrying out a particular behaviour was not entirely dependent upon their intentions, but was affected by personal and /or external constraints. His Theory of Planned Behaviour, shown in figure 5.6 below, incorporated an additional element, 'perceived behavioural control', so that a person's actual behaviour 'depends jointly on motivation (intention) and ability ('behavioural control'), (Azjen 1991). Perceived behavioural control refers 'to people's expectations regarding the degree to which they are capable of performing a given behavior, the extent to which they have the requisite resources and believe they can overcome whatever obstacles they may encounter. (Azjen 2002).

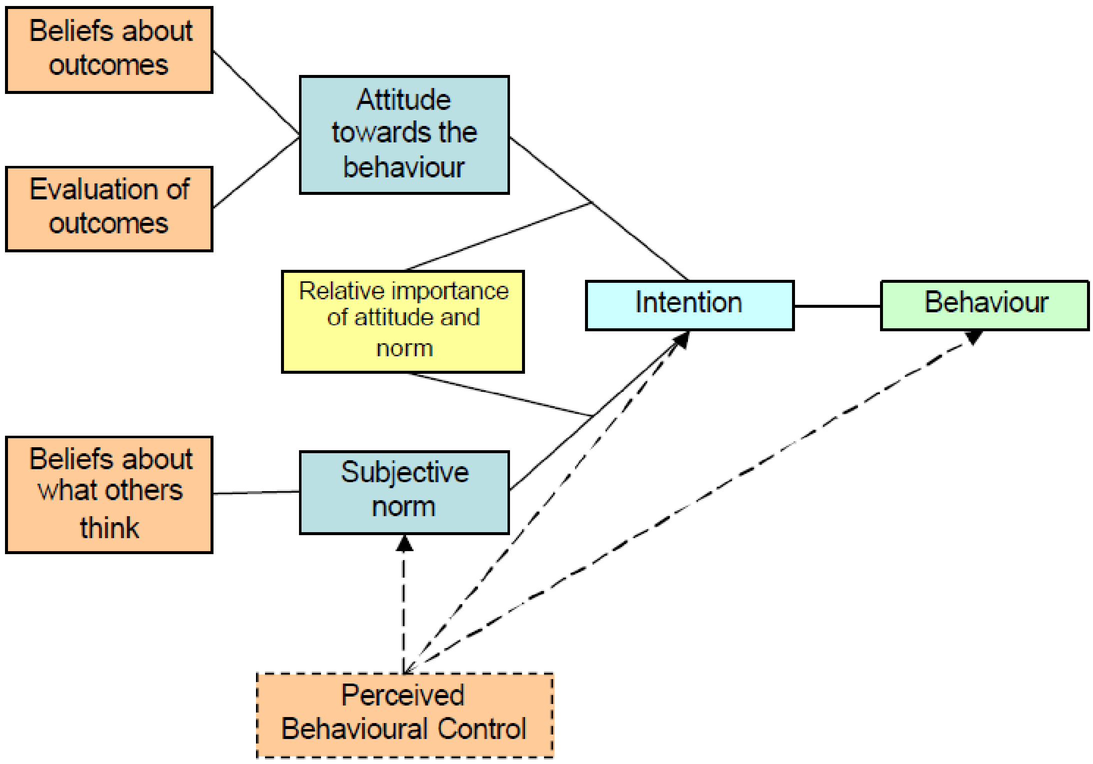


FIGURE 5.6 THEORY OF PLANNED BEHAVIOUR, AZJEN 1985

Source: Jackson, 2005

The Theory of Planned Behaviour has been applied in a wide variety of transport studies. The usual methodology in these studies is to conduct a survey with Likert scale questions to derive a score for each respondent for the TPB variables (attitudes, subjective norms, perceived behavioural control) and their intention to perform the behaviour being studied.

Some studies also survey whether the behaviour is then performed. Walsh et al (2008) studied the intention to use a mobile phone while driving and concluded that attitude was the most significant predictor of intention to use a mobile phone while driving, followed by perceiving the approval of others towards using a phone while driving (subjective norms). Poulter et al (2008) investigated truck drivers and found that general law abiding driving behaviour in truck drivers was related more to attitudes, subjective norms and intentions than perceived behavioural control, however regarding compliance with UK truck regulations, perceived behavioural control had the largest direct effect.

Parker et al (1992) also found that the most significant factor for predicting behavioural intention to commit road offences was perceived behavioural control when they investigated four driver behaviours; intention to speed, drive whilst drunk, close following and dangerous overtaking (Parker et al 1992). Perceived behavioural control (PBC) was also found to be the main predictor for pedestrians crossing roads in dangerous places (Evans et al 1998).

Many studies in the transport area have extended the TPB to include a measure for past behaviour. Chorlton et al (2012) found that past behaviour was the 'most consistent, strong and significant predictor of intention by motorcyclists to ride above speed limits and at inappropriate speeds. Callaghan et al (2006) studied intended and actual cycle helmet use amongst 293 Australian teenagers. They found that past behaviour was the single most important predictor of intended and actual helmet use, as had also been found in similar studies by Quine et al (1998). Brijs (2011) studied the use of seat belts amongst a group of Belgium teenagers and concluded that the main determinants of actual behaviour were 'past behaviour, behavioural intentions and perceived behavioural control'.

Conner and Armitage (1998) reviewed applications of TPB across multiple disciplines and concluded that 'there do appear to be good empirical and theoretical reasons to incorporate habit measures (frequency of past behavior) as predictors of behavior in the TPB alongside intentions and PBC, at least for frequently performed behaviors'. The impact of the frequency of the behaviour on the strength of a habit was also examined in a meta-analysis by Oullette and Woods (1998) of studies into habits in everyday life. This showed that the direct influence of past behaviours on future behaviour was greater for behaviours that were executed often. Gardner et al (2009) used TPB in two studies with car commuters and cycle commuters in the Netherlands. They found that in both groups 'intention predicted behaviour when habit was weak, but where self-reported habit was strong, behaviour was dominated solely by habit and not by intention'.

The efficacy of three alternative behavioural models to predict travel mode was investigated by Bamberg and Schmidt (2003). They studied the use of car for travel to university by 321 students in Germany who completed both an initial survey, with questions designed around the behavioural theories, and a follow up survey three weeks later on their actual travel choices. Using structural equation modelling, they confirmed the position in Triandis' model that 'intention marks the end of the conscious choice process' and that, in the context of mode choice for a regular travel to university, behaviour is affected by both conscious choice and habitual car use. They found that the addition of the 'car use habit' variable from the Triandis model significantly increased the predictive power of Ajzen's Theory of Planned Behaviour model'. Based on this finding, Triandis' Theory of Interpersonal Behaviour, which includes habit, was used as the basis for the conceptual model in this study.

5.3.3 Design decisions

A key design philosophy of agent based modelling is to 'Keep It Simple Stupid' (Gilbert, 2005): starting with a simple model and then adding in complexity. The intention in this project is to build a model that captures the emerging patterns observed in the real world by first modelling the simple processes and then adding further detail and processes incrementally, informed by the understanding created by the original model.

This led to the following design decisions:

- to model each person's mode choice decision over time and observe the resulting mode shares *[Reason: packages of 'Smarter choices' measures are targeted at shifting people away from car to public transport, walking and cycling so a model designed to test 'smarter choices' policy should include the modelling of mode choice decisions].*
- to model the mode chosen as 100% intentional (deliberative) or 100% habitual rather than varying the degree of trade-off between intention and habit *[Reason: this simplifying assumption facilitates the inclusion of habit in the model, with the intention of testing how the inclusion of habitual behaviour could affect model results]*
- a person's choice, whether determined by intention or habit, is affected by facilitating conditions such as whether a person has a car licence or there is a bus route available. *[Reason: the choice set of modes for each person should only represent the modes actually available to them and possible for them to use]*
- to base the initial mode chosen by a person on their intention *[Reason: without this the mode used by people will change more frequently in early time periods as people move away from their randomly assigned initial mode]*
- to base subsequent mode choices for an individual on habit until a 'trigger event' occurs *[Reason: commuting is a frequent and regular activity and so is likely to carried out with minimal thought, as a script-based activity (Garling et al., 2001)]*
- for the occurrence of a 'trigger event' to switch the dominance back to intention over habit for the next mode choice decision *[Reason: the trigger event is assumed to cause a person to actively consider their choice of travel mode]*
- to base the estimation of a person's intended choice on the option that maximises their utility. Each person's preference for time and cost will determine their evaluation of the outcome and hence attitude towards each possible mode. The influence of social factors and emotion will be reflected in the 'utility' value held by them for each mode*. [Reason: behavioural economics seeks to add psychological realism into*

> *standard economic models. The standard rational economic model assumes utility maximising choice behaviour and is used here to predict the mode chosen when an intentional choice is made].*

Trigger events can come from a variety of sources. Potential triggers include key life events, a critical incident (van der Waerden *et al.*, 2003), a change in travel times and costs above a person's threshold, or an act of persuasion from a marketing or information campaign that causes a person to re-assess their travel habits. Van der Waerden reports on a survey of 173 respondents who were asked about life events that influenced their mode choice behaviour. The most common events reported were a house move, entering the workforce, a change in work situation, passing the driving test and getting a car. The latter two events could be considered in this conceptual framework as changes in facilitating constraints, leaving house move and job changes as two key events which trigger a re-assessment of mode choice.

The aim of the model is to investigate the potential of agent based modelling for modelling 'smarter choices' so a visit by a personalised travel planning adviser, representing the effects of a marketing campaign, is included in the model as a trigger event. Other triggers, such as journey costs for a particular mode rising above a person's threshold could be added to the model in future but would require the collection of suitable data.

Returning to the language of Kung, in this agent based model of commuter mode choice, the entities in the model are the individuals. The processes are the 'mechanism' of mode choice based on Triandis' Theory of Interpersonal behaviour and the occurrence of the trigger events which switches the mode choice from habitual to intentional determination.

5.4 Selection of a computing platform

5.4.1 Hardware

The choice of hardware platforms was between using a personal computer or using a main frame computer. For larger simulations, model runs can be carried out using a local main frame computer or by using cloud computing where modellers make use of third party owned processors, store input/output data remotely and access this data using the internet. The widely used Cube transport modelling software now offers a Cloud facility that can be used on the very powerful computers owned by Amazon. By cluster computing, which uses many computers at the same time, the run time for an agent based model for example, can be reduced from 175.13 minutes using 32 processors to 7.17 minutes using 542 processors. (Brown, 2013) The hosting of the model on the third party servers, combined with software optimised to use multiple processors means that run times are reduced considerably and model outputs can be examined by users and model developers at multiple locations. Costs are also lowered as users rent time on the most powerful machines only when model runs are needed and can acquire lower specification machines to use as terminals when preparing input data and reports.

The rapid increases in the speed of computer chips now available in personal computers, 64 bit systems which allow programs access to larger chunks of memory at one time and the simultaneous use of several processors means that it is feasible to run some quite large simulations on personal computers. There is also a wide selection of software available for personal computers to run computer simulations and analyse the results. A personal computer with an Intel i5 processor was used for this project as it was capable of running the model in less than 10 minutes and was the lowest cost option.

5.4.2 Software

'The ideal is a system that requires a minimum of learning, is completely flexible in the models that it will support, and runs efficiently on any hardware', (Gilbert and Bankes, 2002). The decision regarding the choice of computing software for this research was made in two stages, first by comparing the features of a selection of software options and secondly by writing a test model using the shortlisted packages.

Gilbert and Bankes (2002) identify three approaches to coding an agent based model. The first is to write the model directly with a high level programming language such as C++, Java or Python. These object oriented programming languages are ideally suited to coding agent based models as they deal naturally with objects that have attributes and methods that work upon them. A high level language, Netlogo, has been written specifically for coding agent based modelling.

The second approach is to use libraries of routines, or computer code, written to implement common tasks required in agent based models to supplement customised code written in a high level programming language by the user. These ready prepared libraries include code to handle common tasks such as data input, processing tasks, the saving of output results and the user interface. Examples of such libraries for agent based models include Swarm, Mason, Repast and AnyLogic.

The third approach is to use software specifically designed to build agent based models which require the user only to use a visual interface to manipulate symbols and then the software builds the corresponding model. Examples of such programmes include Starlogo and Agentsheets. However the trade-off for the ease of use of such software is the limited range of functions available which can constrict the features of the model.

The selection of the software package is important as much time and effort will be invested in learning the software. The capabilities of the software are also important, as well as the steepness of the learning curve for the software and the ease of building models once climbed, as the lack of capabilities in a particular package could limit the future development and enhancement of a model. In the first stage the software options for this project were considered against the following set of general criteria suggested by Castle and Crooks (2012):

- ‘ease of developing the model/using the system;
- size of the community using the system;
- availability of help or support (most probably from the user community);
- size of the community familiar with the programming language in which the system is implemented (if a programming language is necessary to implement the model);
- is the system still maintained and/or updated;
- availability of demonstration or template models; technical and how-to documentation, etc.’

and criteria regarding the functionality of the software:

- ‘the number of agents that can be modelled;
- degree of interaction between agents;
- ability to represent multiple organisational/hierarchical levels of agents;
- variety of model environments available (network, raster, and vector);
- possible topological relationship between agents;
- management of spatial relationships between agents, and agents with their environment; mechanisms for scheduling and sequencing events, etc.’

The first stage of the selection process was to read published reviews of agent based modelling software. Although these reviews each had their own criteria for grading the software, many of their considerations were similar to those described above. In addition, the recommendations of the reviewers would assist in the selection of three packages for more detailed investigation. The earliest review of software for agent based modelling was undertaken by Serenko *et al.* (2002). He found 20 packages and evaluated them in terms of their suitability as a tool for teaching agent based modelling. As well as considering their functionality, Serenko asked 87 teachers for their user perspective on the software. This study suggests that the search for the software tool for this project should consider user experience as well as software features.

Tobias and Hofmann (2004) limited their review to free software that was based on the Java programming language. They examined four ABM toolkits Repast, Swarm, Quicksilver and VSEit and recommended Repast as 'the clear winner'.

Railsback, Lytinen, and Jackson (2006) examined four packages, NetLogo, Mason, Repast, and Swarm, by building the same example models (the Stupid Model suite) in each package. They recommended Netlogo for learning about ABM and as a tool for building prototype models. They recommended Repast as the most complete package, but noted that it was not very accessible to beginners in programming or agent based modelling.

Castle and Crooks (2006) considered eight toolkits, Swarm, Mason, Repast, StarLogo, NetLogo, Obeus, AgentSheets, and AnyLogic. They were concerned primarily with their geospatial capabilities but also considered their general functionality. They also considered the age of the software, the level of programming experience required and the availability of demonstration models. They did not recommend any particular package, but have used Repast in their own work.

Nikolai *et al.* (2009) examined the documentation for over thirty toolkits and compared them according to five characteristics: the language required to program a model and to run a simulation, the operating system required to run the toolkit, the type of licence that governs the toolkit, the primary domain for which the toolkit is intended, and the types of support available to the user. They did not use any of the software packages but provided a useful long list of software to start the search for a suitable software platform. Allan (2009) also reviews over thirty packages, but not always the same packages as Nikolai. This supplements the Nikola review, as, although it is not as consistent in its comparison of the features of each package, the author provides a commentary on their main distinguishing features and for some packages his experience from using them, although this was sometimes limited to running supplied demonstration models. Neither of these reviews made any specific recommendations on preferred software packages.

After considering these reviews, noting the software used in the papers on agent based modelling covered in the literature review, and comparing the features of the software packages against the Castle and Crooks criteria, the short listed packages for this study were Netlogo, Repast and AnyLogic.

For the second stage of the selection process, the ease of developing a model and the quality of the documentation were assessed by attempting to build a sample model in each of these three packages. This experience provided a fuller understanding of the capabilities of the software and the ease of using them. This approach to evaluating software was used by Railsback *et al.* (2006) who tested the software packages in turn by building the same model, named by them as ‘Stupid Model’.

In this assessment, the software packages were tested by building an implementation of Schnelling’s segregation model (1978). Coding this simple model covered the main tasks involved in coding an ABM: building a user interface, creating agents, implementing a behaviour rule at each time step and recording the results of a simulation run. Implementations of the Schnelling model are available in each of these agent based modelling software packages so after attempting to build the model, a solution by more experienced programmers was available for inspection.

Netlogo is free and was found to be the easiest package to learn. The manual was clearly written, the tutorials were followed successfully and the Schnelling model built. The code for the model on the internet was then examined which pointed to other more efficient ways of coding the model. However the experience showed that the Netlogo manual was incomplete and often lacked information on how to code even basic tasks. The user forum indicated that the dynamic scheduling of events was not well developed at the time the review was undertaken, and was computationally intensive. Although there is an online user forum there is no dedicated software support facility. The graphics output is rather rudimentary and is based on displaying agents in rows and columns in a grid. Interrogation of particular agents is also rather basic and cumbersome.

Repast is also free but was found to be very inaccessible even in the most recent version, Repast Symphony, which has a graphical interface as well as a command line. This experience accorded with the reviewers' opinion that the package has a very steep learning curve and is more suited to experienced Java programmers.

Anylogic has a very comprehensive and detailed manual. There are training materials available and on-line videos, including many from the MSc course in Computer Science taught by Nataniel Osgood at the University of Saskatchewan, Canada, and M.I.T. Boston There is support from the developers and many example models are available. Anylogic also produces high quality output in the form of graphs, summary statistics and information written to text files and databases. It is commercial software but academic licences and a free personal edition are available. The software is very comprehensive and is able to deliver all the required functionality for this project's model. It can also combine agent based and systems dynamics functions in the same model which would make it possible to model some aspects of a system at an aggregate or macro level and others at a disaggregate or individual level.

Following the software review exercise, AnyLogic was selected as the software platform for this study. Training in the software was received and personal experience gained by building small models to test the functions that would be needed in the model built for this study. The software has a graphical interface that links with pre-written libraries of code. These are supplemented with Java code written by the user to meet specific needs of the model being built. The availability of many example programs and extensive documentation was invaluable in learning how to build an agent based model and the first version of the model produced for this study was written in AnyLogic.

During the course of verifying the model it became apparent that the timing of events in the model controlled by closed libraries in the software was not occurring when expected. Castle and Crooks (2012) warned that a risk of using libraries of code is that 'since access to their source code is prohibited, a model developed with proprietary software is essentially black box. A modeller will therefore, to some extent, be left unsure about the internal validity of a model constructed with a proprietary system. This situation is compounded when the output of a model is emergent or unexpected'.

As a result of this experience with the timing of events in the first model and taking heed of the above warning, the second version of the model for this project was written using Python. This programming language was chosen as it is an object oriented language but is also used as the scripting language in several transport modelling software packages. This means that an agent based model of mode choice written in Python could be called as part of a larger transport model built in an industry standard software package such as VISUM. This has great benefits for the eventual use of ABM in multi-modal transport models as they can be built in the same application that can handle other requisites such as routeing algorithms needed to determine the cost of travel on the network.

The next stage in the modelling process, after the selection of computing hardware and software, was the gathering of input data for the model and the preparation of this data into a suitable form for use in the model. This is reported in the next chapter.

6 Data preparation and model overview

6.1 Introduction

The emphasis of this study is on the assessment of the potential for agent based modelling to provide a framework for modelling mode choice which is sensitive to a wider range of policy interventions than the conventional four stage transport models currently used in the UK. The decision was taken to search for a secondary data set rather than collect primary data for the study due to the costs of collecting a sufficiently large data set to use in building the model. The experience of building an ABM would inform the type of data which would need to be collected in any future work to develop the model further from a Mediator to a Predictor model and so help ensure that any future specifically commissioned surveys could be designed to collect the type of data required by an agent based mode choice model.

This chapter describes the search for data required for the model. It then describes the selected dataset and the work undertaken on that data to transform it for use in the model. There was a need for some information on the frequency of house moves and changes in employment for people living in the UK which was not available from the selected dataset. This chapter describes the use of other data sources to provide this data.

6.2 Empirically grounded agent based models

The purpose of many of the first generation of agent based models was to generate and test theoretical propositions in highly stylised artificial worlds with no empirical data, such as the sugar scape models developed by Epstein and Axtell in their use of ABM as 'a generative approach to social science' (1996) and Axelrod's game theory models reported in 'The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration' (1997) Agent based modelling was viewed as a new branch of the experimental sciences; the computer model was considered to be a laboratory where it was possible to compensate for the unavoidable weakness of empirical and experimental knowledge in social science. (Prietula *et al.*, 1998).

There is now a growing recognition of the need to increase the role of empirical data in agent based models and to produce ‘empirically grounded models’ (Rounsevell, 2012). Hassan *et al.* (2008) point to the potential role of data in ABMs to inform the design of the model, set the initial conditions of the model and to validate the model output by comparing the model output with real world observed data.

Boeri and Squazzoni (2005) promote the ‘fruitfulness’ of bringing empirical data into agent based modelling practice and differentiate ABMs into ‘case-based models’, ‘typifications’ and ‘theoretical abstractions’. The difference lies in the target of the model and the intended role of data in the model which has a strong impact on the strategy required to locate suitable empirical data. A ‘case-based model’ relates to a particular time-space domain and in an ABM seeks to capture the micro generative mechanisms that are operating in a very particular domain to produce the observed real world macro-level features. Such a model may be of value, particularly if the model purpose is prediction, as the output provides a forecast of the results of policy changes in a particular location. It is likely to demand high amounts of very specific data on the entities and processes in the target world of the model.

‘Typifications’ are more theoretical models built to investigate processes that may operate in a more general context. These models involve a greater level of abstraction from the real world and aim to produce models that create findings that are relevant to more real world contexts than a case-based model. These models may still demand large amounts of empirical data in order to determine the classification rules and processes and to define the required attributes of the model entities.

‘Theoretical models’ have little direct reliance on empirical data but the initial formulation of the theories used in them will have been informed by at very least general observations of the world.

In this research the role of data was determined by the purpose set for the model. The data to be used in the description of entities and processes, could come from a variety of methods and sources. Robinson *et al.* (2007) reports on an international workshop in which participants shared their experience of empirically grounding ABMs. Five methods of obtaining data for use in building a model at the micro level and validating the model at the macro level were identified:

- Sample surveys
- Participant observation
- Field and laboratory experiments
- Companion modelling
- Remotely sensed spatial data

Each approach has its own strengths and weaknesses; some produce quantitative data and others qualitative and both types of data are useful in the modelling process.

The data used in the model developed for this study could come from either primary data collection or existing, secondary, sources. The benefits of primary data collection are that the data can be collected specifically to meet the requirements of the theoretical constructs captured in the model, and be tailored to the needs of supplying the model inputs, the internal mechanisms in the model such as the parameters of any equations or probability distributions, the timing of events in the model and the needs of the validation task. The main disadvantages are the time, cost of collecting and processing, and the difficulty of obtaining the desired data (Hox and Boeije, 2005). The time and budget constraints of this research study led to the decision to search for a suitable already existing data set to use in the proof of concept model.

6.2.1 Finding a secondary dataset

The difficulties associated with secondary data include the task of identifying and locating suitable data sets and problems encountered in retrieving the data. Compromises may have to be made to the model design to allow for the actual items of data available and limited knowledge of the quality of the data may be available, particularly in terms of how it was collected.

The design specification for the data set was determined by the needs of the conceptual model. The data needed to contain details of individual preferences on elements of a journey such as time and cost as well as details of the agents' age, gender and mobility constraints, so the the search for a dataset was aimed at identifying surveys which contained a stated preference component. It was discovered that few stated preference surveys have been undertaken in the UK in recent years due to budget constraints with modellers using the standard values of time provided for appraisal purposes in the DfT WebTAG guidance and then adjusting mode constants and lambda values in the logit models during the calibration process.

A few recent surveys were identified which included a mode choice stated preference exercise such as the data collected for the Sheffield tram extensions and the Leeds New Generation Transit system. Neither of these data sets were ideal though as they did not include car as one of the possible modes but concentrated only on the choice between public transport modes. An internet search was carried out for possible overseas data sets. However it was difficult to make contact with the owners of the data and when one possible data set was identified the owners wanted to make a considerable charge for use of the data.

The DfT published a report in 2010 with the interim results of a quantitative survey on Climate Change and Transport Options. The report showed that the survey had contained a stated preference exercise for the journey to work that included the options of travelling by car, bus rail or cycling. It had also asked detailed questions about the respondent's actual journeys and their attitudes towards different modes of travel. The report indicated that the study documentation and the individual but anonymised survey responses were available free of charge on the DfT website.

However the stated preference component of the survey was not available on the internet and, when contacted, the DfT discovered they did not have a copy of this part of the survey. The stated preference responses were held by the market research company that had undertaken the surveys for the DfT. I was given permission to contact the survey company and request a copy of the survey data. After tracing the relevant individuals at the survey company it was discovered that the files had been archived in binary format and the company no longer had a valid license for the software that had been used to capture the data. Due to staff redundancies they were also no longer aware of which software package had been used.

The survey company restored the data from their archive and forwarded the files for use in this study. The next step was to undertake an internet search using the suffix in the file names to try and identify the software which had produced the binary files. A number of possible packages were identified. Where possible a demonstration version of the software was downloaded to investigate whether it could read the binary files, otherwise a request was sent to the relevant software help desk. In this way, the files were found to have been produced using Sawtooth software. This is a North American commercial software product used to design, administer and analysis surveys combining quantitative and stated preference elements.

The company provided a complimentary academic license for their software. The binary data files were then re-analysed using the Sawtooth software and it was possible to reproduce the utilities originally calculated by the DfT's survey company, TNS-BMRB, from the same data as reported in the interim report of December 2010. Further details of the surveys used to collect this data set are provided in Appendix A2.

6.2.2 DfT dataset on climate change and transport options

The section of the DfT survey used for this model came from the 626 respondents who completed the stated preference exercise regarding mode choice for a 5 mile journey to work by either car, bus, rail or cycling. The age and gender profile of the respondents is shown in Table 6.1 below. All the respondents were of working age and in employment.

Age	Male		Female		Total
	Number	%	Number	%	
16-20	14	4%	8	3%	22
21-29	47	15%	51	16%	98
30-39	84	27%	67	22%	151
40-49	71	23%	83	27%	154
50-59	62	20%	72	23%	134
60-69	36	11%	27	9%	63
Total	315	100%	311	100%	626

TABLE 6.1 AGE AND GENDER PROFILE OF RESPONDENTS TO THE DFT COMMUTING SURVEY

Most of the respondents (78%) live in urban areas as shown in Table 6.2 below. Less than 12% of the people lived in a rural area. There was a car or van available in 88% of the households and 83% of respondents held a full car driving licence.

Location	Number	%
Urban - London	69	11.0
Urban - Other	421	67.3
Town and Fringe	62	9.9
Village, Hamlet and Isolated Dwellings	74	11.8
Total	626	100.0

TABLE 6.2 HOME LOCATION OF RESPONDENTS

The regular mode of travel to work for the respondents asked is shown in Table 6.3 below. The most common mode is car used by 63%, followed by walk (10%), bus (8%), rail (5%) and cycling (4%)

.

Mode	Number	%
Car/van as driver	363	63.2
Walk	59	10.3
Bus	45	7.8
Railway train	31	5.4
Car/van as passenger	29	5.1
Bicycle	25	4.4
Motorbike/moped/scooter	11	1.9
Tube/metro/light rail/tram	11	1.9
Total	574	100

TABLE 6.3 REGULAR MODE OF TRAVEL TO WORK

6.2.3 Evidence for habits, triggers and constraints

The DfT survey was analysed for all the 3,923 respondents to the main part of the survey in order to identify the main factors that influenced people's choice of mode and the events that prompted them to change mode. The evidence for the role of habits in the DfT data came from answers to the question given to everyone who said they used the car at least once or twice a week, 'When I have to choose how I will travel, choosing the car is something...'. The responses are shown in Table 6.4 below for all respondents, those respondents who were in employment and those respondents who participated in the choice modelling exercise on the regular journey to work.

Choosing to travel by car is something…	All respondents		All employed		All in choice exercise	
a) I do frequently	Freq.	*Percent*	Freq.	*Percent*	Freq.	*Percent*
Yes	2,656	*79.33*	1,586	*84.99*	487	*86.35*
No	683	*20.40*	276	*14.79*	76	*13.48*
Don't know	9	*0.27*	4	*0.21*	1	*0.18*
b) I do automatically	Freq.	*Percent*	Freq.	*Percent*	Freq.	*Percent*
Yes	2,285	*68.25*	1,362	*72.99*	409	*72.52*
No	1,054	*31.48*	498	*26.69*	154	*27.30*
Don't know	9	*0.27*	6	*0.32*	1	*0.18*
c) would require effort not to do	Freq.	*Percent*	Freq.	*Percent*	Freq.	*Percent*
Yes	2,110	*63.02*	1,249	*66.93*	380	*67.38*
No	1,203	*35.93*	601	*32.21*	180	*31.91*
Don't know	35	*1.05*	16	*0.86*	4	*0.71*
d) belongs to my (daily, weekly) routine	Freq.	*Percent*	Freq.	*Percent*	Freq.	*Percent*
Yes	2,688	*80.29*	1,583	*84.83*	474	*84.04*
No	652	*19.47*	278	*14.90*	89	*15.78*
Don't know	8	*0.24*	5	*0.27*	1	*0.18*
e) that's typically me	Freq.	*Percent*	Freq.	*Percent*	Freq.	*Percent*
Yes	2,291	*68.43*	1,356	*72.67*	406	*71.99*
No	1,019	*30.44*	491	*26.00*	153	*27.13*
Don't know	38	*1.14*	19	*1.02*	5	*0.89*
f) that I've been doing for a long time	Freq.	*Percent*	Freq.	*Percent*	Freq.	*Percent*
Yes	2,721	*81.27*	1,576	*84.46*	476	*84.40*
No	618	*18.46*	286	*15.33*	87	*15.43*
Don't know	9	*0.27*	4	*0.21*	1	*0.18*
Total	3,348	*100.00*	1,866	*100.00*	564	*100.00*

TABLE 6.4 INFLUENCE OF HABIT IN CHOICE OF TRAVEL MODE

Although regular use of a car does not necessarily mean that it is habitual, the responses shown above for people in employment indicates that for many using the car is something they do frequently (85%), automatically (73%) and belongs to their routine (85%).

The evidence for the triggers which cause people to re-consider their travel options came from the answers to the question, asked of those respondents who had made regular journeys to work in the last six months and had changed the way they usually travelled to work, what factors had prompted that change. The responses are shown in Table 6.5 below.

Why did you change the way you travelled to work?	Freq.	*Percent*
New method quicker / more convenient	38	*24.02%*
New job or change in place of work	35	*22.29%*
Moved house	25	*16.16%*
New method cheaper / free	20	*12.51%*
I bought a car	18	*11.53%*
Change in season/ weather	10	*6.46%*
Change in family circumstances (e.g. had a baby/got divorced/child left school/etc)	9	*5.67%*
Health reasons	8	*4.92%*
Change in parking arrangements	2	*1.57%*
Public transport overcrowded, unreliable	2	*1.38%*
New method more reliable	2	*1.09%*
I wanted to reduce my CO_2 emissions	2	*1.07%*
I bought a bicycle	1	*0.79%*
Other	9	*6.01%*
Total reasons provided	181	
Total number respondents	157	*100%*

TABLE 6.5 REASONS PROVIDED FOR CHANGING REGULAR MODE USED FOR COMMUTING

Of the 157 respondents who reported that they had changed the mode they regularly used to travel to work, 22% said it was a result of changing their job or place of work and 16% said it was a result of moving house. These two events may have made other people reconsider their commute to work even if they then continued to use the same mode as previously and so would not have been included in these numbers. 24% of respondents said they had changed mode because the new mode was quicker or more convenient and 16% because the new mode was cheaper or free. This explains why they changed mode but may not have been what made them re-consider their current travel arrangements in the first place.

Based on the evidence from the DfT survey, which confirmed the findings of other studies (Verplanken et al, 1997, Van der Waerden, 2003) the triggers to re-consider their current mode incorporated in the model are changes to home or job locations. A visit from a personal travel adviser who provides information on the availability of public transport in their area is also included in the model as it is one of the policy measures which the model is intended to evaluate.

Not all of the four modes included in the model, bus, car, rail and cycle, is a realistic option for each person. Each possible mode needs to be assessed against a set of personal constraints and if that mode is not a feasible option for that person, it is removed from the set of choices available. The constraints incorporated in the basic mode are either due to the characteristics of the individual i.e. they are personal constraints or are due to the characteristics of the transport system i.e. they are constraints arising from the general environment.

Personal constraints cover aspects such as having mobility issues or other disabilities which prevent the use of a particular mode, being required to have a car available for use during the working day and needing to be able to carry work equipment and papers. General constraints cover issues such as the lack of a bus service within a reasonable distance of their house or destination or the lack of parking spaces near the workplace.

Tables 6.6 and 6.7 cover disability issues that restrict the respondent’s ability to use a car, bus, cycle or walk for a journey. This shows that 6% of respondents had a disability or other health issue which meant that they could not use buses. For 16% of respondents cycling was difficult or impossible.

Do you have any disability or other long standing health that makes it difficult for you to do any of the following… [2923 respondents]		
	Freq.	*Percent*
go out on foot: Yes	357	*9.10*
use local buses: Yes	236	*6.02*
Get in or out of a car: Yes	185	*4.72*

TABLE 6.6 HEALTH CONSTRAINTS ON TRAVELLING ON FOOT, BY BUS AND BY CAR

Do you have any disability or other long standing health problem that makes it /would make it difficult or impossible for you to ride a bicycle [3923 respondents]		
	Freq.	*Percent*
Yes – impossible	382	*9.73*
Yes – difficult	238	*6.06*
No	3293	*83.95*
Don't know	10	*0.26*

TABLE 6.7 DIFFICULTIES IN RIDING A BICYCLE

Respondents were asked why they used the car to get to work/school/ college (Table 6.8) and why they did not use public transport (Table 6.9) or cycling (Table 6.10) for their regular journey to work/school/college.

For car drivers or passenger to work/ school/ college What are the reasons for you usually going by car/van? [1331 respondents]		
I can travel when I want to travel	Freq.	*Percent*
No	1060	*79.65%*
Yes	271	*20.35%*
It is quick / quickest way/ other ways take too long		
No	732	*54.97%*
Yes	599	*45.03%*
It is reliable / more reliable than other modes		
No	1167	*87.70%*
Yes	164	*12.30%*
It is cheap / cheapest way		
No	1192	*89.53%*
Yes	139	*10.47%*
It is convenient / most convenient		
No	744	*55.89%*
Yes	587	*44.11%*
it is comfortable / most comfortable		
No	1235	*92.76%*
Yes	96	*7.24%*
I cannot get there any other way		
No	1047	*78.67%*
Yes	284	*21.33%*
I enjoy driving		
No	1298	*97.55%*
Yes	33	*2.45%*
I have to take things (e.g. tools, laptop, luggage etc.) and cannot		
No	1150	*86.40%*
Yes	181	*13.60%*
I usually take my partner with me		
No	1319	*99.08%*
Yes	12	*0.92%*
I usually take my children with me		
No	1261	*94.70%*
Yes	70	*5.30%*

I usually take someone else with me		
No	1302	*97.80%*
Yes	29	*2.20%*
I need my car for work		
No	1175	*88.29%*
Yes	156	*11.71%*
I use my car to make other trips while I'm out		
No	1265	*95.01%*
Yes	66	*4.99%*
It gives me flexibility		
No	1193	*89.65%*
Yes	138	*10.35%*
The weather		
No	1281	*96.22%*
Yes	50	*3.78%*
It is safer		
No	1313	*98.66%*
Yes	18	*1.34%*
Buses do not run at suitable times		
No	1,323	*99.38%*
Yes	8	*0.62%*

TABLE 6.8 REASONS FOR TRAVELLING TO WORK/SCHOOL/COLLEGE BY CAR

The reasons for using a car were grouped into four categories:

1. Modal characteristics
    - Can travel whenever I want
    - Convenient
    - Comfort
    - Enjoyment of driving
    - Flexibility
    - Protected from the weather
    - Safe
2. Journey characteristics
    - Time
    - Reliability
    - Cost

3. Constraints
    - No other way to make the trip
    - Have things to carry
    - Need car for work
    - Buses not run at suitable times
4. Other
    - Take partner
    - Take children
    - Take someone else
    - Make other trips while out

Reviewing these reasons for using the car, the modal characteristics would be captured in the mode preference in the stated preference exercise. The journey characteristics would be captured in the description of the journey options, with the exception of reliability which was not considered in the DfT stated preference exercise. As more (over 12%) respondents stated that reliability was a reason for choosing car than cost (10%), then ideally it should have been included in the stated preference survey. It is more straightforward to include cost than reliability but other studies have included it, particularly for public transport trips, such as the Leeds New Generation Transit surveys.

The 'other' reasons are not strictly constraints which require using the car as it may be possible for all the occupants of the car to make the trip using an alternative mode. The actual constraints identified were the lack of alternatives (21%), need to take things to work (14%), need car for work (12%) and buses not run at suitable times (1%). In the basic model, the need to carry things and use the car for work are taken as constraints that mean the person has to use the car.

The issue of the lack of alternatives to the car was incorporated by introducing a feedback loop, whereby if the number of bus passengers in a time period fell below that needed to run a commercial service, then the number of bus routes are reduced and all the people that might have used the service do not have the choice of a bus. In this way the model itself produces the constraint of the lack of a bus service. When the model starts, everyone has the option of a bus service but the model itself introduces this constraint if the overall number of people choosing to use the bus is too low to maintain the service. The number of people who said that buses did not run at suitable times was low, less than 1%, so this was not included in the model.

The constraints in Table 6.9 also highlight the reasons for not using bus, with the lack of a bus service being the main reason provided for not using the bus and so this constraint was added into the model. Time and cost are also provided as reasons for not using the bus. Again, the need to carry items was mentioned, 10%, but the need to use the car for work was not cited as often.

Respondents who had a journey of less than 10 miles to work/ school/ college were asked why they did not cycle. The main reason provided was distance (30%) followed by safety fears (22%), the weather (17%), not owning a bicycle (17%) and the need to take things (14%). As the model is based on a fixed journey length of 5 miles, the constraint of distance was not included, neither was not owning a bicycle as this is a constraint that can be addressed. The constraint of the weather and having to carry things were included in the model.

What are the reasons why you don’t take the bus to get to work/ school/ college? [1030 respondents]		
	Freq.	*Percent*
Buses do not run when / where I want to travel	389	37.81%
Generally not convenient by bus / easier or more convenient by car	257	24.93%
Bus journey is too slow / infrequent	249	24.16%
I would need to change my bus / no direct route	176	17.10%
Buses are expensive / more expensive / do not offer good value for money/ It’s cheaper by car	117	11.31%
Bus stop is not near to destination	107	10.42%
I have to take things (e.g. tools, laptop, luggage etc) and cannot carry it	101	9.79%
Buses are not reliable and punctual	88	8.54%
Bus stop is not near home	62	6.07%
Can never be sure what time the bus will arrive/how long it will take	62	6.04%
Other reasons	57	5.53%
Buses are uncomfortable / poor condition / not clean / overcrowded /too cold or hot	40	3.87%
No particular reason	31	3.00%
I don’t know what bus services are available	27	2.67%
I do not feel safe on the bus / at bus stations	24	2.33%
Buses are not accessible/easy to get on	19	1.83%
Need to use car for work	18	1.71%
Need car for school run/ lifts for family or friends	10	1.02%
Don't know	9	0.89%

TABLE 6.9 REASONS FOR NOT USING THE BUS FOR TRAVELLING TO WORK/ SCHOOL/ COLLEGE

What are the reasons why you don’t cycle to [work] or [school/college]? [724 respondents]		
	Freq.	*Percent*
It takes too long to cycle / too far away	218	30.04%
Too much traffic / it’s too dangerous	156	21.59%
Weather	125	17.31%
Don’t own / have access to a bicycle	121	16.65%
I have to take things (e.g. tools, laptop, luggage etc.) and cannot carry it all	98	13.60%
Too old / Not fit enough to cycle	43	5.99%
Too hilly round here	39	5.40%
Not my style	36	5.02%
No particular reason	34	4.76%
Worried about crime/personal safety/being attacked	34	4.69%
Can ride a bicycle but not confidently enough to ride to work	28	3.84%
Other	26	3.59%
Cycle lanes/paths are limited / poor quality/unsafe	24	3.34%
Too dark	21	2.92%
Have to take children with me	20	2.70%
No showers	16	2.21%
Nowhere to park a bicycle securely	11	1.54%
Need to use car throughout day	10	1.31%
Too lazy	9	1.24%
Need car for work	8	1.04%
Don't know	7	0.92%
Can’t ride a bicycle	6	0.90%
Worried about bike being stolen	6	0.89%
Have to take other people with me	5	0.73%
Not practical	5	0.64%
Work at night - not like cycling at night	2	0.31%

TABLE 6.10 REASONS FOR NOT CYCLING TO WORK/ SCHOOL/ COLLEGE

6.2.4 Residential mobility

Two of the key events identified as a trigger for the re-assessment of a person's commuting mode are a change in the regular place of residence (22%) and a change in job (16%), as shown in Table 6.5. Data was therefore sought on residential mobility and the length of job tenure in the UK for use in determining the frequency of the occurrence of these trigger events for inclusion in the model.

The data used on residential mobility in the UK came from the British Household Panel survey which is a longitudinal panel study carried out annually in the UK from 1991 and now subsumed within the Understanding Society survey. The British Household Panel Survey follows the same representative sample of individuals in the UK over a period of years. It is a household-based survey and interviews every member of sampled households over 16 years of age. The panel in the first wave in 1991 consisted of around 5,500 households and 10,300 individuals drawn from 250 areas of Great Britain. Additional samples from Scotland and Wales (1,500 households each) were added in 1999 and a further 2,000 households were added from Northern Ireland in 2001.

The data used in the model is the responses to the question on residential mobility answered in the surveys undertaken for waves *b* to *r* of the BHPS, 1992 – 2009. The variable MOVEST records whether the respondent has moved home in the past year. The original BHPS used a clustered stratified sample with an almost equal probability sampling design aimed at being representative of Great Britain in 1990 south of the Caledonian Canal. When additional respondents were added from Wales, Scotland, and Northern Ireland the sample units from the four countries then had unequal selection probabilities and a new weighting variable, XRWTUK1 was introduced to compensate for this. The weights also allow for non-response bias. The data from all waves from 1992 to 2009 was downloaded and analysed using STATA software. Table 6.11 gives the proportion of people, by age and gender, who changed their place of residence in the year preceding the interview.

Age	Male % moved	Female % moved	Age	Male % moved	Female % moved
15	3.44	4.97	43	7.02	4.75
16	4.58	7.35	44	5.50	5.72
17	6.79	8.93	45	6.84	5.16
18	11.08	16.47	46	4.29	4.22
19	17.37	25.88	47	6.31	4.73
20	21.62	28.30	48	4.21	5.10
21	24.88	27.71	49	4.90	4.92
22	25.57	31.49	50	3.79	4.52
23	25.91	29.51	51	5.41	4.86
24	24.44	28.15	52	4.24	4.20
25	26.00	25.42	53	5.09	4.14
26	23.54	21.01	54	4.68	2.33
27	22.56	19.55	55	4.84	5.06
28	19.74	17.94	56	4.07	4.49
29	21.27	17.71	57	2.92	2.80
30	17.72	15.66	58	4.43	4.78
31	14.72	14.62	59	4.04	4.04
32	13.74	12.48	60	3.20	3.09
33	12.51	11.74	61	3.59	2.53
34	13.96	10.08	62	4.18	4.09
35	11.66	10.64	63	3.68	2.78
36	11.14	9.64	64	3.24	3.40
37	11.74	8.97	65	3.50	3.72
38	8.91	6.96	66	2.49	2.96
39	9.76	7.06	67	2.61	2.57
40	7.45	6.20	68	3.86	2.14
41	6.29	6.77	69	2.35	2.89
42	7.33	6.87	70	3.18	3.40

TABLE 6.11 PROPORTION OF PEOPLE MOVING EACH YEAR, BY AGE AND GENDER

This shows a difference in the levels of residential mobility by age and gender. Women up to the age of 25 women are more likely to move than men, and then men become more mobile than women. Older people are less mobile than younger people. After analysing the data the likelihood of moving house in a year was incorporated in the model with separate probabilities for a person depending on their age and gender. This means that the model needs to know the initial age of each respondent (available from the DfT survey) and then track the age of each person over time as the model runs.

6.2.5 Job tenure

Data on job mobility in the UK labour market is available from the Organisation of Economic Co-operation and Development (OECD) for each year from 2002 to 2011. The OECD publishes annual figures on the length of job tenure derived from analysis of data collected by the UK Labour Force Survey. This cross-sectional survey was used in preference to the BHPS as it has a larger sample size, (over 41,000 respondents each quarter) which makes the results, especially when disaggregated by factors such as age and gender more robust than the BHPS. The data on the length of time a person has been in their current job is summarised by the age and gender of respondents for 2011 in Table 6.12 below.

Age group	15 to 24 years				25 to 54 years				55 and over			
Time in current job	Male		Female		Male		Female		Male		Female	
< 6 months	359063	19.5%	310060	18.2%	629029	5.8%	505104	5.3%	75221	2.9%	49620	2.4%
6 to <12 months	359590	19.5%	407391	23.9%	652799	6.0%	601097	6.3%	78694	3.0%	61501	3.0%
1 to <3 years	501031	27.2%	481839	28.3%	1355544	12.6%	1301297	13.7%	191743	7.4%	141976	6.8%
3 to <5 years	373863	20.3%	342458	20.1%	1705655	15.8%	1592344	16.8%	250912	9.6%	212699	10.2%
5 to <10 years	245633	13.3%	162987	9.6%	2747814	25.5%	2610669	27.5%	521968	20.0%	410897	19.7%
10 years and over	870	0.0%	749	0.0%	3703820	34.3%	2888332	30.4%	1487202	57.1%	1206642	57.9%

TABLE 6.12 LENGTH OF JOB TENURE IN THE UK, 2011

Source: (OECD from UK Labour Force Survey)

Survival analysis was used to derive Kaplan-Meier survival functions from this data. These show the probability of an individual surviving longer than a specified time, e.g. three years without experiencing the event of changing their job. The Kaplan-Meier functions from this data are shown in Figure 6.1 by age and Figure 6.2 by gender. A log-rank test for equality of survivor functions showed a substantial statistically significant difference in Kaplan-Meier survival estimates both by gender and age, where p<0.001. For the purpose of building a Mediator model, the expected length of job tenure for an individual is therefore based on a person's age and gender. In more detailed Predictor models, the estimation of length of job tenure could be based on other significant variables such as occupation and whether the employment is in the public or private sector.

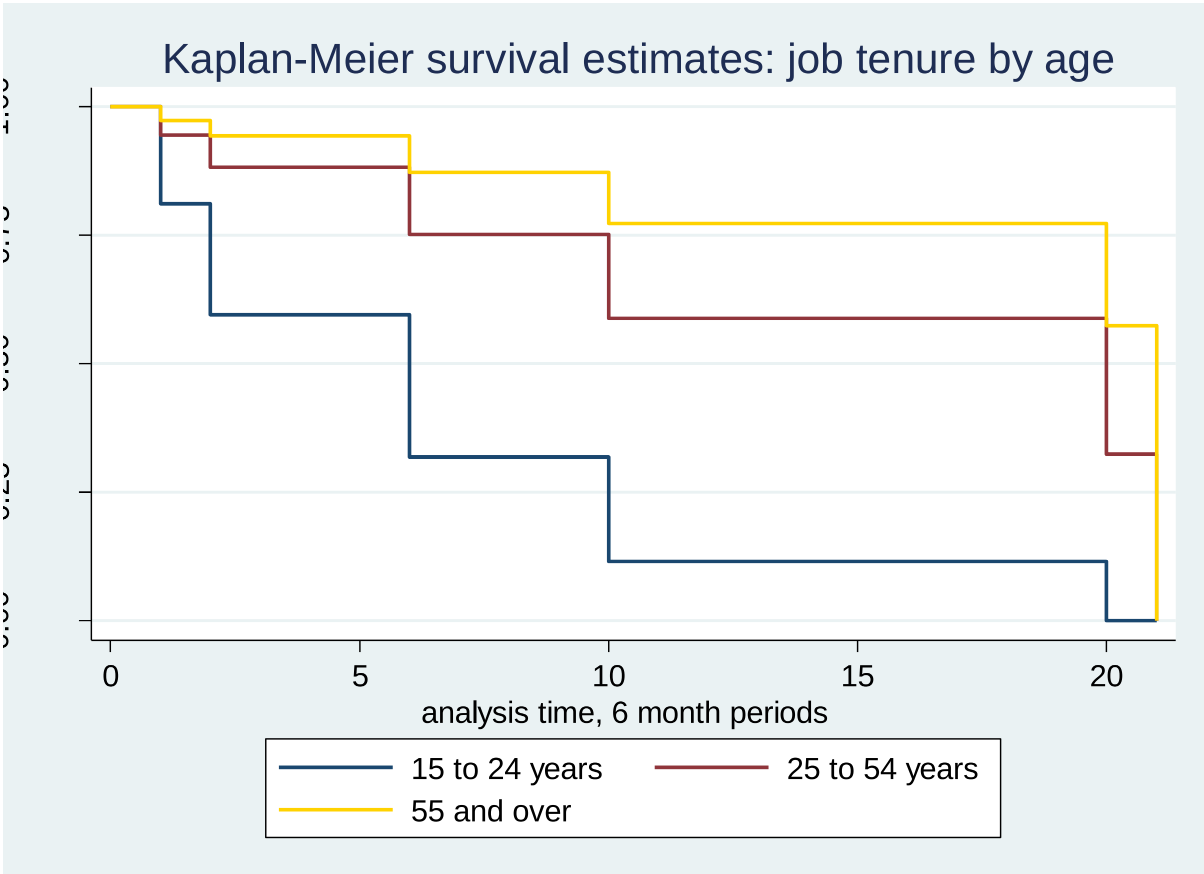


FIGURE 6.1 KAPLAN-MEIER SURVIVAL FUNCTIONS FOR JOB TENURE BY AGE GROUP

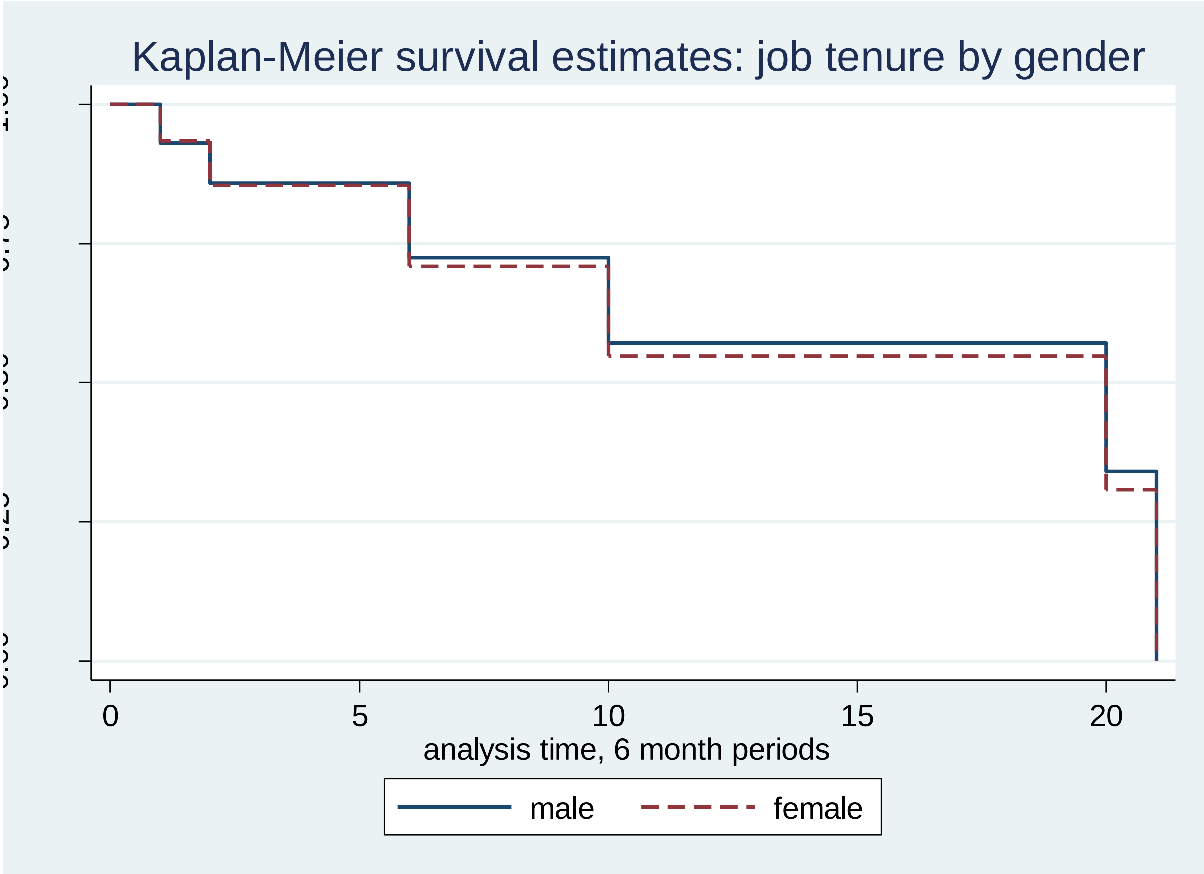


FIGURE 6.2 KAPLAN-MEIER SURVIVAL FUNCTIONS FOR JOB TENURE BY GENDER

These survival functions were then transformed into failure probabilities, which give the probability that an individual will change job in the next six months, given their current age, gender and the length of time they have been continuously employed in the same job. These failure probabilities are given in Table 6.13 below. They are used in the model to predict when a person will change job, given the start date of that job.

The data shows clearly that younger people change job more frequently than older people, with a 20 year old male having a 19.5% probability of leaving a job within 6 months of starting it, compared to 5.3% for a women aged between 25 and 54.

Time Periods	15 to 24 years Men	Women	25 to 54 years Men	Women	55 and over Men	Women
Up to 6 months	0.195	0.182	0.058	0.053	0.029	0.024
6 months to 1 year	0.391	0.421	0.119	0.117	0.059	0.053
1 year to 3 years	0.166	0.176	0.061	0.064	0.033	0.031
3 years to 5 years	0.144	0.151	0.067	0.070	0.038	0.038
5 years to 10 years	0.100	0.100	0.066	0.079	0.043	0.042

TABLE 6.13 PROBABILITY OF CHANGING JOB IN THE NEXT 6 MONTHS

The job tenure data was also used in the pre-model data set up. A current length of time in present job was allocated to people based on their age and gender, using the distribution of job tenure by age and gender presented in Table 6.12 above. The allocation of the initial time to the first job change event could be set within the model run but as it involves a stochastic process, this would increase the variability between model runs. A design decision was made to minimise the number of stochastic processes undertaken **within the model** in order to assist in the comparison of model run results by making them as consistent as possible. Therefore the initial length of job tenure is set outside of the model initialisation process and is the same for a particular individual in all model runs.

6.3 Preparing the data

'More often than not, the choice of an initial base data set will require a trade-off between population representation, data reliability and richness of variables available in the data' (Cassells *et al.*, 2006) The DfT data set scores highly on data reliability and the richness of variables available in the data and moderately well on population representation. The complete stated preference survey for all journey purposes was representative of the UK population but only a subset of these respondents completed the journey to work section of the stated preference survey and this subset was not itself representative of all those in work in the UK.

Some values for the attributes and constraints are not available for every person in the dataset, mainly because not all respondents were asked every question. For example, people who drive more than 5 miles to work were not asked why they do not use a bicycle. There were also respondents who did not respond to questions. In these cases appropriate values were imputed so that in the model input data file all respondents had valid responses in every field.

Three additional fields were created from the survey data. First, the age of the respondent was only recorded in the categories 16-20, 21-29, 30-39, 40-49, 50-59, 50-69 and 70 or over but the model requires an actual age in years, months and weeks for each person. A uniform distribution of respondents by age within each of these categories was assumed and respondents were then allocated an age drawn at random from the relevant age band. When a time related variable such as age is used in a dynamic model it is needed at the same precision as the time steps used in the model. As the model increments time on a weekly basis, the initial age was needed in terms of years and weeks. If all respondents were allocated the median age in the age band then when any event occurred that was triggered by age, such as acquiring a bus pass, then a large number of agents would experience this event at the same time. This would cause an artificial discontinuity in modelled outcomes. Ideally the survey would have asked respondents for their date of birth, but, as this was not the case, the data was smoothed by random allocation of a more precise age to each agent.

Second, when respondents were asked how long they had lived in their current home, the answers were only recorded as either up to a year, up to 2 years, up to 5 years, up to 10 years, up to 20 years or over 20 years. A uniform random distribution within each of these bands was used to allocate a more precise length of residence at the current address expressed in term of years and weeks since they last moved.

Third, respondents were only asked if they had started work or changed job in the last year but the model requires information on both the age of the respondent and the length of time they have been in their current job in order to select the appropriate probability that they change job in the current model time period. The OECD data on labour force job tenure provides a table for the UK in 2011, separately for males and females by five year age groups, of the number of people who had been in their current job for up to 6 months, 6 months to a year, 1 to 3 years, 3 to 5 years, 5 to 10 years and over 10 years. Using this data a distribution of length of job tenure was created for each gender/age group. Then for each respondent a length of job tenure was drawn at random from the appropriate distribution and the number of years and weeks in their current employment assigned to each respondent.

In the preparation of the initial dataset the agents were assigned to one of three destination zones. 25% were allocated to zone 1, 37.5% to zone 2 and 37.5% to zone 3. This allowed the model to include different parking charges in each of the destination zones and for a different proportion of people parking in each zone being liable to pay a parking charge. This was incorporated in the model to simulate the variation found in the real-world of the proportion of free and charged parking spaces available in non-residential places, car parks and on-street parking in different areas.

This illustrates again that ideally the survey would be designed in tandem with the model design so that the data collected would fit the requirements of the model, including data that might be used for model validation. It is often the case though that a model is based on existing data sets, and not all the desired data is available and the imputation of missing data is required.

All these imputed values were added into the input file of agents' attributes that was used in all the scenario runs in order to ensure consistency between model runs. The alternative was to allocate precise ages, length of residence at their current address and time since they last moved job during the initialisation stage of each model run. This approach would increase the variability in the model output as there would be a variation in the values assigned to individuals with missing data which would increase the variation in model output between runs, as the models would vary both in the values used for missing data and the occurrence of random events during the model runs. Ideally the input data would have come from a data set designed specifically for the model and the amount of missing data would be far smaller than in the data set used in this study. For this reason the decision was taken to fix the missing values so that the variation in the model runs came solely from the occurrence of random events during the model run. The resulting amount of variation between runs in the model used in this research is therefore closer to that which would be experienced in a real application with a bespoke data set.

6.4 Overview, Design concepts and Details

6.4.1 The ODD Protocol

The Overview, Design concepts and Details (ODD) protocol (Grimm *et al.*, 2006, revised 2010) is a standardised and widely adopted format for documenting an agent based model. Its purpose is to make the reading and writing of agent based model descriptions more efficient. The protocol starts with the Overview section, which outlines the model's purpose, its entities and provides an overview of the processes incorporated in the model. The Design concepts section, covers typical agent based modelling features such as emergence, learning and interaction. The Details section provides information on initialisation, input data and any sub models. The ODD protocol, by providing a common format for sharing details of the agents and processes in an agent based model, assists in spreading an understanding of the model. This is especially helpful when making a judgement on a model's validation, particularly if the nature of the model is such that it is not amenable to validation by comparing numerical data.

The structure of the implemented model in this research is described here using the ODD protocol. An overview of structure of the model is provided in Figure 6.3 below.

.

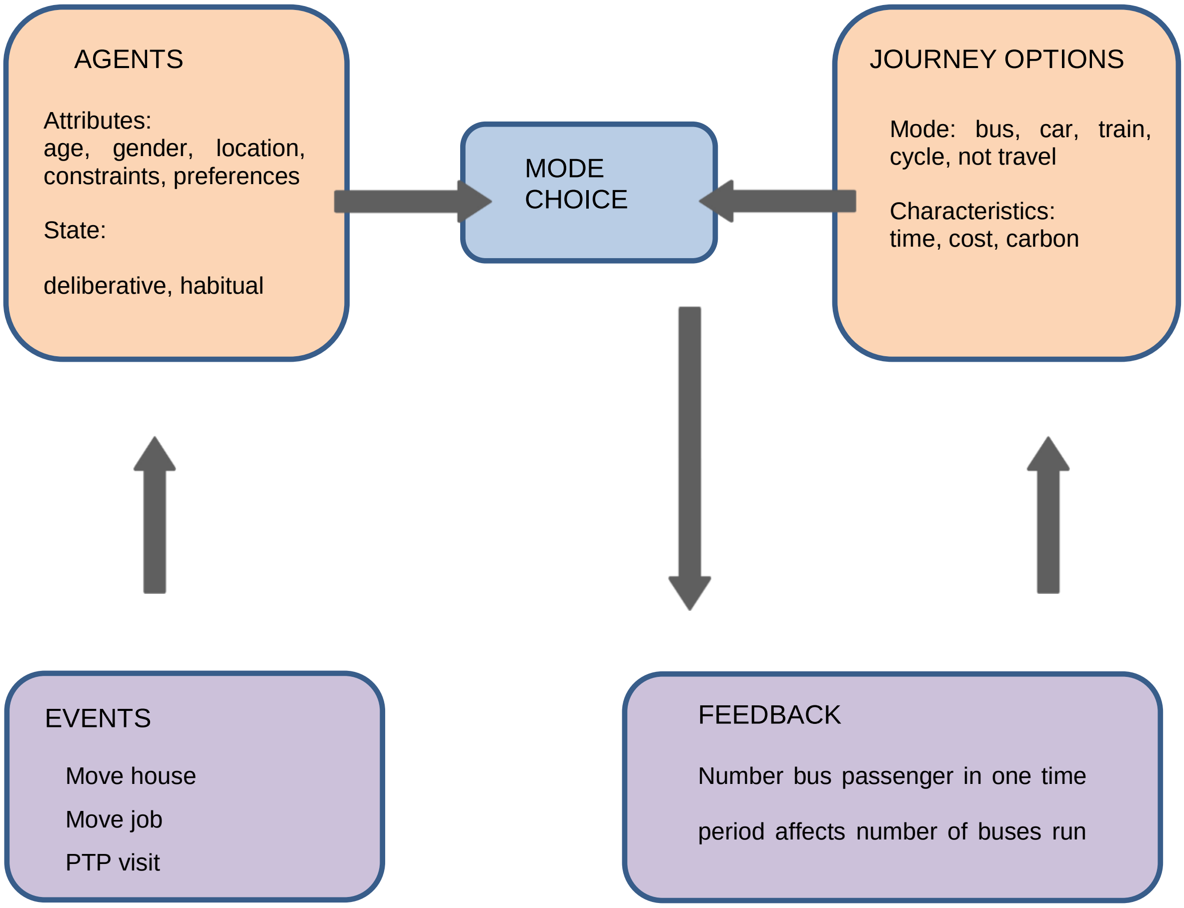


FIGURE 6.3 STRUCTURE OF THE MODEL

6.4.2 Purpose

The purpose of the model is to predict the scale of the impact on mode share and the lags in the achievement of change resulting from policies aimed at influencing the proportion of trips made to work by bus, car, train and cycle. The model shows how the existence of a PTP programme in an area can be observed in changes in mode share.

6.4.3 Entities, state variables, and scales

The model consists of agents and the environment in which they exist. Each agent in the model is a one-to-one match with one of the 626 respondents in the choice modelling section of the DfT dataset. The attributes of each agent are taken, wherever possible, from the answers they provided in the survey. Table 6.14 shows the variables used in the model.

Agent variables

Variable	Variable description	Source of data	Static / Dynamic
Respondent	User ID from choice modelling	DfT	Static
Gender	Gender of respondent	DfT	Static
Age	Age of respondent	DfT	Dynamic
years_house	Years lived in current home	DfT	Dynamic
years_job	Years in current job	Assigned	Dynamic
disability_bus	Has disability which prevents local bus use	DfT	Static
disability_car	Has disability which prevents local car use	DfT	Static
disability_cycle	Has disability which prevents local cycle use	DfT	Static
car_licence	Holds full licence for car valid in England	DfT	Static
need_car	Needs car for work	DfT	Static
bus_things	Can’t use bus as has to take heavy items	DfT	Static
rail_things	Can’t use train as has to take heavy items	DfT	Static
cycle_things	Can’t use bicycle as has to take heavy items	DfT	Static
cycle_combined	Can't ride a bicycle for a variety of reasons	DfT	Static
Destination	Destination zone	Assigned	Static
MODE_BUS	Bus mode constant	DfT*	Static
MODE_CAR	Car mode constant	DfT*	Static
MODE_TRAIN	Rail mode constant	DfT*	Static
MODE_CYCLE	Cycle mode constant	DfT*	Static
COST_150	Part worth cost £1.50	DfT*	Static
COST_200	Part worth cost £2.00	DfT*	Static
COST_250	Part worth cost £2.50	DfT*	Static
COST_300	Part worth cost £3.00	DfT*	Static
CO_2_1	Part worth carbon 1kg	DfT*	Static
CO_2_2	Part worth carbon 2kg	DfT*	Static
CO_2_3	Part worth carbon 3kg	DfT*	Static
CO_2_4	Part worth carbon 4kg	DfT*	Static
TIME_15	Part worth journey time 15 minutes	DfT*	Static
TIME_30	Part worth journey time 30 minutes	DfT*	Static
TIME_45	Part worth journey time 45 minutes	DfT*	Static
TIME_60	Part worth journey time 60 minutes	DfT*	Static
TIME_75	Part worth journey time 75 minutes	DfT*	Static
NONE	Part worth - choose not to travel or use another mode	DfT*	Static
CLASS	Class identifier when using segments	DfT*	Static

*estimated from DfT stated preference responses using Sawtooth software

Environment variables

State Variable	Units	Static / Dynamic
Mode for journey option 1		Static
Time for journey option 1	Minutes	Static
Cost for journey option 1	Pence	Dynamic
Carbon emissions for journey option 1	Kilograms	Static
Mode for journey option 2		Static
Time for journey option 2	Minutes	static
Cost for journey option 2	Pence	dynamic
Carbon emissions for journey option 2	Kilograms	static
Mode for journey option 3		static
Time for journey option 3	Minutes	static
Cost for journey option 3	Pence	dynamic
Carbon emissions for journey option 3	Kilograms	static
Mode for journey option 4		static
Time for journey option 4	Minutes	static
Cost for journey option 4	Pence	dynamic
Carbon emissions for journey option 4	Kilograms	static
Minimum age for bus pass	Years, weeks	static
Proportion people pay to park in zones 1,2,3 six months	%	static
Parking charge in destination zones 1,2,3 months in	Pence	Static
Change in fuel costs	% change	Static
Change in bus fares, years 1 - 10	% change	Static
Change in rail fares, years 1 - 10	% change	Static
Percentage agents receiving a ptp visit each year	%	Static
Retirement age	Years, weeks	Static
% house movers leave area	%	Static
% job changers leave area	%	Static

TABLE 6.14 MODEL VARIABLES

Behavioural strategies

All agents have the same behavioural strategy. When they are initialised they are in a deliberative state and choose their travel mode in the first modelled time period. They then switch to the habitual state and remain in that state until they receive a trigger event. A trigger event causes them switch back to a deliberative state.

When choosing a transport mode, they first reject modes which are not available to them due to constraints and then choose the mode which has the highest total utility level for them. Each agent has their own set of utilities for time, cost, carbon and mode constants.

The model is run for 10 years. Each time period represents one week and the model is run for 520 time periods. It is assumed, in order to simplify the model and reduce run times, that the mode used by an agent is the same on each working day of a modelled week.

6.4.4 Process overview and scheduling

Process overview

Each time period is modelled as a discrete step.

The sub models are performed in this order:

- update the environment (journey) characteristics
- update the agent’s age, individual journey costs and constraints
- check for a trigger event, which changes agent to a deliberative state
- make mode choice decision if agent state is deliberative, otherwise the agent uses mode used in the previous time period.

Scheduling

The regular events in the model are:

- changes in bus fares, rail fares, fuel costs and parking charges, which are applied every 6 months
- the number of people with free parking is updated every 6 months.

The irregular events in the model are:

- people may move in any time period. The probability of someone moving house depends on their age and gender. The model checks for a potential house move event occurring every 4 weeks starting with week 1,

- people may move job in any time period. The probability of someone moving job depends on the number of months since they last changed job, their age and gender. The model checks for a potential job change event occurring every 4 weeks starting with week 2,

- people may receive a personal travel visit at either their home or place of work in any time period. Receiving a visit is a random event, based on the number of visits being undertaken each year. The chances of receiving a visit in any one time period are independent of having received a visit on a previous occasion. The model checks for a potential personalized travel planning visit event occurring every 4 weeks starting with week 3,

6.4.5 Initialization

The initial settings for each agent are read in from an input file. The values are those given by the 626 respondents in the DfT survey. When the agents are initialised, they are all placed in a deliberative state.

The initial state of the environment variables for the time, cost and CO_2 emissions for the four journey options is provided by the user in an input file.

6.4.6 Input data

The model uses data from the DfT survey on climate change and transport choices to set the initial state of the agent variables.

The initial state of the environment variables for the journey options are set by the user and must be within the range of values used in the DfT choice modeling exercise. These are:

- Cost 150 to 300 pence
- Time 15 – 75 minutes
- Carbon 1 – 4 kg

The data on the rate at which people move house and change job is taken from the British Household Panel Survey and OECD job tenure data.

The rates of change in fuel costs, bus fares, rail fares, the change in the number of free parking places in each destination zone and parking charge in each destination zone is read in from an input file prepared by the model user.

The rate of personalised travel visits is provided by the user in the model set up data.

The part worths or preference weightings held by each individual for the attributes of a journey option e.g, mode, time, cost and carbon were estimated from the results of the choice modeling exercise using Sawtooth software. The estimation of these values from the survey data was made using the hierarchical bayes option in the Sawtooth software.

6.4.7 Submodels

There are five sub-models.

The first changes the global journey attributes. For each time period the attributes of the journey options are changed as required by regular events:

- the car costs, bus fares and rail fares are changed every six months
- the parking charges in each zone and the proportion of people in each zone who have to pay to park is changed every six months

The second updates the value of each agent's variables. The attributes of each agent are updated in each time period; their age is raised, eligibility for a free bus pass assessed and their status regarding the availability of free parking updated.

The third reviews each agent's mode choice decision making state. The model checks whether the agent has received a trigger event in this time period. The trigger events are:

- receive personal travel plan visit
- move house
- move job
- receive a free bus pass. This is triggered if the agent reaches the minimum age to receive a concessionary bus pass

If a trigger event has occurred the agent is switched to the 'deliberative' state from the 'habitual' state.

The fourth enacts the decision making. It updates the cost of each option for the individual agent taking into account the destination zone of the agent which affects the parking charge and the cost of bus travel which might be affected by the availability of a free bus pass. The total utility of each option is calculated using the agent's own preferences. The agent rejects any modes which are not available to them due to constraints and then chooses between the remaining options or not travelling at all, the option which has the highest utility for them.

The fifth model controls agent replication. When an agent reaches the age of 68 it leaves the model and is replaced by an agent with a random age between 17 and 60, with the same attributes and personal preferences as the agent they replace, but with zero years since they moved house or left their job.

6.5 Conclusion

This chapter reported on the development of the agent based model of commuting mode choice developed for this study. This is an empirically grounded agent based model and the data for each agent comes from a DfT survey which reported the attributes of each respondent and their preference weightings for time, cost and carbon emissions when selecting a mode for a five mile commute to work.

The model is based on Triandis' theory of Interpersonal Behaviour which states that a person's observed behaviour is influenced by their intention and their habits and depends upon the facilitating conditions. In each time period an agent's choice is made on the basis of intention or habit. The actual modes available to a person are constrained by the facilitating conditions. These may be personal constraints, such as the lack of a car or mobility issues or external conditions, such as the lack of a bus service. The model is run for each week for ten years.

The next chapter presents the results of the modelling work undertaken with this model. The aim of the modelling exercise was to investigate the potential of ABM for modelling smarter choices and this aim influenced the design of the model runs. The work was based on the approach of starting with a simple model and then adding further processes to illustrate the features of agent based modelling that are relevant to the modelling of 'smarter choices' measures.

7 Model results

7.1 Introduction

This chapter presents the results from using the agent based model of commuter mode choice described in the previous chapter. A variety of scenarios were run which varied by the assumed changes in the time and cost of the journey options over time and whether a programme of personalised travel planning was undertaken or not. Each scenario was run 100 times to allow for analysis of the variability in results between model runs due to the operation of the stochastic processes within the model. These processes govern the timing and frequency of the main trigger events in the model. The trigger events are moving house, changing job, gaining a free bus pass and receiving a visit from a personal travel advisor.

The aim of this research is to investigate the merits of agent based modelling as a technique to provide the information needed for an appraisal of the impact of a package of measures intended to reduce car use and promote 'smarter choices'. This aim informed the choice of tests undertaken with the model. The tests concentrated on features of agent based modelling which could improve the accuracy of the model outputs required for use in the business case for the type of 'smarter choices' interventions which might be funded by the developer of a large development such as Alconbury Weald or a local authority. Typical example of such interventions designed to reduce car use are the subsidy of a bus service and the funding of a personalised travel planning campaign.

The results from the model runs are presented here under themes which illustrate features of the modelling approach that demonstrate its suitability for this task.

7.2 Modelling the trajectory over time of the number of people using each mode

It is valuable for the appraisal of a transport intervention to know how the number of people using each mode will evolve over time. This will assist in the planning of sufficient capacity in the transport network, such as ensuring that enough train carriages are provided to meet demand. Detailed passenger forecasts over time will improve the estimate of the revenue that will be received by public transport operators and the level of any subsidy required to ensure that a particular level of service is maintained.

In a conventional four stage transport model there is an assumption that the transport system is in a state of equilibrium and the model is run for a particular point of time in the future with the intention of forecasting the state of each element of the transport system at the equilibrium point for that future date. Models are set up to run for one particular moment in time and separate model runs are required if results for other moments in time are required.

Agent based modelling does not make any presumption about the existence or otherwise of a state of equilibrium in the transport network. The output that comes from the model of mode choice developed in this research provides a set of forecasts of the number of users of each mode over time rather than for a single moment in the future. It shows the trajectory towards the mode share forecast in the final modelled time period. This is illustrated by the output from scenario 1 which was run for every week for 10 years and produces a demand forecast for the transport system in each of these time periods as well as for the model year 10 years away from the model's base year.

In scenario 1 every aspect in the model that might affect mode choice remains constant over time except for the changes in the cost of travel by each mode. The input values for this scenario are presented in Figure 7.1 below.

The model is run with all 626 agents who each have the choice between four modes of travel: bus, car, train and cycle or the choice not to travel. Initially the bus journey costs £1.60, takes 35 minutes and has CO_2 emissions of 1 kg. The car journey costs £1.00, takes 20 minutes and has CO_2 emissions of 3 kg. The train journey costs £2.40, takes 20 minutes and has CO_2 emissions of 1 kg. The cycle journey is free, has no CO_2 emissions and takes 60 minutes.

The age at which agents become eligible for a free bus pass is 61 years and 48 weeks. Current government policy in England is for this age to rise over time but in this model the age is kept constant.

Agents are replaced when they reach retirement age, which is set to 68 years. The replacement agent inherits all the current attributes of the agent they replace, with the exception of age, which is set at random between 17 and 60 years. When an agent moves house, 30% of the agents move out of the area and are replaced. When an agent changes job, 20% of the agents change to a job out of the area and are replaced.

There are no personalised travel plan visits in scenario 1, so the triggers for people to re-assess their travel behaviour are moving house, moving job, entering the area as a replacement for a person who has retired or left the area, and receiving a bus pass.

In scenario 1 the percentage of drivers going to zone 1 who have to pay to park is 50%, in zone 2 it is 0% and in zone 3, 50%. The parking charge in zones 1 and 3 is 150 pence and parking is free in zone 2. The allocation of agents to a destination zone, the proportion of agents in each zone that pay to park and the parking charge in each zone remain the same in each time period.

SCENARIO 1

INITIAL JOURNEY DETAILS

MODE	Cost (pence)	CO2 (kg)	Time (mins)
BUS	160	1	35
CAR	100	3	20
TRAIN	240	1	20
CYCLE	0	0	60

PTP % households visited a year	0%	
% house movers leave area	30%	
% job changers leave area	20%	
	Years	Weeks
Minimum age for bus pass	61	48
Minimum age for new agents	17	0
Maximum age for new agents	60	0
Retirement age	68	0

	Yr1 Q1	**Yr1 Q3**	**Yr2 Q1**	**Yr2 Q3**	**Yr3 Q1**	**Yr3 Q3**	**Yr4 Q1**	**Yr4 Q3**	**Yr5 Q1**	**Yr5 Q3**
COST CHANGES										
BUS	1.0000	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198
CAR	1.0000	0.9909	0.9909	0.9941	0.9941	0.9919	0.9919	0.9925	0.9925	0.9923
TRAIN	1.0000	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050
TIME CHANGES										
BUS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
CAR	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TRAIN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
% PAY TO PARK										
ZONE 1	50	50	50	50	50	50	50	50	50	50
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	50	50	50	50	50	50	50	50	50	50
PARK CHARGE (pence)										
ZONE 1	150	150	150	150	150	150	150	150	150	150
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	150	150	150	150	150	150	150	150	150	150

	Yr6 Q1	**Yr6 Q3**	**Yr7 Q1**	**Yr7 Q3**	**Yr8 Q1**	**Yr8 Q2**	**Yr9 Q1**	**Yr9 Q3**	**Yr10 Q1**	**Yr10 Q3**
COST CHANGES										
BUS	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198
CAR	0.9923	0.9920	0.9920	0.9914	0.9914	0.9902	0.9902	0.9898	0.9898	0.9899
TRAIN	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050
TIME CHANGES										
BUS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
CAR	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TRAIN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
% PAY TO PARK										
ZONE 1	50	50	50	50	50	50	50	50	50	50
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	50	50	50	50	50	50	50	50	50	50
PARK CHARGE (pence)										
ZONE 1	150	150	150	150	150	150	150	150	150	150
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	150	150	150	150	150	150	150	150	150	150

FIGURE 7.1 INPUT VALUES FOR SCENARIO 1

The journey time by each mode also remains constant over time in this scenario.

All costs in the model runs are expressed in real terms. Bus fares are set to rise at 1.98% every six months which is equivalent to a 4% rise in bus fares each year. (Bus fares rose in Wales in real terms by 3.9% in 2013 and 4.9% in 2014 according to the DfT Transport Statistics table BUS0415). Car costs decline over time at the rate set out in WebTAG unit 3.5.6 (DfT, January 2014). The rate of decline in car costs increases over time, that is car use becomes increasing cheaper over time in real terms. This is a result of the predicted increases in fuel efficiency in car engines outweighing the real increases in fuel cost, resulting in a net reduction in fuel costs of a car per kilometre driven. Train fares are assumed to rise at 1% per annum reflecting the current regulatory regime on rail fares which limits the overall rise in regulated fares to the retail price index plus 1%. This means that over time the cost of travelling by public transport rises, with the rise in bus fares being greater than train fares, while the cost of car travel in the model declines. The concessionary bus pass is included in this scenario so travelling by bus becomes free for everyone when they reach the qualifying age. The qualifying age used in this model run is 61 years and 48 weeks which was the age for receiving the concessionary bus pass when the scenario was run. This means that some more elderly workers can commute for free by bus as the retirement age in the model is set at 68 which is higher than the qualifying age for the concessionary bus pass.

All agents use the decision rule that they chose either the mode or the option not to travel that maximises their individual utility. At this stage in the development of the model the mode chosen is based 100% on their intentions and is not influenced by habit. In the remainder of this report a decision that is based on intentions is described as ‘deliberative’. The scenario is run weekly for 10 years, resulting in 520 weeks or modelled time periods.

Table 7.1 below shows the number of users of each mode at the start of each year, assuming that no agent has any personal constraints that influence the mode they can use and that every agent makes a conscious decision about the mode they will use in every time period i.e. they are always in a deliberative state. The full output provided from the model gives the number of people using each mode every week. This is used to plot mode shares over time, as presented In Figure 7.2 overleaf.

Time period	Week	Bus		Car		Train		Cycle		Not travel	
beg year 1	**1**	46	7.3%	410	65.5%	112	17.9%	52	8.3%	6	1.0%
beg year 2	**53**	53	8.5%	409	65.3%	108	17.3%	50	8.0%	6	1.0%
beg year 3	**105**	55	8.8%	409	65.3%	105	16.8%	51	8.1%	6	1.0%
beg year 4	**157**	59	9.4%	412	65.8%	98	15.7%	51	8.1%	6	1.0%
beg year 5	**209**	60	9.6%	410	65.5%	97	15.5%	53	8.5%	6	1.0%
beg year 6	**261**	60	9.6%	416	66.5%	91	14.5%	53	8.5%	6	1.0%
beg year 7	**313**	62	9.9%	417	66.6%	88	14.1%	53	8.5%	6	1.0%
beg year 8	**365**	63	10.1%	416	66.5%	87	13.9%	53	8.5%	7	1.1%
beg year 9	**417**	68	10.9%	414	66.1%	84	13.4%	53	8.5%	7	1.1%
beg year 10	**469**	68	10.9%	414	66.1%	84	13.4%	53	8.5%	7	1.1%
end year 10	**520**	65	10.4%	415	66.3%	86	13.7%	53	8.5%	7	1.1%

TABLE 7.1 SCENARIO 1: NUMBER OF PEOPLE USING EACH MODE, UNCONSTRAINED CHOICE, ALWAYS DELIBERATIVE

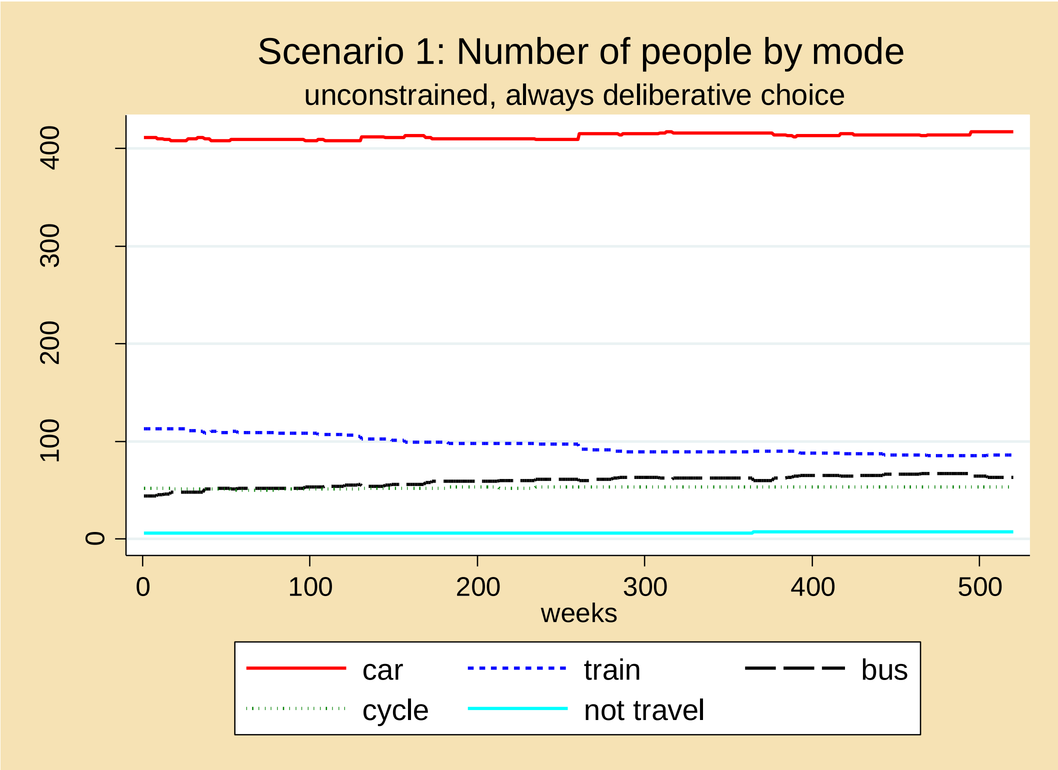


FIGURE 7.2 SCENARIO 1: NUMBER OF PEOPLE USING EACH MODE, UNCONSTRAINED CHOICE, ALWAYS DELIBERATIVE

The information gained from seeing the trajectory of mode share over time can be especially important if there is a significant change in journey time or costs for a mode at a particular point in time, for example after introducing a new charge for on-street parking or raising existing parking charges.

Such a situation was tested in Scenario 2, which has exactly the same inputs as scenario 1, except for a parking charge policy that periodically increases the proportion of people paying to park and the parking charge. At the beginning of year 3, the proportion of people paying to park increases from 50% to 75% in zones 1 and 3 and from 0% to 50% in zone 2. The parking charge in zone 2 is £1. At the beginning of year 6 everyone in zones 1, 2 and 3 has to pay to park. The parking charge rises to £3 in zones 1 and 3 and £1.50 in zone 2. At the beginning of year 8 the parking charge in zone 2 rises to £3

The inputs used for scenario 2 are shown in Figure 7.3 below.

SCENARIO	2						
INITIAL JOURNEY DETAILS							
MODE	Cost	Time	CO2		PTP % households visited a year	0%	
	(pence)	(mins)	(kg)		% house movers leave area	30%	
BUS	160	35	1		% job changers leave area	20%	
CAR	100	20	3			Years	Weeks
TRAIN	240	20	1		Minimum age for bus pass	61	48
CYCLE	0	60	0		Minimum age for new agents	17	0
					Maximum age for new agents	60	0
					Retirement age	68	0

	Yr1 Q1	**Yr1 Q3**	**Yr2 Q1**	**Yr2 Q3**	**Yr3 Q1**	**Yr3 Q3**	**Yr4 Q1**	**Yr4 Q3**	**Yr5 Q1**	**Yr5 Q3**
COST CHANGES										
BUS	1.0000	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198
CAR	1.0000	0.9909	0.9909	0.9941	0.9941	0.9919	0.9919	0.9925	0.9925	0.9923
TRAIN	1.0000	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050
TIME CHANGES										
BUS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
CAR	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TRAIN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
% PAY TO PARK										
ZONE 1	50	50	50	50	75	75	75	75	75	75
ZONE 2	0	0	0	0	50	50	50	50	50	50
ZONE 3	50	50	50	50	75	75	75	75	75	75
PARK CHARGE (pence)										
ZONE 1	150	150	150	150	150	150	150	150	150	150
ZONE 2	0	0	0	0	100	100	100	100	100	100
ZONE 3	150	150	150	150	150	150	150	150	150	150

	Yr6 Q1	**Yr6 Q3**	**Yr7 Q1**	**Yr7 Q3**	**Yr8 Q1**	**Yr8 Q2**	**Yr9 Q1**	**Yr9 Q3**	**Yr10 Q1**	**Yr10 Q3**
COST CHANGES										
BUS	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198
CAR	0.9923	0.9920	0.9920	0.9914	0.9914	0.9902	0.9902	0.9898	0.9898	0.9899
TRAIN	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050
TIME CHANGES										
BUS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
CAR	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TRAIN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
% PAY TO PARK										
ZONE 1	100	100	100	100	100	100	100	100	100	100
ZONE 2	100	100	100	100	100	100	100	100	100	100
ZONE 3	100	100	100	100	100	100	100	100	100	100
PARK CHARGE (pence)										
ZONE 1	300	300	300	300	300	300	300	300	300	300
ZONE 2	150	150	150	150	300	300	300	300	300	300
ZONE 3	300	300	300	300	300	300	300	300	300	300

FIGURE 7.3 INPUT VALUES FOR SCENARIO 2

The number of users of each mode at the start and end of each year are shown in Table 7.2 below. There is a step change in the cost of car journeys at the beginning of years 3, 6 and 8 which is reflected in the sudden changes in the number of people using the car when these changes in costs occur. At the beginning of year 3 the number of car users falls by 36 people from 408 to 372, a 9% decrease. 32 of these people switch to train and the number of train users rises from 108 to 140 which is a 30% increase. The number of bus users increases by 3 from 53 to 56, a 6% increase.

The increase in parking costs at the beginning of years 6 and 8 also causes a sudden change in car usage which is in contrast to the steady change in numbers using each mode seen in scenario 1, when the relative cost of travel by each mode changed steadily over time. Figure 7.4 shows the number of users of each mode in each week for scenario 2. The sudden changes in mode shares at the beginning of years 3, 6 and 8 can be clearly seen.

Time period		Bus		Car		Train		Cycle		Not travel	
beg year 1	**1**	44	7.0%	411	65.7%	113	18.1%	52	8.3%	6	1.0%
end year 1	**52**	52	8.3%	408	65.2%	109	17.4%	51	8.1%	6	1.0%
beg year 2	**53**	51	8.1%	409	65.3%	110	17.6%	50	8.0%	6	1.0%
end year 2	**104**	53	8.5%	408	65.2%	108	17.3%	51	8.1%	6	1.0%
beg year 3	**105**	56	8.9%	372	59.4%	140	22.4%	53	8.5%	5	0.8%
end year 3	**156**	58	9.3%	377	60.2%	132	21.1%	54	8.6%	5	0.8%
beg year 4	**157**	58	9.3%	379	60.5%	130	20.8%	54	8.6%	5	0.8%
end year 4	**208**	62	9.9%	377	60.2%	128	20.4%	54	8.6%	5	0.8%
beg year 5	**209**	61	9.7%	378	60.4%	128	20.4%	54	8.6%	5	0.8%
end year 5	**260**	63	10.1%	378	60.4%	126	20.1%	54	8.6%	5	0.8%
beg year 6	**261**	79	12.6%	282	45.0%	200	31.9%	60	9.6%	5	0.8%
end year 6	**312**	82	13.1%	284	45.4%	195	31.2%	60	9.6%	5	0.8%
beg year 7	**313**	79	12.6%	286	45.7%	196	31.3%	60	9.6%	5	0.8%
end year 7	**364**	81	12.9%	286	45.7%	194	31.0%	60	9.6%	5	0.8%
beg year 8	**365**	82	13.1%	255	40.7%	222	35.5%	61	9.7%	6	1.0%
end year 8	**416**	88	14.1%	254	40.6%	217	34.7%	61	9.7%	6	1.0%
beg year 9	**417**	86	13.7%	260	41.5%	213	34.0%	61	9.7%	6	1.0%
end year 9	**468**	91	14.5%	260	41.5%	208	33.2%	61	9.7%	6	1.0%
beg year 10	**469**	91	14.5%	261	41.7%	207	33.1%	61	9.7%	6	1.0%
end year 10	**520**	91	14.5%	261	41.7%	207	33.1%	61	9.7%	6	1.0%

TABLE 7.2 SCENARIO 2: NUMBER OF PEOPLE USING EACH MODE, UNCONSTRAINED CHOICE, ALWAYS DELIBERATIVE

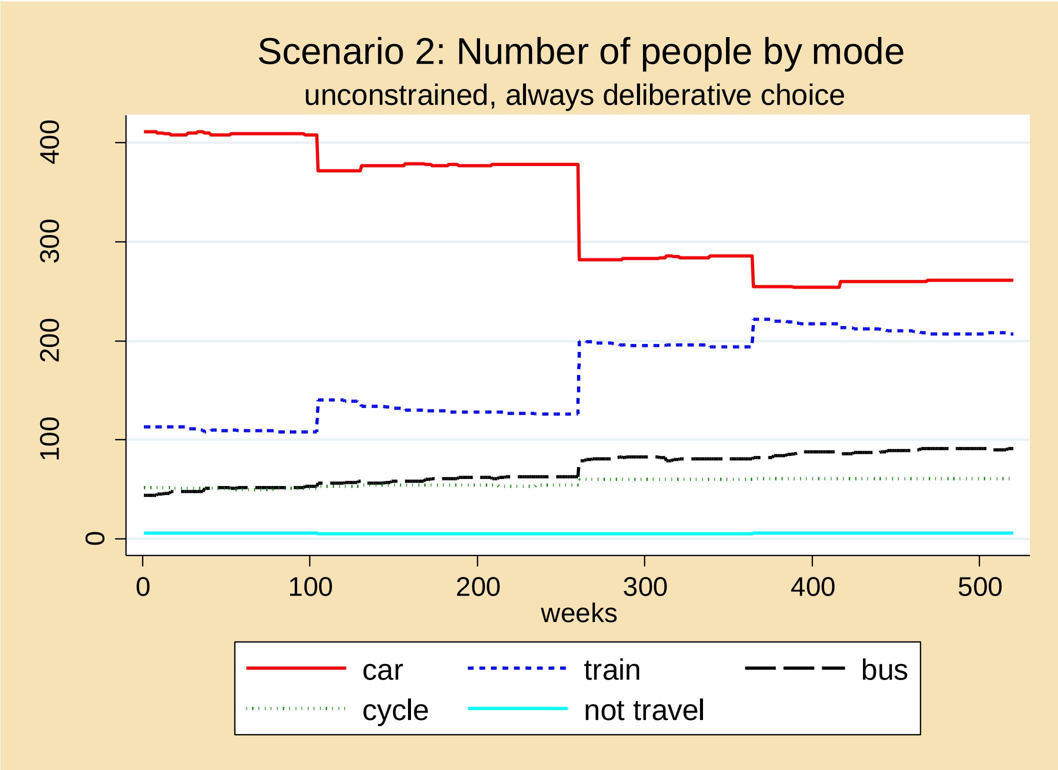


FIGURE 7.4 SCENARIO 2: NUMBER OF PEOPLE USING EACH MODE

Knowing the trajectory of the number of users over time for each mode is particularly important for the transport business case when there are sudden changes in journey times and costs for particular modes over the appraisal period. In the UK, the costs and benefits of an intervention are customarily appraised over 60 years in accordance with the DfT guidance in WebTAG, although for some schemes a shorter time period is used. The transport model is usually run for the year of scheme opening and then for one or occasionally two future years, with demand assumed constant thereafter for the remaining years in the appraisal period. The change in the number of users in each of the intervening years is taken as a straight line between the modelled points.

Figure 7.5 below shows the modelled number of train users if the model is run every week (regular reporting as in an agent based model) and if the model is run only at opening and after 6 years (occasional modelling).

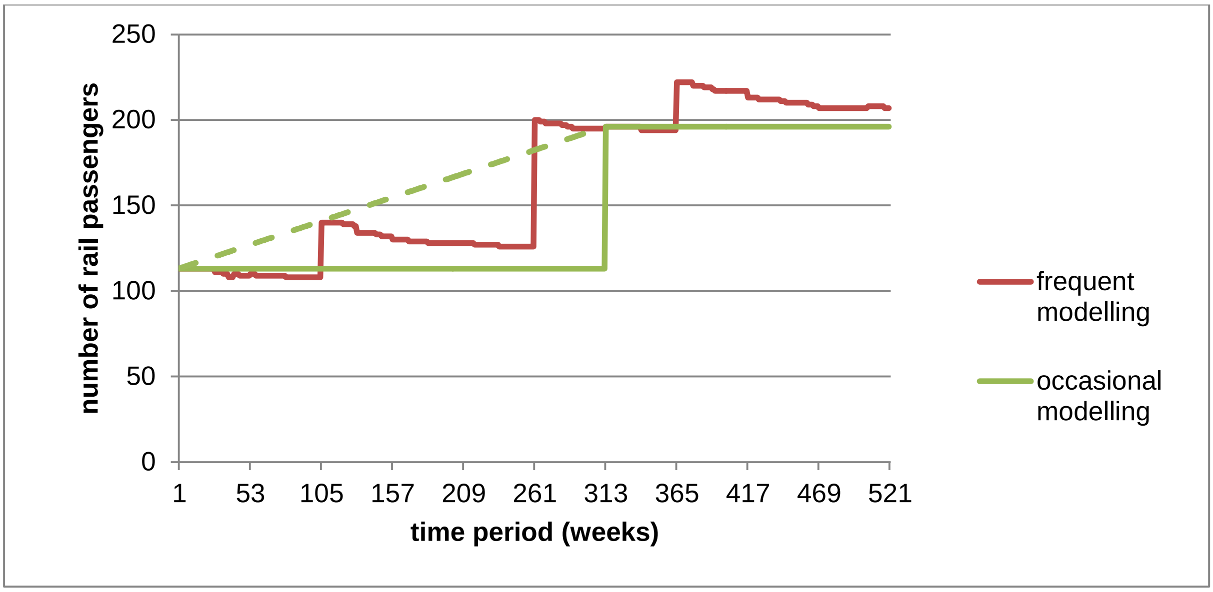


FIGURE 7.5 TRAJECTORY OF NUMBER OF RAIL PASSENGER

The red line shows the number of rail users when the model is run for every week up to ten years. The green line shows the number of rail users if the model is only run at the beginning of year 1 and at the end of year 6, with demand assumed constant thereafter. At certain points in time, for example at the beginning of year 6 the discrepancy between the two numbers is considerable, with the regular modelling showing rail passengers numbers at 200 passengers and the occasional model estimates being 57% lower as the opening year estimate of 113 is still being used. This discrepancy arises because the selected point in time for the occasional modelling (the end of year 6) misses the moment when the car park charges rose considerably causing a switch of users from car to rail.

In the appraisal of transport schemes, a straight line is drawn between modelled time points and interpolated results are used for each year in the appraisal (shown as a dashed line in Figure 7.5 above). This reduces, but does not remove, the discrepancy between the actual and modelled number of users of each mode in the non-modelled years. With the agent based model the timing of change points can be identified and more accurate estimates of the number of users in each year can be used in the business case.

When the benefit cost ratio of schemes is calculated using the DfT's TUBA software, the benefits are calculated separately for each year and then added together, so the underestimation of passenger numbers each year will lead to a miscalculation of the actual benefits. In TUBA the discounted value of the benefits is calculated using a discount rate of 3.5% for the first thirty years and 3.0% for the next thirty years. This means that differences in passenger forecasts in the early years will have a larger impact on the correct calculation of the total discounted value of benefits than differences in future years.

It is possible to run a conventional model many times but each model run often takes a considerable length of time, so the model is only run for a selected number of forecast years. In standard four stage transport models each of these forecast years is modelled separately and in each case the forecast year is modelled with reference to the base year. This means that when for instance the base year is 2014 and the model is run for 2029 and 2036, the results for 2036 are not dependent in any way on the results for 2029. In some implementations of the four stage model, for example in the Cambridge Strategic Regional Model, the model is run every five years and the first forecast year becomes the base year for the second forecast year. Agent based modelling, because of its dynamic nature, carries this further, with much shorter time periods, and the modelling of each time period is affected by the state of the model in the previous time period.

The convention in the UK is to run the transport model for opening year and fifteen years after opening year but these may not be the best choice of model years in order to produce the best estimates of passenger numbers in each year in the appraisal period. An agent based model runs the model for every time step and so reports the results for multiple model years.

An agent based model could be run to inform the selection of the model years to run in a full conventional model. Such an agent based model could use simplified assumptions on the time and cost of travel between zones and the expected changes in time and cost of travel by mode over time. Analysis of the trajectories of trip numbers by mode produced by the agent based model, especially having regard to the timing of external events such as changes in parking charges, could be used to inform the selection of the number of model years and which model years to run in the full conventional model.

7.3 Reporting at an individual level assists in understanding the causes of output trends reported at the aggregate level

In the results for scenario 1 (trend changes in mode costs) the model outputs show a steady increase in bus users over time as presented in Figure 7.6 below. The rise in the number of bus users seems counter-intuitive, as bus costs are rising steadily at a higher rate than train fares and car costs are generally declining. This means that over time bus travel becomes more expensive relative to both rail and car travel and so the expectation would be that the number of bus users would decline.

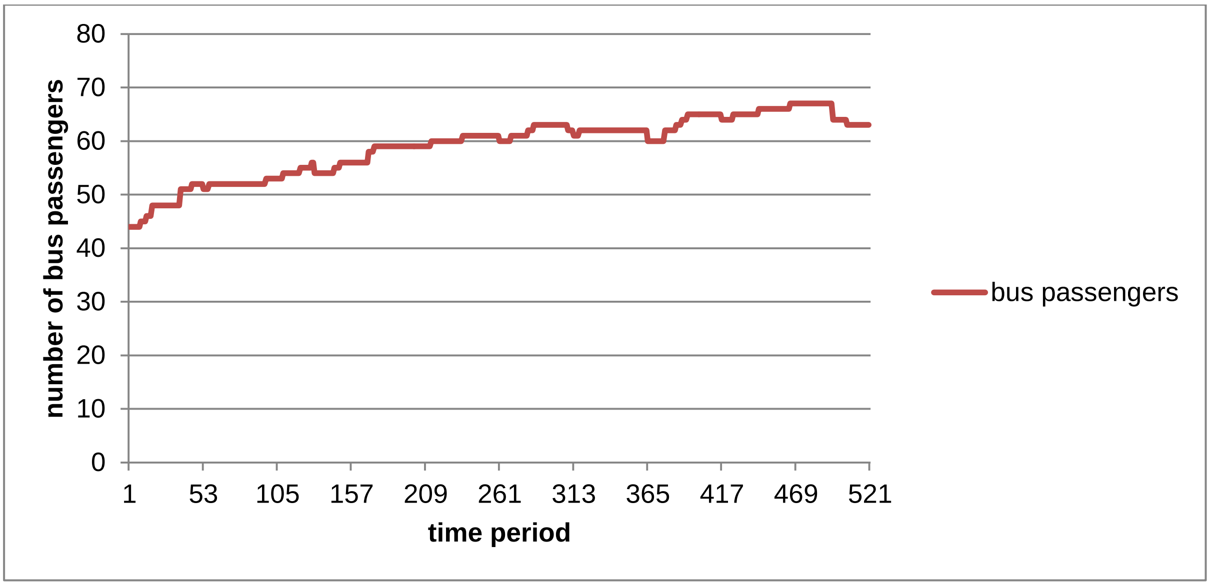


FIGURE 7.6 NUMBER OF BUS PASSENGERS

A possible explanation is that the concessionary bus pass, which makes bus travel free for elderly passengers, introduces a counter-acting mechanism into the model with the availability of free travel encouraging those holding a concessionary bus pass to use the bus when they would not have done so if they had had to pay the fare. As the population in the model ages during the model run, an increasing number of people fall into the group of people who have a concessionary bus pass but are still working and therefore included in this model of commuting trips.

As the model operates at the individual level it is possible to record results at this most disaggregate level and investigate whether the effect of the ageing of the working population is having an impact on the number of bus users. The model was run for scenario 1, recording the mode chosen by people with and without the availability of concessionary bus passes. The number of bus users over time is shown in Figure 7.7 below.

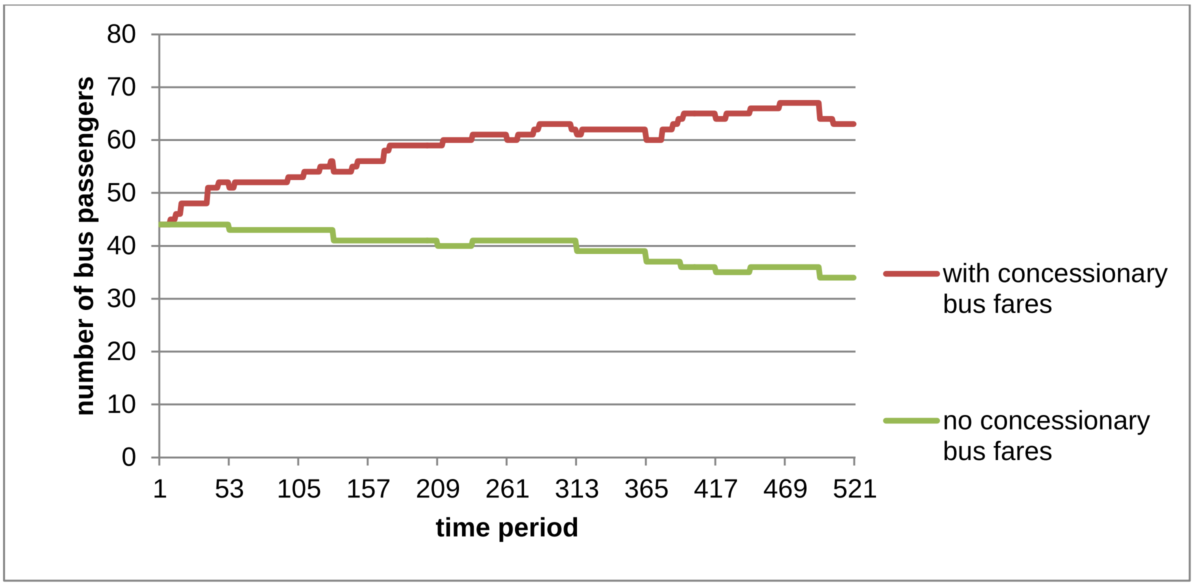


FIGURE 7.7 NUMBER OF BUS USERS OVER TIME WITH AND WITHOUT CONCESSIONARY BUS PASSES AVAILABLE FOR ELDERLY PASSENGERS

This shows that without the availability of concessionary bus passes for the elderly the number of bus users does go down over time as expected, from 44 bus users initially to 34 bus users after 10 years, a fall of 23%. With bus passes for the elderly the number of bus users rises from 44 initially to 63 after 10 years, a rise of 43%.

The difference in the number of bus users with and without concessionary fares is 29 people. Investigating the details for these passengers shows that they are the holders of a concessionary bus pass and the mode that they would have used if they didn't have a concessionary bus pass is shown in Table 7.3 below. 17 of them (59%) would have used a car instead, 9 (31%) would have used rail and 3 (10%) would have cycled.

Car	Rail	Cycle	Total
17	9	3	29
59%	31%	10%	100%

TABLE 7.3 ALTERNATIVE MODE PASSENGERS WOULD HAVE USED IF THEY HAD NOT POSSESSED A CONCESSIONARY BUS PASS

The impact of the concessionary bus fares policy in terms of the percentage change in users of each mode at the end of year 10 is shown in Table 7.4 below. The availability of concessionary bus passes results in a 46% increase in the number of people using the bus to get to work, a 4% decline in the number of commuters driving, a 10% decline in the number using rail and a 6% decline in the number cycling. Although the greatest absolute switch is from bus to car, 17 passengers, this has a smaller proportionate effect on the highway network (4%) as there were already 434 people driving if there were no bus passes. The effect is more noticeable on the rail network with the availability of the concessionary bus pass reducing rail patronage by 10%. The model also shows an impact on active travel with 3 people switching from cycling to the bus, a 6% reduction, if they become eligible for a concessionary bus pass.

	Bus	Car	Rail	Cycle	Not travel
Without bus passes	34	434	95	56	7
With bus passes	63	417	86	53	7
Change	29	-17	-9	-3	0
% change	46%	-4%	-10%	-6%	0%

TABLE 7.4 IMPACT OF CONCESSIONARY BUS FARES ON MODE SHARES

The numbers of car and rail passengers for the model runs with and without the availability of concessionary bus fares are shown in Figures 7.8 and 7.9 respectively. The number of car and rail users is always lower when bus concessionary fares are available. This is because the working population is ageing in the model and the percentage of the workforce holding a concessionary bus pass rises from an initial 8% to 16% in the middle of year 10 as shown in Figure 7.10 below. The slight variations in the percentage of people holding a concessionary bus pass between time periods is due to the relatively small number of people in the model. This means that there are some clusters of people of similar ages in the model which leads to some variation in the number of people reaching the qualification age for a bus pass in each time period.

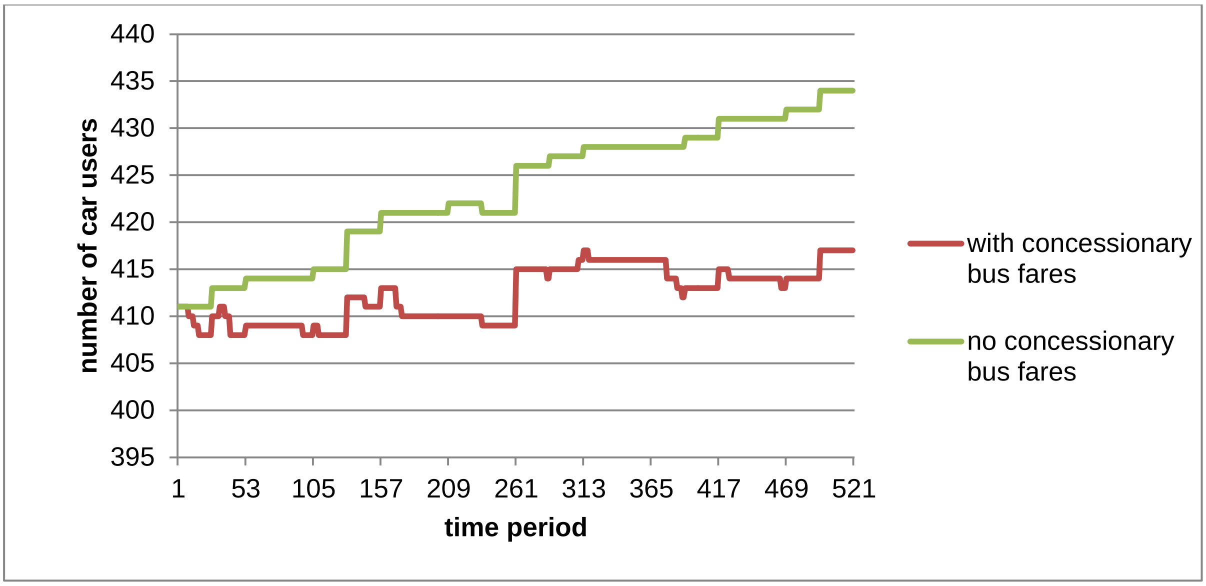


FIGURE 7.8 NUMBER OF CAR USERS OVER TIME WITH AND WITHOUT CONCESSIONARY BUS PASSES FOR ELDERLY PASSENGERS

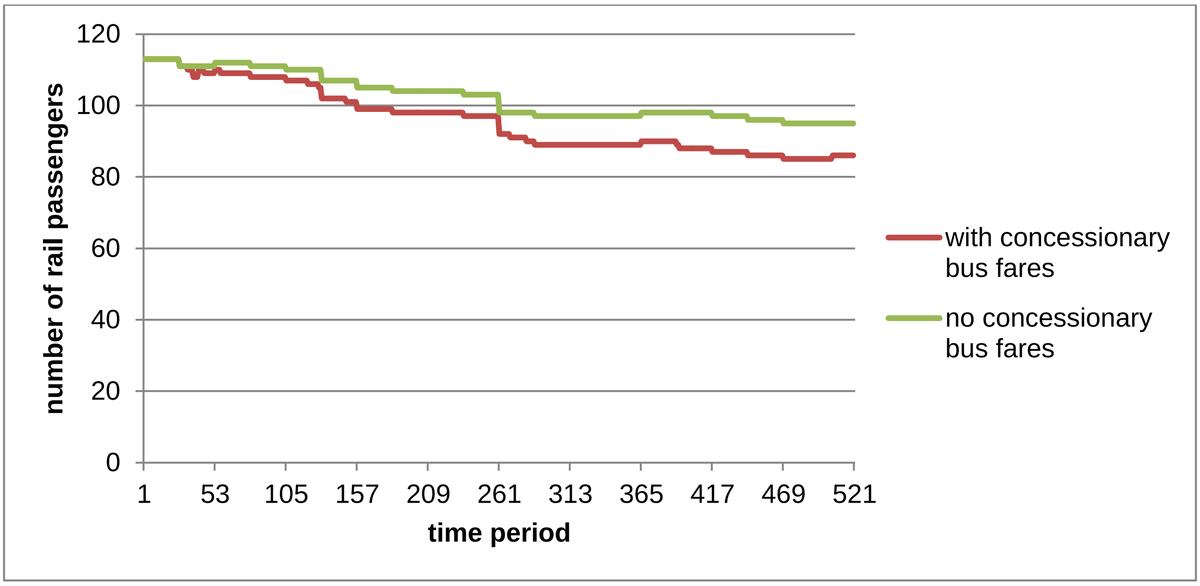


FIGURE 7.9 NUMBER OF RAIL USERS OVER TIME WITH AND WITHOUT CONCESSIONARY BUS PASSES FOR ELDERLY PASSENGERS

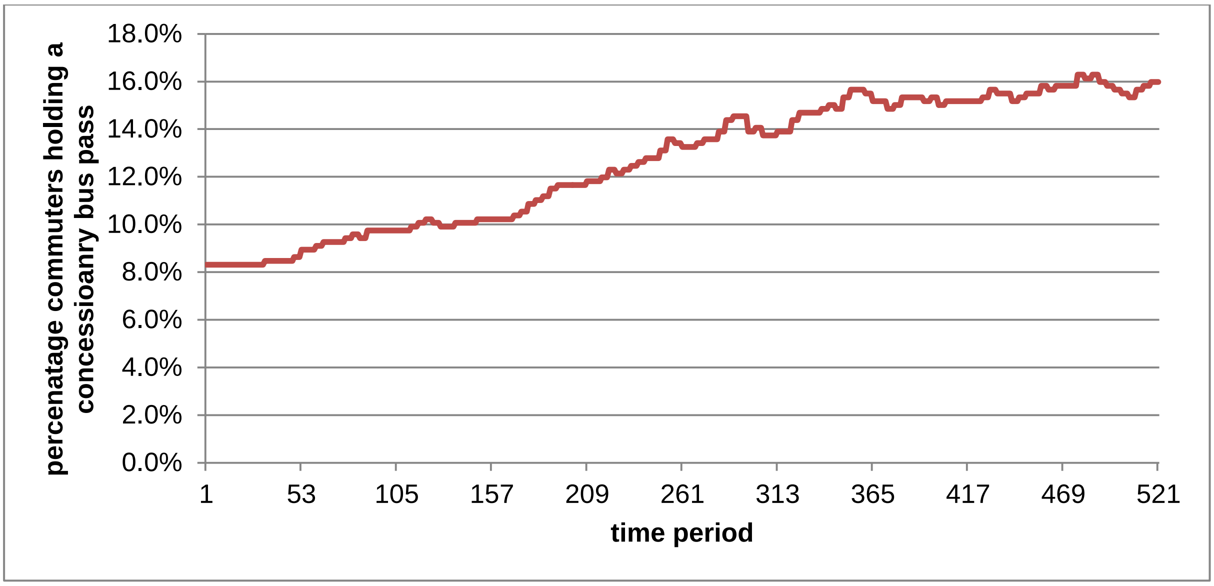


FIGURE 7.10 PERCENTAGE OF COMMUTERS HOLDING A CONCESSIONARY BUS PASS

These results came from a model where there were few people eligible for bus passes, as most commuters are below the age of eligibility for bus passes. Agent based modelling could be applied to trips for all journey purposes in an area to test the wider impacts on the transport system of a change in the qualifying age for concessionary bus passes as the population ages or a reduction in the size of the discount it provides from say 100% to 50% as has been discussed in some policy circles.

This ability to interrogate the model results at an individual level, for example by having a record of the characteristics and attributes of each person, such as their age, in the model at each model time period, enables the modeller to carry out detailed investigations of the model results. In this case the results support the conclusion that the rise in bus usage over time is a result of the availability of concessionary bus passes for the elderly and an increase in the number of people working for longer and so having a bus pass available for use for their commuting trip.

7.4 Modelling at the individual level allows more accurate handling of personal and external constraints

With an agent based modelling approach the model is able to handle detailed information about each person and their circumstances. In traditional transport models the data is held in matrices and trips are usually classified into different matrices according to their current mode, journey purpose, time of day and sometimes income. These divisions are based on the factors that are believed to affect people's preference weighting for time and cost, with all people whose trips are held in the same matrix having the same weighting for time and cost.

The only common explicit treatment of constraints in the matrices comes through the division of people into car available and car not available matrices. Other constraints, such as the lack of a nearby bus service are dealt with indirectly in the model in as far as a long walk time to the nearest bus service results in a high journey time and hence a high generalised journey cost for that option. When a logit model is used for mode choice, a small proportion of trips are still allocated to this mode and transport modellers sometimes take this into account in the modelling process by then over-riding the allocation of trips to that mode by the logit model by setting the number of trips to zero.

In an agent based mode choice model information can be held on the individual circumstances of each person which makes it possible to apply information on personal constraints during the mode choice modelling. The model can have a more accurate assessment of the actual choice set available to each person and exclude those modes which are not feasible, (for example because there is no public transport service available in that particular area at that time of day), or because of the personal circumstances of the individual, (for example they may have mobility issues which means they cannot walk to a public transport stop).

In the survey which provided the dataset for the model used in this research, respondents were asked about some of their personal constraints which could affect their travel choices.

This information is held within the model as part of each agent's characteristics and used to personalise the choice set available to each person. Table 7.5 below shows the personal constraints covered in the survey and how they affected the choice set for individuals in the model. The most significant personal constraints were the lack of a car and the lack of a driving licence. In this model these constraints are assumed to remain constant over time but it would be possible, given sufficient data, to add additional functionality into the model that would change a person's individual constraints over time.

Constraint	Impact on choice set	Number of people	
No car in the household	Car removed from choice set	76	12%
Do not have a car licence	Car removed from choice set	105	17%
Cannot use bus as need to carry things	Bus removed from choice set	31	5%
Cannot use train as need to carry things	Train removed from choice set	22	4%
Cannot use cycle as need to carry things	Cycle removed from choice set	26	4%
Cannot ride a bicycle	Cycle removed from choice set	25	4%
Will not cycle because of the weather	Cycle removed from choice set	33	5%
Need to use car for work	All non-car modes removed from choice set	10	2%

TABLE 7.5 PERSONAL CONSTRAINTS APPLIED TO CHOICE SETS

In the dataset the greatest personal constraint was the lack of a driving licence which affected 17% of people. This was a greater constraint than the measure customarily used in transport models, car availability, as only 12% of people lived in a household with no car. The number of people who said they needed to use their car for work was 2%. Only a small percentage of people (4%) couldn't use a bus, train or cycle as they needed to carry things. 4% of people could not cycle and 5% would not cycle because of the weather so more people had the cycle option removed from their choice set than had the bus or train removed because of personal constraints.

The model was run with the application of personal constraints for scenario 1, which has a trend only change in costs for each mode over time. Table 7.6 below shows the number of users of each mode for scenario 1 without applying any personal constraints and with the application of constraints. Table 7.7 shows the mode shares with and without personal constraints.

Time period	Week	Bus		Car		Train		Cycle		Not travel	
		unc*	con**	unc	Con	unc	Con	unc	con	unc	con
beg year 1	**1**	44	47	411	349	113	177	52	42	6	11
beg year 2	**53**	52	56	408	347	109	170	51	42	6	11
beg year 3	**105**	51	55	409	348	110	171	50	41	6	11
beg year 4	**157**	53	58	408	347	108	168	51	42	6	11
beg year 5	**209**	53	58	409	348	107	167	51	42	6	11
beg year 6	**261**	56	60	411	351	101	161	52	43	6	11
beg year 7	**313**	56	60	413	352	99	160	52	43	6	11
beg year 8	**365**	59	62	410	350	98	159	53	44	6	11
beg year 9	**417**	59	62	410	350	98	159	53	44	6	11
beg year 10	**469**	61	66	409	349	97	156	53	44	6	11
end year 10	**520**	60	65	415	354	92	152	53	44	6	11

* unconstrained ** constrained

TABLE 7.6 SCENARIO 1: NUMBER OF PEOPLE USING EACH MODE WITH AND WITHOUT PERSONAL CONSTRAINTS

Time period	Week	Bus		Car		Train		Cycle		Not travel	
		unc*	con**	unc	Con	unc	con	unc	con	unc	con
beg year 1	**1**	7.0%	7.5%	65.7%	55.8%	18.1%	28.3%	8.3%	6.7%	1.0%	1.8%
beg year 2	**53**	8.3%	8.9%	65.2%	55.4%	17.4%	27.2%	8.1%	6.7%	1.0%	1.8%
beg year 3	**105**	8.1%	8.8%	65.3%	55.6%	17.6%	27.3%	8.0%	6.5%	1.0%	1.8%
beg year 4	**157**	8.5%	9.3%	65.2%	55.4%	17.3%	26.8%	8.1%	6.7%	1.0%	1.8%
beg year 5	**209**	8.5%	9.3%	65.3%	55.6%	17.1%	26.7%	8.1%	6.7%	1.0%	1.8%
beg year 6	**261**	8.9%	9.6%	65.7%	56.1%	16.1%	25.7%	8.3%	6.9%	1.0%	1.8%
beg year 7	**313**	8.9%	9.6%	66.0%	56.2%	15.8%	25.6%	8.3%	6.9%	1.0%	1.8%
beg year 8	**365**	9.4%	9.9%	65.5%	55.9%	15.7%	25.4%	8.5%	7.0%	1.0%	1.8%
beg year 9	**417**	9.4%	9.9%	65.5%	55.9%	15.7%	25.4%	8.5%	7.0%	1.0%	1.8%
beg year 10	**469**	9.7%	10.5%	65.3%	55.8%	15.5%	24.9%	8.5%	7.0%	1.0%	1.8%
end year 10	**520**	9.6%	10.4%	66.3%	56.5%	14.7%	24.3%	8.5%	7.0%	1.0%	1.8%

* unconstrained ** constrained

TABLE 7.7 SCENARIO 1: PERCENTAGE OF PEOPLE USING EACH MODE WITH AND WITHOUT PERSONAL CONSTRAINTS

The biggest impact from the application of constraints is on the number of people using car, which at the beginning reduces the number of car users from 411 to 349, a reduction of 15%. At the end of the modelled time period it reduces the number of car users from 415 to 354, also a reduction of 15%. This highlights the importance of considering the possession of a car licence and car availability in the modelling of mode choice. Models often rely on census data which only identifies which households have a car available rather than the personal level of car availability. Details on the personal level of car availability and whether an individual has a car licence could be obtained from household surveys where respondents are asked to complete a travel diary and answer personal questions but this information is not always requested.

The high number of people without a car licence and/or personal access to a car suggests that this should be included in mode choice modelling as it has a significant impact on mode choice decisions, and imposes a more severe constraint on car usage than the level of household car availability. As the level of car licence acquisition in the UK is falling the inclusion of this constraint in modelling will become even more important.

Some people are also constrained in their choice of mode, not for reasons that are personal to them but because of external constraints, for example because there isn't a public transport service in their area. An agent based model allows for the context or environment of each individual to be considered and for example, public transport can be excluded from their choice set if it is not available within a certain distance of a person's house. Some agent based modelling software packages, such as AnyLogic, allow for the integration of these models with Geographical Information Systems which can for example, make accurate measurements of the distance to bus services. The AnyLogic software update in 2014 brought many common GIS functions into the software.

7.5 Modelling reflects interactions between agents and the environment

An aspect of agent based modelling which distinguishes it from microsimulation is the interaction between agents and the environment in which they exist. This feature means that the model can include feedback between the agents and their world. This is illustrated in a model run in which the available bus service is considered to be part of the environment.

Every six months the bus operator reviews the financial performance of the bus network and if the minimum number of bus passengers required for them to continue running each bus service is not met, then a bus is withdrawn.

The reduction in the number of buses operated may mean that the frequency on a bus route is reduced. This increases the waiting time for passengers, which effectively increases the time of making a journey by bus relative to other modes, leading to some people switching away from bus to other modes. Alternatively the removal of the bus may mean that an entire bus service is withdrawn and the people served by that bus route may no longer have a bus service within a reasonable walking distance of their house.

An agent based model can record spatial data for each agent and this can be used to assess which agents are affected by the reduction in frequency or removal of a bus service. In this study the feedback between the level of provision of bus services in the study area and the availability of the option of a bus for a particular agent is incorporated in the model by specifying:

- the number of bus services at the start of the model run
- the minimum average number of users required for all the bus services to be retained
- the number of agents who no longer have access to a bus service if a bus is withdrawn.

At the end of each time period the number of bus users is counted. If the average number of bus users per service is less than the minimum required, one service is removed. The number of agents who no longer have access to a bus is calculated and starting with the agent with the lowest ID number, that number of agents are set as having no bus available due to this environmental or external constraint. In the mode choice modelling, both personal and environmental constraints are taken into consideration when determining which modes are available in each agent's choice set.

The impact of this feedback between the environment and agents is illustrated in the results for Scenario 3. The input data for this scenario is given in Figure 7.11 below. There is:

- a trend change in bus, rail and car costs without any sudden change in car parking costs as in scenario 1
- the initial cost of travel by bus, car and rail is as set in scenario 1 (£1.60, £1.00 and £2.40 respectively)
- the initial bus time is set to 20 minutes which is the same time as for car and rail journeys. This results in a higher number of bus users than in scenario 1, where the bus time was 35 minutes
- there are concessionary bus fares

The feedback between the number of bus users and the number of bus services was modelled by setting the initial number of buses to 18, the average minimum number of users required for the bus service to remain the same was set at 10 and the number of agents who lost their bus service if a bus was withdrawn was set to 25.

Figure 7.11 below shows the forecast number of bus passengers with and without the feedback between the number of bus passengers (agents) and the number of bus services run (environment). It shows that with bus, car and rail all having the same initial journey time of 20 minutes, and the costs and trend changes in costs as in scenario 1, the number of bus users over time declines slightly even if the number of bus services remains constant.

If though, there is a feedback loop between the number of bus passengers and the number of bus services provided, once the falling number of bus passengers triggers the removal of a bus service, a ‘vicious’ circle of decline in bus use occurs. In the middle of year 4 the number of bus passengers falls below 180, which as there are 18 buses running, means that the average number of passengers per bus is less than 10 and the bus operator withdraws one service. As a consequence 25 people no longer have the bus option available to them for their journey to work. Some of these people would have used the bus if there had been one, so the number of bus users falls further. Combined with the general decline in the level of bus use amongst those who still do have a bus available, as a result of the relative increase in bus fares compared to car costs and rail fares, the overall number of bus users continues to decline and triggers the removal of another bus service. This then in turn accelerates the decline in the number of bus users.

SCENARIO	3									
INITIAL JOURNEY DETAILS										
MODE	Cost	CO2	Time		PTP % households visited a year				0%	
	(pence)	(kg)	(mins)		% house movers leave area				30%	
BUS	160	1	20		% job changers leave area				20%	
CAR	100	3	20						Years	Weeks
TRAIN	240	1	20		Minimum age for bus pass				61	48
CYCLE	0	0	60		Minimum age for new agents				17	0
					Maximum age for new agents				60	0
					Retirement age				68	0

	Yr1 Q1	**Yr1 Q3**	**Yr2 Q1**	**Yr2 Q3**	**Yr3 Q1**	**Yr3 Q3**	**Yr4 Q1**	**Yr4 Q3**	**Yr5 Q1**	**Yr5 Q3**
COST CHANGES										
BUS	1.0000	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198
CAR	1.0000	0.9909	0.9909	0.9941	0.9941	0.9919	0.9919	0.9925	0.9925	0.9923
TRAIN	1.0000	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050
TIME CHANGES										
BUS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
CAR	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TRAIN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
% PAY TO PARK										
ZONE 1	50	50	50	50	50	50	50	50	50	50
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	50	50	50	50	50	50	50	50	50	50
PARK CHARGE (pence)										
ZONE 1	150	150	150	150	150	150	150	150	150	150
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	150	150	150	150	150	150	150	150	150	150

	Yr6 Q1	**Yr6 Q3**	**Yr7 Q1**	**Yr7 Q3**	**Yr8 Q1**	**Yr8 Q2**	**Yr9 Q1**	**Yr9 Q3**	**Yr10 Q1**	**Yr10 Q3**
COST CHANGES										
BUS	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198	1.0198
CAR	0.9923	0.9920	0.9920	0.9914	0.9914	0.9902	0.9902	0.9898	0.9898	0.9899
TRAIN	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050	1.0050
TIME CHANGES										
BUS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
CAR	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TRAIN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
% PAY TO PARK										
ZONE 1	50	50	50	50	50	50	50	50	50	50
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	50	50	50	50	50	50	50	50	50	50
PARK CHARGE (pence)										
ZONE 1	150	150	150	150	150	150	150	150	150	150
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	150	150	150	150	150	150	150	150	150	150

FIGURE 7.11 INPUT VALUES FOR SCENARIO 3

The number of bus passengers, with and without this feedback of the bus operator withdrawing a bus service when the average number of bus passengers falls below his financial viable number of passengers, is shown in Table 7.8 below. This table also shows the number of buses that are running.

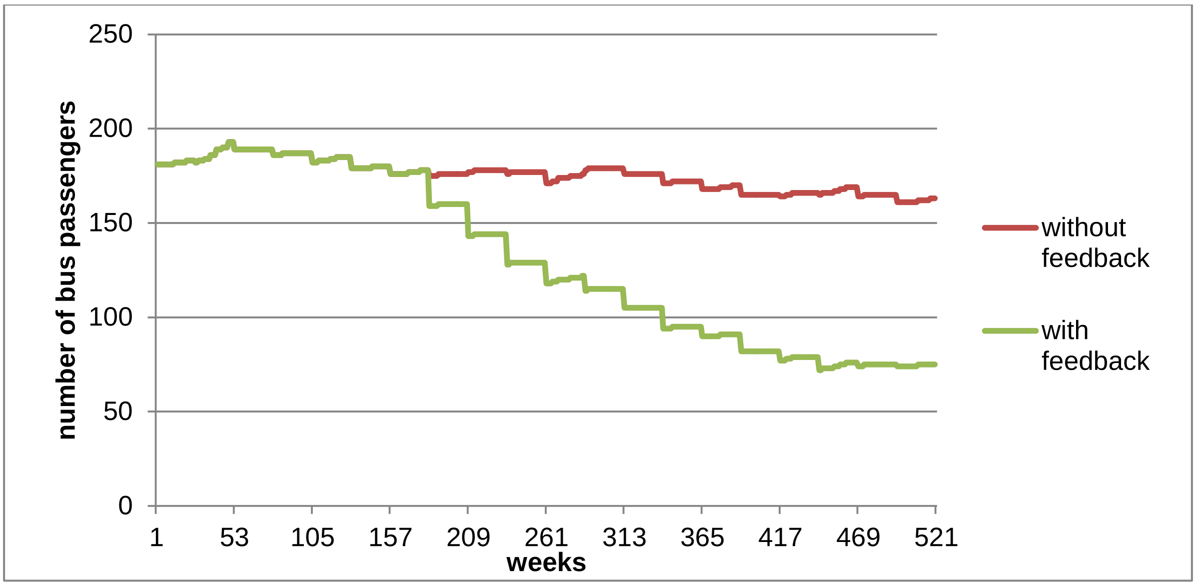


FIGURE 7.12 NUMBER OF BUS PASSENGERS WITH AND WITHOUT THE NUMBER OF BUS SERVICES RUNNING BEING DEPENDENT UPON THE NUMBER OF BUS PASSENGERS, WITH CONCESSIONARY FARES

Time period		Without feedback		With feedback	
		Passengers	Buses	Passengers	Buses
beg year 1	**1**	181	18	181	18
end year 1	**52**	193	18	193	18
beg year 2	**53**	189	18	189	18
end year 2	**104**	187	18	187	18
beg year 3	**105**	182	18	182	18
end year 3	**156**	180	18	180	18
beg year 4	**157**	176	18	176	18
end year 4	**208**	176	18	160	16
beg year 5	**209**	177	18	143	16
end year 5	**260**	177	18	129	14
beg year 6	**261**	171	18	118	14
end year 6	**312**	179	18	115	12
beg year 7	**313**	176	18	105	12
end year 7	**364**	172	18	95	10
beg year 8	**365**	168	18	90	10
end year 8	**416**	165	18	82	8
beg year 9	**417**	164	18	77	8
end year 9	**468**	169	18	76	7
beg year 10	**469**	164	18	74	7
end year 10	**520**	163	18	75	7

TABLE 7.8 NUMBER OF BUS PASSENGERS, WITH AND WITHOUT FEEDBACK, WITH CONCESSIONARY FARES

Without the presence of concessionary fares the decline in bus numbers is even more severe. Figure 7.13 and Table 7.9 below show the number of bus passengers, with and without feedback, when there are no concessionary fares.

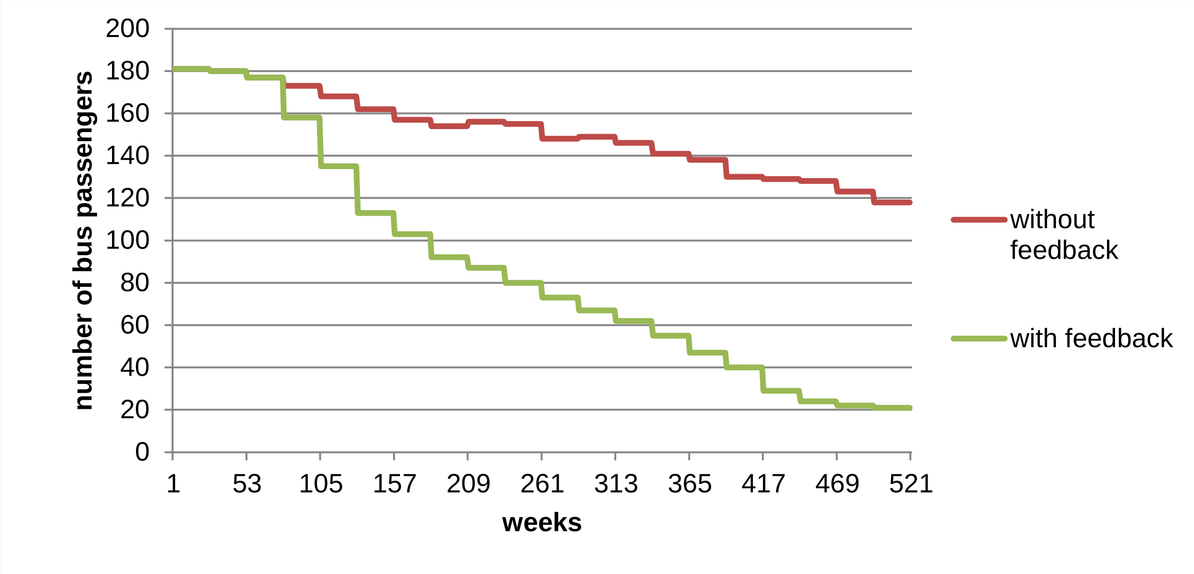


FIGURE 7.13 NUMBER OF BUS PASSENGERS WITH AND WITHOUT FEEDBACK, WITHOUT CONCESSIONARY FARES

Time period		Without feedback		With feedback	
		Passengers	Buses	Passengers	Buses
beg year 1	**1**	181	18	181	18
end year 1	**52**	180	18	180	18
beg year 2	**53**	177	18	177	18
end year 2	**104**	173	18	158	16
beg year 3	**105**	168	18	135	16
end year 3	**156**	162	18	113	14
beg year 4	**157**	157	18	103	14
end year 4	**208**	154	18	92	12
beg year 5	**209**	156	18	87	12
end year 5	**260**	155	18	80	10
beg year 6	**261**	148	18	73	10
end year 6	**312**	149	18	67	8
beg year 7	**313**	146	18	62	8
end year 7	**364**	141	18	55	6
beg year 8	**365**	138	18	47	6
end year 8	**416**	130	18	40	4
beg year 9	**417**	129	18	29	4
end year 9	**468**	128	18	24	2
beg year 10	**469**	123	18	22	2
end year 10	**520**	118	18	21	2

TABLE 7.9 NUMBER OF BUS PASSENGERS, WITH AND WITHOUT FEEDBACK, WITHOUT CONCESSIONARY FARES

The difference that the presence of concessionary fares makes in slowing down the decline in the number of bus passengers and the number of bus services run is shown in Figure 7.14. This shows the number of bus passengers when there is feedback between the number of bus passengers and the number of bus services, with and without the availability of concessionary fares.

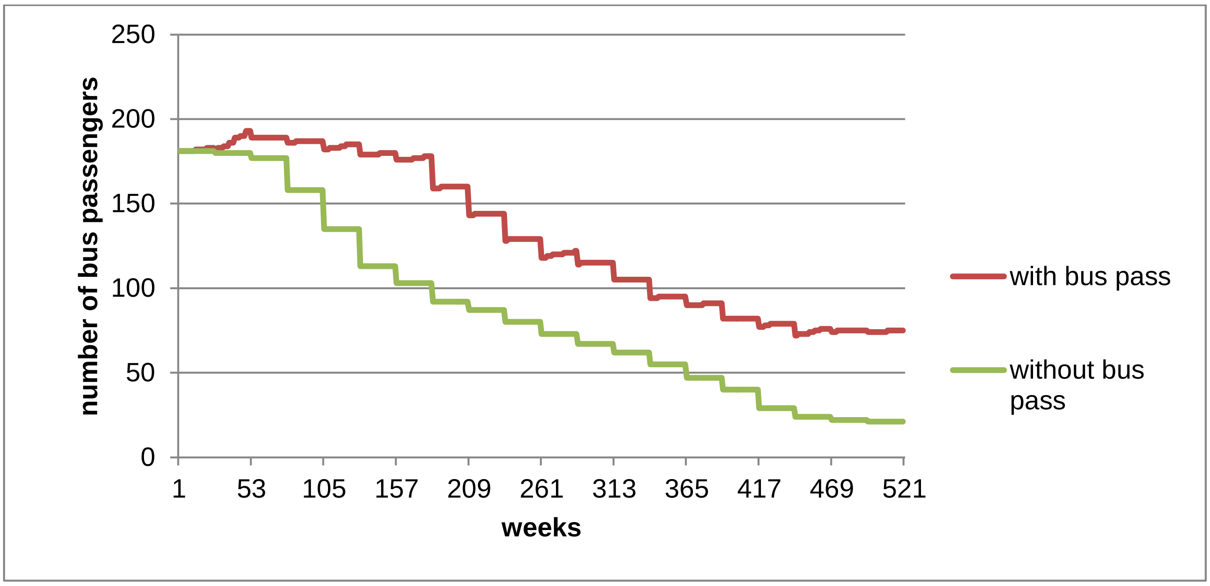


FIGURE 7.14 NUMBER OF BUS PASSENGERS, WITH FEEDBACK, WITH AND WITHOUT CONCESSIONARY FARES

The model also shows the impact of the decline in the number of bus services on the number of people using other modes. Tables 7.10 and 7.11 show the total number of people using each mode, with and without concessionary fares.

In the early years, more of the people who would have used the bus if it were available switch to rail rather than car. For example at the end of year 2 the number of bus passengers falls by 23, from 158 to 135. Three (13%) of these people switch to car, 18 (78%) to train and 2 (9%) to cycling. In later years though, when rail has become relatively even more expensive than car than it was in earlier years, a greater proportion of the people who would otherwise have used bus switch to car rather than rail. For example at the end of year 7, of the 8 people who leave bus, 3 (37.5%) switch to car, 3 (37.5%) switch to rail, 1 (12.5%) switches to cycling and 1 (12.5%) person chooses not to travel.

Time period		Bus		Car		Train		Cycle		Not travel	
beg year 1	**1**	181	28.9%	317	50.6%	81	12.9%	36	5.8%	11	1.8%
end year 1	**52**	180	28.8%	320	51.1%	78	12.5%	37	5.9%	11	1.8%
beg year 2	**53**	177	28.3%	323	51.6%	79	12.6%	36	5.8%	11	1.8%
end year 2	**104**	158	25.2%	325	51.9%	91	14.5%	38	6.1%	14	2.2%
beg year 3	**105**	135	21.6%	328	52.4%	109	17.4%	40	6.4%	14	2.2%
end year 3	**156**	113	18.1%	332	53.0%	126	20.1%	41	6.5%	14	2.2%
beg year 4	**157**	103	16.5%	338	54.0%	130	20.8%	41	6.5%	14	2.2%
end year 4	**208**	92	14.7%	340	54.3%	137	21.9%	43	6.9%	14	2.2%
beg year 5	**209**	87	13.9%	342	54.6%	140	22.4%	43	6.9%	14	2.2%
end year 5	**260**	80	12.8%	343	54.8%	144	23.0%	43	6.9%	16	2.6%
beg year 6	**261**	73	11.7%	345	55.1%	149	23.8%	43	6.9%	16	2.6%
end year 6	**312**	67	10.7%	350	55.9%	150	24.0%	43	6.9%	16	2.6%
beg year 7	**313**	62	9.9%	354	56.5%	151	24.1%	43	6.9%	16	2.6%
end year 7	**364**	55	8.8%	355	56.7%	156	24.9%	44	7.0%	16	2.6%
beg year 8	**365**	47	7.5%	358	57.2%	159	25.4%	45	7.2%	17	2.7%
end year 8	**416**	40	6.4%	359	57.3%	165	26.4%	45	7.2%	17	2.7%
beg year 9	**417**	29	4.6%	363	58.0%	172	27.5%	45	7.2%	17	2.7%
end year 9	**468**	24	3.8%	366	58.5%	174	27.8%	45	7.2%	17	2.7%
beg year 10	**469**	22	3.5%	367	58.6%	175	28.0%	45	7.2%	17	2.7%
end year 10	**520**	21	3.4%	367	58.6%	176	28.1%	45	7.2%	17	2.7%

TABLE 7.10 NUMBER OF USERS BY MODE, WITH FEEDBACK AND NO CONCESSIONARY FARES

Time period		Bus		Car		Train		Cycle		Not travel	
beg year 1	**1**	181	28.9%	317	50.6%	81	12.9%	36	5.8%	11	1.8%
end year 1	**52**	193	30.8%	309	49.4%	77	12.3%	36	5.8%	11	1.8%
beg year 2	**53**	189	30.2%	313	50.0%	78	12.5%	35	5.6%	11	1.8%
end year 2	**104**	187	29.9%	313	50.0%	79	12.6%	36	5.8%	11	1.8%
beg year 3	**105**	182	29.1%	316	50.5%	81	12.9%	36	5.8%	11	1.8%
end year 3	**156**	180	28.8%	317	50.6%	81	12.9%	37	5.9%	11	1.8%
beg year 4	**157**	176	28.1%	321	51.3%	81	12.9%	37	5.9%	11	1.8%
end year 4	**208**	160	25.6%	320	51.1%	93	14.9%	39	6.2%	14	2.2%
beg year 5	**209**	143	22.8%	319	51.0%	109	17.4%	41	6.5%	14	2.2%
end year 5	**260**	129	20.6%	319	51.0%	124	19.8%	40	6.4%	14	2.2%
beg year 6	**261**	118	18.8%	323	51.6%	131	20.9%	40	6.4%	14	2.2%
end year 6	**312**	115	18.4%	322	51.4%	135	21.6%	40	6.4%	14	2.2%
beg year 7	**313**	105	16.8%	328	52.4%	139	22.2%	40	6.4%	14	2.2%
end year 7	**364**	95	15.2%	331	52.9%	144	23.0%	40	6.4%	16	2.6%
beg year 8	**365**	90	14.4%	331	52.9%	147	23.5%	41	6.5%	17	2.7%
end year 8	**416**	82	13.1%	334	53.4%	152	24.3%	41	6.5%	17	2.7%
beg year 9	**417**	77	12.3%	339	54.2%	151	24.1%	42	6.7%	17	2.7%
end year 9	**468**	76	12.1%	337	53.8%	152	24.3%	44	7.0%	17	2.7%
beg year 10	**469**	74	11.8%	338	54.0%	153	24.4%	44	7.0%	17	2.7%
end year 10	**520**	75	12.0%	338	54.0%	153	24.4%	43	6.9%	17	2.7%

TABLE 7.11 NUMBER OF USERS BY MODE, WITH FEEDBACK AND CONCESSIONARY FARES

7.6 Modelling at the individual level allows the inclusion of habitual behaviour

A feature of agent based modelling is that the history of each agent can be recorded and is available to inform future decisions on mode choice. The agent based model of commuting trips developed in this research is used to consider the impact of habitual behaviour on the number of people using each mode. In Triandis' Theory of Interpersonal Behaviour (1977) intention and habit are separate pathways leading, after the impact of facilitating conditions, to the final behaviour undertaken. As experience of a behaviour is acquired, the influence of habit increases, and that of intention declines. For a behaviour such as commuting, which is a regular activity performed frequently, the current behaviour becomes routine and is undertaken automatically (Darnton, 2008).

Placing this theory in the language of Critical Realism, the observed travel mode is in the domain of the empirical. It is the result of two named underlying mechanisms, intention and habits, which can be re-inforcing or counter-acting, and the context of their commuting mode decisions, through which facilitating conditions affect the choice of mode. These mechanisms exist in the real world, are activated by events in the actual world and the resulting mode choice observed in the empirical world.

The basic ABM in this study was extended to include habitual behaviour. The model assumes that commuters do not re-assess their travel mode in every time period, but rather, by default display habitual behaviour, using the same mode as in the previous time period, unless an event occurs which causes them to re-consider their mode choice. The events in this model which trigger a re-appraisal of their mode choice are moving house, changing job and acquiring a concessionary bus pass.

Consideration was given to the inclusion of thresholds in the model, for example assigning to each agent a tolerance level for the size of fare increase or increase in journey time they would accept. When these levels exceed the threshold for each agent this would act as a trigger event to prompt the reconsideration of the mode they used for commuting. The design decision was taken to exclude threshold trigger events due to the lack of available evidence on the levels that should be set for each agent.

When a commuter is triggered in the model to re-consider their travel mode they calculate the time, cost and carbon impact of each of the four alternative modes, car, bus, train and cycle and the option of not travelling at all. They then choose the mode which gives them the greatest utility, given their personal weightings or preferences for time, cost, carbon impacts and the general characteristics of each mode. It is quite possible that the mode that gives them the maximum utility once they re-evaluate their options will be the same mode that they are currently using. Only if an alternative mode has a higher utility than the mode they are currently using will they change mode.

The impact of this assumption of habitual behaviour on the number of people using each mode is seen in Tables 7.12 below. These results are based on the availability of concessionary bus passes and with personal constraints on mode choice. In Table 7.12 the results are shown for the inclusion of habitual behaviour in scenario 1, which has trend changes in costs by mode.

The number of train users, when there is habitual behaviour, is slightly higher than when there is an intentional, deliberative mode choice decision made by everyone in every time period (called deliberative behaviour in this report). This is because users continue to travel by rail even though the cost of commuting by car is becoming relatively cheaper, until a trigger event causes them to re-consider their mode choice. The forecast number of rail users with and without habitual behaviour for scenario 1 is shown in Figure 7.15 below.

Time period		Bus		Car		Train		Cycle		Not travel	
		delib	habit	Delib	habit	delib	habit	delib	habit	delib	habit
beg year 1	**1**	47.0	47.00	349	349.00	177	177.00	42	42.00	11	11.00
end year 1	**52**	56.0	49.67	347	348.96	170	174.37	42	42.00	11	11.00
beg year 2	**53**	55.0	49.71	348	348.93	171	174.36	41	42.00	11	11.00
end year 2	**104**	58.0	53.61	347	347.09	168	172.41	42	41.89	11	11.00
beg year 3	**105**	58.0	53.67	348	347.12	167	172.31	42	41.90	11	11.00
end year 3	**156**	60.0	60.49	351	345.35	161	167.20	43	41.96	11	11.00
beg year 4	**157**	60.0	60.48	352	345.41	160	167.14	43	41.97	11	11.00
end year 4	**208**	62.0	62.54	350	344.75	159	165.39	44	42.32	11	11.00
beg year 5	**209**	62.0	62.62	350	344.76	159	165.27	44	42.35	11	11.00
end year 5	**260**	66.0	66.23	349	345.66	156	161.42	44	41.69	11	11.00
beg year 6	**261**	65.0	66.18	354	345.80	152	161.28	44	41.74	11	11.00
end year 6	**312**	67.0	68.63	354	346.52	150	157.73	44	42.12	11	11.00
beg year 7	**313**	66.0	68.55	354	346.65	151	157.65	44	42.15	11	11.00
end year 7	**364**	68.0	68.66	354	347.97	149	155.91	44	42.46	11	11.00
beg year 8	**365**	66.0	68.63	354	348.08	150	155.83	44	42.46	12	11.00
end year 8	**416**	70.0	72.32	353	347.49	147	151.79	44	43.25	12	11.15
beg year 9	**417**	69.0	72.25	355	347.59	146	151.76	44	43.25	12	11.15
end year 9	**468**	71.0	74.09	355	348.02	144	148.55	44	43.34	12	12.00
beg year 10	**469**	71.0	74.10	356	348.07	143	148.49	44	43.34	12	12.00
end year 10	**520**	70.0	72.37	356	350.44	144	147.77	44	43.42	12	12.00

TABLE 7.12 NUMBER OF USERS BY MODE, HABITUAL AND ALWAYS DELIBERATIVE CHOICES FOR SCENARIO 1

Conversely the number of car users is lower with habitual behaviour as people remain on public transport rather then switch to cars as the cost of motoring falls. This is illustrated in Figure 7.16 below. There are fewer car users when habitual behaviour is included in the model because people do not respond immediately to the declining cost of driving relative to using public transport. The impact of the reduced cost of driving does not affect their mode choice until they receive a trigger event and reconsider the mode they use whereas with always deliberative behaviour the impact of the reduction in car costs is reflected immediately in the number of people using car.

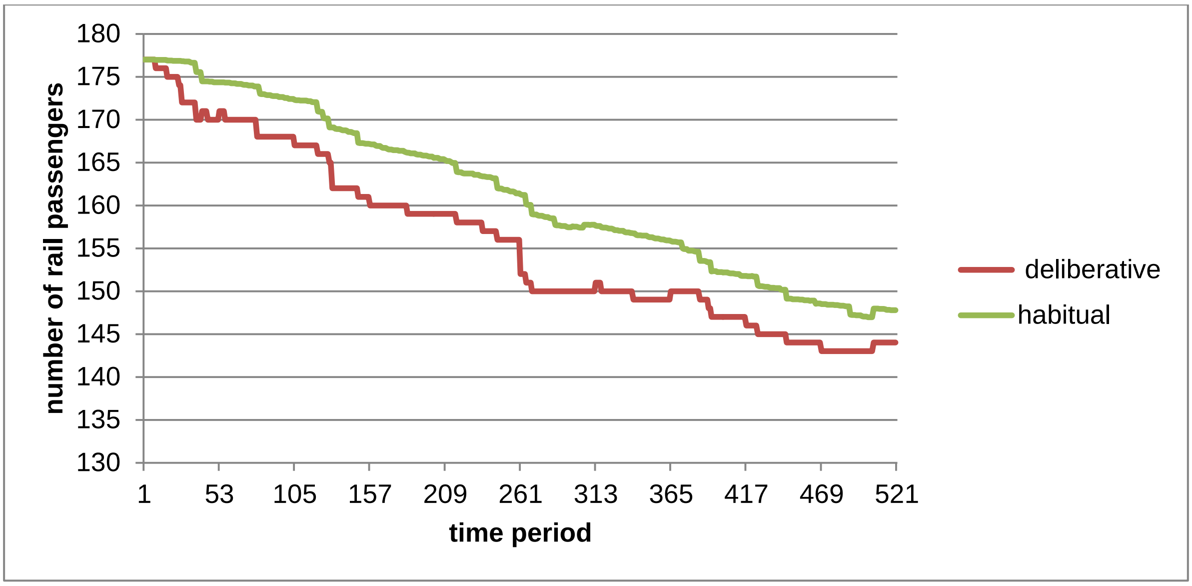


FIGURE 7.15 NUMBER OF RAIL PASSENGERS WITH DELIBERATIVE AND HABITUAL MODE CHOICE

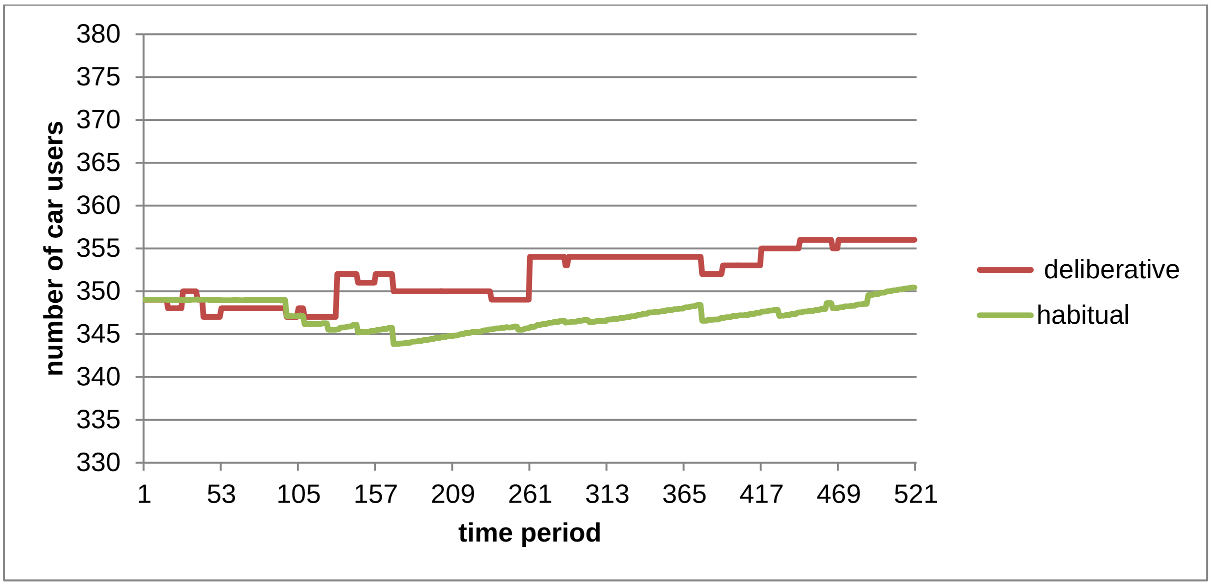


FIGURE 7.16 NUMBER OF CAR USERS WITH DELIBERATIVE AND HABITUAL MODE CHOICE

When there are discontinuities in the trend rate of change of the cost of using modes, the differences between the estimates of the number of users of each mode become much more pronounced when habitual behaviour is introduced into the model. This is illustrated in Table 7.13 below and in Figures 7.17 to 7.19 for rail, car and bus users which show the results for the inclusion of habitual behaviour for scenario 2, which has periodic increases in parking costs.

These differences can be large and significant as can be seen from the results for scenario 2, which has periodic increases in car park charges and in the proportion of people paying to park. For example, at the beginning of year 6 when parking charges are increased, the always deliberative model has 245 car users but with habitual behaviour there are still 345 users (a difference of 35%). A few months later the difference is still apparent but not so high (27%) as car users experience trigger events and some of them decide to switch from car to other modes.

Time period		Bus		Car		Train		Cycle		Not travel	
		delib	habit	delib	habit	delib	habit	Delib	habit	delib	habit
beg year 1	**1**	47.0	47.00	349	349.00	177	177.00	42	42.00	11	11.00
end year 1	**52**	56.0	49.52	347	348.96	170	174.52	42	42.00	11	11.00
beg year 2	**53**	55.0	49.57	348	348.89	171	174.54	41	42.00	11	11.00
end year 2	**104**	58.0	53.36	347	347.20	168	172.56	42	41.88	11	11.00
beg year 3	**105**	62.0	53.53	319	346.78	191	172.77	44	41.92	10	11.00
end year 3	**156**	63.0	60.67	325	339.82	183	172.41	45	42.29	10	10.81
beg year 4	**157**	63.0	60.71	326	339.64	182	172.54	45	42.30	10	10.81
end year 4	**208**	66.0	63.98	325	334.45	180	174.02	45	42.84	10	10.71
beg year 5	**209**	65.0	64.01	326	334.32	180	174.10	45	42.86	10	10.71
end year 5	**260**	68.0	67.52	326	332.02	177	172.95	45	42.91	10	10.60
beg year 6	**261**	84.0	67.75	245	331.15	239	173.54	48	42.97	10	10.59
end year 6	**312**	86.0	74.60	246	312.46	236	184.51	48	43.90	10	10.53
beg year 7	**313**	83.0	74.73	247	311.70	238	185.10	48	43.95	10	10.52
end year 7	**364**	86.0	77.66	247	297.87	235	195.32	48	44.66	10	10.49
beg year 8	**365**	84.0	78.42	225	295.63	257	196.78	49	44.68	11	10.49
end year 8	**416**	88.0	84.35	225	280.48	253	204.65	49	45.94	11	10.58
beg year 9	**417**	87.0	84.35	230	279.92	249	205.18	49	45.96	11	10.59
end year 9	**468**	90.0	88.59	230	267.53	246	212.76	49	46.43	11	10.69
beg year 10	**469**	90.0	88.54	231	267.19	245	213.13	49	46.45	11	10.69
end year 10	**520**	90.0	89.57	231	259.74	245	219.06	49	46.88	11	10.75

TABLE 7.13 NUMBER OF USERS BY MODE, HABITUAL AND ALWAYS DELIBERATIVE CHOICES FOR SCENARIO 2

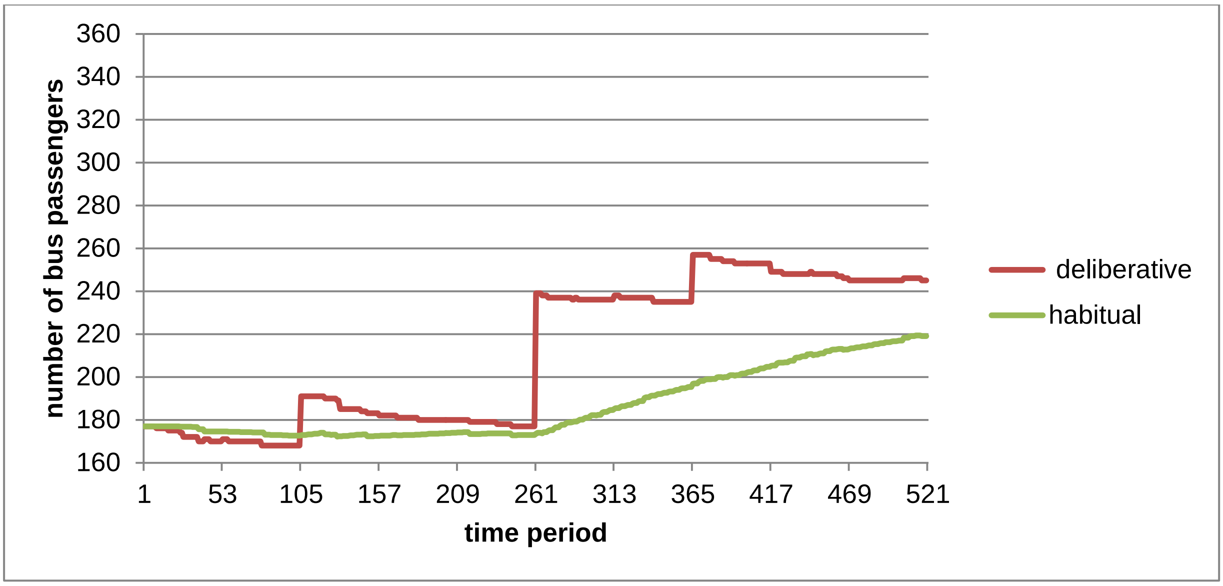


FIGURE 7.17 NUMBER OF BUS USERS WITH DELIBERATIVE AND HABITUAL MODE CHOICE FOR SCENARIO 2

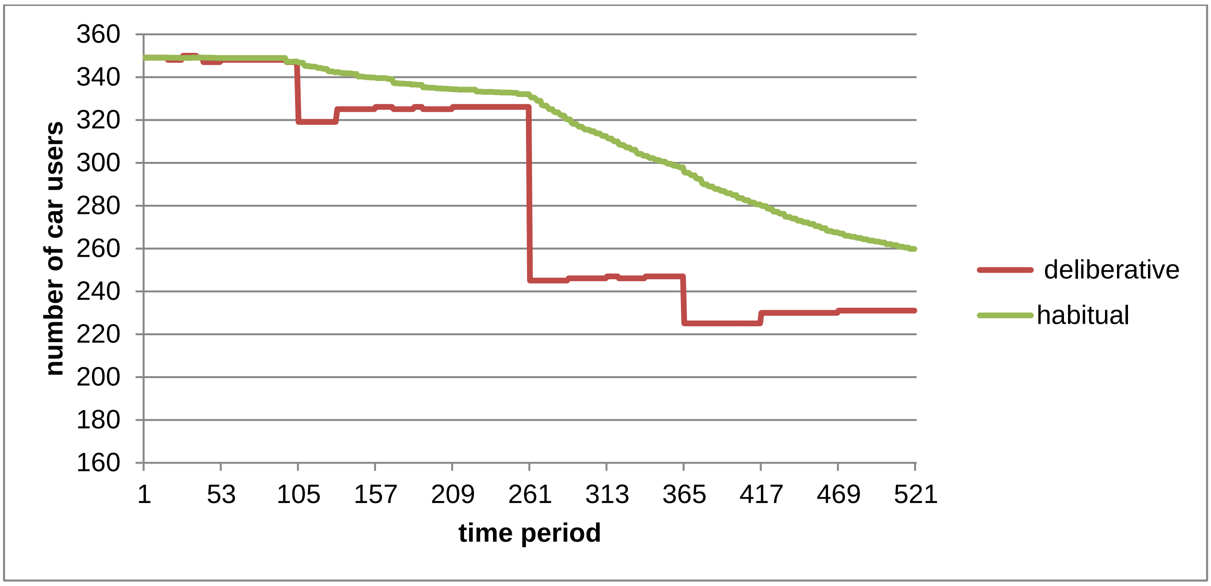


FIGURE 7.18 NUMBER OF CAR USERS WITH DELIBERATIVE AND HABITUAL MODE CHOICE FOR SCENARIO 2

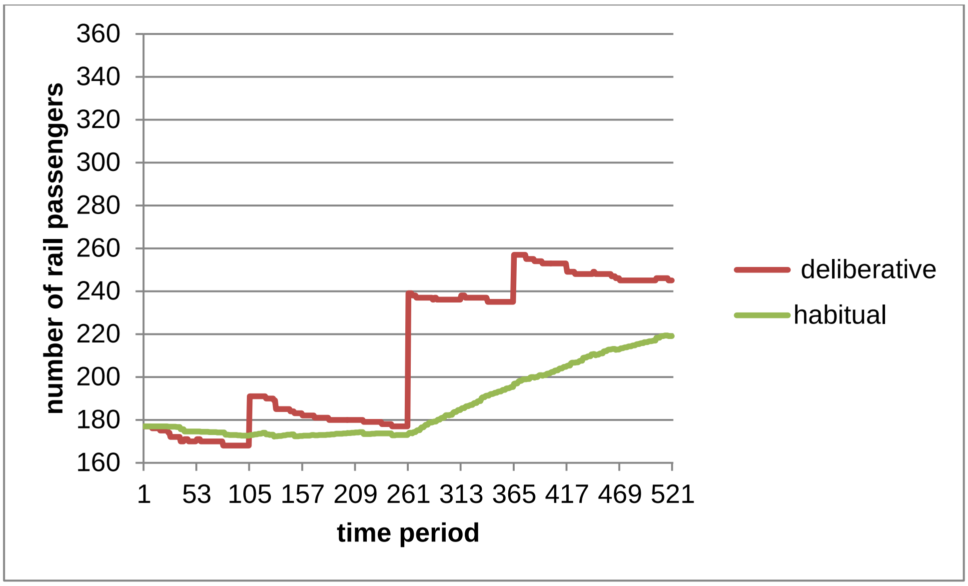


FIGURE 7.19 NUMBER OF RAIL USERS WITH DELIBERATIVE AND HABITUAL MODE CHOICE FOR SCENARIO 2

7.7 Modelling shows lags in response to changes in cost

The incorporation of habitual behaviour in the mode choice modelling allows for the emergence of information about the strength of lags in the response of overall mode share to changes in the relative cost of travel by each mode. This is illustrated in Figure 7.18 above for car users, where after the first increase in car costs at the end of year one the switch away from car use is more gradual with a habitual mode choice model than with a deliberative model. By the end of year 5 the numbers are similar as most people have experienced a trigger event and re-assessed their mode choice. However the large increase in car costs due to the increase in car park charges at the end of year 6 again causes a difference in the number of car users depending on whether a habitual or deliberative mode choice model is assumed. Over time, the number of car users rises with the habitual model, but the model shows a lag in response to the change in parking costs.

The modelling of the lag in responses makes it possible to estimate both the short and long term cost elasticities implied by the model results. Hanly, Dargay and Goodwin (2002) reviewed published estimates of price elasticities in the demand for road travel and found consistent evidence that there is a difference in short and long term elasticities and 'that studies using methods which allow explicit estimation of short run and long run elasticities separately, nearly always find that the long run effect is substantially higher than the short run effect'. This is the same effect as shown here; the long run impact of the change in car costs is greater than the short term effect.

The comparison of the elasticities produced by the model with those estimated from longitudinal studies provides a means of validating this agent based model by testing whether the patterns produced by the model replicate those observed in the real world. The model incorporating habitual behaviour was run to test the impact on bus patronage over time from a change in bus fares.

Dargay and Hanly (1999) estimated bus fare elasticities from observed data using a dynamic econometric approach to estimate elasticities over time. Figure 7.20 below shows how the change in the number of bus passengers following a change in bus fares varies over time. Dargay and Hanly (1999) estimated the short run elasticity at -0.4 and the long run elasticity at -0.9. The short term was around two years after the fares increase, with a 10% increase in bus fares leading to a 4% decrease in patronage. The long term was around seven years after the fares change, with the 10% increase in bus fares leading to a 9% decrease in patronage.

Balcombe *et al.* (2004) provided a short run elasticity of demand with respect to bus fares for fare paying passengers of -0.42 based on 33 UK studies, a medium term elasticity of between -0.5 and -0.6 based on 2 studies and a long term elasticity of -1.01 from 3 studies. They acknowledged that their long term elasticity was a higher value than other studies and considered this might be caused by the use of different methodologies for estimating the fares elasticity.

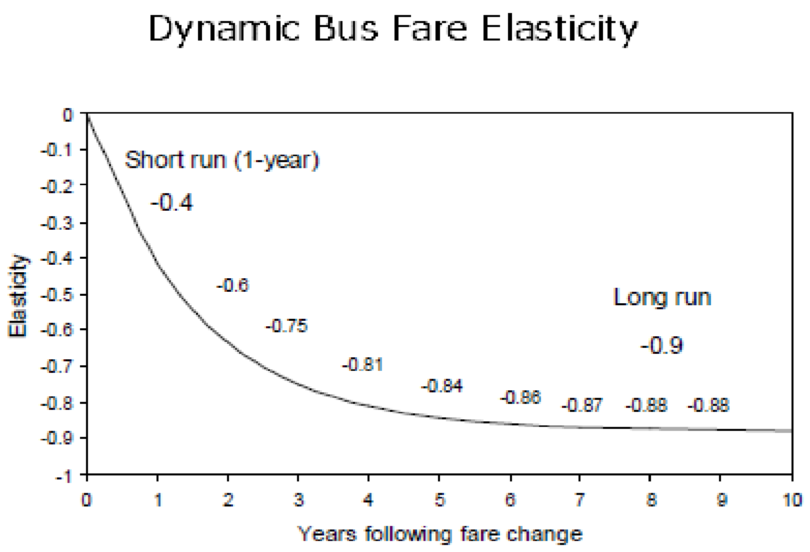


FIGURE 7.20 SHORT AND LONG TERM BUS FARE ELASTICITIES

Source: Dargay and Hanly, 1999

Scenario 4 was run to test the responsiveness of the number of bus passengers to a change in bus fares. The input values used in this scenario are shown in Figure 7.21 below. The initial times and costs of travel by bus, car, train and cycle are the same as for Scenario 1, but all journey times and costs are kept constant over time except for bus cost where there is a 10% increase in bus fares at the beginning of year 2. The model was run without any concessionary bus passes.

SCENARIO	4									
INITIAL JOURNEY DETAILS										
MODE	Cost (pence)	Time (mins)	CO2 (kg)		PTP % households visited a year				0%	
					% house movers leave area				30%	
BUS	160	35	1		% job changers leave area				20%	
CAR	100	20	3						Years	Weeks
TRAIN	240	20	1		Minimum age for bus pass				101	0
CYCLE	0	60	0		Minimum age for new agents				17	0
					Maximum age for new agents				60	0
					Retirement age				68	0

	Yr1 Q1	**Yr1 Q3**	**Yr2 Q1**	**Yr2 Q3**	**Yr3 Q1**	**Yr3 Q3**	**Yr4 Q1**	**Yr4 Q3**	**Yr5 Q1**	**Yr5 Q3**
COST CHANGES										
BUS	1.0000	1.0000	**1.1000**	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
CAR	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TRAIN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TIME CHANGES										
BUS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
CAR	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TRAIN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
% PAY TO PARK										
ZONE 1	50	50	50	50	50	50	50	50	50	50
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	50	50	50	50	50	50	50	50	50	50
PARK CHARGE (pence)										
ZONE 1	150	150	150	150	150	150	150	150	150	150
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	150	150	150	150	150	150	150	150	150	150

	Yr6 Q1	**Yr6 Q3**	**Yr7 Q1**	**Yr7 Q3**	**Yr8 Q1**	**Yr8 Q2**	**Yr9 Q1**	**Yr9 Q3**	**Yr10 Q1**	**Yr10 Q3**
COST CHANGES										
BUS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
CAR	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TRAIN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TIME CHANGES										
BUS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
CAR	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TRAIN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
% PAY TO PARK										
ZONE 1	50	50	50	50	50	50	50	50	50	50
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	50	50	50	50	50	50	50	50	50	50
PARK CHARGE (pence)										
ZONE 1	150	150	150	150	150	150	150	150	150	150
ZONE 2	0	0	0	0	0	0	0	0	0	0
ZONE 3	150	150	150	150	150	150	150	150	150	150

FIGURE 7.21 INPUT VALUES, SCENARIO 4

The number of bus passengers at each time period, with and without personal constraints on travel choices, is shown in Table 7.11 below.

Time period		Unconstrained		Constrained	
		Habit	Delib	Habit	Delib
beg year 1	**1**	44.00	44.00	47.00	47.00
end year 1	**52**	44.00	44.00	47.00	47.00
beg year 2	**53**	43.96	39.00	46.96	42.00
end year 2	**104**	42.95	39.00	45.95	42.00
beg year 3	**105**	42.91	39.00	45.91	42.00
end year 3	**156**	42.17	39.00	45.17	42.00
beg year 4	**157**	42.15	39.00	45.15	42.00
end year 4	**208**	41.55	39.00	44.55	42.00
beg year 5	**209**	41.54	39.00	44.54	42.00
end year 5	**260**	41.00	39.00	44.00	42.00
beg year 6	**261**	40.96	39.00	43.96	42.00
end year 6	**312**	40.55	39.00	43.55	42.00
beg year 7	**313**	40.54	39.00	43.54	42.00
end year 7	**364**	40.31	39.00	43.31	42.00
beg year 8	**365**	40.28	39.00	43.28	42.00
end year 8	**416**	39.70	39.00	42.70	42.00
beg year 9	**417**	39.70	39.00	42.70	42.00
end year 9	**468**	39.51	39.00	42.51	42.00
beg year 10	**469**	39.51	39.00	42.51	42.00
end year 10	**520**	39.41	39.00	42.41	42.00

TABLE 7.14 NUMBER OF BUS USERS AFTER A 10% INCREASE IN BUS FARES

When the model is run in deliberative mode, the reaction to the 10% increase in bus fares, at the beginning of year 2 is instantaneous as shown in table 14 above. For unconstrained demand, the number of bus passengers falls from 44 to 39 passengers (11%) and for constrained demand the number of passengers falls from 47 to 42 (11%), implying an elasticity of -1.1%. This elasticity value is a long term elasticity but as the model implements the full effect of the fares increase immediately it is also the short term elasticity which means that in deliberative mode the model over-estimates the size of the response to a fares increase in the short term.

When the model incorporates habitual behaviour, the number of bus passengers falls more gradually. The implied fares elasticities are shown in Table 7.15 below. After one year the fares elasticity is -0.26 and -0.24 for unconstrained and constrained demand respectively, after two years the fares elasticity is -0.45 and -0.42, after seven years it is -0.98 and -0.91 and finally after eight years is -1.00 and -0.94 .

	Unconstrained		Constrained	
Years	Passengers	Elasticity	Passengers	Elasticity
1	42.85	-0.26	45.85	-0.24
2	42.02	-0.45	45.02	-0.42
3	41.45	-0.58	44.45	-0.54
4	41.02	-0.68	44.02	-0.63
5	40.49	-0.80	43.49	-0.75
6	40.97	-0.86	43.97	-0.81
7	40.49	-0.98	43.49	-0.91
8	40.20	-1.00	43.20	-0.94

TABLE 7.15 IMPLIED FARE ELASTICITIES OVER TIME, WITH AND WITHOUT PERSONAL CONSTRAINTS, WITH HABITUAL BEHAVIOUR

The number of bus passengers over time for the model run with personal constraints and habitual behaviour is shown in Figure 7.22 below.

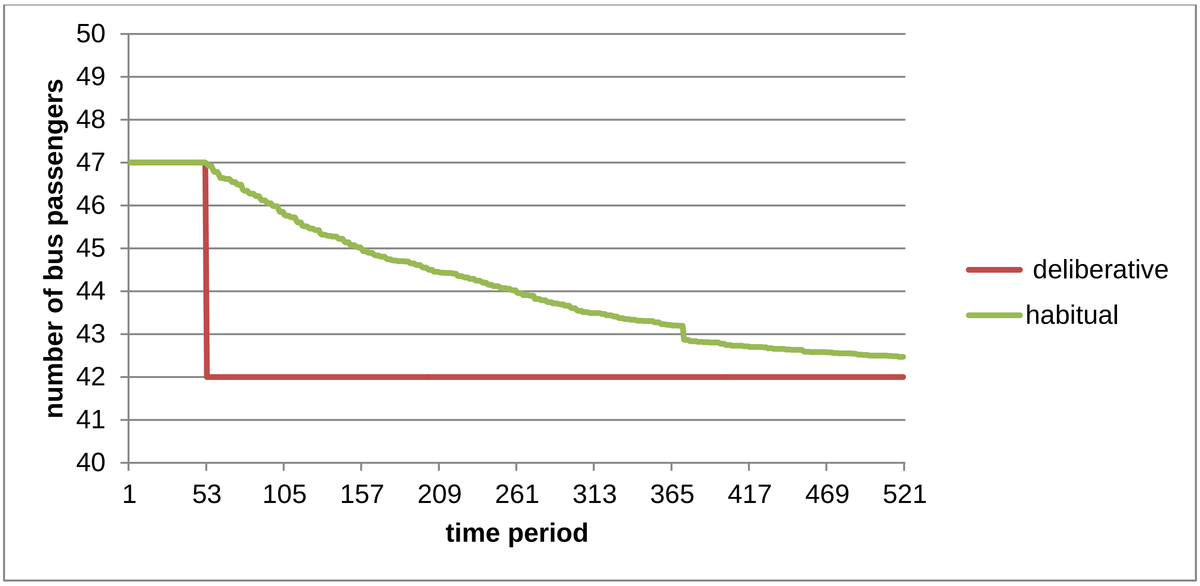


FIGURE 7.22 NUMBER OF FULL FARE PAYING BUS PASSENGERS OVER TIME WITH PERSONAL CONSTRAINTS

The shape of the curve in Figure 7.22 shows a close similarity to the curve of the bus fare elasticity over time shown in Figure 7.20.

In the graph above there is a kink at time period 370 where there is a small but sudden drop in the number of bus users. In a model of individuals records can be kept on the circumstances of each person in each time period and the decisions they make. An examination of these records shows that there is a cluster of individuals who reach retirement age in time period 370. The age distribution of people in the model at the start of the model period is shown in Figure 7.23 below. This shows a small cluster of people who are 62 when the model starts. When they reach the model's retirement age of 68 they are replaced with younger agents; this replacement is treated as a trigger event and these individuals assess their own travel mode rather than use the mode used by the agent they have replaced. The replacement agents are in effect placed into the deliberative decision mode in that time period.

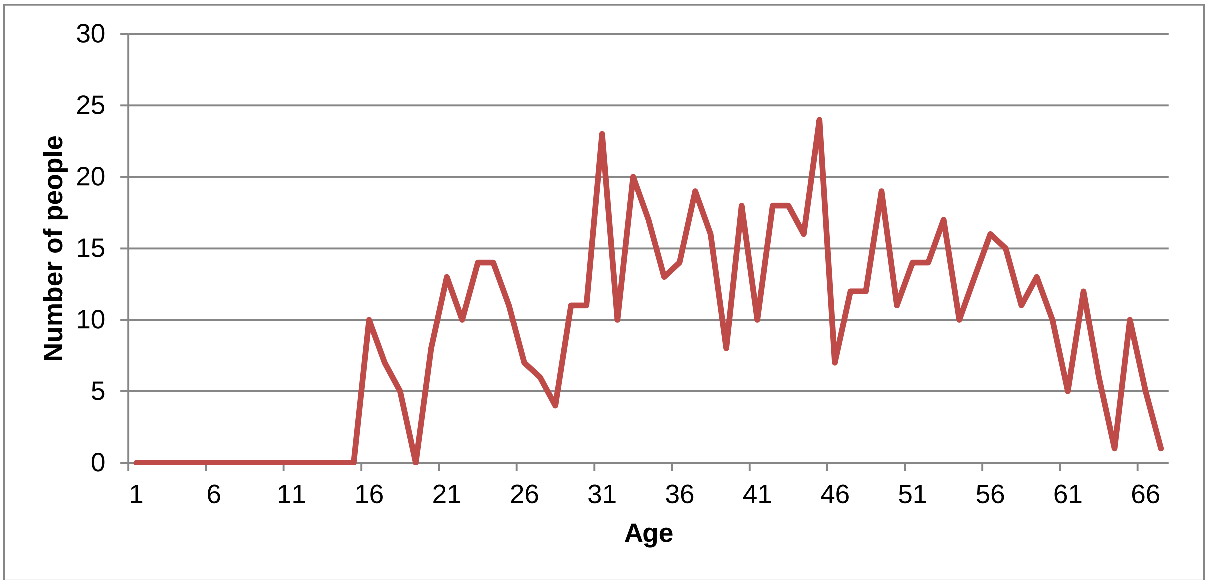


FIGURE 7.23 AGE DISTRIBUTION OF AGENTS IN THE MODEL AT THE START OF THE MODELLING PERIOD

A test was undertaken with all the agents having a random age between 16 and 68 allocated to them rather than an age within the age band reported in the survey. Figure 7.24 shows the plot of the number of bus passengers in each time period and this shows a smoother decline in bus use over time as fares increase over time with a more even spread of ages amongst the workforce. This shows that the deviation from the trend decline in bus user numbers was caused by a cluster of older people in the DfT dataset receiving a bus pass in the same year. With a larger number of individuals used in a model there is likely to be a more uniform distribution of ages amongst the agents. The investigation of the kink in the curve seen in Figure 7.23 illustrates that modelling individuals makes it is possible to interrogate the model at a very detailed level to seek an understanding of phenomena observed at the aggregate level; in this case the slightly larger than expected drop in the number of bus passengers in a particular year.

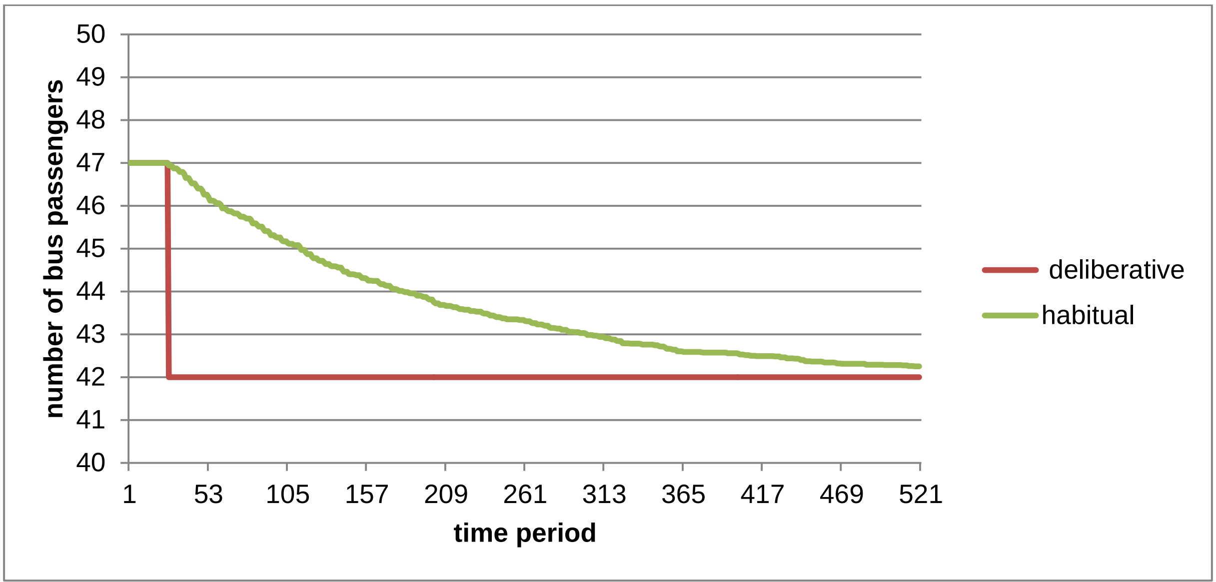


FIGURE 7.24 NUMBER OF FULL FARE PAYING BUS PASSENGERS OVER TIME WITH RANDOM AGES

7.8 Modelling shows variability in the predicted outcomes

The inclusion of trigger events, the timing of which are affected by stochastic processes in the model, means that the outputs from the model will vary between model runs even for the same initial input values. Running the model a multitude of times provides understanding into the degree of possible variability in the forecasts. This is of particular relevance for a business case as it provides an indication of the likely variability in the final benefit to cost ratio calculated for a scheme and the amount of subsidy that might be required or profits made.

The monitoring and evaluation of schemes is often dependent upon the comparison of before and after counts. It is sometimes erroneously assumed that, as a conventional deterministic four stage model run to convergence provides a single forecast of the number of users of each mode, that this single forecast figure can be compared to observed counts after implementation of the scheme. However when carrying out the comparison of 'after' modelled flows with observed counts, practitioners should be aware of the variability in the observed counts; WebTAG guidance for traffic counts for example recommends that at least two weeks of data is collected. There is less awareness of the variability around the modelled flows and the model used in this research shows that, even for the same inputs, model outputs can vary.

This model highlights that the context of a scheme is important and that other factors, other than the intervention being assessed may affect the modelled flows. For example, in this model the impact of the change in parking costs (the intervention) varies depending on the rate at which people move out of the area when they move house or change job.

The model also illustrates that the degree of response varies over time. This should influence when the evaluation takes place. If it takes place too early after the intervention it may not pick up the full extent of the changes that will occur as a consequence of that change. The later the evaluation occurs though, increases the chances of other aspects of the transport system changing, so that the impact of the particular intervention of interest has been affected by these other factors that have changed as well.

The scenarios tested in this research were each run 100 times, with exactly the same inputs. When the mode choice is deliberative the results are the same for each run. This is because the timing of trigger events is irrelevant; they only influence the forecasts in the habitual model where the occurrence of a trigger event, (the timing of which is affected by stochastic processes in the model), causes a re-assessment of the mode chosen by the agent.

The graphs below in Figures 7.25 to 7.29 show the spread of the forecasts of the number of bus, car, rail and cycle users for Scenario 1 (trend only changes in bus, car and rail costs) from 100 model runs. The green line shows the number of users for that mode when the model is run with always deliberative mode choice behaviour. The red line shows the mean of the forecast number of users of that mode from 100 runs of the model incorporating habitual behaviour. The dark blue area shows the boundary of the forecast number of users for that mode for the 100 model runs and indicates the degree of variability. The graphs show that the variability in the forecast number of users tends to increase over time.

The range of the forecast number of users of each mode from the 100 model runs are shown in Table 7.16 below. The table shows the mean, minimum and maximum number of users of each mode in the first time period and then after every two years. The table also gives the standard deviation of the forecast number of users of each model in the 100 model runs which confirms that the variation between the runs tends to increase over time.

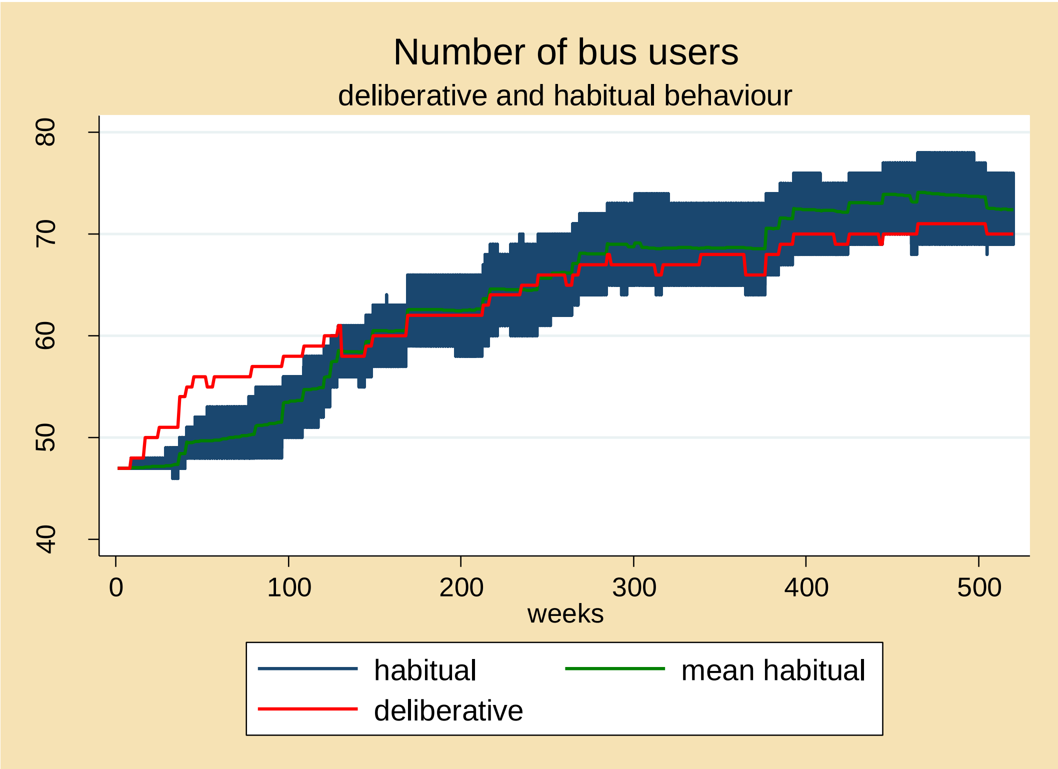


FIGURE 7.25 VARIABILITY IN THE NUMBER OF BUS USERS, DELIBERATIVE AND HABITUAL BEHAVIOUR, SCENARIO 1

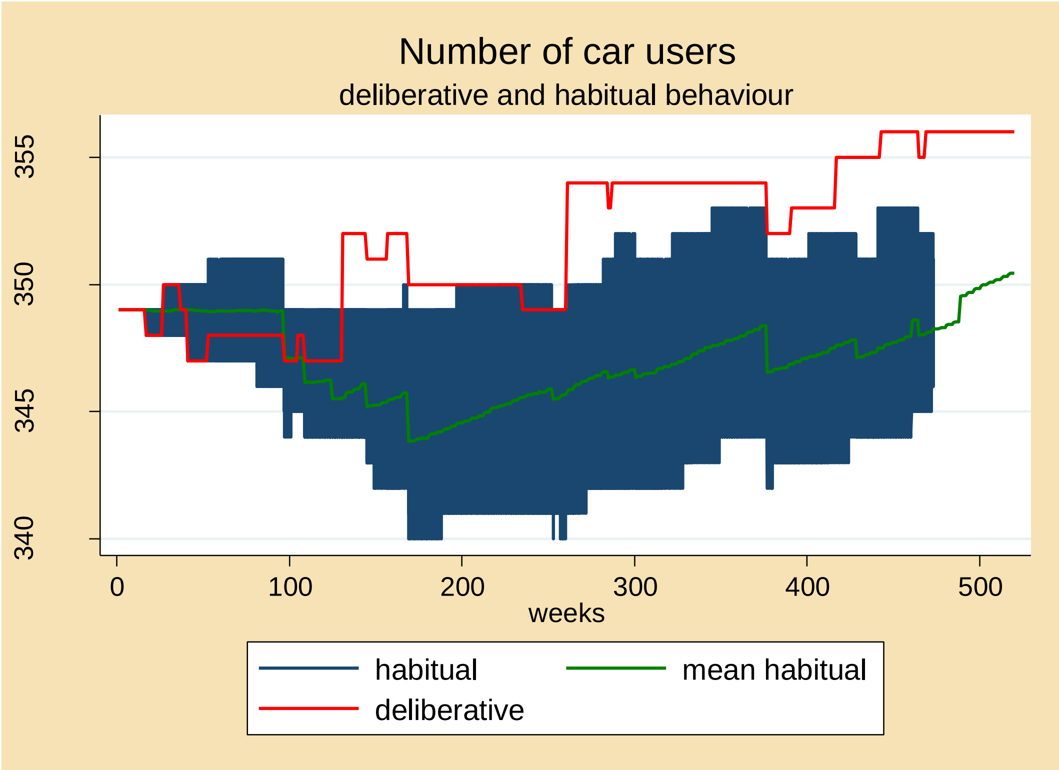


FIGURE 7.26 NUMBER OF BUS USERS WITH AND WITHOUT PTP, SCENARIO 2, HABITUAL BEHAVIOUR

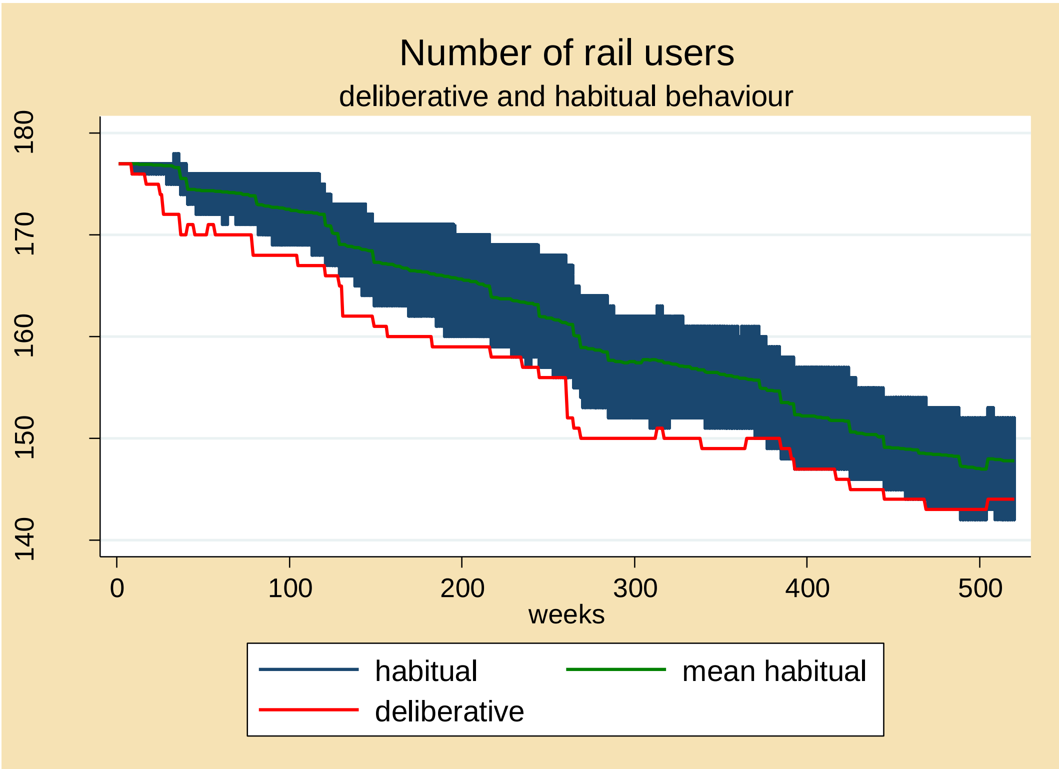


Figure 7.27 Variability in the number of rail users, deliberative and habitual behaviour, scenario 1

Mode	Time period	Unconstrained Mean	Std. Dev.	Min	Max	Constrained Mean	Std. Dev.	Min	Max
Bus	1	44.00	0.00	44	44	47.00	0.00	47	47
Bus	104	49.45	1.39	46	53	53.61	1.36	50	56
Bus	208	59.28	1.68	55	63	62.54	1.65	58	66
Bus	312	61.97	1.65	58	66	68.63	1.69	65	74
Bus	416	67.40	1.56	63	71	72.32	1.70	68	75
Bus	520	66.79	1.65	61	72	72.37	1.68	69	76
Car	1	411.00	0.00	411	411	349.00	0.00	349	349
Car	104	408.60	1.18	406	411	347.09	1.01	345	349
Car	208	405.11	1.93	400	410	344.75	1.79	341	350
Car	312	408.54	1.96	404	412	346.52	1.84	342	351
Car	416	407.02	1.85	401	412	347.49	1.78	343	352
Car	520	409.07	2.12	404	416	350.44	1.70	347	355
Train	1	113.00	0.00	113	113	177.00	0.00	177	177
Train	104	110.30	1.06	107	113	172.41	1.29	169	176
Train	208	104.29	1.95	100	108	165.39	2.08	160	170
Train	312	98.37	2.30	93	103	157.73	2.30	151	162
Train	416	93.18	2.07	89	99	151.79	2.17	147	157
Train	520	90.72	2.20	85	95	147.77	2.37	142	152
Cycle	1	52.00	0.00	52	52	42.00	0.00	42	42
Cycle	104	51.65	0.67	50	53	41.89	0.57	41	43
Cycle	208	51.32	0.93	50	54	42.32	0.93	41	45
Cycle	312	51.12	1.04	49	53	42.12	1.04	40	44
Cycle	416	52.25	0.88	50	54	43.25	0.88	41	45
Cycle	520	52.42	0.89	50	54	43.42	0.89	41	45

TABLE 7.16 STANDARD DEVIATION OF MODELLED NUMBER OF USERS OF EACH MODE, SCENARIO 1

When the model is run with step changes in costs as in Scenario 2, then there is a greater variation in the forecast number of users of each mode between the model runs. This is shown in Table 7.17 below which shows the standard deviation of modelled number of users of each mode for scenario 2. For example, for car users after ten years, the standard deviation for scenario 1 with personal constraints is 1.70 but for scenario 2 it is 3.98. The greatest variation in results between model runs is for the number of car users after 8 years in scenario 2 without personal constraints.

		Unconstrained				Constrained			
Mode	**Time period**	**Mean**	**Std. Dev.**	**Min**	**Max**	**Mean**	**Std. Dev.**	**Min**	**Max**
Bus	1	44.00	0.00	44	44	47.00	0.00	47.00	47.00
Bus	104	49.20	1.29	46	52	53.36	1.31	50.00	58.00
Bus	208	60.18	1.90	54	64	63.98	1.92	58.00	68.00
Bus	312	69.53	2.48	63	75	74.60	2.29	70.00	80.00
Bus	416	81.89	2.39	77	87	84.35	2.24	79.00	89.00
Bus	520	88.80	2.43	83	95	89.57	2.06	84.00	93.00
Car	1	411.00	0.00	411	411	349.00	0.00	349.00	349.00
Car	104	408.96	1.33	406	412	347.20	0.97	344.00	350.00
Car	208	392.04	3.35	383	400	334.45	2.95	327.00	342.00
Car	312	365.28	5.32	352	383	312.46	4.46	301.00	326.00
Car	416	323.70	6.09	306	339	280.48	5.05	270.00	294.00
Car	520	297.50	5.11	286	313	259.74	3.98	251.00	271.00
Train	1	113.00	0.00	113	113	177.00	0.00	177.00	177.00
Train	104	110.30	1.09	107	113	172.56	1.35	168.00	176.00
Train	208	116.23	3.19	108	125	174.02	2.97	167.00	183.00
Train	312	131.97	5.10	118	144	184.51	4.32	173.00	194.00
Train	416	158.17	5.67	144	174	204.65	4.77	191.00	215.00
Train	520	175.53	5.10	160	188	219.06	4.26	206.00	229.00
Cycle	1	52.00	0.00	52	52	42.00	0.00	42.00	42.00
Cycle	104	51.54	0.74	50	53	41.88	0.52	41.00	43.00
Cycle	208	51.84	1.06	50	54	42.84	1.06	41.00	45.00
Cycle	312	53.69	1.22	50	56	43.90	1.11	41.00	47.00
Cycle	416	56.66	1.61	53	60	45.94	1.20	43.00	49.00
Cycle	520	58.42	1.41	55	61	46.88	1.24	44.00	49.00

TABLE 7.17 STANDARD DEVIATION OF MODELLED NUMBER OF USERS OF EACH MODE, SCENARIO 2

The range in the forecast number of users of bus, car and rail in scenario 2 are shown in Figures 7.28 to 7.30 below.

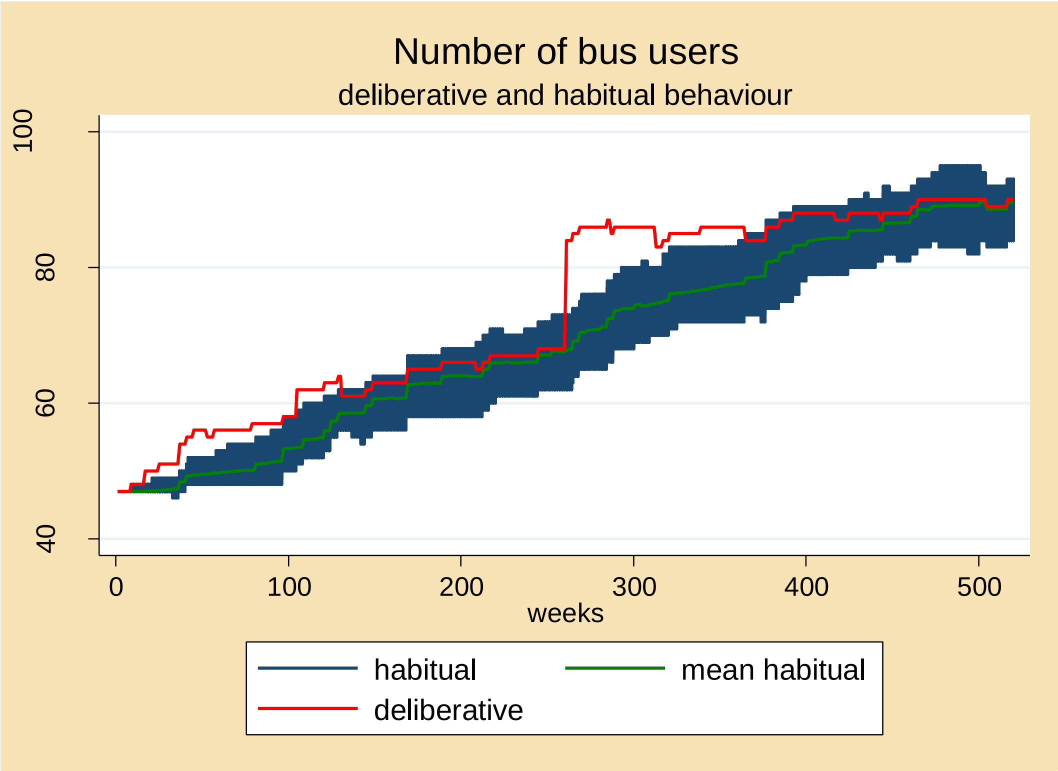


FIGURE 7.28 VARIABILITY IN NUMBER OF BUS USERS, DELIBERATIVE AND HABITUAL BEHAVIOUR, SCENARIO 2

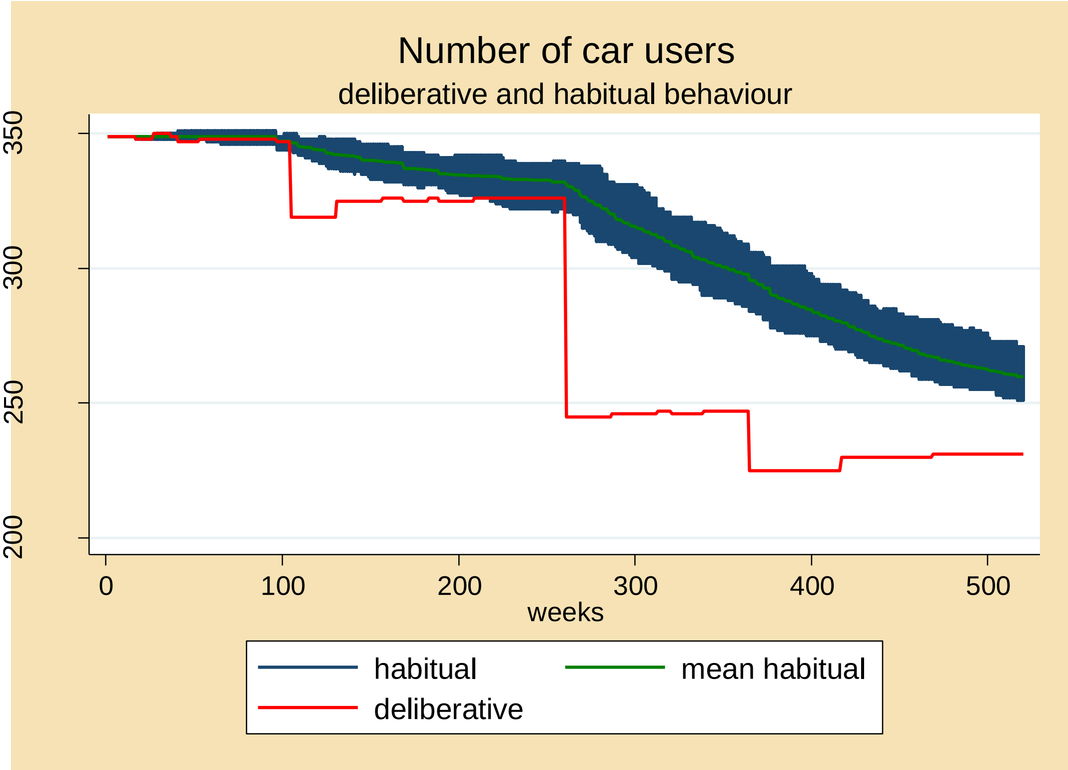


FIGURE 7.29VARIABILITY IN NUMBER OF CAR USERS, DELIBERATIVE AND HABITUAL BEHAVIOUR, SCENARIO 2

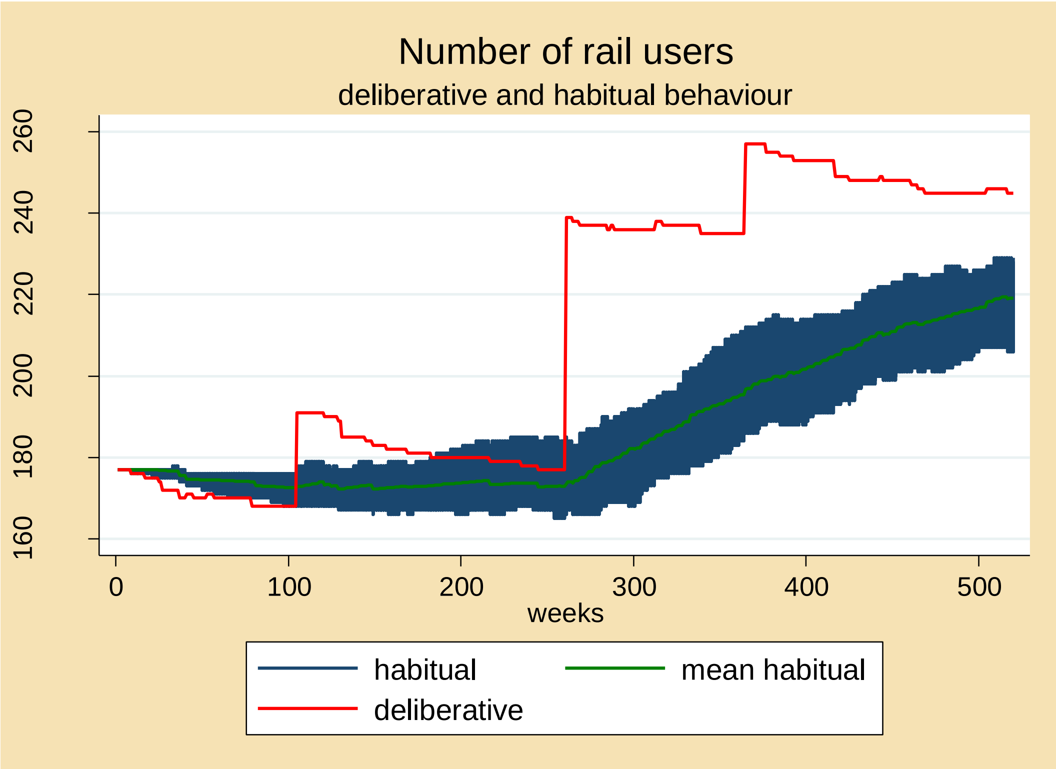


FIGURE 7.30 VARIABILITY IN NUMBER OF RAIL USERS, DELIBERATIVE AND HABITUAL BEHAVIOUR, SCENARIO 2

In the graphs above the range of the forecast values is shown but there is no detail on the distribution of values within these ranges. This is provided below in Figures 7.31 to 7.33 which show the frequency distribution for the forecast number of users of bus, car and rail after the model has run for 10 years for scenario 2. The graphs show a normal distribution curve for comparison purposes. For bus and rail users many of the model runs have a forecast value which lays above the mean but for car users the forecasts from the 100 model runs show a more normal distribution.

A benefit of agent based modelling is that the model allows for the inclusion of stochastic processes and so provides a method for building models that can replicate the variation that is observed in the real world. A comparison of the variability observed in the real world and that predicted from the model could also be used for validating the model. This would mean a change in current practice away from collecting counts for a single day for public transport services towards collecting data for multiple days. For cars this is achieved by using automatic traffic counters and it could be possible to collect data on the daily variation in public transport usage from ticket sales data and ticket gates.

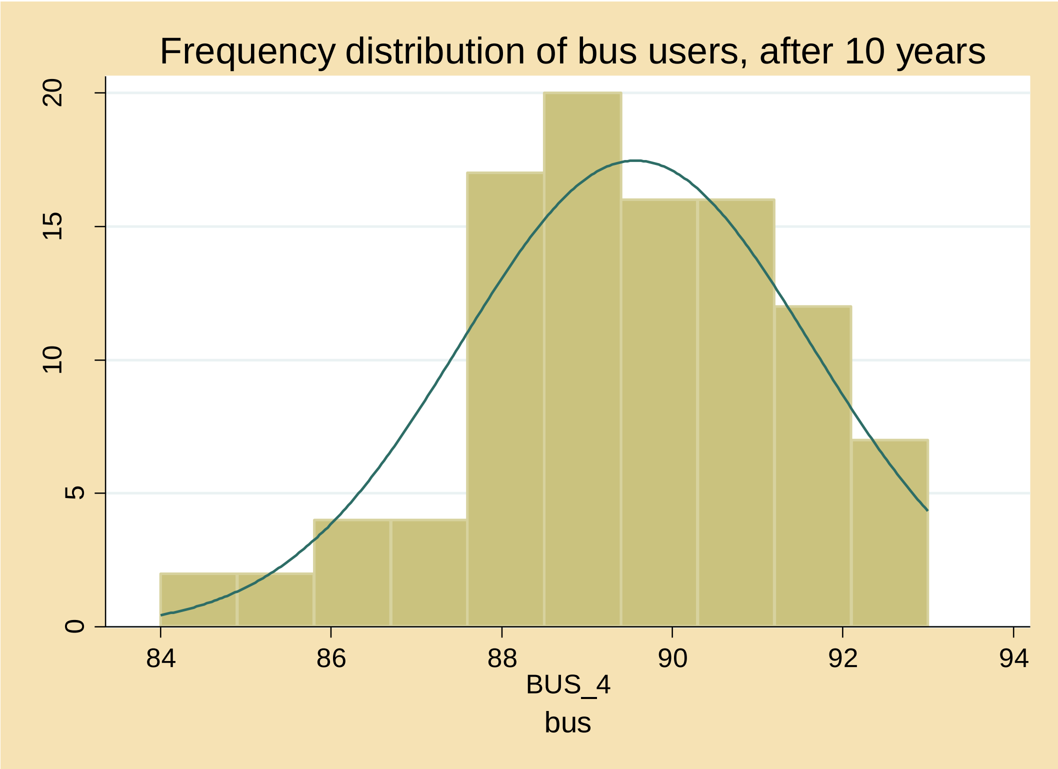


FIGURE 7.31 FREQUENCY DISTRIBUTION OF FORECAST NUMBER OF BUS USERS, SCENARIO 2

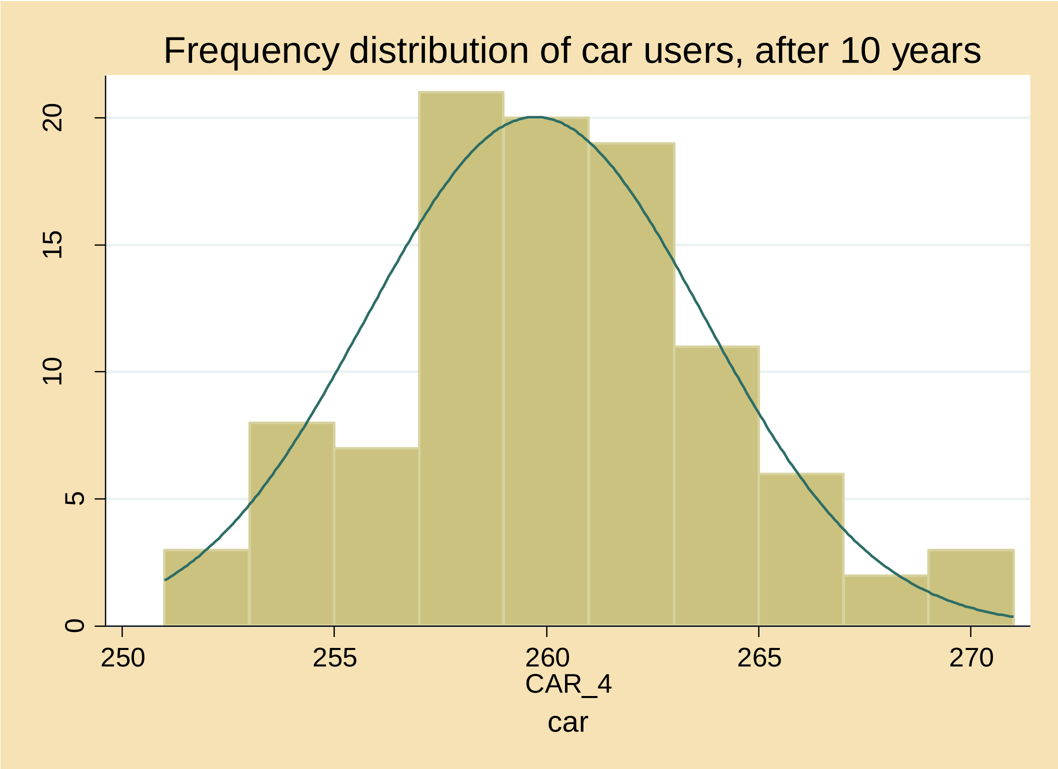


FIGURE 7.32 FREQUENCY DISTRIBUTION OF FORECAST NUMBER OF CAR USERS, SCENARIO 2

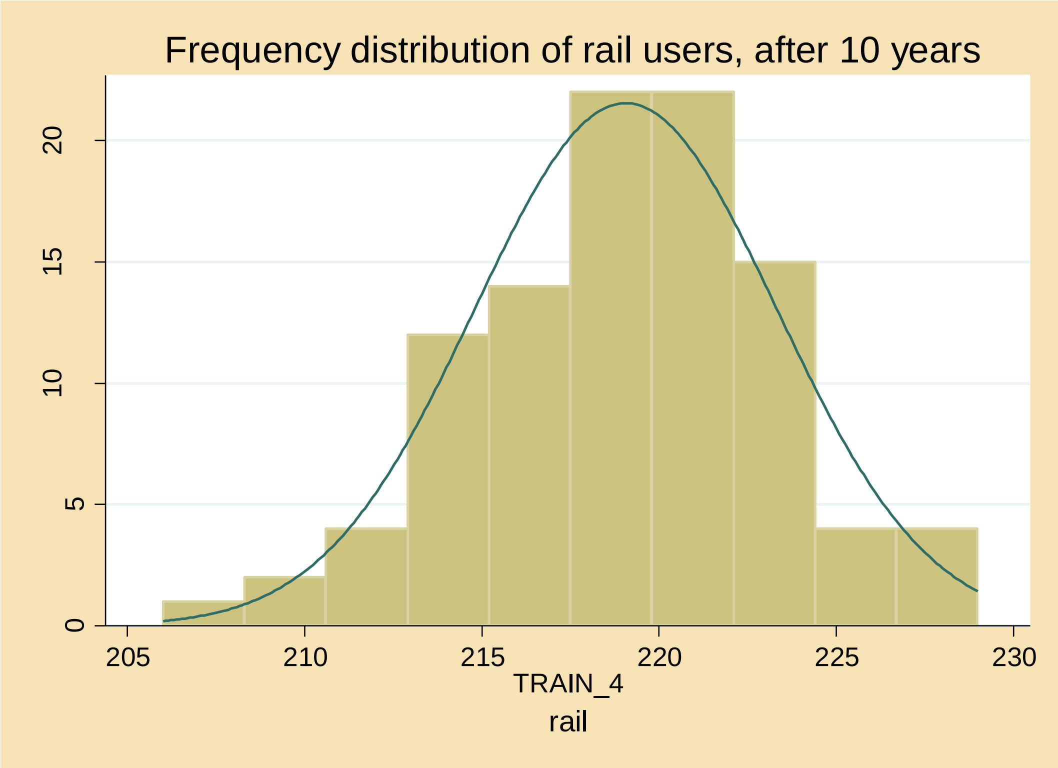


FIGURE 7.33 FREQUENCY DISTRIBUTION OF FORECAST NUMBER OF RAIL USERS, SCENARIO 2

7.9 Modelling at the individual level allows for testing a wider range of policies

The effect of a personalised travel planning campaign can be included in the model by assuming that a personalised travel planning (PTP) visit acts as a trigger event. Taking Scenario 2 as the base case, PTP is modelled by assuming that each year 50% of households have a PTP visit and that as a result of the visit they became fully aware of the time and cost of each of the travel options available to them. Table 7.18 below shows the number of users of each mode in each time period, with and without the PTP programme.

In the early years, without any sudden changes in parking charges there is only a small difference between the number of users of each mode with and without the PTP visits. This is because the model assumes that at the start of the model time period everyone has made their initial mode choice decision based on knowledge of the time and cost of their journey options. With a gradual change in costs, there is not a marked difference between people's habitual choice and their optimal choice, as in many cases the small change in costs means that a person's habitual choice is still their optimal choice, whether they have re-assessed it or not recently.

Once there is a sudden change in costs, in this model through an increase in parking costs, then the divergence between the mode used if the users exhibit habitual behaviour and if they re-consider their mode choice becomes much greater. A PTP visit makes them aware of these changes and causes them to re-consider their mode choice earlier than they would have done otherwise. In this case, because they become aware of the increased cost of driving compared to using public transport, they switch to public transport earlier than they would have done without the PTP visit.

Time period		Bus		Car		Train		Cycle		None	
		No PTP	PTP	No PTP	PTP	No PTP	PTP	No PTP	PTP	No PTP	PTP
beg year 1	**1**	47.0	47.0	349.0	349.0	177.0	177.0	42.0	42.0	11.0	11.0
end year 1	**52**	49.5	50.8	349.0	349.0	174.5	173.2	42.0	42.0	11.0	11.0
beg year 2	**53**	49.6	50.8	348.9	349.0	174.5	173.2	42.0	42.0	11.0	11.0
end year 2	**104**	53.4	55.5	347.2	347.1	172.6	170.7	41.9	41.7	11.0	11.0
beg year 3	**105**	53.5	55.6	346.8	346.8	172.8	171.0	41.9	41.7	11.0	11.0
end year 3	**156**	60.7	62.9	339.8	332.6	172.4	176.9	42.3	43.1	10.8	10.5
beg year 4	**157**	60.7	62.8	339.6	332.6	172.5	176.9	42.3	43.1	10.8	10.5
end year 4	**208**	64.0	65.8	334.5	327.6	174.0	178.3	42.8	44.1	10.7	10.2
beg year 5	**209**	64.0	65.8	334.3	327.6	174.1	178.4	42.9	44.1	10.7	10.2
end year 5	**260**	67.5	68.6	332.0	326.7	173.0	176.5	42.9	44.1	10.6	10.1
beg year 6	**261**	67.8	68.8	331.2	325.8	173.5	177.1	43.0	44.1	10.6	10.1
end year 6	**312**	74.6	80.1	312.5	282.3	184.5	207.7	43.9	45.9	10.5	10.0
beg year 7	**313**	74.7	80.1	311.7	281.9	185.1	208.1	44.0	45.9	10.5	10.0
end year 7	**364**	77.7	83.4	297.9	263.0	195.3	222.7	44.7	46.9	10.5	10.0
beg year 8	**365**	78.4	83.8	295.6	261.9	196.8	223.4	44.7	46.9	10.5	10.0
end year 8	**416**	84.4	87.8	280.5	242.8	204.7	236.8	45.9	48.2	10.6	10.5
beg year 9	**417**	84.4	87.7	279.9	242.7	205.2	237.0	46.0	48.2	10.6	10.5
end year 9	**468**	88.6	91.1	267.5	236.0	212.8	239.6	46.4	48.6	10.7	10.8
beg year 10	**469**	88.5	91.1	267.2	236.0	213.1	239.6	46.5	48.6	10.7	10.8
end year 10	**520**	89.6	90.5	259.7	233.5	219.1	242.4	46.9	48.7	10.8	10.9

TABLE 7.18 NUMBER OF AGENTS BY MODE WITH AND WITHOUT PERSONALISED TRAVEL PLANNING

The number of people using public transport is higher when there is a PTP programme than without, particularly after there has been a recent rise in car costs. This is also shown in Figures 7.34 to 7.36 below, which show the difference in the number of people using bus, car, train and cycling, with and without the PTP programme. The model was run 100 times and the mean values from these model runs are shown.

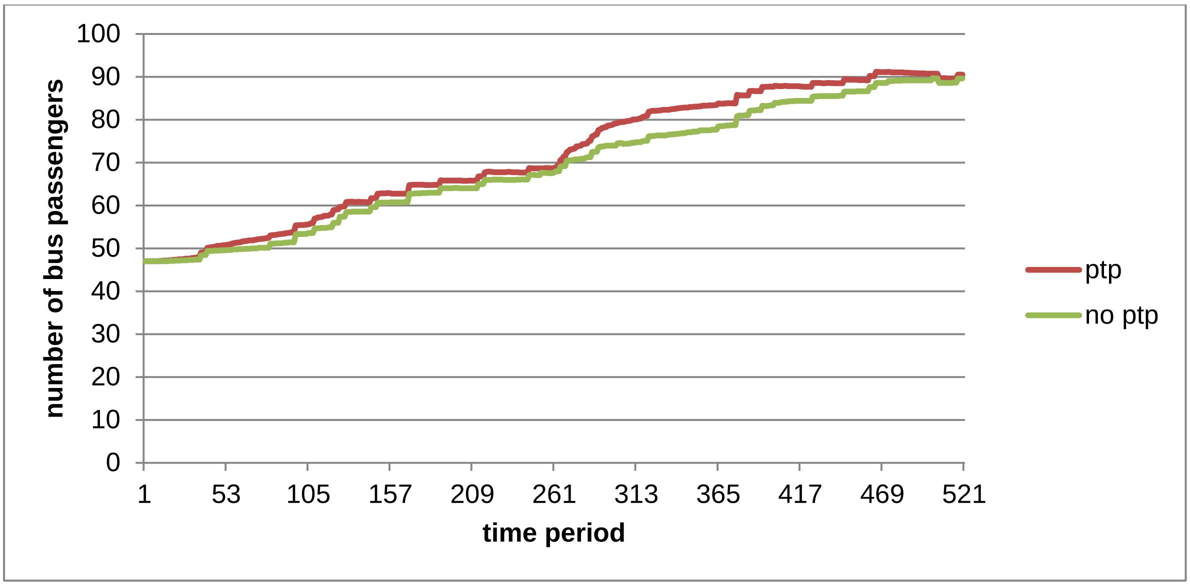


FIGURE 7.34 NUMBER OF BUS USERS WITH AND WITHOUT PTP, SCENARIO 2, HABITUAL BEHAVIOUR

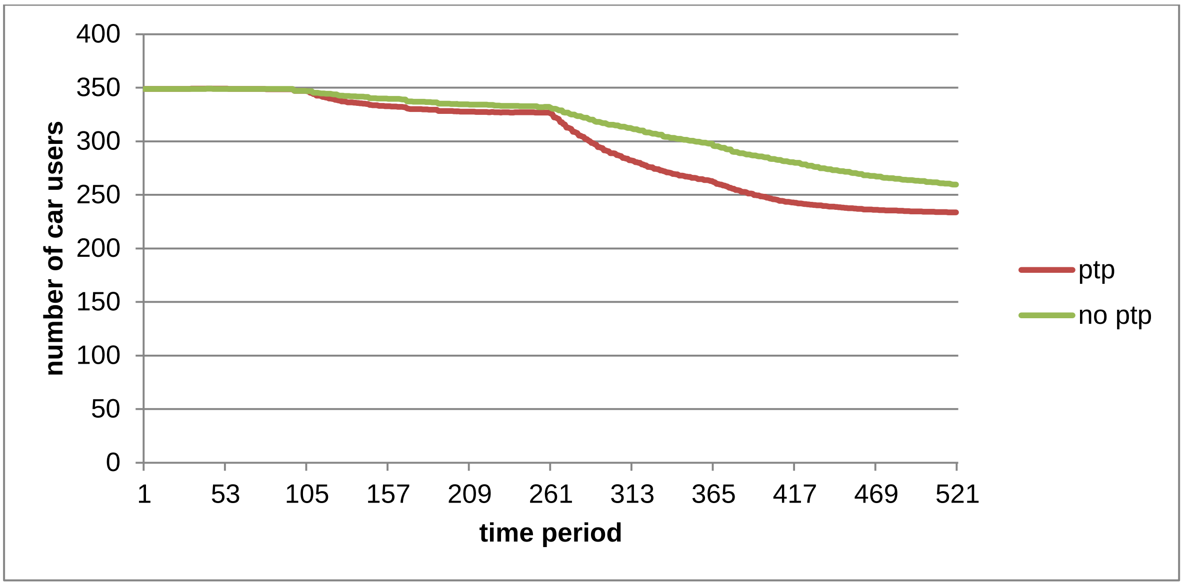


FIGURE 7.35 NUMBER OF CAR USERS WITH AND WITHOUT PTP, SCENARIO 2, HABITUAL BEHAVIOUR

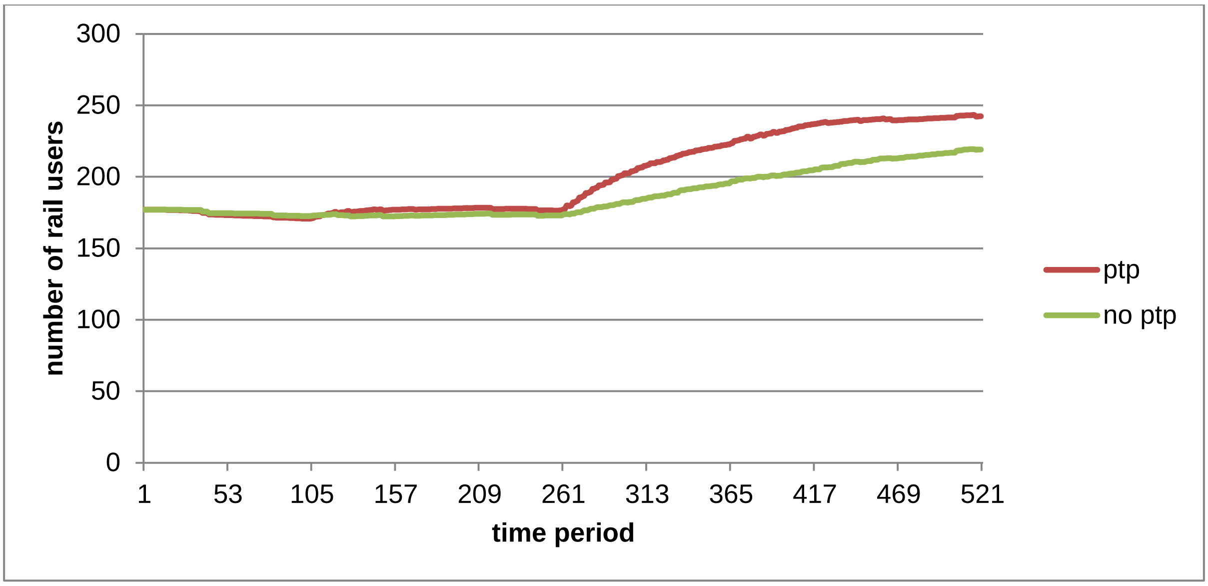


FIGURE 7.36 NUMBER OF RAIL USERS WITH AND WITHOUT PTP, SCENARIO 2, HABITUAL BEHAVIOUR

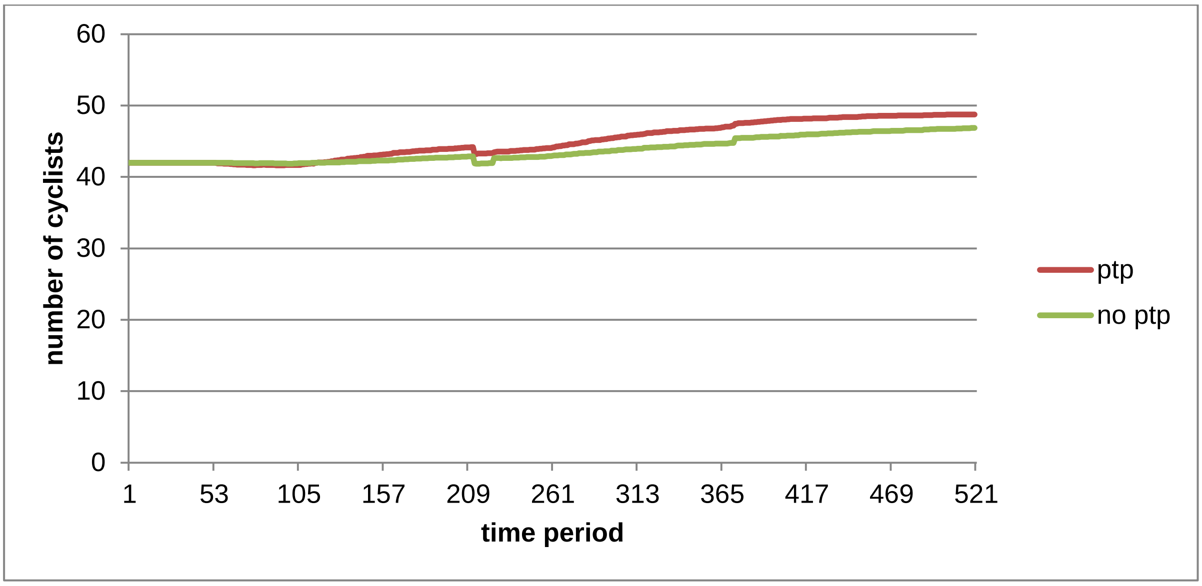


FIGURE 7.37 NUMBER OF CYCLISTS WITH AND WITHOUT PTP, SCENARIO 2, HABITUAL BEHAVIOUR

The impact of the PTP visits is to bring the number of users of each model closer to the number there would be, if everyone chose their mode of travel in every time period i.e. always in deliberative mode. This is illustrated in Figures 7.38 to 7.41 below which show the forecast number of people travelling by bus, car, train and cycling with habitual behaviour and no PTP, habitual behaviour and PTP and always deliberative behaviour. For bus users, the PTP programme increases the number of bus users. The difference between the number of bus users with and without PTP visits becomes smaller over time. As the change in car costs returns to trend, after the step change in car costs, then more people are already using the most optimal mode for them when they have a PTP visit and so do not change mode as a consequence of the visit. This suggests that a PTP programme will be more effective in changing people's mode if it happens shortly after a step change in transport costs for a mode, either an increase in car costs or an improvement in the public transport options.

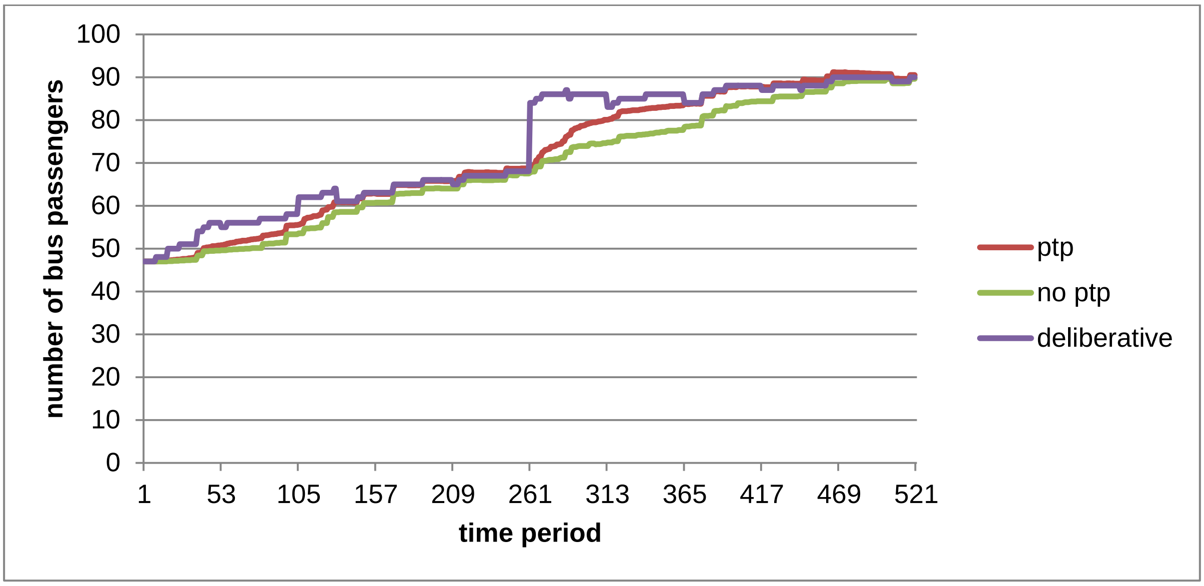


FIGURE 7.38 NUMBER OF BUS USERS IN SCENARIO 2

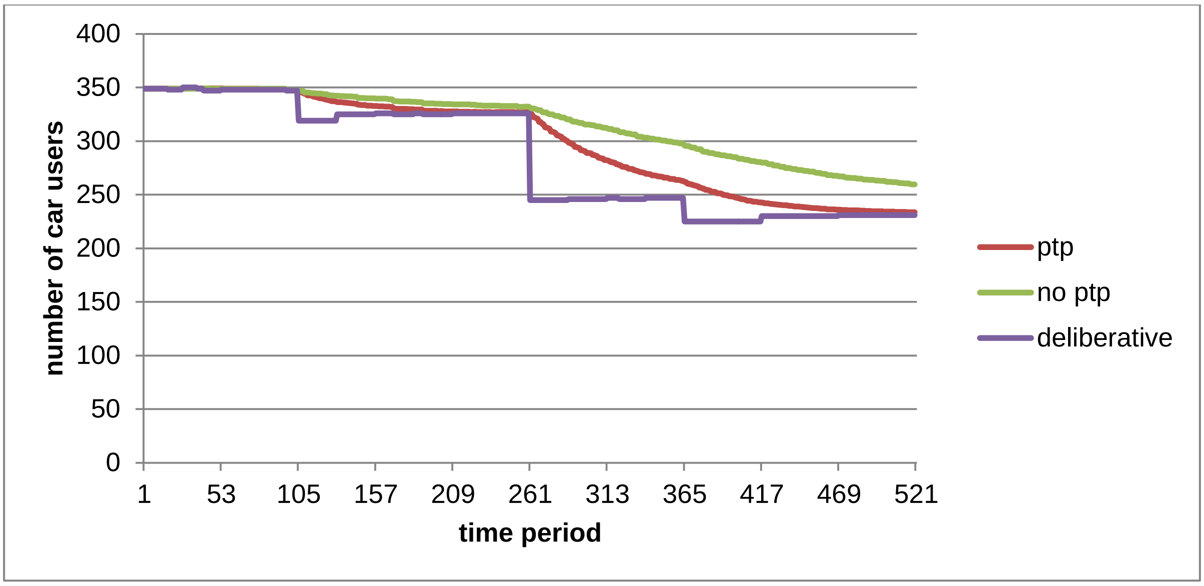


FIGURE 7.39 NUMBER OF CAR USERS IN SCENARIO 2

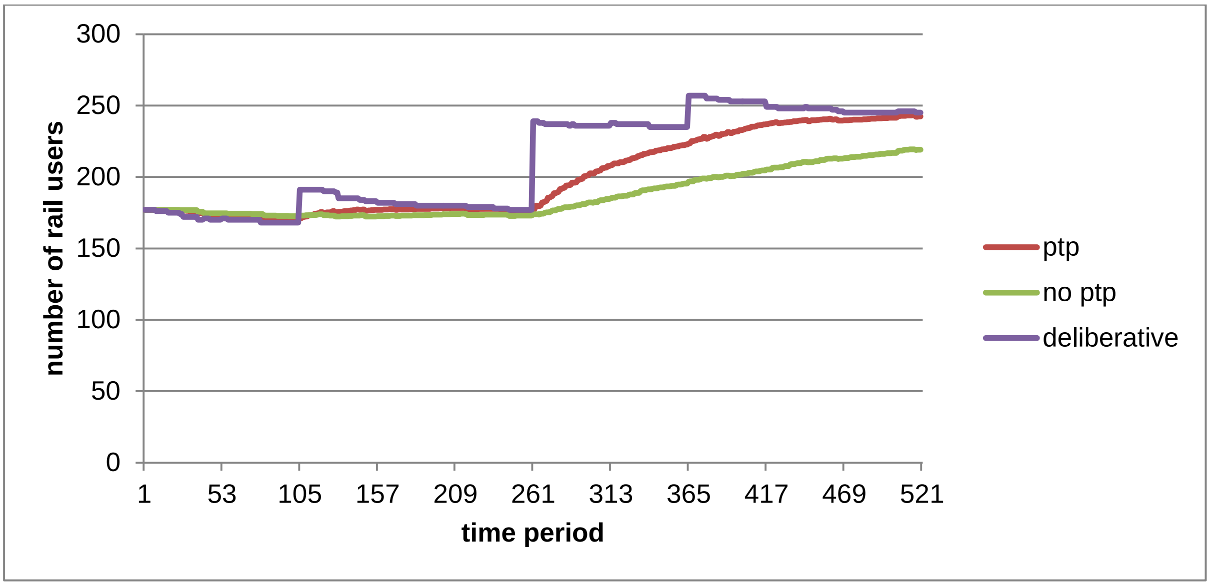


FIGURE 7.40 NUMBER OF RAIL USERS IN SCENARIO 2

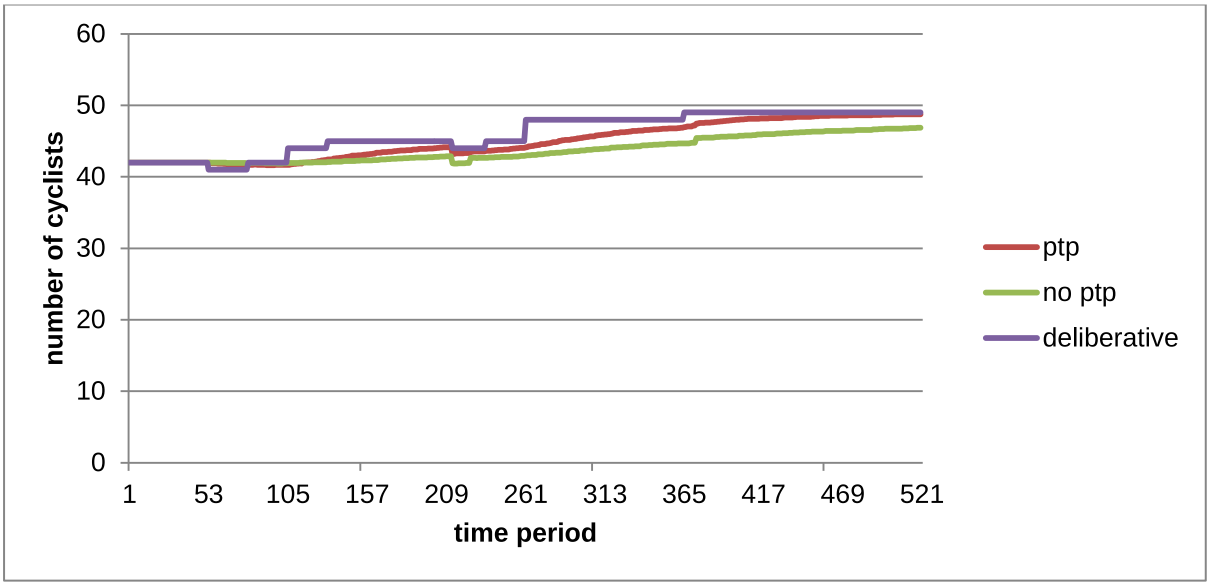


FIGURE 7.41 NUMBER OF CYCLISTS IN SCENARIO 2

It is not always the case that a PTP programme will result in an increase in public transport use. For example if bus costs are rising, a PTP visit may cause people to re-consider their mode choice and switch to car once they realise that it has become relatively cheaper to bus than it was previously. In scenario 1, the model was run with habitual behaviour and a PTP programme that visited 50% of houses in every year. Initially the number of bus users rises with the PTP programme but as bus fares continue to rise and car costs begin to decline, the number of bus users becomes lower with the PTP programme in place as people realise the emerging cost difference between bus and car use. The number of rail users is always lower with a PTP programme, again because of the greater rise in rail fares than car costs.

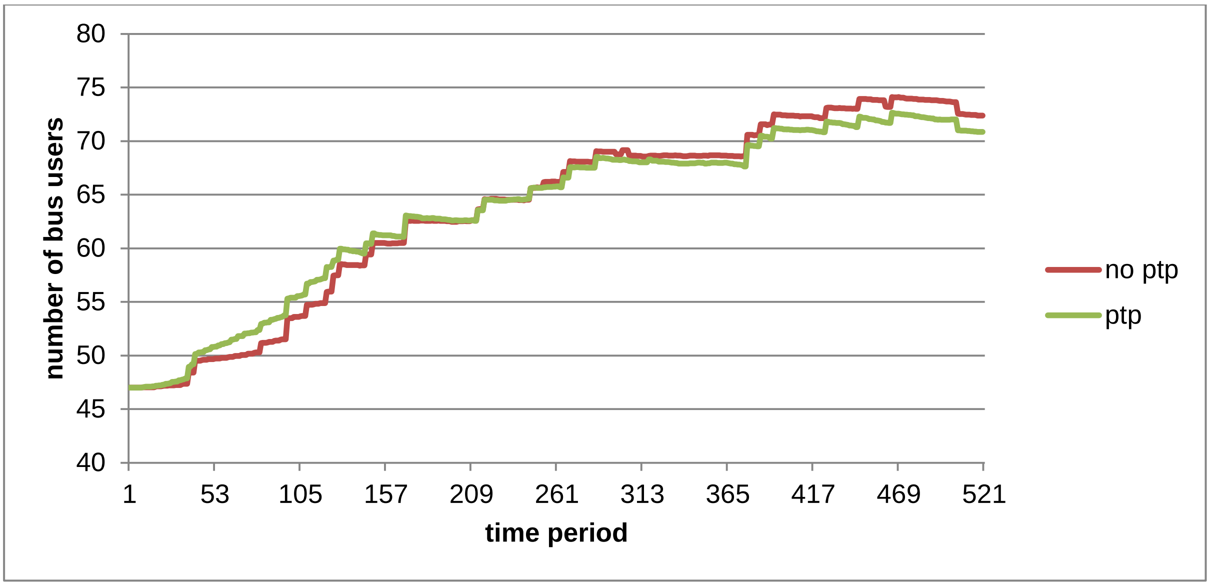


FIGURE 7.42 NUMBER OF BUS USERS IN SCENARIO 1, WITH AND WITHOUT PTP

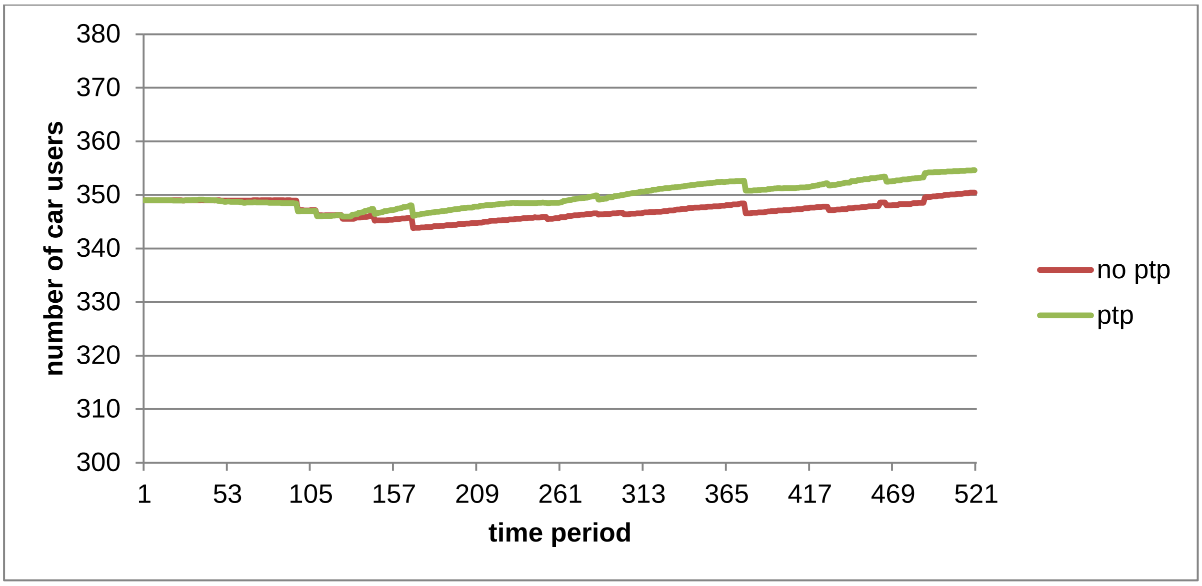


FIGURE 7.43 NUMBER OF CAR USERS IN SCENARIO 1, WITH AND WITHOUT PTP

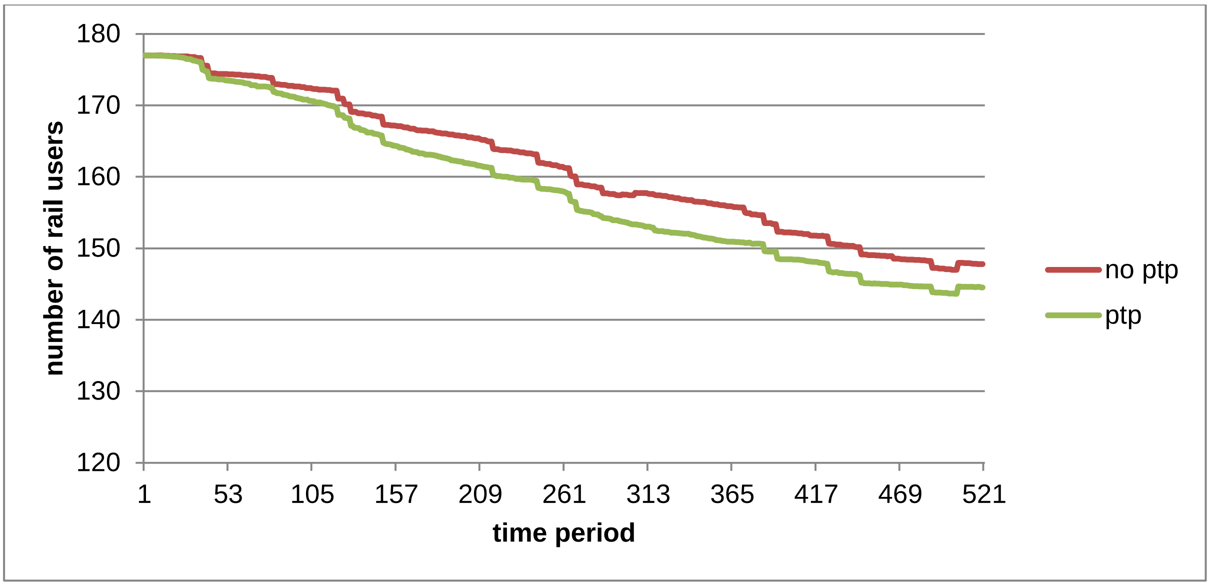


FIGURE 7.44 NUMBER OF RAIL USERS IN SCENARIO 1, WITH AND WITHOUT PTP

7.10 Conclusion

This chapter has presented the results from a variety of runs with an agent based model of commuter mode choice. The personal characteristics and constraints for the 626 agents in the model came from a survey conducted for the DfT. This survey also contained a stated preference section which provided details of the personal preferences held by each person for four basic characteristics of a journey: mode, time, cost and carbon emissions. These preferences were used in a maximum utility model of mode choice for each agent.

The model was run for 10 years and 100 replications were run for each scenario. The model results were used to illustrate the importance of the presence of concessionary bus fares for the elderly in maintaining the number of bus users, which would otherwise have fallen over time as bus fares rise relative to rail fares and car costs. When the bus operators adjust the number of bus routes to reflect the actual patronage on bus services, the availability of concessionary fares helps maintain more bus routes than would otherwise be provided. It slows down the vicious circle of bus decline, with lower bus patronage leading to a reduction in services which removes the option of using a bus for some people and/or results in longer wait times which in turn leads to a further reduction in bus numbers.

The model also shows the importance of personal constraints in mode choice and that the inclusion of more precise choice sets for agents in the model affects the overall forecasts of users of each mode. The lack of a driving licence is a key constraint on mode choice which may be a result of not wishing to drive but also may also be the result of lacking the money to afford driving lessons and insurance. This points to the possibility of using policies that affect the number of people experiencing personal constraints on travel choices as well as policies that change the time and cost of travel if the intention is to affect the number of users of each travel mode.

The use of an agent based modelling approach allows the mode choice model to potentially reflect a wider range of policy mechanisms that can currently be modelled using the four stage models used in the UK such as the role of personalised travel planning campaigns that make people more fully aware of the characteristics of their travel options.

In an agent based model, the decisions of each person can be affected by the decisions they have made in the past and their personal histories. This allows for the introduction of habitual behaviour in the model with people only re-considering their mode when they experience a trigger event. The inclusion of habitual behaviour results in a model that produces lags in the change in the number of users of a mode following a change in journey costs. The output elasticity of demand with respect to bus fares in the model was similar to that observed in the real world, and far more realistic than the instantaneous response produced by conventional four stage models.

This chapter documented the use of a proof-of-concept agent based model for commuting mode choice and showed that this modelling technique has potential to capture a wide range of transport measures, such as changing parking costs and public transport fares, subsidising fares for particular groups of people and the provision of information about travel options.

The challenge for using this model for a particular geographically defined study area is considered in the next chapter. The chapter also discusses methodological issues that would be encountered when scaling up the model to include all the agents in a particular area.

8 Extending the proof-of-concept model to a specific area

8.1 Introduction

This chapter considers a range of methodological issues that will arise when applying this general agent based model for mode choice to a particular study area. It considers the challenges that will be faced in obtaining data on the attributes and preferences for each agent in the model. Data on attributes for the agents can be achieved by using population synthesis techniques. Appropriate segmentation is offered as a means to achieve the preference data for individuals by using the preferences of the group for each person. Latent class analysis is examined as a method of grouping agents with similar preference and so reducing aggregation bias in the model results.

In order to apply this model to a specific area, information will be required on the characteristics of the agents and the environment in which they exist, as shown in Figure 8.1 below.

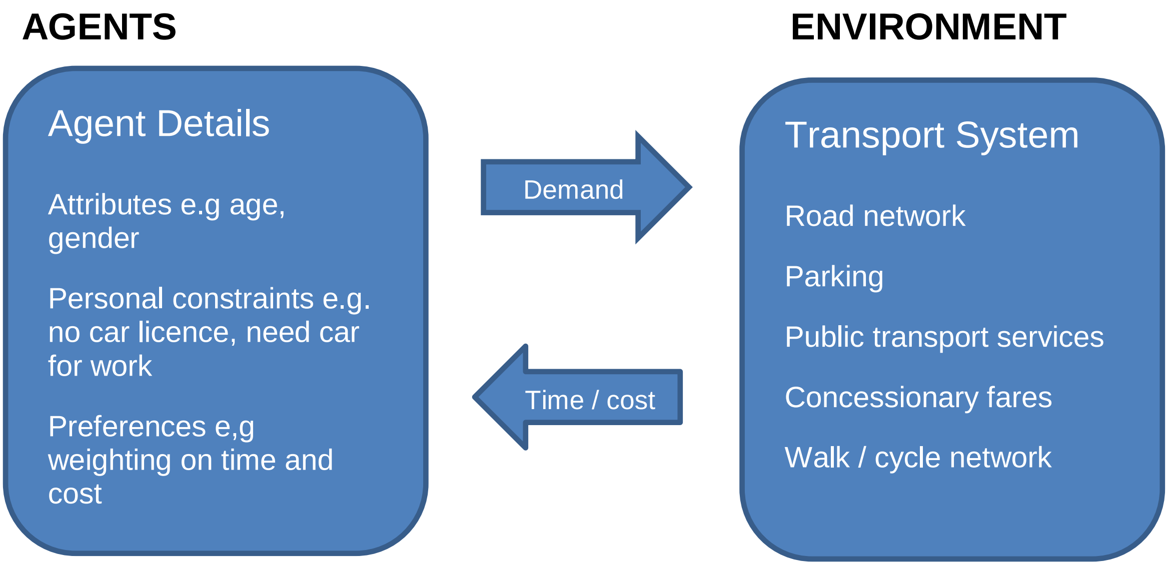


FIGURE 8.1 DATA REQUIREMENTS FOR THE AGENT BASED MODEL OF COMMUTER MODE CHOICE

8.2 Data on the environment

Many of the required details of the transport network can be readily attained from digital data sources. In the UK the Ordnance Survey publishes very detailed information on the road network. If road junctions are being modelled, the geometry of road junctions can be measured from mapping and junction layouts viewed from aerial photography. The timing settings of traffic signals can be obtained from the appropriate local authority or from site visits. Public transport timetables are available digitally and regularly updated. Commercial transport modelling software packages can be used to take the demand for travel as output from the agent based model, assign it onto the transport networks and then calculate travel times by mode which can be fed back into the mode choice agent based model. For highway times, the number, origin and destination of other trips on the network will be required, so either the transport assignment model can be pre-loaded with these trips and / or the agent based model be expanded to cover trips for all journey purposes. For example, all person trips may be modelled using an agent based model with the goods vehicle matrices produced independently and combined with the agent base model outputs before trip assignment.

The addition of any feedback loops in the agent based model brings with it a requirement for data on these feedbacks. For example, a feedback loop which has the commercial decisions of the bus operators affecting the number of bus services being run needs information on the strength and timing of this feedback. The decision making behaviour of the bus companies could be estimated from observations of their past behaviour or by interviews with bus operators and experienced analysts of the bus industry.

The focus of this chapter is on a consideration of the issues involved in constructing a complete set of agents for a particular area, with details of their **personal attributes and constraints** such as age and gender, location of home and work, personal mobility issues and proximity to the public transport network and their **personal preferences,** as reflected in their weightings for time and cost and their mode preference.

8.3 Data on agent attributes and constraints

Microsimulation models have always been very demanding in the amount and quality of data they require, as they need information on the characteristics of each agent and their transition probabilities, i.e. the probability for each person that one or more of their characteristics will change in the next model time period. Agent based models initially had less modest data requirements as they were built to test theories, often using an artificial data set of agents, to examine the emerging output at the top level from the operation of a simple set of behavioural rules at the lower level. In recent years, more agent based models have been designed as 'empirically grounded' models (Rounsevell *et al.*, 2012) and the need for much more specific data on each agent has grown.

The experience of the microsimulation community is useful in providing methodologies to meet the data needs of agent based models. They recognise the importance of high quality data for a model, the difficulties in obtaining the breadth and depth of data required for such models and the frequent necessity of using multiple data sets to construct the input dataset required. 'A model cannot exist without data, and the quality of the initial dataset in a dynamic microsimulation model is critical to the overall strength and sophistication of the model. Dynamic microsimulation is extremely demanding in terms of data requirements, and it is not conceivable that any one data set will contain all the information required for the base population (Scott *et al.*, 2003).

It is rare that, after designing a model, a single dataset can be found that already contains all the required information on agent characteristics or that can be collected within a project's budget. As noted in Section 6.3 above, available data sets are likely to vary in the degree of detail they contain about the population, the breadth of variables recorded and the reliability of the data. (Cassells *et al.*, 2006). The model may be redesigned to fit the dataset available or a synthetic dataset created with the required variables based on those present in a number of available datasets.

The need to create an appropriate input dataset is particularly acute in models with a spatial dimension as detailed information on agents' characteristics is seldom available alongside detailed spatial information, in order to preserve the anonymity of the individuals in the data set. A common approach to this problem in microsimulation modelling is to start with the largest available cross-sectional representative population of acceptable quality and to supplement this with data from other sources (Collins *et al.*, 2006). There is a wealth of literature in the field of spatial microsimulation, which covers static and dynamic microsimulation models with a spatial dimension, on techniques for creating the initial dataset of agent characteristics required for these models. Rossiter *et al.* (2009) observe that 'in geography, most of the effort in microsimulation models had been in constructing good quality geographically disaggregated population microdata'.

SimCrime and SimBritain are typical examples of spatial microsimulation models. SimCrime (Kongmuang, 2006) is a static spatial microsimulation model of crime in Leeds. Its data set comes from merging information from the census with data from the British Crime Survey. The census provides consistent data across the whole of the study area by output area (around 150 households). The British Crime Survey contains detailed information on the victims of crime, including demographic and socio-economic characteristics of crime victims, details of the crime and its precise location. The two datasets were used together to produce a synthetic data set of individuals which fitted data from both donor data sets i.e. a dataset of individuals from which cross-tabulations could be built which matched the cross tabulated results from both of the donor data sets.

SimBritain (Ballas *et al.*, 2007) is a dynamic spatial microsimulation model of the population of Britain used to predict the geographical and socio-economic impact of policies such as the introduction of the minimum wage and winter fuel payments. The model base year was 1991 and the dataset for the model was created by merging data from the 1991 census with longitudinal data from the British Household Panel Survey (BHPS) and trend data derived from the 1971, 1981 and 1991 census. The 1991 census was used to create an initial population for the model with variables for age, gender, household composition, social class, car ownership, housing type and employment status. The trend data from the census and BHPS longitudinal information was used to forecast how the number of individuals in each of these categories would change over time. This produced the synthetic populations in each of the model's forecast years which were used when forecasting the impact of policy options with the model.

In the transport modelling world, the task of producing a dataset for individuals by merging data sources has been tackled by the developers of activity based models where the process is known as 'population synthesis'. The task is to build a dataset that contains the socio-economic, demographic and geographical details of individuals in order to then estimate the daily activity plans of each person. An activity plan is an outline of their activities for the day described in terms of Hagerstrand's time - space prisms. Peter Vovsha, an experienced practitioner in the USA, states that 'synthetic populations are **essential** to simulating individual activity-travel patterns' (Vovsha, 2012). From the activity plan built for each person, the time, start and finish point of each trip undertaken in the course of executing this activity plan is constructed and fed into the transport model.

8.4 Methods for producing a population dataset

The two main methods for producing full datasets of the attributes for each person for use in microsimulation models are Combinatorial Optimisation and Synthethic Reconstruction using Iterative Proportional Fitting (Ryan, 2007).

In both approaches two datasets are used. The first dataset is often described as being 'short but wide'. It contains detailed information but is only available for a few agents in the study area. Examples of this type of dataset include the Sample of Anonymised Records available from the UK Census or data obtained from specifically commissioned surveys such as household travel diaries. The second dataset contains information on every person but only for a limited set of variables. This is a 'long but narrow' dataset which often comes from general census data. In both techniques at least one attribute, such as age, needs to be common to both datasets.

8.4.1 Combinatorial Optimisation

In the Combinatorial Optimisation (CO) approach, the study area is divided into a set of mutually exclusive zones that completely cover the modelled area. A synthetic population is created separately for each zone in the study area, by randomly selecting people from the short and wide dataset until the required number of people (known from the long but narrow dataset) have been selected. It is likely that some people will be selected from the short but wide dataset multiple times. The distribution in the synthetic dataset of any or several common attributes with the first dataset in each zone, is compared with the actual distribution of these attributes in the first dataset, for example the number of people by age bands in the synthetic dataset for each zone is compared with the observed number of people in each age band from the census data. The fit between the synthetic and observed distribution is measured by the $RSSZ_m$ statistic, Relative Sum of Squared Z-scores (Huang and Williamson, 2001). The $RSSZ_m$ equals zero when the two distributions are an exact match.

After comparing the synthetic and observed distribution, one of the individuals is removed from the synthetic dataset and replaced by another randomly selected individual. If the fit statistic is improved with the replacement of this individual, this person is kept. Otherwise the replacement is rejected and an alternative replacement selected. This procedure is repeated until the match between the two distributions on the selected comparator attributes meets a pre-defined threshold set by the user.

The CO approach produces an output dataset that consists of 'whole' agents as required for microsimulation. This is the long and wide dataset of 'whole' agents also required for an agent based model.

8.4.2 Synthetic Reconstruction

Synthetic Reconstruction using Iterative Proportional Fitting (IPF) is the most commonly used method in activity based models, for example by Beckman (1996) for models built using TRANSIMS, by Frick and Axhausen (2004) for MATSIM models and Arentze *et al.* (2007) in the Dutch ALBATROSS model.

The IPF technique is similar to the Furness technique widely used in transport planning to update a base year trip matrix to a future year using estimates of the row and column totals for the future year. The starting point is the short and wide dataset. The number of people in this dataset in each zone, for a certain characteristic (e.g. age band), is compared to the target number taken from the long but narrow dataset and a factor calculated that is applied to each person in the short/wide dataset. Then the number of people in the short/wide dataset for a second required attribute (e.g. car ownership) after applying the factor to match the first attribute, (e.g. age), is compared to the target number in the long/narrow dataset. A new factor is calculated and applied to each person in the short/wide dataset.

This means that the age band totals in the short/wide dataset no longer match the age band totals in the long/narrow dataset, so for each zone the observed and target totals for this attribute are divided again to produce a new factor that is then applied to the short/wide dataset. This then means that the car ownership totals differ between the short/wide and the long/narrow datasets so another factor is calculated and applied in order to match these again. This procedure is carried out in turn for all the attributes which are being used to match the short/wide with the long/narrow dataset. Sufficient iterations are carried out until the match between the target totals for each attribute in the long/narrow dataset match those in the short/wide dataset within a tolerance level selected by the user.

The output of the process is the short/wide dataset with a weighting attached to each record so that tabulations from the short/wide dataset using this weighting factor closely match the tabulations from the long/narrow dataset for the shared attributes. The benefits of the IPF technique include the speed of computation, its simplicity and guarantee of convergence (Pritchard and Miller, 2012) but it produces non-integer weights which results in some complete and some fractional individuals when the dataset is converted into persons for use in an agent based model. Techniques such as Truncate, Replicate, Sample (Lovelace, 2013) address this issue and convert a dataset produced using IPF into a dataset of whole people for use in an agent based model.

8.5 Data on agent preferences

The model of commuter mode choice developed in this study assumes that each person makes their mode choice decision individually, that when they make a choice they are able to trade-off their preferences for time, cost and carbon emissions, and that they always act rationally choosing the mode which gives them the maximum utility. The model requires that each individual has a stable set of known preferences. For the proof-of-concept model used in this study, which modelled a world consisting of the 626 people interviewed in the DfT study, the information on individual preferences was available. However, if the model were to be applied to a wider area it is unlikely that data on each person's preferences would be available.

There are a number of practical difficulties with collecting data on each person's preferences. First, there is the expense of carrying out a stated preference survey to deduce a person's preferences. These surveys are expensive and budget constraints limit the number of surveys that can be carried out. Second, it is extremely unlikely that all the people in an area will agree to participate in the survey or be available when the survey is conducted so there will be some people who should be in the model but for whom no preference data is available. Third, if the model is used to forecast a scenario in the future, it is not known which people will be living there in the future. Some of the current residents will remain but it is not known who will move into the area and who will leave. Finally, there are methodological concerns with deducing preferences for a person from the limited number of questions that can be presented to a single person in a stated preference survey.

Respondent fatigue when completing surveys means that often only up to around 10 or 12 choice sets are presented to a respondent, but even with three variables and only 2 or 3 levels for each variable it is not possible to capture sufficient combinations to ensure that the survey elicits the preferences correctly. In addition, the respondent may not pay full attention to some of the choice sets, particularly if the survey is lengthened to include more questions. This has led to the common practice of dividing the total set of questions needed to produce an orthogonal design amongst a series of complementary surveys. Each respondent is given one of the sub-set of survey questions and the responses are combined together to produce a single set of weightings for a 'representative' person.

8.6 Segmentation

Common practice, built into the design of the software used for conventional transport models such as Saturn, Visum and Cube Voyager, is for travellers to be segmented into groups as discussed previously in section 3.2.4, with the trip origins and destinations of members of each group stored in a single matrix. The segregation of people used in transport models is based on:

- journey purpose, often employer's business, commuting and other
- car available, not available
- time of day
- income (particularly in models designed to test toll road and road pricing proposals)

The segmentation is designed in advance of an examination of the data and is built into the design of the model as trips are grouped into separate matrices, one for each segment. A segmentation based on observed characteristics such as those listed above may not be the most successful at dividing people into groups with shared preferences nor completely accurate in grouping people so that they share the same costs and constraints.

By using an agent based modelling approach, the unit of analysis is at the most dis-aggregate level possible, the individual. As information is recorded throughout the modelling process at the level of the individual they can be aggregated into the groupings most appropriate for each particular modelling task being undertaken. For example, people could be segmented into one set of groups for trip generation modelling, another set for modelling mode choice decisions and yet another for route assignment. As long as the agent attributes has the relevant fields needed to keep track of the segment to which the individual belongs for each of the modelling processes, the same individual can be assigned to the most appropriate group for each of the modelling processes.

There are numerous methodologies that can be employed to set up a segmentation system. Schreiber and Pekarik (2014) distinguish between groupings based on *constructed composites* which are derived directly from the original variables (e.g. journey purpose, socioeconomic status) and *latent constructs* which are derived indirectly. Latent constructs point to underlying characteristics (e.g. preference for time saving) which is not directly measured but can be identified through mathematical modelling. For the purposes of this model the ideal method for grouping respondents is to use groups which share the same preferences for time/cost/carbon emissions and then to examine the observable characteristics of these people so as to construct a way of allocating agents in the model into the relevant preference group. Latent class analysis is a methodology ideally suited to the task of segmenting agents by their preferences.

8.7 Latent class analysis

8.7.1 Latent class analysis and mode choice preferences

The feasibility of using latent class analysis to identify groups for use in an agent based mode choice model was investigated using a dataset provided by South Yorkshire Passenger Transport Executive. The data comes from a survey designed to inform decision making over the provision of improved public transport between Sheffield and Rotherham. The options under consideration were to improve the current bus service, extend the tram to cover the whole route or to construct a bus rapid transit (brt) route. A bus rapid transit system has bus lanes for a high proportion of its route, bus priority at junctions and modern high quality vehicles.

The survey was conducted in December 2007. A mixture of people who currently use car, bus, rail or tram within the Sheffield to Rotherham corridor and are not exempt from paying their own fare on public transport fare were interviewed. Respondents were recruited on rail services between Sheffield and Rotherham, at key bus stops between Sheffield and Rotherham, on Supertram services between Sheffield and Meadowhall (which lies between Sheffield and Rotherham) and at car parks in Sheffield and Rotherham city centres. The survey was a paper exercise and respondents were asked to complete the survey and post it back. The number of completed surveys is shown in Table 8.1 below.

Current mode	Area	Number of responses
Bus	On Sheffield – Rotherham - Meadowhall corridor	136
Tram	On Sheffield – Meadowhall routes	92
Rail	On services between Sheffield, Meadowhall and Rotherham	90
Car	Users of central car parks in Sheffield and Rotherham, travelling from within the proposed brt / tram corridor	71

TABLE 8.1 NUMBER OF COMPLETED SURVEYS BY CURRENT MODE OF TRAVEL

The survey consisted of four parts. The first part asked respondents about their current journey between Sheffield and Rotherham. The second section presented a set of scenarios for alternative journey options for travel by bus, bus rapid transit or tram and asked the respondent their preferred choice. The third section asked attitude questions towards public transport. The final section asked about the respondent’s attributes such as age, gender income and employment status.

The second section, where respondents were asked to choose between alternative options for making a trip between Sheffield and Rotherham contained a set of nine exercises. Each time the respondent was asked to make a choice between three options. Each option was described with a picture of the vehicle, the journey time, the cost of the journey and frequency of service. An example from the survey is shown in Figure 8.2 below.

The bus vehicle to the left was described to the respondent as being diesel powered with uneven ride and acceleration. The bus rapid transit vehicle, in the centre, was described as being diesel powered with a smooth ride and acceleration. The tram, on the right, was described as being electrically powered and having very smooth ride and acceleration.

The design of the survey had 27 choice sets in order to ensure orthogonality (Kocur *et al.*, 1982). In order to keep the length of each survey to within a respondent’s expected attention span, the scenarios were divided into three sets. This produced three different versions of the survey each containing 9 exercises. The survey was completed by 369 fare paying passengers.

FIGURE 8.2 A TYPICAL SET OF ALTERNATIVES PRESENTED IN THE STATED PREFERENCE SURVEY

Latent class analysis was carried out using Latent Gold software to investigate if it was possible to discern distinct groupings amongst these respondents based on the relative value they placed on the type of public transport, journey time, cost, and frequency of service. The latent class analysis was run to examine the groupings achieved for up to six classes. The BIC statistic, which indicates how well a model with a particular number of groups describes the natural number of segments among the respondents, was used to judge which number of classes produced the best match with the data. Collins *et al.* (2010) recommend that the number of classes selected for the final model is based on a low BIC value, the nature of the clusters produced and an assessment of whether they create a useful typology of the situation to be modelled. Based on both the BIC statistic and a review of the differences between the segments in each of the models, the four class model was selected. The number of respondents allocated to each class is given in Table 8.2 below.

	Class1	Class2	Class3	Class4
Percentage	34%	26%	23%	16%
Frequency	127	97	86	59

TABLE 8.2 NUMBER OF RESPONDENTS IN EACH CLASS

The relative importance accorded by members of each class to each of the journey attributes for the 4 class model, is shown in Table 8.3 and Figure 8.3 below. Members of class 1 (34% of respondents) are primarily concerned with the frequency of public transport services and much less concerned with the actual journey time or type of public transport mode. As tram services carry more people per vehicle but run less frequently, these people are likely to prefer bus or bus rapid transit services as these would run with higher frequency than the tram.

The members of class 2 (26%) gave most importance to fare and less to journey time and frequency. Members of both classes 3 (23%) and 4 (16%) had public transport type as their most important attribute. For members of class 3, although public transport type was the most important attribute, fare was almost as important. But for members of class 4, the 'tram lovers', the mode was significantly more important in their decision making than journey time, frequency or fare.

Journey characteristic	Class1	Class2	Class3	Class4
FARE	0.2808	**0.4797**	0.3237	0.0558
FREQUENCY	**0.5269**	0.1301	0.0777	0.0109
TIME	0.1594	0.0913	0.2248	0.2459
TYPE	0.0329	0.2989	**0.3737**	**0.6873**

TABLE 8.3 RELATIVE IMPORTANCE OF ATTRIBUTES BY CLASS

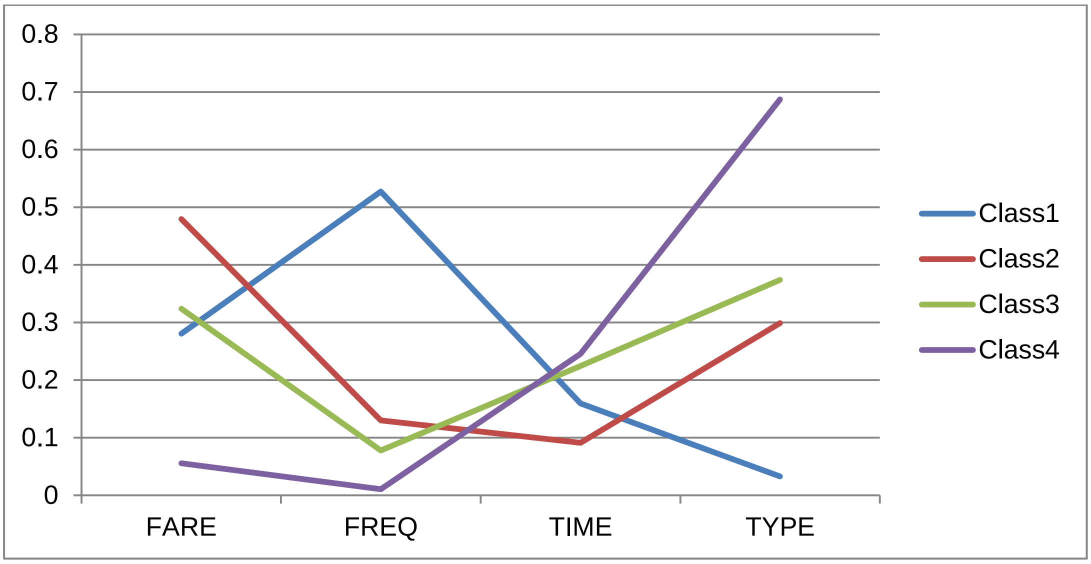


FIGURE 8.3 RELATIVE IMPORTANCE OF ATTRIBUTES BY CLASS

Further information on the public transport type valued by each of the four classes is revealed by considering the actual vehicle type which was valued highest by the members of each class. These values are given in Table 8.4 below. Respondents in Class 3, for whom fare as well as public transport type is important, strongly favour bus rapid transit or tram over the bus. Class 4, who are overwhelmingly concerned with public transport type, strongly favour the tram.

Vehicle type	Class 1	Class 2	Class 3	Class 4
Class size	34%	26%	23%	16%
Bus	0.323	0.254	0.045	0.024
BRT	**0.350**	**0.743**	**0.528**	0.076
Tram	0.327	0.003	0.427	**0.900**
Total	1.000	1.000	1.000	1.000

TABLE 8.4 RELATIVE PREFERENCES AMONG PUBLIC TRANSPORT TYPES BY CLASS

8.7.2 Placing agents into classes

As well as uncovering distinct classes amongst respondents, latent class analysis can be used to predict the probability of a person belonging in a particular class. Table 8.5 below shows the probability of respondents belonging to a particular class, together with the socio-economic data collected about each respondent. For example respondent 1002, given the choices they made in the survey, has a 99% probability of belonging to class 2. The next respondent, 1056, has a 1% probability of belonging in class 1, a 89% probability of belonging in class 2 and a 11% probability of belonging in class 4. The final column of Table 8.5 shows the most likely class membership forthe respondent.

Serial	Current Mode	Gender	Age	Income	Employment	Class 1	Class 2	Class 3	Class 4	Allocated Class
1002	tram	female	45-54	20,000 - 34,999	full time	0.00	0.99	0.00	0.00	2
1056	tram	male	55-64	prefer not to say	full time	0.01	0.89	0.00	0.11	2
1063	bus	female	16-24	10,000 - 19,1009	Education	1.00	0.00	0.00	0.00	1
1065	tram	female	25-34	35,000 - 49,999	full time	0.00	1.00	0.00	0.00	2
1071	tram	male	35-44	35,000 - 49,999	full time	0.19	0.81	0.00	0.00	2
1073	tram	female	16-24	less than 10,000	Education	0.94	0.05	0.00	0.00	1
1076	tram	female	35-44	20,000 - 34,999	full time	0.00	1.00	0.00	0.00	2
1102	car	male	45-54	20,000 - 34,1066	full time	0.01	0.99	0.00	0.00	2

TABLE 8.5 ALLOCATION OF INDIVIDUALS TO A CLASS

The information on the socio-economic characteristics of each respondent and their most likely class membership can also be used to assess, using CHAID analysis, which items of information about a person, such as their income or education, will provide a good indicator of their class membership. This enables class membership to be assigned to people not interviewed in the survey but for whom this information is available, possibly from census data.

8.7.3 Determining the characteristics used to allocate people to classes

The analysis of the Sheffield data showed that the only statistically significant socio-economic indicators of a respondent's class membership were current mode used at the time of the survey, income and age. The non-significant indicators were employment status, gender and journey purpose. The strongest indicator of class membership was the current mode used, with 29% of current car and rail users belonging to class 4 (tram lovers). Only 5% of current bus users belonged to class 4. A higher income was an indicator of membership of the 'tram lovers' class.

8.7.4 Latent class analysis with the DfT Climate Change dataset

Having demonstrated the feasibility of using latent class analysis to generate segments with distinctive behavioural preferences using the Sheffield dataset, the technique was used with the DfT dataset described in chapter 6 to segment the agents into distinct groups of people who shared similar preferences. The latent class analysis was conducted using Sawtooth software as the survey data had been collected using this software package and it was held in the software's proprietary binary file format. The results are presented below for the five class segmentation. The weightings for each class, re-scaled for comparability, are presented in Table 8.6 below. The factor with the strongest influence on mode choice is highlighted in yellow.

		Class 1	Class 2	Class 3	Class 4	Class 5
Number of respondents		**14.80%**	**6.50%**	**33.00%**	**32.70%**	**13.10%**
Mode	**Bus**	-81.17	117.28	20.85	-18.28	-38.42
	Car	213.31	-103.56	67.65	18.76	-24.07
	Train	-57.47	74.17	38.33	-10.56	-52.40
	Cycle	-74.66	-87.89	-126.83	10.08	114.90
Cost	**£1.50**	-2.72	24.21	16.09	24.51	29.01
	£2.00	5.85	-19.08	8.03	7.83	19.56
	£2.50	5.84	16.74	-8.45	-10.32	-17.25
	£3.00	-8.98	-21.87	-15.67	-22.02	-31.33
CO2	**1 kg**	8.98	15.25	11.18	13.75	18.36
	2 kg	-1.58	4.04	1.21	5.21	9.69
	3 kg	9.76	-23.97	-4.87	-0.44	-14.15
	4 kg	-17.16	4.68	-7.52	-18.52	-13.90
Time	**15 min**	36.46	42.24	80.41	146.55	63.87
	30 min	16.54	25.37	32.45	74.16	43.07
	45 min	-20.66	1.19	0.05	-6.83	16.68
	60 min	-5.05	-17.17	-38.26	-76.27	-47.64
	75 min	-27.30	-51.62	-74.65	-137.61	-75.98

TABLE 8.6 FIVE CLASS GROUPING OF RESPONDENTS TO DFT SURVEY

Class 1 (14.8%) are Car Lovers. The factor with the strongest influence over their choice of mode is the actual mode itself with a strong preference for car (+213). Their least preferred mode is bus. The second factor affecting their mode choice is time with cost and carbon emissions as the least important influencing factors.

Class 2 (6.5%) are Public Transport Fans, with a strong preference for bus and train and a dislike of car and cycle. Their next strongest consideration was journey time which was more important than either cost or carbon. This group took the level of carbon emissions into greater consideration than any of the other groups.

Class 3 (33.0%) are Time Savers. Their strongest consideration is for a quick journey time. They are moderately in favour of car but show a strong dis-like of cycling. Cost and carbon emissions are not important factors in their choice of mode.

Class 4 (32.7%) are Extreme Time Savers. Their choice of mode is dominated by the journey time and the mode itself is of little importance to them.

Class 5 (13.10%) are Cyclists. Their preferred mode is cycling and this exerts a strong influence on their choice of mode. However time is also an important consideration in their mode choice decision.

8.7.5 Using the maximum utility rule with segments

The agent based model presented in chapter 7 used the rule that each agent choses the mode that gives them the highest or maximum utility. If the preferences of each individual agent are not known, due to the practical and financial constraints of collecting this data for each person in the study area, the agents could be modelled using the set of preferences common to their segment or latent class. A model run was carried out to investigate the impact on model results of using the **maximum utility** rule with segment rather than individual preferences.

The model was run for Scenario 2 which has periodic increases in car parking charges, without the availability of bus passes or constraints on modes available to each person, and in the first instance, without habitual behaviour. The input values for this scenario were provided in chapter 7 in Figure 7.3. The model recorded the mode choice decision of each agent using their own preference set and the preference set of their latent class.

The number of people predicted to use bus, train and car is shown in Figures 8.4 to 8.7 below.

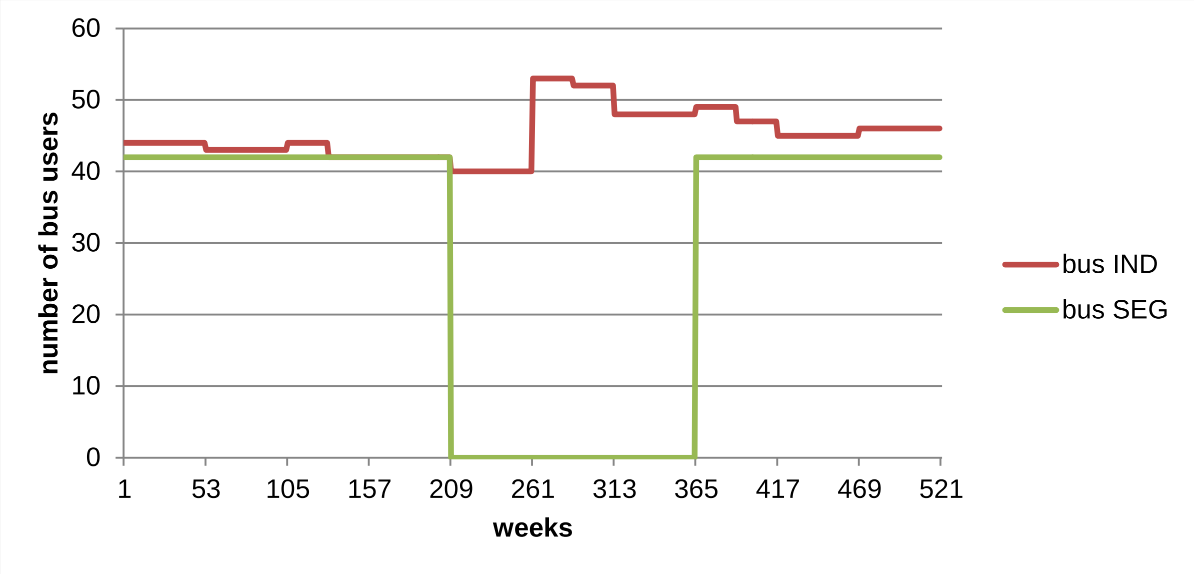


FIGURE 8.4 NUMBER OF BUS USERS WITH SEGMENTATION

When the agents each have their own set of preferences, the biggest change in the number of bus users occurs at the start of year 5, rising from 40 to 53. Until the increase in car park charge and the number of people having to pay to park, which occurs at the start of year 5, the number of bus users had ranged between 40 and 43. After the increase in parking costs causes the number of bus users to rise at the start of year 5, the number of bus users generally falls as bus fares continue to rise.

In contrast, when the common preference function is used for each agent in a latent class or segment, the number of bus users is steady until the start of year 4 and then drops suddenly to zero, as a whole group of agents switch away from bus. They then all return in year 7. The sudden swings in the number of people using each mode and the difference between the forecast number of users of each mode when using individual or segment preference functions, is also seen in Figures 8.5 to 8.7 for car, train and cycle. With the use of segments, at the end of year 10, the number of bus and car users is lower while the number of rail users and cyclists is higher than when using individual preference functions.

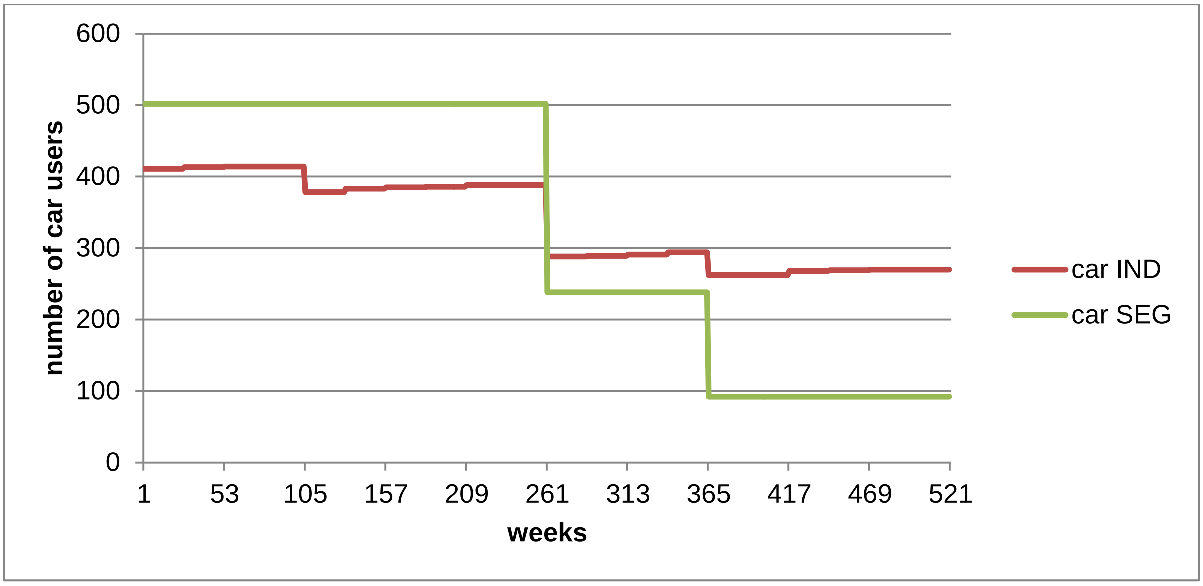


FIGURE 8.5 NUMBER OF CAR USERS WITH SEGMENTATION

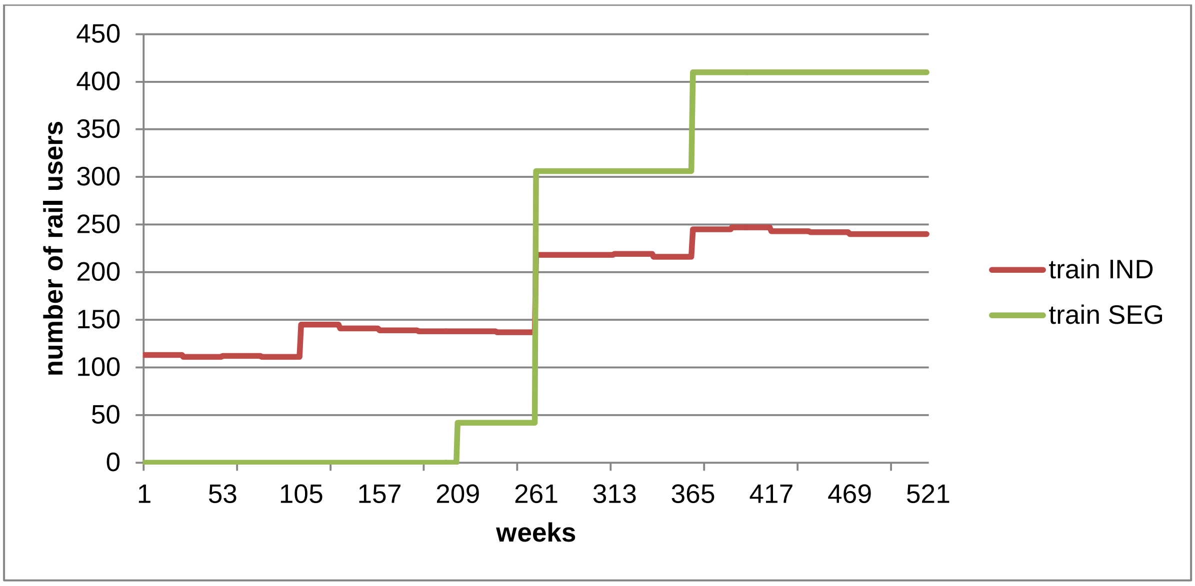


FIGURE 8.6 NUMBER OF RAIL USERS WITH SEGMENTATION

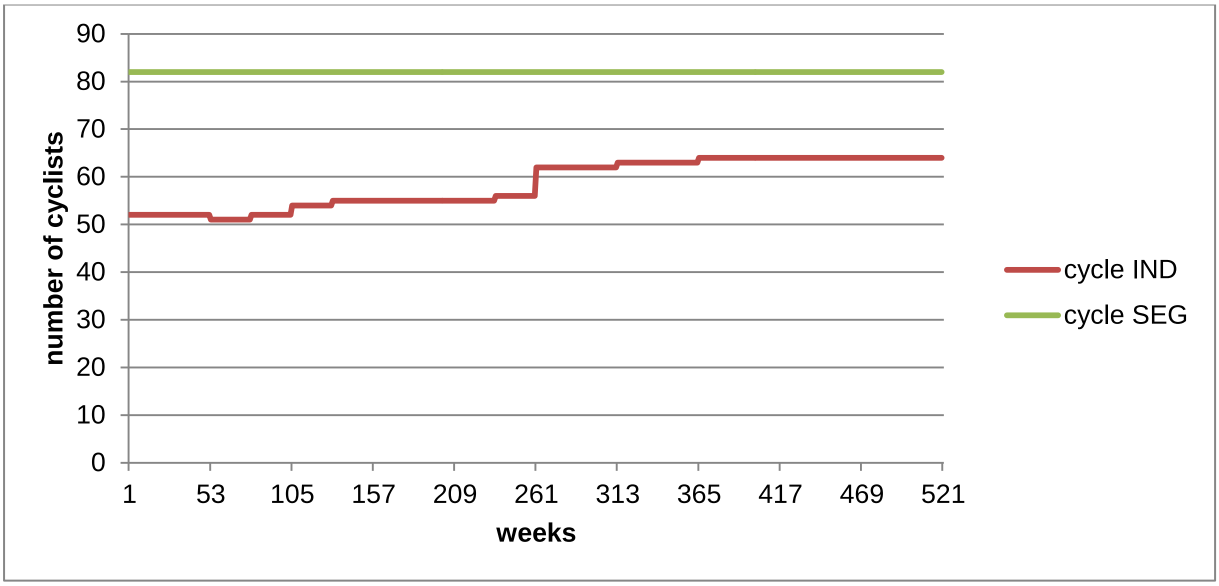


FIGURE 8.7 NUMBER OF CYCLE USERS WITH SEGMENTATION

These results show that with the use of shared preferences for each group of agents rather than individual preferences, the number of users of each mode varies considerably when there is a change in the relative costs of modes of sufficient magnitude to change which mode has the highest utility. This then causes a whole block of agents to switch at the same time.

The introduction of habitual behaviour and constraints on mode choices reduces the impact in each time period of using segment rather than individual preferences as not all agents will make a mode choice decision at the same time and so not all agents in the same segment would change in the same time period. Figures 8.8 to 8.11 below show the number of users of each mode for the same model run but with habitual behaviour and constrained choice sets.

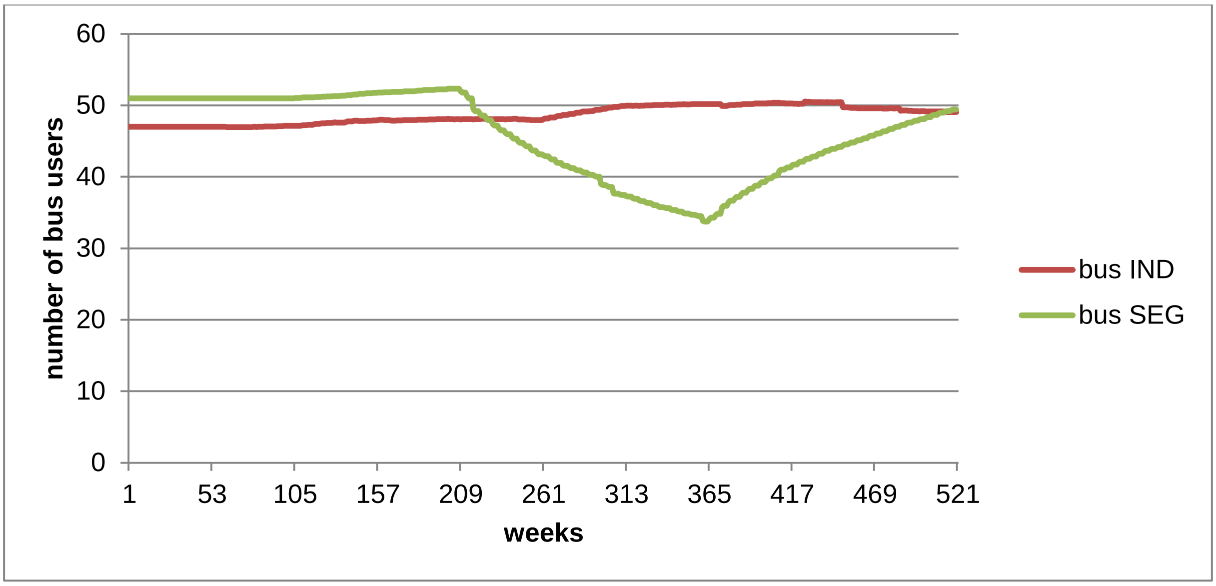


FIGURE 8.8 NUMBER OF BUS USERS WITH HABITUAL BEHAVIOUR AND CONSTRAINED CHOICE SET

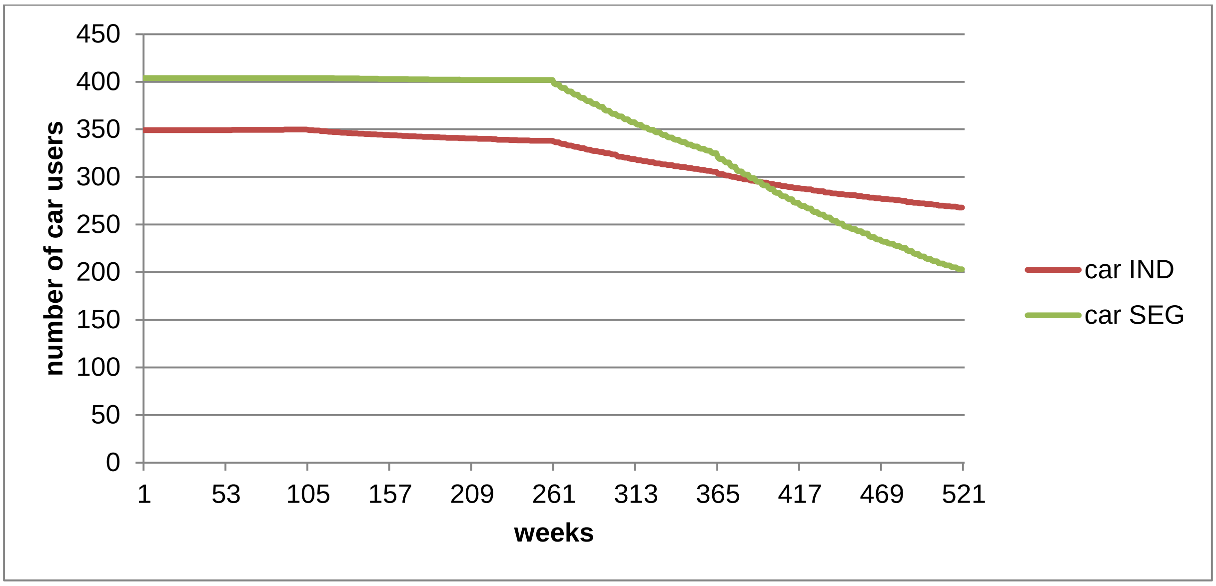


FIGURE 8.9 NUMBER OF CAR USERS WITH HABITUAL BEHAVIOUR AND CONSTRAINED CHOICE SET

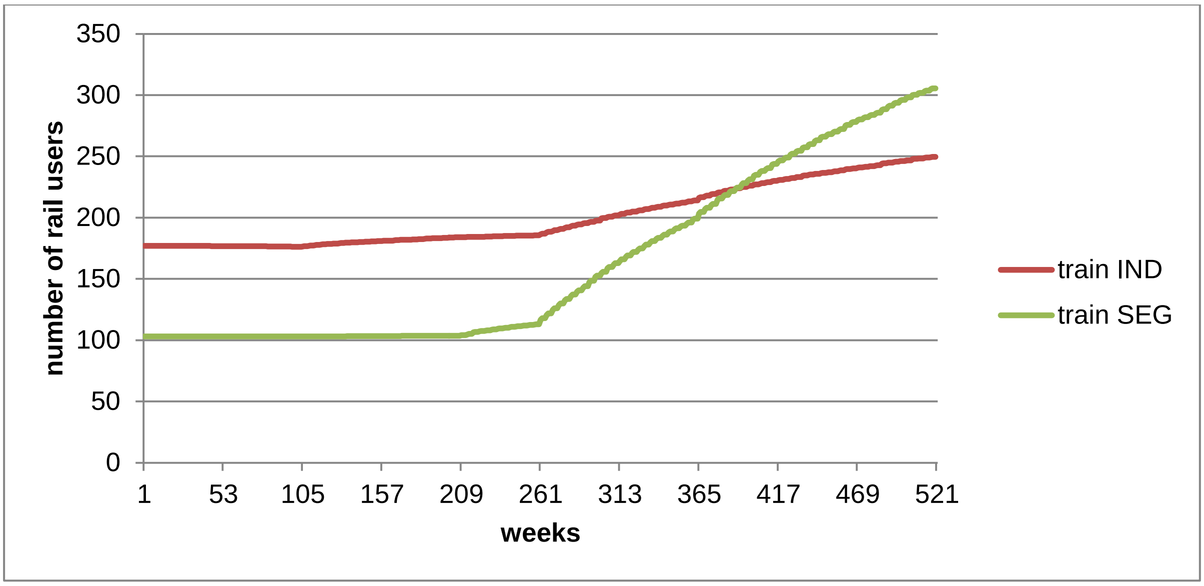


FIGURE 8.10 NUMBER OF RAIL USERS WITH HABITUAL BEHAVIOUR AND CONSTRAINED CHOICE SET

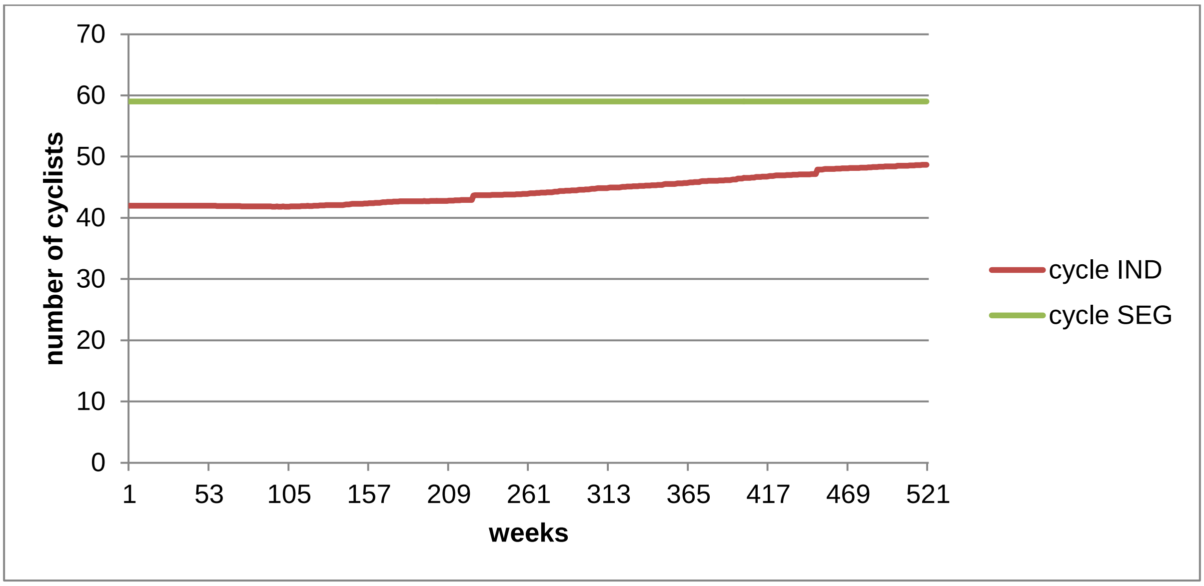


FIGURE 8.11 NUMBER OF CYCLISTS WITH HABITUAL BEHAVIOUR AND CONSTRAINED CHOICE SET

The introduction of habitual behaviour has moderated the sudden swings in the number of people using each mode with choices based on the maximum utility rule but there is still a sometimes substantial and varying difference between the forecast number of users of each mode depending on whether an agent uses their own individual preference function or that of their group.

This suggests that the maximum utility rule may not be appropriate to use when the agents are placed into segments. A possible alternative approach was explored, in which a logit model (a linear-in-parameters random utility model) was used to determine the probability of an agent choosing a particular mode based on the utilities of each of the options. For each agent, whenever they make a mode choice decision, a random number is drawn and compared to the calculated probabilities of their using each of the alternative modes in order to assign the agent to one particular mode.

Figures 8.12 to 8.15 show the number of users of each mode with individual preferences following the maximum utility rule (IND MAX UTIL) and with segment preferences using the logit model (SEG LOGIT). The results are taken from the same model run, with habitual behaviour and constrained choice, so that the events such as moving house and changing job, occur at the same time for each individual.

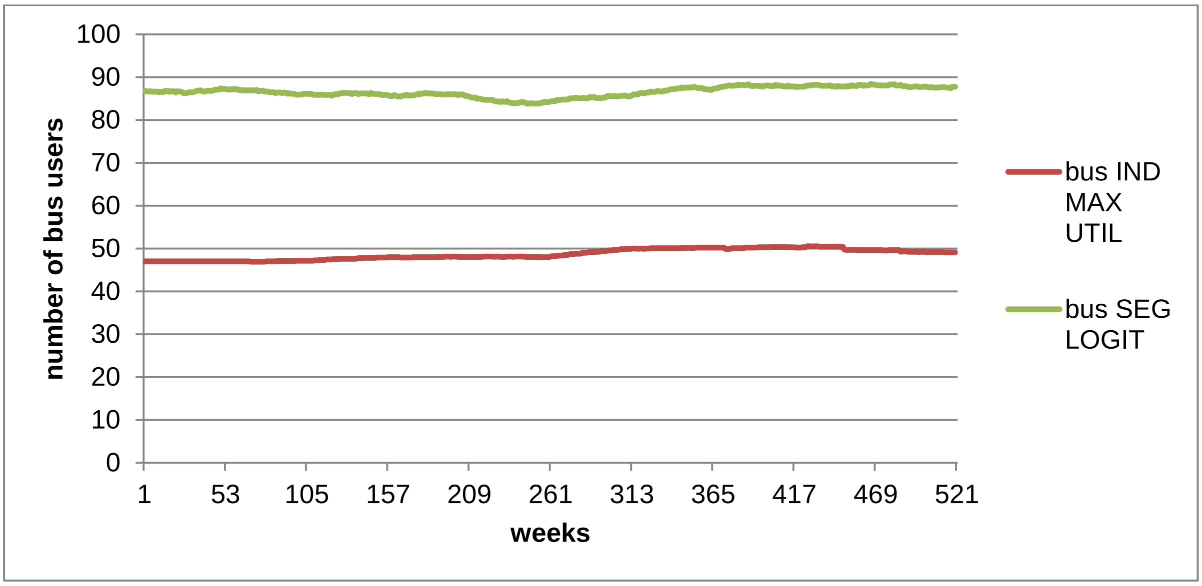


FIGURE 8.12 NUMBER OF BUS USERS, INDIVIDUAL PREFERENCES WITH MAXIMUM UTILITY RULE AND SEGMENT PREFERENCES USING A LOGIT MODEL

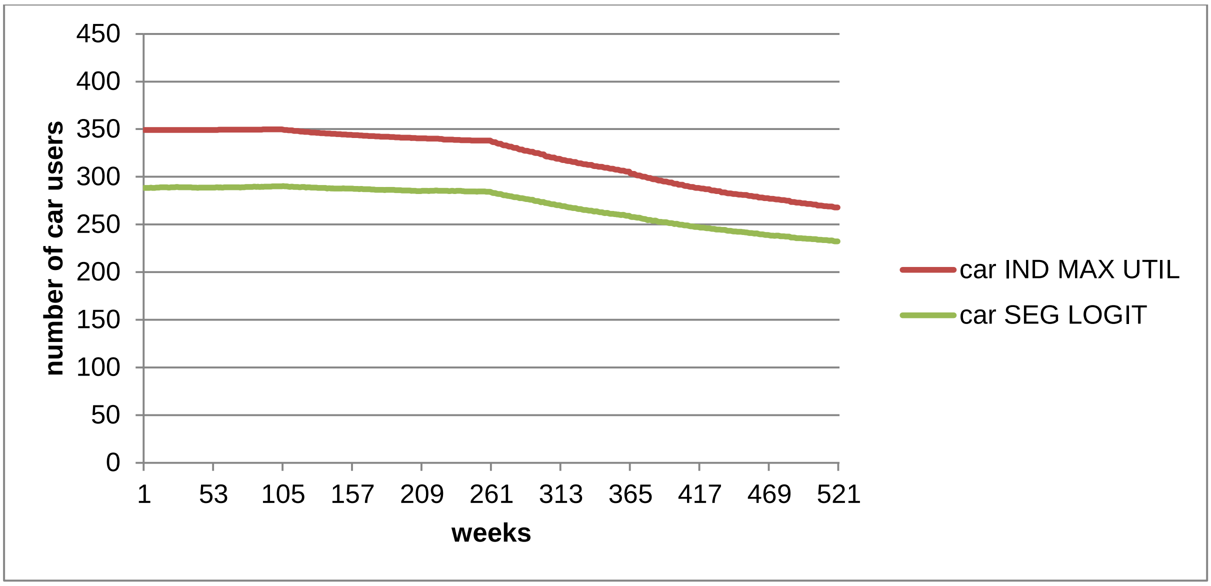


FIGURE 8.13 NUMBER OF CAR USERS, INDIVIDUAL PREFERENCES, WITH MAXIMUM UTILITY RULE AND SEGMENT PREFERENCES USING A LOGIT MODEL

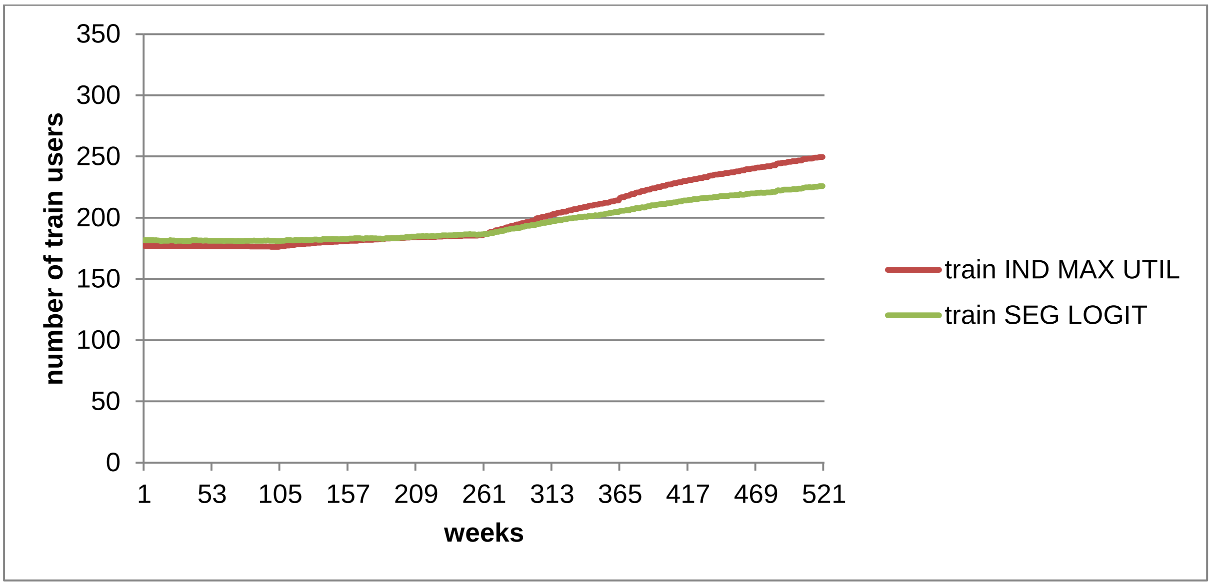


FIGURE 8.14 NUMBER OF RAIL USERS, INDIVIDUAL PREFERENCES, WITH MAXIMUM UTILITY RULE AND SEGMENT PREFERENCES USING A LOGIT MODEL

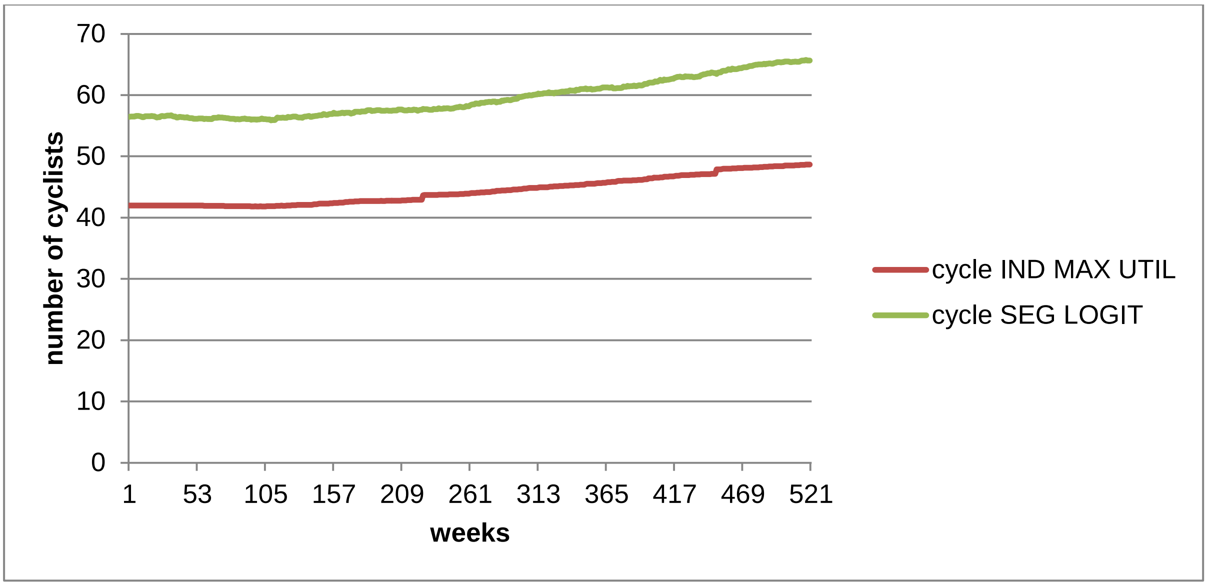


FIGURE 8.15 NUMBER OF CYCLISTS, INDIVIDUAL PREFERENCES, WITH MAXIMUM UTILITY RULE AND SEGMENT PREFERENCES USING A LOGIT MODEL

Using the logit model with the shared preferences for each segment removed the discontinuity in the number of bus users but gave a consistently higher estimate of the number of bus users. For car, the use of the logit model resulted in a trend in the number of car users that closely reflected that obtained from using individual preferences, but consistently underestimated the number of car users (Figure 8.13) against the benchmark of the number of users predicted using individual preferences and the maximum utility rule. For train users (Figure 8.14) the use of the logit model meant that the forecast number of train users was similar using segments as for individual modelling in the early years and then it underestimated the number of rail users. For cycling (Figure 8.15) the use of the logit model for the segments over estimated the number of cyclists but showed the same trend in usage as the individual maximum utility rule.

Using the logit model for mode choice when using shared preferences is preferable to using the maximum utility rule as it removes the sudden changes in the results but it will require the use of multiple model runs to produce stable results because of the use of a random number for comparison with the mode probabilities to assign a person to a mode.

8.7.6 Aggregation and the representative agent

The transparency of the agent based modelling approach highlights the issue that affects, but is seldom recognised in reviews of the standard four stage model, that modelling each individual will produce different overall results than if groups are modelled. This topic is discussed in the field of economics when considering the links between microeconomic theory, which studies individual economic entities, and macroeconomics, which looks at the behaviour of the market as a whole.

In neo-classical economics the link between the behaviour of individuals and the behaviour of consumers as a whole is achieved by aggregating the individual demand curves to a single aggregated market demand curve. However Gorman (1953) showed that the forecast of change in the quantity of a good demanded as price changes, taken from the aggregate demand curve, is only the same as the forecast change in demand for a similar price change predicted by using each of the individuals' demand curves under very restrictive conditions. This is because a person's preferences, as captured in their demand curve, show the relationship between the quantity of a good consumed and its price, **if everything else is kept constant**, including the income of that person.

Gorman (1953) showed that the conditions under which the results from using the aggregated demand curve match the results from using individual demand curves are that the preferences of every individual person must remain constant as their income changes and that the income distribution of the people represented by the aggregate demand curve remains constant.

The aggregate demand curve is the summation of all the individual demand curves, at a particular point in time and given the income of each of those individuals at that time. If an individual suffers a significant drop in income, following for example the loss of their job, this change in income is likely to cause the individuals' preference function (or the relative weights applied to a unit change in time or cost of a journey) to change as they become more sensitive to the same increase in the cost of travel as their income falls. The shape of their individual demand curve has changed as a result of their change in income and the shape of the aggregate demand curve, calculated using their original individual demand curve, is no longer an accurate representation of the current aggregate demand curve.

Economists responded to Gorman's work by switching from the use of the aggregate demand curve to the use of the demand curve of a representative agent. Demand forecasts are based on the preference function of a single 'representative' agent. This does not however remove the problem. Alan Kirman (1992) explained that 'this reduction of the behaviour of a group of heterogeneous agents *even if they are all themselves utility maximizers*, is not simply an analytical convenience as often explained, but is both unjustified and leads to conclusions which are usually misleading and often wrong'. He showed that the use of the representative agent can not only produce a different forecast of the absolute size of change but can forecast the change in the wrong direction compared to forecasts based on using each individual's demand curve.

Chakrabarty and Schmalenbach (2002) illustrate this using data from the UK Family Expenditure Survey for the time period 1974 -1993 to test the impact of forecasting demand using a representative agent rather than aggregating the result of modelling individual's demand for five commodity groups: food; fuel and light, services, clothing and footwear, and non-durable goods. They tested the impact of changes in income level over time by forecasting the demand for certain goods each year using first a mean income for the representative agent each year, and then using each of the actual individual incomes which together produced the mean income level used. They found that the 'representative agent' with the mean income was a reasonable approximation in the case of food, fuel and light, clothing and footwear in this particular data set, but it produced statistically significant different results in the case of services and non-durable goods. They also found that changes in the income distribution over time affected the forecasts of demand when using the representative agent for services and non-durable goods.

The strongest effect was on services, which was the category with the highest income elasticity (where demand is most sensitive to changes in income). Their work shows that income effects can affect the accuracy of forecasts produced using a single representative demand curve for all people rather than individual demand curves. The varying income elasticities of different modes of transport means that this is an issue for transport forecasts. The level of discrepancy in the estimate of total demand between using an aggregate rather than individual demand curves is likely to vary by mode.

8.7.7 Applying a logit model using aggregate and individual preference functions

The difference in forecast results from using either individual preferences, shared preferences for members of each of the five latent classes or from the standard division of travellers into those with and without a car available for the trip was tested using the model software prepared in this project. Additional code was added to the agent based model to run a standard logit mode choice model to calculate the proportion of people using each mode, with people allocated to two segments (car available and car non available) and using individual preferences. The logit model was run for each of 520 time periods. The times and cost of travel are those used in scenario 2.

It should be noted that in this model the incomes of each person remained constant as far as exogenous changes in income are concerned. The only change in income is the effective change caused by changes in the cost of travel. In the real world the influences of the change in income and income distribution would be expected to accentuate the differences found here in the forecast of the number of users of each mode from using a single preference function rather than individual preference functions.

The results of the model run for each mode are shown in Figures 8.16 to 8.19 below. The red line in Figure 8.16 shows the number of people who use the bus to commute, obtained by multiplying the proportion of people using each mode from the logit model by the total number of travellers, when individual preference functions are used, (IND). The green line shows the number of people who use bus when commuters are segmented into five latent classes (SEG) and the blue line when commuters are segmented by car available and not available, as is common practice in four stage transport models, (ALL).

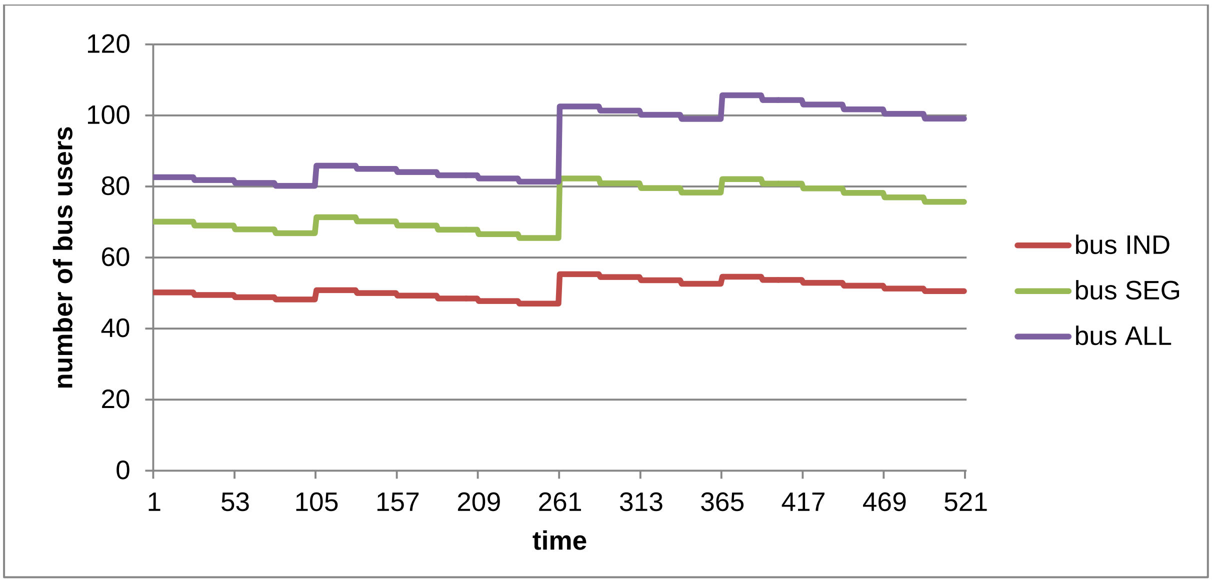


FIGURE 8.16 NUMBER OF BUS USERS FROM A LOGIT MODEL OF MODE CHOICE SHOWING THE IMPACT OF SEGMENTATION

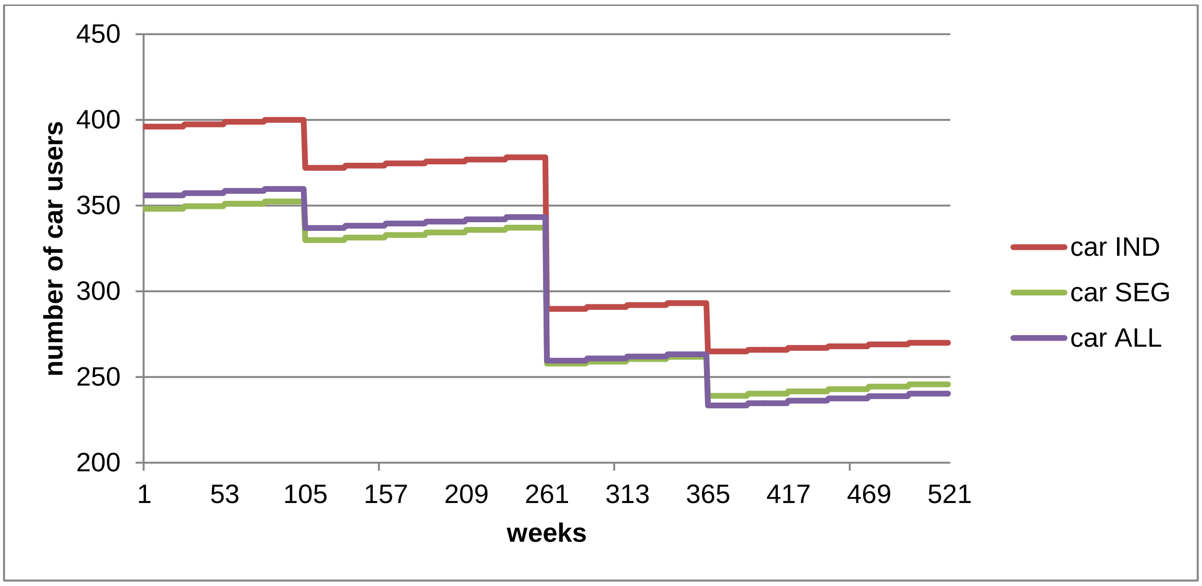


FIGURE 8.17 NUMBER OF CAR USERS FROM A LOGIT MODEL OF MODE CHOICE SHOWING THE IMPACT OF SEGMENTATION

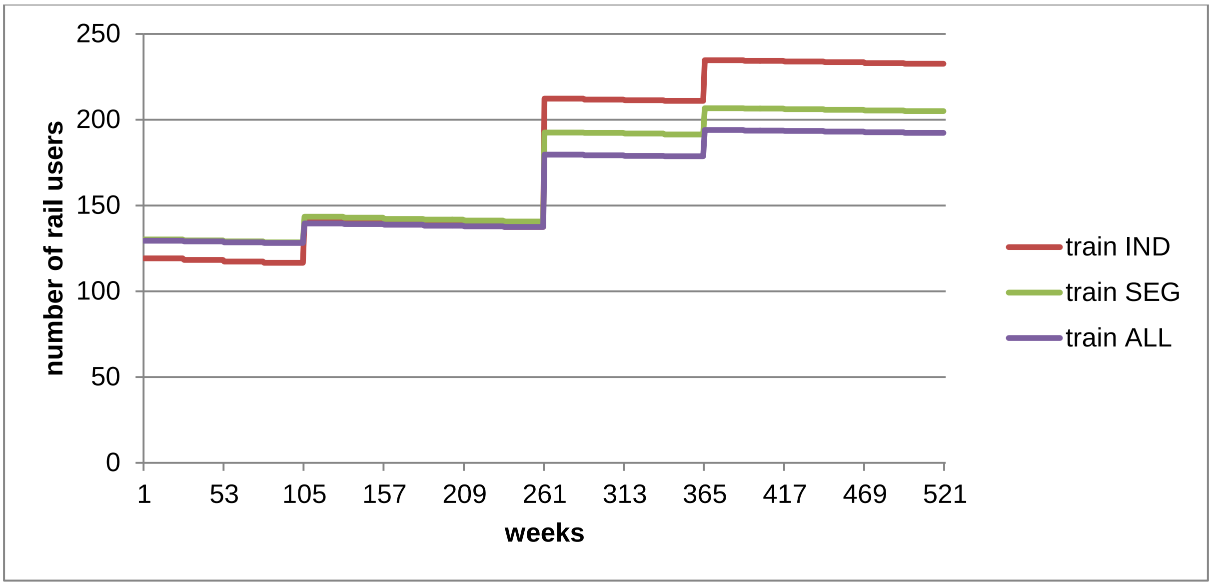


FIGURE 8.18 NUMBER OF RAIL USERS FROM A LOGIT MODEL OF MODE CHOICE SHOWING THE IMPACT OF SEGMENTATION

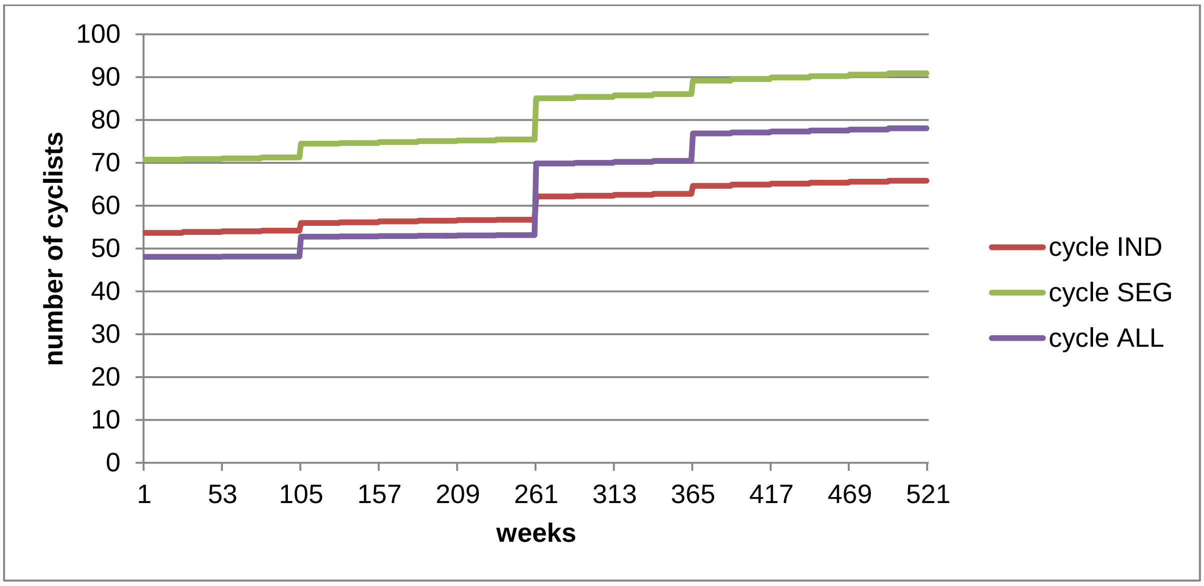


FIGURE 8.19 NUMBER OF CYCLISTS FROM A LOGIT MODEL OF MODE CHOICE SHOWING THE IMPACT OF SEGMENTATION

For the forecasts of the number of bus users, the logit model gives the highest forecast passenger numbers if the population is only split between car available and car non available people and the lowest forecast when individual preferences are used. The situation is reversed for car, with car available / non available giving the lowest forecasts of car use and the use of individual preferences giving the highest forecasts. It is a common criticism of multi-model models that they over-estimate bus use and under-estimate car use and this work suggests that the very coarse segmentation used in four stage models and the use of a single preference function for each of those segments can have a material impact on the forecasts. This has implications for the quality of the business cases built from these forecasts and the advice provided to decision makers. When dealing with modes with a low number of users, such as is often the case with bus and cycling, the variation in numbers produced by the model as a result of the amount of segmentation used in the model will be of particular concern as a small difference in absolute terms is a large difference in percentage terms.

8.8 Conclusion

Modelling mode choice at the individual level makes greater demands on the data required to build the model, as details are needed on the attributes of each person in the model and their preference function. Techniques have been developed to build a population, with information on each of their personal characteristics, for use in microsimulation models that can be applied in the development of agents in agent based models.

Establishing the preference function for each agent poses another challenge. Applying a standard logit mode choice model to the 626 respondents in the DfT dataset, showed that there is considerable variation in forecast mode shares when using individual preference functions rather than a shared preference function derived from the same data. Practical difficulties in both the cost and feasibility of collecting sufficient information to derive individual preferences means that some segmentation is required whichever modelling methodology is adopted. The potential size of the variance in mode forecasts as a result of the use of shared rather than individual preference shows that attention should be paid to the number of segments used and the allocation of individuals to those segments.

An agent based modelling approach to transport modelling, with software designed to store data at an individual level rather than in a spatially aggregated form in matrices, facilitates the creation and handling of more segments than is possible in current conventional transport modelling software packages. The error introduced into forecasts by the use of a shared preference function can be reduced by increasing the number of segments used and by using groupings based on the similarities in the individuals' preference functions. Segmentation using latent class analysis, which makes use of both observed and unobserved information, is recommended as a technique to achieve more consistent groupings of people by their preferences rather than the current practice of grouping people by observed characteristics alone such as the time of travel and journey purpose.

9 Conclusion

9.1 Motivation for this research

The four stage framework for transport models (trip generation, distribution, mode choice and assignment) was first developed in the 1950s for the purpose of planning new highways. By the 1970s and 80s the primary purpose for some transport models switched to the planning of public transport schemes, such as the extension of the Jubilee underground line in London. The replacement of diversion curves with the random utility logit model in the mode choice stage enabled the original four stage model framework to better meet this new requirement.

Recently the purpose of transport modelling has changed again with some models, particularly for mature cities in developed countries, being built to assess packages of transport implementations. These often include public transport changes and ‘smarter choices measures’ and only sometimes changes to the highway network. This has brought to transport modellers the need to again develop ways to improve the ability of their models to accommodate this new task.

This challenge provided the motivation for this research, which concentrated on the modelling of commuter trips, but the issues covered and the findings of this research are applicable to the modelling of mode choice for all journey purposes.

9.2 First research question

The research project was based around three related research questions. The first of these was ‘What has been the experience of transport modellers extending their current four stage models to include the impact of ‘smarter choices’ programmes on the mode used for commuting trips?’

This question was addressed by a review of the documentation of state of the art four stage models developed in the UK and interviews with the builders of those models to investigate whether practitioners had found satisfactory methods to include ‘smarter choices’ into their existing model frameworks. This work showed that practitioners had failed to fully integrate ‘smarter choices’ within the current four stage framework. They had been unable to implement a methodological improvement to their current modelling framework that could handle these measures, and instead applied an adjustment to the matrices coming out of the mode choice process in the last iteration of the models, before the final assignment of trips to the transport networks.

These interviews showed that part of the difficulty of modelling ‘smarter choices’ in the current framework arose from the feature that the measures were often targeted at particular groups of travellers and/or specific geographical areas. This makes them hard to incorporate in the software packages used to implement four stage models as these still incorporate the design principle of grouping trips into matrices. The use of matrices had been developed as a way of storing information on a large number of trips within the small storage capacity of the microcomputers of the 1970s. These matrices could be readily accessed and manipulated using the array functions in Fortran, a procedural high level programming language widely used at that time for technical and scientific computing.

The use of matrices means that trips are aggregated geographically into zones. Many walk and cycling trips are lost from the model altogether as they start and finish within the same zone. The use of a common set of times and costs for trips between each origin and destination zone pair ignores deviations arising from the actual location of the trip end within those zones. These deviations are of increased importance for short distance trips, which are often the trips targeted by ‘smarter choices’ as they can potentially use walk and cycling for the entire trip. Accurate information on walk access times to public transport is also essential for assessing whether public transport is a realistic alternative to the car for commuters.

Using a matrix structure for modelling also creates difficulties when ‘smarter choices’ measures are implemented in only some parts of the study area. For example Birmingham, like many other cities, intended to implement personalise travel planning combined with public transport improvements along particular corridors into the city centre. This is particularly difficult to handle within a matrix based structure that applies the same utility weightings and scaling parameter to all trips within the same matrix. In order to be able to vary these parameters geographically the matrices would have to be subdivided into matrices of trips that would use the corridor and those that would not; this would increase the number of matrices and it is difficult at the periphery of the corridors to know which trips to place in which matrix as the routing of trips could be affected by the measures introduced.

The current four stage modelling software also struggles in terms of performance and ease of file management when the number of matrices increases. This limits the amount of segmentation it is possible to implement in these models. In the UK, the DfT’s software programme TUBA which is used for the economic evaluation of transport schemes had to be re-written to handle the increased number of matrices used in the TIF models once they expanded the usual time of day, vehicle class and journey purpose segmentation to include income segmentation by low, medium and high income bands. Further segmentation, for example by demographic or other personal characteristics, was considered too impractical to implement. This means measures that affect particular groups (e.g. employees at certain companies with travel to work plans) could not be incorporated. The modellers particularly struggled to model car park policies at employment sites, yet these are known to have a significant impact on the mode choices of commuters.

Apart from the matrix structure of the current software used for four stage modelling, the standard utility functions used in logit models have also created issues for the modellers. Once the scaling parameter and mode constant had been calibrated, the logit models when used in forecasting mode were only sensitive to changes in the time and cost of travel by a mode. However some of the ‘smarter choices’ measures, such as providing information, marketing campaigns and the provision of safer cycle routes, were expected to change mode choice but do not act directly on the time or cost of travel by alternative modes and therefore their impacts could not be represented.

The experience of transport modellers when extending their current four stage models to include the impact of ‘smarter choices’ measures on the mode chosen for commuting trips was that it proved to be far more difficult than they anticipated. They failed to be able to directly model many ‘smarter choices’ initiatives and resorted to coarsely adjusting the final output matrices from the model using the findings of the Sustainable Travel Towns Studies (Sloman *et al.*, 2010) to guide the size and direction of these adjustments.

9.3 Second research question

This led on to further investigation in this research as to whether alternative modelling approaches would be better suited to the task of incorporating ‘smarter choices’ into multi-modal transport models. The second research question asked was ‘What modelling approaches could be used to model the impact of ‘smarter choices’ programmes on the mode chosen for commuting trips?’

Modelling approaches used in other disciplines were investigated to consider whether they could usefully be transferred into the transport field to tackle the modelling of mode choice. The review primarily looked at models used in the fields of marketing, health and environmental studies as these are concerned with choices and behaviour, be they made by humans or animals. Systems dynamics looks at the area to be modelled as a whole system, and is useful for examining the breadth and extent of re-inforcing and counter-acting first, second and even third order impacts of changes in either the transport system or its wider context. It is not so well suited to ‘smarter choices’ measures which often need a fine grain of spatial positioning to identify potentially affected travellers and the ability to distinguish between the circumstances and behaviour of many individuals or small groups of people.

An alternative modelling approach is to build a model from the bottom up, starting with each individual human or animal. Microsimulation and agent based modelling are dis-aggregate modelling techniques which hold considerable promise as potential approaches for mode choice modelling. The software developed to implement these types of models is well suited to handling detailed information on a very large number of individuals and tracking them through the modelling process. Microsimulation models simulate the development of the target system over time using a set of transition probabilities that govern the timing of a change in each object, and a set of rules which determine what change occurs. Microsimulation models can have a very fine spatial resolution for each agent and can handle a multitude of changes, coming from the potential transitions and possible states included in the model.

The transport modelling community already uses the microsimulation approach in dynamic traffic assignment software and activity based models. By modelling at the individual level over the course of a day these models can ensure consistency between a person's activity and transport decisions throughout the day. They can use more accurate spatial data for each person and better replicate the actual choice set of travel modes facing each person.

Agent based modelling shares these advantages and has the added benefit of being able to incorporate more behavioural realism into the modelling of choices. It provides software platforms and a modelling approach that breaks away from the constraints imposed by current four stage transport modelling software from its requirement to group people into a limited number of groups held in matrices and with the influences on the mode choice coming only from the those variables which are inputs into the utility functions used in logit models.

For these reasons agent based modelling was selected as a modelling approach for use in a practical exercise aimed at investigating the potential of this method for bringing 'smarter choices' into multi-model transport models.

9.4 Third research question

The third research question posed was ‘What are the strengths and limitations of using an agent-based approach for modelling the impact of a ‘smarter choices’ programme on the mode chosen for commuting trips?’

Agent based modelling was selected as the methodology to be used for a more in-depth investigation into its potential for modelling ‘smarter choices’. The ability of agent based models and the software that implements them to hold and track data at the individual level means that the precise geographical location of individuals can be used which gives the opportunity to reduce errors caused by geographical aggregation bias. Aggregation bias can also be reduced by modelling the actual time and cost of travel faced by each individual rather than those of the group into which they have been placed.

This research has shown that a further source of aggregation bias is present in transport models when people are grouped together into large segments and group weightings on the time and cost of travel are used when modelling the travel choices of each person rather than their individual preferences. The agent based modelling approach enables the use of individual preferences, if these are known, or the segmentation of people into groups on the basis of the similarity of their preference functions rather than, quite possibly, unrelated but observable features such as journey purpose. This research has shown that this source of aggregation bias can have a non-trivial impact on model results. (Section 8.7.7). The modelling community needs to be more aware of this weakness in the current standard implementation of logit based models for mode choice.

9.4.1 Strengths of agent based modelling

The experience of building an agent based model informed the following assessment of the strengths and limitations of this modelling approach for the task of incorporating smarter choices into multi-modal modelling. The key strengths of ABM are the ability to hold and use data on the attributes and circumstances of each individual which provides the opportunity to model measures that apply only to certain individuals, to customise the actual choice set facing each person and to have a more accurate description of the time, cost and other features of each of the mode options in their choice set.

The ability to handle individual preferences, or a shared preference set for a large number of groups of people, provides a way of minimising the aggregation bias introduced by using group rather than individual preferences. By holding details of each individual separately they can be re-grouped in ways that are most appropriate for each stage in the modelling process, for example, they may be combined together in a different way for mode choice and then for route assignment.

The recording of the choices made by each individual assists in the analysis of the social and distributional impacts of a transport intervention as it provides detailed information on which people are affected by a proposed intervention and how.

The software can record the history of each agent; this means that the modelling approach can incorporate more aspects into the mode choice decision than conventional four stage models. It can for instance, record the actual journey times and costs experienced by agents on previous journeys and use this to influence their decisions in the present. As shown in this research, it can also be used to incorporate habitual behaviour in the model. (Section 7.6).

The ability to incorporate feedback loops within the model enables the modeller to increase the richness of the modelled system with built-in feedback between the environment and the agents. (Section 7.5).

The flexibility of agent based modelling means that the four stages of conventional modelling do not have to be applied in the same order for all agents. There is a debate in variable demand transport modelling as to whether destination choice should come before or after mode choice. The default order recommended by the DfT in WebTAG Unit M2, that destination choice is made before mode choice, is different from the order commonly used in

Europe. With agent based modelling a mixed approach could be used, with for instance destination choice preceding mode choice for work trips but mode choice preceding destination choice for leisure trips.

Agent based modelling offers the potential to enrich our models with greater behavioural realism and the option to replace the logit model with alternative methods for modelling how a person will choose their travel mode. The challenge lies in determining alternative theories for how these decisions are made and collecting the data required to implement them, but agent based modelling does provide a way of modelling many of these alternatives and provides a versatile platform for the future development of transport models.

The fine level of detail that can be handled in an agent based model makes it capable of modelling more ‘smarter choices’ measures than current four stage models even without further advances in the development of more behaviourally realistic choice models. Table 9.1 below is based on a summary table from the WSP report for the DfT into modelling smarter choices with four stage models. An additional column has been added to show how an agent based modelling approach enables the modelling of many of the ’smarter choices’ measures they considered.

Many of these, such as running bus services for a company’s workers, subsidized fares and restricted car parking can be handled by the greater level of detail that is possible in agent based models on the precise origin and destination of commuting trips and the attributes of workers. Agent based modelling offers greater potential to model these measures than can be achieved by microsimulation alone as it offers a way of modelling the networks that connect people. This could, for example, be used to develop models for car sharing and responses to marketing campaigns in the workplace that may be targeted to building a critical mass of people cycling to work.

Possible component measures within an initiative	Is there a general need to enhance operational models?	Relevant ABM characteristic
New conventional bus or rail services linking to site/area	Effects can be represented only if the model zones are detailed enough	More accurate representation of walk times to new bus service
‘Work buses’ between site and town centre that can be used by any traveller	Effects can be represented only if the model zones are detailed enough to identify the site	More accurate representation of walk times to new bus service
Dedicated ‘work buses’ between site and town centre	Effects can be represented only if the site and the workers from the site can be identified in demand modelling	Can restrict bus services to particular eligible users
Subsidised fares; Interest-free season ticket loans	Not for reflecting the effective reduction in commuting fares	More accurate representation of fare costs for each users
Special deals to reduce the cost of bus and rail commuting	Not for reflecting the effective reduction in commuting fares	Can change costs for a short period of time
Parking ‘cash out’ (payment on days of not driving)	No	Can reflect this benefit only for those receiving it
Car parking restricted to essential users	Most models would not be able to segment between essential and non-essential users, and would need to approximate the effects	Individual recording of agent characteristics e.g. permission to park at certain locations
Parking charges	Most models do not represent parking well and would require enhancement	More accurate representation of a person’s actual parking charge
Car sharing Scheme	Yes	Can track people working in some office and at similar times
Preferential car parking for sharers	Yes	Individual recording of agent characteristics e.g. permission to park at certain locations
Demand activated bus services (variable routes, bus-taxis)	Yes	Individual recording of waiting time for public transport service
Giving all staff public transport information	Additional work required to quantify implicit mode specific constants in the case of incremental demand models	Recorded when each individual is aware of their travel options
Offering personalised journey plans to staff	Additional work required to quantify implicit mode specific constants in the case of incremental demand models	Recorded when each individual is aware of their travel options
Secure cycle Parking	Additional work required to quantify implicit mode specific constants in the case of incremental demand models	Records who has access to secure cycle parking

Changing/shower facilities at work	Additional work required to quantify implicit mode specific constants in the case of Incremental demand models	Records who has access to these facilities
Business cycle mileage Allowance	Yes, if the scale of effects is found to be significant	Individual recording business cycle mileage allowance
Services on site to reduce need to travel (cafeteria, convenience shop, cash dispenser)	Most models would not be able to identify convenience shopping from other personal trips, and would need to approximate the effects	Can't model trip generation at the individual level
Encouraging Teleworking	Yes	Individual recording of whether a person permission can work at home
Flexible working hours / compressed working hours to e.g. a 4-day week, or commuting during inter-peak or off-peak periods	Yes	Individual recording of a person's ability to vary working hours

TABLE 9.1 'SMARTER CHOICES' MEASURES AIMED AT REDUCING CAR TRIPS TO WORKPLACES

Source: adapted from WSP, 2010

9.4.2 Limitations of agent based modelling

Even with using agent based modelling methods and software some challenges will remain when modelling 'smarter choices'. Adopting a new modelling approach does not overcome the problems that arise for any modelling methodology resulting from a lack of knowledge and understanding of the processes involved in the real world system being modelled. The WSP report acknowledges that for some measures, such as cash payments to reward those not driving to work and flexible working hours there is a lack of data as to how people respond to these policies. Such data is needed to develop and validate the modelling of these measures.

Running ABMs can be time consuming, especially if a model contains many agents and stochastic events. It can take a considerable amount of time to model each time step and so to cover the overall time period desired by the modeller. These models can also be demanding in terms of the computer processing power they require in order to deal with a large number of agents.

An ABM will often require the collection of more data than would be required of an aggregate model. This is because data is needed for each of the agents. An ABM is also likely to need data on a wider set of agent attributes. For example in a standard transport model segmentation is usually only by time of day, journey purpose and occasionally income group. An ABM can handle a much more detailed segmentation which, if included, could require additional data such as gender, age, employment status and number of dependent children. There is also the problem that the data required for the ABM may not have already been collected as historic data collection exercises have been built around the needs of the logit models embedded in traditional four stage transport models or activity based models.

The ability of ABMs to model bounded rationality and any number of alternative ways of replicating human decision making, leads to the criticism that human cognitive processes are still unknown and the subject of academic debate, and modellers cannot be sure therefore that they have correctly replicated them in the agent based model. In addition, the interactions in the model between agents and between agents and their environment may not be a true representation of reality and the model results misleading. In defence, current models are already based on theories of human decision making such as rational choice behaviour which have been shown by behavioural economists to be flawed, and ABM does provide a framework in which the influence of relaxing some of these standard economic assumptions on model results can be explored.

A task facing the ABM community is the development of standard procedures for the verification, calibration and validation of the models. The ODD protocol used in Chapter 6 is an attempt to provide standards for the approach to, and level of detail expected, when documenting an ABM. Verifying the correct coding of a model is often handicapped by the use of proprietary software where the code covering some of the processes is hidden from inspection by others. These software packages make it easier to build ABMs but at the expense of full knowledge of the code that is implementing the model.

There are currently no formal standards for model validation as exist for four stage transport models in the UK. The potential of ABMs to offer multiple outcomes is both a help and a hindrance in this respect. It is more realistic to show a range of outcomes, as provided by an ABM, but it requires more observed data to validate the model results. Further work is required on developing approaches to the verification and validation of ABMs. Carley (1996) presents eight validation methods that can be applied to ABMs and may assist in addressing this issue:

- Pattern validity: the pattern of results from the model match the pattern of the real data, judged by using statistical tests, matching by eye or ordinal pattern analysis (Thorngate, 2013)
- Point validity: when changing one variable at a time, the mean of the output matches the mean of the target observed data
- Distributional validity: the distributional characteristics of the model output match those of the observed data, e.g. have similar means, standard deviations and shape.
- Value validity: the specific results obtained from the model match the target data on a point by point basis.
- Face validity: when the model is running it looks like reality and that the values of the agent attributes and any parameter values used look reasonable.
- Parameter validity: the parameters used in the model match real data, collected for example by surveys and experiments
- Process validity: the processes described in the model match the processes observed in the real world
- Theoretical validity: the underlying theoretical constructs in the model provide better results than a simple linear model

It should also be noted that difficulties in the verification, calibration and validation of models are not unique to ABMs. This is also a challenge for four stage models, particularly when they include variable demand processes, which change the overall number of trips as well their mode, destination and sometimes time of travel. The standard validation reports required by the DfT concentrate on the model's ability to replicate current conditions and comparisons of current traffic or passenger model flows with counts, and modelled journey times with observed journey times. Validation tests on a model's ability to forecast change are restricted to testing the realism of the aggregate level of the demand change in the model as a result of changes in car fuel cost, car journey times and public transport fares. (DfT WebTAG Unit M2).

The analysis of results from an ABM also poses a challenge due to the volume of results generated. Many ABM models include stochastic events and responses; so each model run will produce a different set of results. Many runs are therefore needed in order to determine the stability of the outputs, the range of the outputs and the average values. A criticism of ABM in practice is that practitioners do not perform sufficient runs or do not present the evidence to show that sufficient runs have been carried out. In addition, the complex situations which are often the subject of ABM may show a dependence on the initial conditions i.e. the final outcome is a result not just of the process of the model but of the initial state of the agents. This adds a further requirement that multiple model runs are carried out from a range of initial states to show if the model results are stable over a variety of starting conditions.

9.5 Contribution to knowledge

This research has contributed in a number of ways to the current knowledge of suitable methods for modelling 'smarter choices'. The four stage transport modelling framework will always have limitations in its ability to handle 'smarter choices' measures because it operates at the level of matrices. Although continued enhancements to the logit model are welcome, they will not provide a way for this modelling framework to model many 'smarter choices' measures. This will require models with a higher degree of behavioural realism, operating at the level of the individual, and able to handle the very specific details of each person's constraints and travel costs. (Section 7.4).

With many ‘smarter choices’ measures the policy maker is often deliberating on measures which will affect a small number of people, for example the number of people on a new bus service or the number of people who transfer to cycling. This increases the importance of using modelling methods which can produce good forecasts of changes in small numbers. Many of these interventions are short lived, so estimates of the change in these numbers over time are important for the production of accurate estimates of the value for money they will deliver.

This study has shown that the agent based modelling approach provides an extremely useful method for working at the level of the person and can be used to build mode choice models which contain more realistic portrayals of the processes affecting travel decisions, such as the impact of habits. The simple model developed in this study is a contribution to knowledge as it shows that it is possible to introduce habitual behaviour into the modelling of mode choice. (Section 7.6). The introduction of habits into the mode choice modelling produces a model that forecasts the change in the elasticity of demand over time as a model output and shows that it is possible to model the lag in responses to a change in the transport system.

This research shows that agent based modelling allows the building of models which can challenge the standard assumptions of neo-classical economics and bring the findings of behavioural economists and psychologists into transport planning. This research, for example, used the features of agent based modelling to build a model of mode choice that implements Triandis’ Theory of Interpersonal Behaviour. The greater realism of the patronage forecasts that such a model can produce, together with model results at far more points in time should, in turn, improve the quality of the business and financial cases produced for a range of ‘smarter choices’ measures, such as the time-limited funding of new bus services by developers.

The proof of concept model shows that with agent based modelling it is possible to code feedbacks directly into the model. (Section 7.5). This is important, as there is a tendency for mode choice models to over-estimate public transport patronage as it includes routes and services which will not actually be available as they are not going to be financially viable.

This study has highlighted that aggregation bias occurs in demand forecasts, not only when average rather than specific costs are used for people (Bowman *et al.*, 2006) but when common rather than specific preference functions are used. (Section 8.7.6). The model has shown that the variation in demand forecasts when using personal preference functions compared to functions for the typical segmentation employed in transport models can be non-trivial. It has shown that a segmentation based on latent classes can reduce the level of aggregation bias in a model as it groups together people with similar preference functions. This also points to the need to model transport at the individual level, so that a greater match can be achieved between each person's modelled and actual preferences, in order to reduce the impact of aggregation bias on the model results.

9.6 Future work

Agent based modelling offers the immediate possibility of contributing to the improvement of the ability of transport models to model the impacts of many 'smarter choices' measures. The model in this project was written using python, the scripting language used in several commonly used four stage model software packages such as VISUM and EMME/4. This offers a way in which the benefits of ABM can be used to incrementally extend current models by storing the data on travellers at an individual level and aggregating it when required into an appropriate segmentation system for each stage in the model. Further work arising from this study is the writing of new routines within these current four stage modelling programmes to implement a change in the way data is stored and to enable the use of alternative agent based models for mode choice alongside the other modules available within these software packages.

The increased use of agent based models will require the collection of more and different data. Population synthesis techniques can be used to supplement current methods to build a database of all the travellers in an area and their characteristics. The use of more spatially accurate information on each traveller brings with it a requirement for more specific data on the journey options facing each traveller, covering their personal constraints and travel costs. Techniques need to be developed to build these more detailed datasets and to match the characteristics of agents with appropriate preference functions. The availability of new big data sources, such as mobile phone data which can track the movements and mode choices of the same individual over several weeks, may assist in this task.

The inclusion of more behavioural aspects into transport models requires an understanding of the processes at work in the real world and ways of coding these processes into models. For example, this research has shown how coding habitual behaviour into the model can replicate the response to a change in bus fares observed over time with aggregate patronage data. More research is required into the trigger events which cause people to re-evaluate their mode choice. In particular, the model could be extended to include the impact of the experience of delays experienced on recent past journeys on current choices or a rise in fares or parking charges above a personally tolerated threshold.

Agent based models introduce the possibility of modelling the impact of the choices made by other people known to the agent on their own travel choice. For example it may be that a person's willingness to cycle to work may be affected by the opinions and experience of people known to that agent. Research into the impact of social networks and social norms on personal decisions is now relevant to transport modelling as, with the use of ABM, the findings could be used to inform ways of incorporating these considerations into transport models.

Further work is recommended on the impact of aggregation bias on the results from transport modelling work and the contribution of alternative segmentation techniques, such as latent classes, in minimising the errors introduced by this often overlooked issue. The reduction of the error introduced into demand forecasts by the way the model is designed and operates is important, particularly for public transport schemes where the assessment of the financial viability of a service is very sensitive to the forecast number of passengers.

Alongside the development of models using agent based techniques and more behaviourally realistic models, further work is required on ways to document, verify, calibrate and validate the models. These efforts would be assisted by further work on methods for analysing the large quantity of outputs generated by these models and visualising this material so as to assist both the modellers who develop and use the tools and the stakeholders with whom the results need to be communicated.

9.7 Conclusions

The four stage model has been used in transport modelling for over sixty years. Its endurance is due to the fact that it provides a way to tackle four questions which are central to the transport system:

- How many trips will be made? (generation)
- Where will they go? (destination)
- How will they be made? (mode choice)
- Which route will they use? (assignment)

It has provided a way to consider the major changes to the supply side of the transport system such as the construction of new roads and public transport schemes. Now though these models are required to answer other questions, such as how will people respond to changes in the conditions on the transport network, will they change when they travel and where they travel to, and how can policy makers influence people's travel choices? These questions have been tackled by others by extending the four stage model to include, for example, time of day choice and providing feedback loops between the modelling stages so that the output of each stage can influence the preceding stages. The ability of modellers to further extend the four stage framework to meet the ever more sophisticated and detailed questions asked of transport models is restricted by:

- the design of current software products, as they are based on a matrix system to store travel data;
- the logit models used for mode, destination and time of day choices and the elasticities used for trip frequency responses are based on observed mathematical relationships rather than the underlying behavioural processes; and
- the lack of capability to model the detail of the circumstances of individuals and the external and personal constraints on their transport choices.

This research has shown that the agent based modelling approach can increase the ability of transport models to predict the potential impacts of 'smarter choices' measures. By building models with the individual as the key model unit, it becomes possible to build models that can take well-established tools developed within the four stage modelling framework and extend these to take advantage of the additional detail, processes and relationships that can be captured in agent based models as a means of bringing 'smarter choices' into multi-modal models.

References

Aarts, H., and Dijkesterhuis, A., (2000) The automatic activation of goal-directed behaviour: the case of travel habit. *Journal of Environmental Psychology, 20, 76-82.*

Aarts, H., Verplanken, B. and van Knippenberg, A., (1998) Predicting behaviour from actions in the past: Repeated decision making or a matter of habit? *Journal of Applied Social Psychology, 28, 1356-1375.*

Adamowicz, W. and Swait, J. (2013). Are Food Choices Really Habitual? Integrating Habits, Variety-seeking, and Compensatory Choice in a Utility-maximizing Framework. *American Journal of Agricultural Economics. 95(1): 17-41.*

Allan, R. (2009) Survey of Agent Based Modelling and Simulation Tools*. DL Technical Reports DL-TR-2010-007.*

Ariely, D. (2008) *Predictably Irrational,* HarperCollins Publishers Ltd.

Arthur, W. (1999) Complexity and the Economy. *Science. Vol 284.*

Ajzen, I. (1991) The theory of planned behaviour. *Organisational Behaviour and Human Decision Processes. 50, 179–211.*

Arentze, T., Hofman, F., van Mourik, H., Timmermans, H. (2007) ALBATROSS: Multiagent, Rule-Based Model of Activity Pattern Decisions. *Journal of the Transportation Research Board.*

Ahmed F., Yawson D., and Pappoe F. (2010) A Systems Dynamics Approach to Explore Traffic Congestion and Air Pollution Link in the City of Accra, *Sustainability Vol 2 pp 252 - 265*

Arthur, W. (1995), Complexity in economic and financial markets: Behind the physical institutions and technologies of the marketplace lie the beliefs and expectations of real human beings. *Complexity, 1: 20–25.*

Atasoy, B., Glerum, A. and Bierlaire, M. (2011) Mode choice with attitudinal latent class: A Swiss case-study. *Paper presented at the 2nd International Choice Modelling Conference, Leeds.*

Axelrod, R. (1997) *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration* Princeton, NJ: Princeton University Press.

Ajzen, I. (1985) From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action control: From cognition to behavior*. Berlin, Heidelber, New York: Springer-Verlag.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 50 (2): 179–211.

Ajzen, I. (2002). Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior. *Journal of Applied Social Psychology, 32, 665-683.*

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., White, P., (2004). The demand for public transport: a practical guide, *TRL Report TRL 593, Crowthorne, UK.*

Ballas, D., Clarke, G. P., Dorling, D. and Rossiter, D. (2007). Using SimBritain to Model the Geographical Impact of National Government Policies. *Geographical Analysis 39(1), 44- 77.*

Balmer, M. and Axhausen, K.W. and Nagel, K. (2005) An Agent Based Demand Modeling Framework for Large Scale Micro-Simulations. *Working paper, 329, Institute for Transport Planning and Systems (IVT), ETH Zurich, Switzerland.*

Bamberg, S. (2006) Is a residential relocation a good opportunity to change people's travel behaviour? Results from a theory-driven intervention study. *Environment and Behaviour 38 820 – 840.*

Bamberg, S. and Schmidt (2003) Incentives, Morality, Or Habit? Predicting Students' Car Use for University Routes With the Models of Ajzen, Schwartz, and *Triandis Environment and Behavior 35:264-285,*

Barnsley, M. (2007) Environmental Modelling

Bhaskar, R.A., (1975), A Realist Theory of Science, *London: Verso.*

Bhaskar, R.A., (1979), The Possibility of Naturalism, *London: Routledge.*

Beckman, R. , Baggerly, K. , & McKay, M. (1996). Creating synthetic base-line populations. *Transportation Research A, Policy and Practice, 30, 415-429.*

Beimborn, E. (1995) A Transportation Modeling Primer *(Milwaukee: Centre for Urban Transportation Studies, University of Wisconsin).*

Bekhor, S., Dobler, C. and Axhausen, K. (2010) Integration of activity-based with agent-based models.an example from the Tel Aviv model and MATSim. *ETH Zürich, InstitutfürVerkehrsplanung, Transporttechnik, Strassen- und Eisenbahnbau (IVT) .doi:10.3929/ethz-a-006121544*

Ben-Akiva. M., McFadden, D., Train, K., Walker, J., Bhat C., Bierlaire, M., Boldue, S., Boersch-Supan, A., Brownstone, D., Bunch, D., Daly, A., de Plama, A., Gopinath, D., Karstrom, A. and Munizaga, M. (2002) Hybrid choice models: Progress and Challenges. *Marketing Letters Volume 13 Issue 3 p 163-175.*

Benenson, I., Martens, K. and Birfir, S. (2008) PARKAGENT: An agent-based model of parking in the city. Comp*uters, Environments and Urban systems, Vol. 32, 431-439.*

Bhaskar, R. (1975) A Realist Theory of Science

Bhaskar, R. (1979) The Possibility of Naturalism

Bhatta, B. and Larsen, Odd I. (2011) Are intrazonal trips ignorable? *Transport policy*, 18(1), p13-22.

Bonabeau, E. (2002) Agent-based modeling: Methods and techniques for simulating human systems.*Proceedings National Academy Science* 993: 7280–7287.

Bonsall, P. (2009) Do we know whether personal travel planning really works? *Transport Policy*, *Vol 16.*

Bourguignon, F. and Spadaro, A. (2006). Microsimulation as a tool for evaluating redistribution policies. *Journal of Economic Inequality, 4, 77-106.*

Bowman, J., Bradley, M. and Gibb, J. (2006) The Sacramento Activity-Based Travel Demand Model: Estimation and Validation Results, *European Transport Conference, Strasbourg, France.*

Brody, H., Rip. M., Peter Vinten-Johansen, P., Paneth, N. and Rachman S., (2000) Map-making and myth-making in Broad Street: the London cholera epidemic, 1854. *The Lancet Vol 356.*

Brown, B., Werner C. and Kim, N., (*2003)* 'Personal and contextual factors supporting the switch to transit use: evaluating a natural transit intervention', *Analyses of Social Issues and Public Policy, 3 (1), pp. 139 – 160,*

Bruce, D. and Williams, H. (2003) Urban Travel Forecasting in the USA and UK. *Presented at the 43rd Congress of the European Regional Science Association, Finland.*

Cairns, S., Sloman, L., Newson, C., Anable, J., Kirkbride, A. and Goodwin, P. (2004) Smarter Choices: Changing the Way We Travel.

Cambridgeshire County Council, Cambridge Transport Innovation Fund Package Outline Proposal for Funding, 2007

Carley, K. (1996) Validating Computational Models. *Carnegie Mellon University*

Carlson, T., Dulla, R., Kuzmyak, J. and Porter, C. (2005) COMMUTER model v2.0- User manual. *U.S. Environmental Protection Agency, 420-B-05-017*

Cassells, R., Harding, A. and Kelly, S. (2006) Problems and Prospects for Dynamic Microsimulation: A Review and Lessons for APPSIM. *Discussion Paper no. 63 University of Canberra.*

Castle, C. and Crooks, A. (2006) Principles and Concepts of Agent-Based Modelling for Developing Geospatial Simulations. *UCL Working Paper 110.*

Castiglione J., Bradley M., and Gliebe J. (2014) Activity-Based Travel Demand Models: A Primer. *The Strategic Highway Research Program of the Transportation Research Board.*

Casti, J. (1997) *Would-Be Worlds: How Simulation is Changing the World of Science.* Wiley: New York.

Chakrabarty, M., and Schmalenbach, A. (2002). The Representative Agent Hypothesis: An Empirical Test *Bonn economic discussion papers.*

Chatterjee, K. (2009) A comparative evaluation of large-scale personal travel planning projects in England. *Transport Policy*, Vol 16.

Chatterjee, K., Sherwin, H. and Jain, J. (2013) Triggers for changes in cycling: The role of life events and modifications to the external environment. *Journal of Transport Geography*, *30. pp. 183-193.*

Chatterton, T. and Department of Energy and Climate Change (2011) An introduction to thinking about 'energy behaviour': A multi-model approach. *Department of Energy and Climate Change, Department of Energy and Climate Change website.*

Chattoe-Brown, E., (2009) Agent-based modelling and microsimulation: Ne'er the twain shall meet? *ESRC Microsimulation Seminar Series Seminar 2: 'Adding behaviour' London.*

Clark, B., Chatterjee, K., Melia, S., Knies, G. and Laurie, H. (2014) Life events and travel behaviour: Exploring the interrelationship using UK household longitudinal study data. *Transportation Research Record, 2413. pp. 54-64.*

Collins, L. and Lanza, S. (2010) *Latent Class and Latent Transition Analysis.* Wiley.

Connectics Transportation Group (2008) Washtenaw Area Transportation Study Travel Demand Model Improvements Technical Memorandum 12 – Non Motorized Assignment Techniques.

Conner M. and Armitage C. (1998) Extending the Theory of Planned Behavior: A Review and Avenues for Further Research. Journal of Applied Social Psychology, 28, 15, pp. 1429-1464.

Crooks, A. and Castle, C. (2012), The Integration of Agent-Based Modelling and Geographical Information for Geospatial Simulation, in Heppenstall, A., Crooks, A., See, L. and Batty, M. (eds.), *Agent-based Models of Geographical Systems*, Springer, New York, NY, pp. 219-252.

Danermark, B., Ekstrom, M., Jakobsen, I., Karlsson, J. (2001) Explaining Society: An Introduction to Critical Realism in the Social Sciences (Critical Realism: Interventions).

Dargay, J. and Hanly, M. (1999) Bus Fare Elasticities. Report to Department of the Environment, Transport and the Regions.

Darnton, A. (2008) Reference Report: An overview of behaviour change models and their uses *Government Social Research Behaviour Change Knowledge Review.*

Developing an Effective Evaluation Plan (2011) *Atlanta, Georgia: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health*

Promotion, Office on Smoking and Health; Division of Nutrition, Physical Activity, and Obesity.

Department of Transport (2006) Transport Innovation Fund Guidance.

Department of Transport (2010) Presentation on draft for consultation DfT WebTAG guidance on modelling smarter choices by Robin Canberry. Unpublished.

Department of Transport (2010) DIADEM User Manual.

Department of Transport (2011) WebTAG unit 3.10.4: Variable Demand Modelling.

Department of Transport (2011) Draft WebTAG guidance on modelling smarter choices.

Department of Transport (2011) Local Sustainable Transport Fund - Supplementary Guidance for Local Authorities Shortlisted for Large Projects.

Desyllas, J., Duxbury, E., Ward, J., and Smith, A. (2003). Pedestrian demand modelling of large cities: an applied example from London.

Dia, H. (2002) An agent-based approach to modelling driver route choice behaviour under the influence of real-time information. *Transportation Research Part C10 p. 331–349.*

Dieussaert K., Aerts K., Steenberghen T., Maerivoet S. and Spitaels K. (2009) SUSTAPARK: An Agent-based Model for Simulating Parking Search. *12th AGILE International Conference on Geographic Information Science Leibniz Universität Hannover, Germany.*

Dobson, P. (2002) "Critical Realism and information systems research: why bother with philosophy?" *Information Research* **7**(2).

Eash, R. 1999. Destination and Mode Choice Models for Non-Motorized Travel, *Transportation Research Record 1674, 1999, pp. 1-8.*

Epstein J. (2008) Why Model? *Journal of Artificial Societies and Social Simulation* vol. 11, no. 4 12.

Epstein, J. and Axtell, R. (1996) Growing Artificial Societies: Social Science from the Bottom Up. *The MIT Press.*

Erhardt, G., Sall, E., Zorn, L., Tischler, D., Alsup, R., Nassir, N. (2013) Development and Application of a Dynamic Traffic Assignment Model for San Francisco*. 41st European Transport Conference, Frankfurt, Germany.*

Evans, R., Burke, M., and Dodson, J. (2007) Clothing the Emperor?: Transport modelling and decision-making in Australian cities.

Evans, D., and Norman, P. (1998) Understanding pedestrians' road crossing decisions: An application of the theory of planned behaviour. *Health Education Research, 13,* 481-489.

Fleetwood, S. (2005) Ontology in Organization and Management Studies: A Critical Realist Perspective, *Organization* 12(2): 197–222.

Flyvbjerg, B., Holm, M., and Buhl, S. (2006) Inaccuracy in traffic forecasts. *Transport Reviews* 26.

Franklin, S. and Graesser, A. (1996) Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. *Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Springer-Verlag*

Forrester, J. (1969) Urban Dynamics *M. I. T. Press, 50 Ames Street, Cambridge, Massachusetts.*

Forrester, J. (1989) The Beginning of Systems Dynamics. *Banquet talk at the international meeting of the Systems Dynamics Society, Stuttgart, Germany.*

Frick, M. and Axhausen, K. (2004) Generating Synthetic Populations using IPF and Monte Carlo Techniques 4th Swiss Transport Research Conference, Monte Verità, Ascona, Switzerland.

Fujii, S., Garling, T. and Kitamura, R. (2001) Changes in Drivers' Perceptions and Use of Public Transport during a Freeway Closure: Effects of Temporary Structural Change on Cooperation in a Real Life Social Dilemna. *Environment and Behavior Vol 33.*

Gardner, B. (2009) Modelling motivation and habit in stable travel mode contexts. *Transportation Research Part F. 12 p 68 – 76.*

Garling, T., Fujii, S., and Bie, O. (2001) Empirical tests of a model of determinants of script-based driving choice. *Transportation Research Part F* 4 (2) 89 – 102.

Genesereth, M. and Ketchpel, S. (1994) Software Agents *CIFE Working Paper 32. Stanford University.*

Gilbert N. and Bankes, S. (2002) Platforms and methods for agent-based modeling *Proceedings of the National Academy of Sciences Volume 99.*

Gilbert, N. and Troitzsch, K. (2005). *Simulation for the Social Scientist* (2 ed.). Open University Press.

Gilboa. (2010) *Rational Choice.* The MIT Press, Cambridge Massachussets.

Girengezer, G., and Goldstein D. (1996) Reasoning the fast and frugal way. Models of bounded rationality. *Psychological Review 103 (4) 650 – 669.*

Givoni, M., Beyazit, E., Shiftan, Y., Ishaq, R. and Tzur, O. (2012) The use of state-of-the-art models by policy makers to address global energy and environment challenges: the case of transport policy *TSU Working Paper Series, Ref. 1060. School of Geography and the Environment, Oxford.*

Gladwell M. (2000) The Tipping Point: How Little Things Can Make a Big Difference. *Little Brown.*

Gliebe, J. and Picado, R. (2012) Seminar 2 Institutional Issues for Managers. Activity Based Modeling *TMIP Webinar Series.*

Gliebe, J. and Vovsha, P, (2012) Seminar 5 Population Synthesis and Household Evolution. Activity Based Modeling *TMIP Webinar Series.*

Goodwin, P. (1977). Habit and hystersis in mode choice. *Urban Studies*, 14, 95-98.

Gordon, A. (2010) Diadem User Manual.

Gorman, W. (1953) Community Preference Fields *Econometrica 21*:63-80.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., rand, T., Heinz, S., Huse, G., Huth, A., Jepsen, J., Jørgensen, C., Mooij, W., Müller, B., Railsback, S., Robbins, A.M., Robbins, M., Rossmanith, E., Rüger, N., Strand, E., Stillman, R., Vabø, R., Visser, U. and DeAngelis, D., 2006. *A standard protocol for describing individual-based and agent-based models.* Ecological Modelling. 198, 115–126.

Grimm, V., Berger, U., DeAngelis, D., Polhill, J., Giske, J. and Railsback, S. 2010 *The ODD Protocol: A review and first update.* Ecological Modelling, 221, 2760-68.

Hägerstrand, T. (1970) “What About People in Regional Science*?”, Papers of the Regional Science Association*, *Vol XXIV.*

Han, Q., Arentze, T., Timmermans, H., Janssens, D. and Wets, G., (2011) The effects of social networks on choice set dynamics: Results of numerical simulations using an agent based approach. *Transportation Research part A Vol 45 p. 310 – 322.*

Hanly, M., Dargay, J., & Goodwin, P. (2002). Review of income and price elasticities in the demand for road traffic. *Department for Transport, London*.

Haining, R. (2003) Spatial Data Analysis: Theory and Practice. *Cambridge University Press.*

Harding, A. (2007). APPSIM: The Australian dynamic population and policy microsimulation model. In *Presentation to the Technical Workshops of the 1st General Conference of the International Microsimulation Association, Vienna, Austria* (pp. 20-22).

Harland, Kirk, Heppenstall, Alison, Smith, Dianna and Birkin, Mark (2012) 'Creating Realistic Synthetic Populations at Varying Spatial Scales: A Comparative Critique of Population Synthesis Techniques' *Journal of Artificial Societies and Social Simulation 15 (1) 1.*

Hassan, S., Pavon, J. and Gilbert, N. (2008) Injecting Data into Simulation: Can Agent-Based Modelling Learn from Microsimulation? In *Proceedings of the World Congress of Social Simulation 2008*. Washington, D.C.

Heath, B., Hill, R., and Ciarallo, F. (2009). 'A Survey of Agent-Based Modeling Practices (January 1998 to July 2008)'. *Journal of Artificial Societies and Social Simulation* 12(4)9

Hensher, D. and Button, K. (2007) *Handbook of Transport Modelling.* Elsevier Science Ltd

Hess, S. and Polak, J. (2004) An analysis of parking behavior using discrete choice models calibrated on SP datasets. *ERSA conference papers, European Regional Science Association.*

Hess, S., Rose, J. and Polak, J. (2010), Non-trading, lexicographic and inconsistent behaviour in stated choice data, *Transportation Research Part D, 15(7), pp. 405-417.*

Hess, S., Stathopoulos, A. & Daly, A. (2012), Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies, *Transportation 39(3).*

HM Treasury (2014) Impact on households: distributional analysis to accompany Budget.

Holland, J., Holyoak, K., Nisbett, R. and Thagard, P., (1986) Induction: Processes of Inference, Learning and Discovery. *Cambridge, MA: MIT Press*

Holland, J. (1995) Hidden Order: How Adaptation Builds Complexity. *Addison-Wesley: Reading, MA*

Hox, J. and Boeije, H. (2005). Data collection, primary versus secondary. *in K. Kempf-Leonard (Ed.). Encyclopedia of Social Measurement. San Diego, CA: Academic Press.*

Huang Z and Williamson P (2001) Comparison of synthetic reconstruction and combinatorial optimisation approaches to the creation of small-area microdata. *Working Paper 2001/2, Population Microdata Unit, Department of Geography, University of Liverpool, Liverpool.*

Jackson, T. (2005) Motivating Sustainable Consumption*. Centre for Environmental Strategy University of Surrey.*

Jotisankasa, A. and Polak, J. (2008) Modelling and Appraisal of Smarter Choices: A Literature Review. *Produced for the Department of Transport.*

Jones, P. M. (1977) Assessing policy impacts using the Household Activity-Travel Simulator. *(No. TSU/REF-039/WP18).*

Kirman, A. (1992) Whom or What Does the Representative Individual Represent? *The Journal of Economic Perspectives, Vol. 6, No. 2.*

Klöckner, C. (2004) How single events change travel mode choice – a life span perspective. *Paper presented at the 3rd International Conference on Traffic and Transport Psychology, Nottingham.*

Klockner, C. and Matthies, E. (2004) How habits interfere with norm-directed behaviour: A normative decision-making model for travel mode choice. *Journal of Environmental Psychology 24, 319-327.*

Kocur, G., Adler, T., Hyman, W. and Audet, E. (1982) Guide to Forecasting Travel Demand with Direct Utility Measurement. *UMTA, USA Department of Transportation, Washington D.C*

Kongmuang, C., Clarke, G., Evans, A., and Jin. J. (2006). SimCrime: A Spatial Microsimulation Model for the Analysing of Crime in Leeds.

Kung, C. and Solvberg, A. (1999) Activity Modeling and Behavior Modeling, In: T. Ollie, H. Sol, A. Verrjin-Stuart, Proceedings of the IFIP WG 8.1 working conference on comparative review of information systems design methodologies: improving the practice. North-Holland.

Laufer, J., (2007) Freeway capacity, saturation flow and the car following behavioural algorithm of the VISSIM microsimulation software *30th Australasian Transport Research Forum.*

Law, A. 2006. How to build valid and credible simulation models. *Proceedings of the 2006 Winter Simulation Conference.*

Lay-Yee, R. and Cotterell, G. (2012) The role of microsimulation in the development of public policy. *COMPASS Research Centre The University of Auckland.*

Lovelace, R., and Ballas, D. (2013). Truncate, replicate, sample: A method for creating integer weights for spatial microsimulation. *Computers, Environment and Urban Systems, 41, 1–11.*

Lunn, P. (2008) Basic Instincts, Human Nature and the New Economics. *Marshall Cavendish.*

Macmillan, A., Connor, J., Witten, K., Kearns, R., Rees, D. and Woodward, A. (2014). The societal costs and benefits of commuter bicycling: simulating the effects of specific policies using system dynamics modeling. *Environmental Health Perspectives 122:335–344;*

MacLeod, M. and Nersessian, N. (2013). Building Simulations from the Ground-Up: Modeling and Theory in Systems Biology. *Philosophy of science 80: 533-555.*

Martini, A., and Trivellato, U. (1997). The role of survey data in microsimulation models for social policy analysis. *Labour, 11, 83-112.*

Marshall, N. and Grady, B. (2006) Sketch Transit Modeling Based on 2000 Census Data *Transportation Research Board Conference,*

McCutcheon, A. (1987) *Latent class analysis*. Beverly Hills: Sage Publications.

McDonnell, S. and Zellner M., (2011) Exploring the effectiveness of bus rapid transit a prototype agent-based model of commuting behaviour. *Transport Policy Vol 18 p. 825–835.*

McFadden, D.(1974) The Measurement of Urban Travel Demand *Journal of Public Economics Volume 3 issue 4p 303- 328.*

McFadden D. (2007) The behavioural science of transportation. *Transport Policy Vol 14 p. 269–274.*

McNally, M. and Rindt, C. (2007) The Activity-Based Approach in *The Handbook of Transport Modeling.*

Miller, E. (2009) Simulating Cities: the ILUTE Project*. 2nd General Conference of the International Microsimulation Association "Microsimulation: Bridging Data and Policy".*

Miller, J. and Page, S. (2007) Complex Adaptive Systems: An Introduction to Computational Models of Social Life.

Morrison, M. and Morgan, M. (eds.) (1999). *Models as Mediators.* Cambridge University Press.

Mylopoulos, J., (1992) Conceptual Modelling and Telos, *University of Toronto.*

Nikolai, C., and Madey, G. (2008). Tools of the Trade: A Survey of Various Agent Based Modeling Platforms, *Journal of Artificial Societies and Social Simulation, 12 (2).*

North, M. and Macal, C. (2007) Managing Business Complexity: Discovering Strategic Solutions with Agent-Based Modeling and Simulation. *Oxford University Press: Oxford, UK.*

OECD (2013), "Labour Market Statistics: Employment by job tenure intervals: average tenure", *OECD Employment and Labour Market Statistics* (database).
doi: 10.1787/data-00294-en (Accessed on 28 September 2013).

Orcutt, G. (1957). A new type of socio-economic system*. Review of Economics and Statistics, 39, 116–123.*

O'Sullivan, D. and Haklay, M. (2000) Agent-based models and individualism: is the world agent-based? *Environment and Planning 328 1409 – 1425.*

Ouellette, J. and Wood, W. (1998). Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. *Psychological bulletin*, *124*(1)

Outhwaite, W. (1987) New Philosophies of Social Science: Realism, Hermeneutics and Critical Theory. *Theoretical Traditions in the Social Sciences.*

Outwater, M. and Vovsha, P. (2012) Seminar 7 Long term and mobility choice models. *Activity Based Modeling TMIP Webinar Series.*

Panwai, S. and Dia, H. (2005) Comparative Evaluation of Microscopic Car-Following Behavior *IEEE Transactions On Intelligent Transportation Systems Vol 6 No.3.*

Parker, D., Manstead, A., Stradling, S., Reason, J. and Baxter, J. (1992.) Intention to commit driving violations: an application of the theory of planned behavior. *Journal of Applied Psychology 77 (1)*

Parsons Brinckerhoff (2011) Activity-Based Model for Highway Pricing Studies at Chicago Metropolitan Agency for Planning.

Prietula, M., Carley, K. and Gasser, L. (1998) *Simulating Organizations: Computational Models of Institutions and Groups.* Cambridge, Massachusetts: The MIT Press.

Prillwitz, J. and Lanzendorf M., (2006) The Importance of Life Course Events for Daily Travel Behaviour – A Panel Analysis, *paper presented at the 11th International Conference on Travel Behaviour Research, Kyoto, August 2006.*

Pritchard, D. and Miller, E. (2012) Advances in population synthesis: fitting many attributes per agent and fitting to household and person margins simultaneously. *Transportation Vol 39 (3).*

Poulter D., Chapman P., Bibby P., Clarke D., Crundall D. (2008) An application of the theory of planned behaviour to truck driving behaviour and compliance with regulations. *Accident Analysis and Prevention 40, 2058–2064.*

Quine, L., Rutter, D., and Arnold, L. (1998). Predicting and understanding safety helmet use among schoolboy cyclists: A comparison of the theory of planned behavior and the health belief model. *Psychology and Health*, *13*, 251-269.

Railsback, S. and Lytinen, S. (2006). Agent-based simulation platforms: review and development recommendations. *Simulations* 82: 609–623.

Railsback, S. F., & Grimm, V. (2011). *Agent-based and Individual-based Modeling: A Practical Introduction*. Princeton university press.

Robinson, D., Brown, D ., Parker, D., Schreinemachers, P., Janssen, M., Huigen, M., Wittmer, H., Gotts, N., Promburom, P., Irwin, E., Berger, T., Gatzweiler, F., Barnaud, C. (2007) Comparison of empirical methods for building agent-based models in land use science. *Journal of Land Use Science Volume 2, Issue 1.*

Rounsevell, M., Robinson, D., and Murray-Rust, D., (2012) From actors to agents in socio-ecological systems models. *Philosophical Transactions of the Royal Society.*

Rossiter, D., Ballas, D., Clarke, G., and Dorling, D. (2009). Dynamic spatial microsimulation using the concept of ghosts. International Journal of Microsimulation, 2(2):15–26.

Russell, S. and Norvig, P. (1995) Artificial Intelligence: A Modern Approach, *Englewood Cliffs, NJ: Prentice Hall.*

Ryan, J., Maoh, H., Kanaroglou, P., (2007) Populaion Synthesis: Comparing the Major Techniques Using a Small, Complete Population of Firms*. Working Paper 026, Centre for Spatial Analysis, McMaster University.*

SACTRA (1994) Transport and the economy: full report. DfT.

Samuelson, P. (1938) A Note on the Pure Theory of Consumers' Behaviour. *Economics 5 (17): 61-71.*

Samuelson, P. (1947) *Foundations of Economic Analysis*, Harvard University Press.
Samuelson, P. (1948) *Economics: An Introductory Analysis*. McGraw–Hill.

Sayer, R. (2000). *Realism and social science*. London: Sage.

Scheibehenne, B., Miesler, L. and Todd, P. (2007) Fast and frugal food choices. Uncovering individual decision heuristics. *Appetite 49 587-589.*

Schelling, T. (1978) Micromotives and Macrobehavior. *Norton:New York.*

Schreiber, J., and Pekarik, A. (2014) Technical Note: Using Latent Class Analysis versus K-means or Hierarchical Clustering to Understand Museum Visitors. *The Museum Journal 57(1).*

Scott, A., Evandrou, M., Falkingham, J., Johnson, P. and Zaidi, A. (2003) Beyond 2020: Lessons from the Development of the SAGE Dynamic Microsimulation Model for Britain, *ESRC SAGE Discussion Paper no.14, London School of Economics.*

Serenko, A. and Detlor, B. (2002). 'Agent Toolkits: A General Overview of The Market and an Assessment of Instructor Satisfaction with Utilizing Toolkits in the Classroom'. *Working Paper #455.*

Shmueli, G. (2010) To Explain or to Predict? *Statistical Science, Vol. 25, No. 3, 289–310*

Schwanen, T. and Lucas, K. (2011) Understanding auto motives. *In, Lucas, K., Blumenberg, E. and Weinberger, R. eds. Auto Motives: Understanding car use behaviours Bradford: Emerald.*

Simmonds, D., Bates, J. and May, A. (2001) A New Look at Multi-Modal Modelling. *Department of Environment, Transport and the Regions.*

Simon, H. (1955) A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *Vol. 69, No. 1. pp. 99-118*

Sivakumar, A. (2007) Modelling Transport: A Synthesis of Transport Modelling Methodologies, *Imperial College London.*

Sloman, L., Cairns, S., Newson, C., Anable, J., Pridmore, A. and Goodwin, P. (2010) The Effects of Smarter Choice Programmes in the Sustainable Travel Towns: Research Report.

Spielauer, M. (2011). What is social science microsimulation? *Social Science Computer Review, 29(1), 9-20.*

Stanbridge, K., Lyons, G., and Farthing, S. (2004). Travel behaviour change and residential relocation. *Paper presented at the 3rd International Conference on Traffic and Transport Psychology, Nottingham.*

Stopher, P., Clifford, E., Swann, N. and Zhang, Y. (2009) Evaluating voluntary travel behaviour change: Suggested guidelines and case studies. *Transport Policy, Vol 16.*

Swanson, J. (2008) Transport and the Urban Economy: The Urban Dynamic Model.

Teahan, W. (2010) Artificial Intelligence – Agents and Environments. Ventus Publishing.

Temme, D., Paulssen. M, and Dannewald, T. (2007) Integrating latent variables in discrete choice models – How higher-order values and attitudes determine consumer choice. *SFB 649 Discussion Paper 2007-065.*

Thalheim (2011) Handbook of Conceptual Modeling: Theory, Practice, and Research Challenges. *Springer.*

Thorngate, W. and Edmonds, B. (2013) 'Measuring Simulation-Observation Fit: An Introduction to Ordinal Pattern Analysis'. *Journal of Artificial Societies and Social Simulation 16 (2).*

Thornton, A., Evans, L., Bunt, K., Simon, A., King, S. and Webster, T. (2011). Climate Change and Transport Choices. Segmentation Model - A framework for reducing CO_2 emissions from personal travel.

Timms P. (2008) Transport models, philosophy and language, *Transportation Volume 35.*

Tobias R. and Hofmann, C. (2004) Evaluation of free Java-libraries for social-scientific agent based simulation". *Journal of Artificial Societies and Social Simulation* **7** (1).

Triandis, H. (1977) Interpersonal Behaviour. *Monterey, CA: Brooks/Cole.*

Urban and Civic (2012) Alconbury Weald Transport Assessment.

Van der Waerden, P., Timmermans, H., and Borgers, A., (2003) The influence of key events and critical incidents on transport mode switching behaviour: a descriptive analysis. *Presented at the 10th international conference on travel behaviour research. Lucerne August 2003.*

Verplanken, B., Aarts, H., and van Knippenberg, A. (1997) Habit and information use in travel mode choices. *Acta psychological.* 96 1-14.

Vigar G. (2006) Deliberation, participation and learning in the development of regional strategies: Transport policy making in North East England. *Planning Theory and Practice* , **7**(3), 267-287.

Vovsha, P. and Gliebe, J. (2012) Activity Based Modeling. *TMIP Webinar Series.*

Walker, J. and Ben-Akiva, M. (2002), "Generalized Random Utility Model," *Mathematical Social Sciences, 2002, Vol. 43, No. 3, pp. 303-343.*

Walsh, S. P., White, K. M., Hyde, M. K., & Watson, B. (2008). Dialling and driving: Factors *influencing intentions to use a mobile phone while driving. Accident Analysis & Prevention, 40, 1893-1900*.

Weiner, E. (1002) Urban Transportation Planning In The US - A Historical Overview. *Office of the Secretary of Transportation. Washington, D.C.*

Wilensky, U. (1999) Netlogo, *Center for Connected Learning and Computer-Based Modeling. Northwestern University: Evanston,IL.*

Williamson, P. (2007) The Role of the International Journal of Microsimulation. *International Journal of Microsimulation Vol 1.*

Williamson, P. (2009) Microsimulation in the UK: the current state of play. *ESRC Microsimulation Seminar Series Seminar 4: 'Bridging the Gaps; Setting the Agenda'.*

Wooldridge, M and Jennings, N. (1995) Agent Theories, Architectures, and Languages: a Survey, *in Wooldridge and Jennings Eds., Intelligent Agents, Berlin: Springer-Verlag, 1-22.*

WSP (2008) Modelling and Appraisal of Smarter Choices: Review of empirical data for practical modelling. *Department of Transport.*

Yanez, M., Raveau, S., Rojas, M. and de D'Ortuzar, J. (2009) Modelling and forecasting with latent variables in discrete choice panel models. *European Transport Conference.*

Appendices

Appendix 1: Modelling workplace travel plan initiatives

This appendix presents and extends a table produced by WSP (2008) which considers how standard four stage transport models can be extended to include workplace travel plan initiatives. The table considers four groups of measures:

Group 1: Measures that are already represented by existing models

Group 2: Measures that are already represented or approximated by best practice models only

Group 3: Effects can be modelled specially using existing techniques, if supported by robust employee surveys

Group 4: Individual effects not well understood to support transport modelling.

Each measure is considered with regard to current modelling guidance from the DfT, known as WebTAG, and whether more data is needed to understand the scale and nature of the response to be modelled. The final column has been added here to illustrate how agent based modelling may address the issues raised for the modelling of each measure.

Group	.Possible component Measures within an initiative	Expected demand Responses	Can use standard Modelling methods covered by WebTAG?	Is there a general need to enhance operational models?	Is there a need for further empirical evidence?	Relevant ABM characteristic
1	New conventional bus or rail services linking to site/area	Increase in bus/rail mode share	Yes	Effects can be represented only if the model zones are detailed enough	Not essential	Individual recording of agent characteristics e.g. location relative to transport services
1	'Work buses' between site and town centre that can be used by any traveller	Increase in bus mode share	Broadly yes	Effects can be represented only if the model zones are detailed enough to identify the site	Not essential	Individual recording of agent characteristics e.g. location relative to transport services
1	Dedicated 'work buses' between site and town centre	Increase in bus mode share	Broadly yes	Effects can be represented only if the site and the workers from the site can be identified in demand modelling	Not essential	Individual recording of agent characteristics e.g. location relative to transport services
1	Subsidised fares; Interest-free season ticket loans	Increase in PT mode share	Broadly yes - Can reflect the effective reduction in commuting fares	Not for reflecting the effective reduction in commuting fares	Yes	Individual recording of agent characteristics e.g. fare paid for transport services
1	Special deals to reduce the cost of bus and rail commuting	Temporary deals have probably no extra effects over and above PT information provision	Yes, as part of commuting fares	Not for reflecting the effective reduction in commuting fares	Not essential	Individual recording of agent characteristics e.g. fare paid for transport services. Can change costs for a short period of time.
1	Parking 'cash out' (payment on days of not driving)	Reduction in car mode share	As reduction in non-car travel costs	No	Not essential	Individual recording of agent characteristics e.g. fare paid for transport services
	Car parking restricted to	Reduction in	As restraint in	Most models would not	Yes	Individual recording of

2	essential users	car mode share	capacity, although this assumes that parking is well represented in demand modelling	be able to segment between essential and non-essential users, and would need to approximate the effects		agent characteristics e.g. permission to park at certain locations
2	Parking charges	Reduction in car mode share	No – the general principles of demand modelling applies but there is no comprehensive advice on parking	Most models do not represent parking well and would require enhancement	Not essential	Individual recording of agent characteristics e.g. fare paid for transport services
3	Car sharing scheme	Increase in car occupancy And reduction of single occupancy car trips – likely to have higher take-up rates within a large pool of workers who have nearly identical arrival and departure times	Not covered as a matter of course	Yes	Yes	Individual recording of agent characteristics e.g. location of home and work place, proximity to others with similar details, networks modelled
3	Preferential car parking for sharers	Increase of car sharing probably – although the potential scale of the effect is not known	No	Yes	Yes	Individual recording of agent characteristics e.g. permission to park at certain locations
4	Demand activated bus services (variable routes, bus-taxis)	Increase in bus mode share	No	Yes	Yes	Individual recording of agent characteristics, e.g. waiting time for pt service
4	Giving all staff public transport information	Increase in PT mode share, if associated with improved services/facilities	Bluntly as part of the mode specific constant	Additional work required to quantify implicit mode specific constants in the case of incremental demand models	Yes	Networks, history of agent, recorded whether each individual is aware of their travel options
4	Offering personalised journey plans to staff	Possible increases in PT walk and cycle share, if associated with improved services/facilities	Bluntly as part of the mode specific constant	Additional work required to quantify implicit mode specific constants in the case of incremental demand models	Yes	Networks, history of agent, recorded whether each individual is aware of their travel options
	Secure cycle	More cycle mode share,	Bluntly as part of the	Additional work	Yes	Individual recording of

4	parking	although minimal effects if implemented in isolation	mode specific constant (if cycle is identified as a separate mode)	required to quantify implicit mode specific constants in the case of incremental demand models		agent characteristics e.g. access to secure cycle parking
4	Changing/shower facilities at work	More cycle mode share, although minimal effects if implemented in isolation	Bluntly as part of the mode specific constant (if cycle is identified as a separate mode)	Additional work required to quantify implicit mode specific constants in the case of Incremental demand models	Yes	Individual recording of agent characteristics e.g. access to showers
4	Business cycle mileage allowance	More cycle mode share, although limited by daily activity patterns	Limited - as a reduction in cycling cost, although no empirical model is known to indicate the trade-off between cycle mileage costs and travel time/convenience	Yes, if the scale of effects is found to be significant	Yes	Individual recording of agent circumstances e.g. business cycle mileage allowance
4	Services on site to reduce need to travel (cafeteria, convenience shop, cash dispenser)	Reduction of some convenience shopping trips from the site	Yes, if distribution modelling cover these trips	Most models would not be able to identify convenience shopping from other personal trips, and would need to approximate the effects	Yes	Individual recording of agent characteristics e.g. location relative to services
4	Encouraging Teleworking	Reduction in commuting trips in the short term and possible longer term changes in residential location and daily activity patterns; increase in other trips during the day	Not completely: Yes for reductions in commuting trips, if the model can identify staff categories that can telework; No for forecasting take-up rates, frequencies and saturation levels of telecommute, home location changes, and associated	Yes	Yes - The effects of teleworking on commuting patterns are believed to be significant but there is a need for substantial further research to ascertain the nature and scale of effects	Individual recording of agent characteristics e.g. permission and ability to tele-work

			increases in other trips.			
4	Flexible working hours / compressed working hours to e.g. a 4-day week, or commuting during inter-peak or off-peak periods	Reduction in peak time travel	Not completely: Yes for reductions in commuting trips, if the model can identify staff categories that can flexiwork. No for associated changes in daily activity patterns	Yes	Yes	Individual recording of agent characteristics eg ability to vary working hours

Appendix 2: DfT Climate Change and Transport Choices Dataset

The main dataset used in this research project was collected for a study commissioned by the UK Department of Transport (DfT) to 'identify and quantify groups or segments within the population that differ in terms of the factors relevant to reducing CO_2 emissions from personal transport use'. The motivation for the study was the belief that a better understanding of the segments within the adult population in England would assist the DfT and others 'to develop more targeted and effective sustainable behaviour initiatives'. The segmentation analysis produced by the study team was made available for use by local authorities bidding for large projects and small projects (Tranche 2) to the DfT's Local Sustainable Transport Fund.

The study was undertaken as a response to the findings of a series of reviews commissioned by the DfT to understand if and how individual's awareness of and attitude towards climate change affected their travel behaviour. A study of the evidence on public attitudes to climate change and travel (Anable *et al.,* 2006) had noted that the UK population was not homogeneous in respect of its attitudes and desire to reduce carbon emissions from personal travel and that attempts by central government to influence travel choices would be more effective if, rather than a 'one size fits all' campaign the marketing was more targeted at distinct segments.

They noted that demographic characteristics alone were insufficient to define these targets and that a segmentation exercise should also take into attitudes towards climate change, motivations for and barriers against changing travel behaviour and how other relevant psychological factors also vary between people.

A subsequent qualitative study for the DfT (King et al., 2009) explored in more depth public attitudes to climate change and transport, considering their motivations for and barriers against changing personal travel behaviour. They reported that although increasing a person's awareness of climate change appeared to increase their willingness to change their own travel behaviour this was not translated into an actual change in their behaviour. The findings of the evidence review (2006) and the qualitative research (2009) were used to inform the design of the DfT Climate Change and Transport Choices segmentation study (2010, 2011) which produced the dataset used in the agent based model of commuter mode choice.

Quantitative survey

The DfT segmentation study consisted of two parts; a quantitative survey was used to collect the data for use in the construction of a segmentation model for the UK population and a qualitative survey was held to better understand the attitudes, behaviours and characteristics of each of the segments. The quantitative survey consisted of over 3900 face to face interviews carried out between November 2009 and June 2010, except for between 5 March and 21 May 2010 when contact with the public was suspended due to the general election held on 6 May 2010. The survey was designed to explore the respondents' attitudes and actual behaviour regarding travel and climate change. It concentrated on their current trips by car to work, school or college, business travel and regular and smaller food shopping trips as the evidence reviews indicated that these types of trips offered the most potential for reducing carbon use in transport through changes such as trip reducing the number of trips made, switching to the use of modes which had lower carbon emissions or purchasing cars which produced less emissions.

The questionnaire consisted of four sections. The first section started with questions to record the respondents current travel behaviour and their attitudes towards different modes of transport. The second section of the survey was a choice modelling exercise designed to uncover which of a set of factors (mode, time taken, cost and CO2 emissions) most affected a respondents' mode choice, the relative value placed on each of these factors and level of inherent like or dislike of each mode. In the choice modelling section a respondent was shown three options for a specific journey. Each option gave the mode, journey time, cost and CO_2 emissions for that option. The respondent picked their most favoured option or could choose not to travel at all. Each respondent was always asked about the same journey but the characteristics of each of the available options varied every time. In all each respondents was asked carry out this task for 10 sets of options.

The third section asked about the respondent's attitudes towards climate change issues. The questions were asked in this order to reduce the tendency of people to answer in a more environmentally favourable way if the issue of climate change had recently been raised with them. In order to reduce the risk of social desirability bias i.e. respondents producing the answer they think would be received favourably by the interviewer and others, more sensitive questions were answered directly by the respondent onto a computer privately. The final section asked about key demographic factors and other factors which might be used to classify respondents in the segmentation model and to calculate survey weights. A segmentation model was built using the results of sections 1, 2 and 4 of the quantitative survey in autumn 2010.

Qualitative survey

The qualitative research was then undertaken in November and December 2010. It consisted of 14 focus groups, the members of each group belonged to the same segment, as identified in the segmentation model. The aim of the focus groups was 'to test the segmentation and further understand the barriers and motivations towards using various modes of transport or sustainable travel behaviours'. The results of the focus groups have not been reported separately but were used together with the results of the quantitative survey in the writing of the final report on the segmentation model (DfT July 2011) which describes the demographic, behavioural and attitudinal factors of each of the segments in turn.

Choice modelling data

The survey questions and the responses are available on the DfT Website for sections one, three and four of the quantitative sections. The choice modelling exercise and responses are not publicly available. This section of the survey was carried out using proprietary market research survey produced by Sawtooth. The interim report on the quantitative surveys (2010) contains a chapter on the choice modelling. As a check on the data files, the files retrieved from the project archive were re-analysed using the Sawtooth software and the results matched with those reported in the interim report. The re-run results were found to be perfect replications of those recorded in the interim report.

The choice modelling results for each person were used to calculate the value they placed on each element of a journey option i.e. the mode, time, cost and CO_2 emissions. The value, or utility, placed on these elements for each individual was then merged into the data set that contained their responses to the other three sections of the quantitative survey.