Modelling Travellers' Choice of Information Sources and of Mode

AGHA FAISAL HABIB PATHAN

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

Institute for Transport Studies

March, 2010

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

The author wishes to express his earnest gratitude, deepest respect and heartfelt thanks to his thesis supervisors, Prof. Peter Bonsall and Prof. Gerard de Jong for their valuable guidance, keen attention, and indefatigable encouragement throughout the study. Without their support and generous help, this thesis might not have attained the present shape.

The author is especially indebted to the Mehran University of Engineering and Technology and the Institute for Transport Studies for providing a good venue to further his study, without which this study would not have been possible.

Special thanks must be conveyed to Network Rail and WY Metro for permitting and allowing the author to collect data, the author is deeply impressed with their kindness.

Much appreciation also goes to the people who helped me from time to time during my PhD, Dr. Marco Kouwenhoven who gave invaluable help whenever I needed; Dr. Phani Kumar Chintakayala from whom I have learnt much. The author also appreciates his friend Miss Kaushali Dev for her help in data collection and questionnaire surveys.

Finally, the author is honoured to dedicate this effort to his father, Sardar Agha Habibullah Khan in Heaven who made it possible where he is and made him to study in U.K and to finish his PhD and mother, Mrs. Naseem. The author understands and appreciates their contribution and efforts.

Abstract

This study investigates the travellers' choice of information sources and their subsequent mode choice decisions. The goal of this study is to develop a comprehensive choice model that can capture the information acquisition process by predicting the choice of information sources together with its effects on modal choices of the travellers. A decision making framework for travel Information acquisition is developed and the abstract terms, necessary to be tested in the models, are identified. A Stated Preference experiment is developed based on the complicated decision making process and an interactive CATI questionnaire is designed to cope with it. Utility functions are formulated by expanding travellers' choice set to include different combinations of the viable sources of information and with the inclusion of policy sensitive variables. The research employs a wide range of modelling methodologies and examines a range of traditional and newly developed calibration and estimation procedures including Mixed Logit models with individual specific parameters and the newly developed Random Regret Minimisation framework. The study also analyses the effects of travel planning websites on travel decisions and establishes a link between content, design, advertisements, and presentation of information on overall modal shift.

The results indicate that travellers give credence to government owned sources and give more importance to their own previous experiences followed by multimodal websites, train websites, friends and coach websites respectively. A website with less search time, specific information on users' own criteria, and real time information is regarded as most attractive by the travellers. The study also found that the market share of the modes increases when information sources show decreased travel time and cost values and the maximum results are achieved when different information sources give the same information to the travellers. These results show that information sources could be used to influence the mode choice of the travellers.

Contents

Tables

Figures

Chapter 1
Introduction

1.1 Introduction

Traveller information has long been used to improve the traffic conditions and many types of information services and products are already in market while research continues to introduce second generation of these systems by using portable devices of personal communication. Abundant bodies of literature are available on the impacts of these technologies since last decade or so (see among many others, Arnott *et al.*1991; Ben-Akiva *et al.* 1991; Mahmassani and Jayakrishan 1991; Emmerink *et al.* 1995 a,b; Bonsall 1995; Al-Deek *et al.* 1998; Mahmassani and Liu 1999; Stough 2001).

Most of the studies focused on the usage of these Advanced Traveller Information System (ATIS) technologies and their impact on the behaviours of the people. There are substantial empirical studies that study the travellers' process of information acquisition and its effect on their travel decisions (especially Polak and Jones, 1993; Emmerink *et al.* 1996; Pollydoropoulo & Ben-Akiva 1998; Hato *et al.* 1999; Khattak *et al.* 2003, Kenyon & Lyons 2003; Chatterjee & McDonald, 2004; Abdel-Atty & Abdella 2004; Bogers *et al.* 2005).

The literature review, carried out in this study, on the information need and type suggests that, in most of the cases the people require information about travel time and travel cost for different modes (Hague Consulting Group, 1991). The literature also suggests that age, sex, income level and education are the key factors that influence the use of traveller information (see among others Bonsall 1992a, 1993, 1995; Khattak *et al.*, 1993a, Mannering *et al.*, 1994; Emmerink *et al.*, 1996; Petrella and Lappin, 2004, Caplice and Mahmassani, 1992, Allen *et al.*, 1991, Wardman *et al.*, 1997, Khattak *et al.*, 1993b, Petrella and Lappin, 2004, Hato *et al.*, 1999, Polydoropoulou & Ben-Akiva, 1998, Yim & Khattak, 2001, Goulias *et al.*, 2004). Trip purpose has been found to be a very important factor that drives individuals to use traveller information. It has been reported by various studies that different trip purposes have produced different responses towards traffic information. It was found that commuters were less likely to divert to alternative route under information as compared to other trip purposes. Literature suggests that combination of prescriptive and quantitative information influence travellers more in their decisions as compared to only qualitative information. The credibility of the information source is found to be an important determinant which influences the travellers'' decisions. Bonsall and Parry (1991) found that travellers tend to give less credence to the information in comparison to what they actually observe with their own eyes. Bonsall *et al.* (1991) also found that travellers prefer to test the credibility of the information randomly before considering using that source. Bonsall (1992a) reported that the influence of traffic information on route choice depended on whether the information was credible, relevant and clear. He also reported that the credibility of the information source heavily influences its compliance.

Very few studies (notably Chorus *et al.* 2006a) actually focused on the process of information acquisition together with its effects simultaneously in one behavioural network. But the study only focused on the attributes of the alternative mode choices assuming information source as fully reliable. Moreover attributes of information sources available now were not considered. This study takes into account the properties of available information sources in the market, their attractiveness and people's choice of these information sources. Chorus *et al.* (2006a) has carried out detailed and comprehensive literature review of the use and effects of ATIS and have identified large gaps of empirical studies on behavioural mechanisms in a multimodal context.

The motivation behind this research is to investigate the factors that people do consider in selecting different sources of information in information rich travel environment. It is important to explicitly model the abstract terms involved before incorporating the remaining process of information acquisition and subsequent travel decisions. Moreover, these stages also require market segmentation of the people for example their income, economic factors etc so as to better comprehend degrees of market penetration and willingness to pay for these sources. These understandings will benefit information service providers, manufacturers and suppliers of these products to understand the impacts of their ATIS and predict their profitability. It will also help government agencies and policy makers to stimulate changes in the travel decisions of travellers and influence their mode choice decisions. Thirdly it will also benefit public transport providers to use these sources in attracting and retaining their customers. The goal of this study is to develop a comprehensive choice model that can capture the information acquisition process by predicting the choice of information sources together with its effects on travel choices of the travellers.

1.2 Objectives of the Thesis

Most of the previous research on information acquisition and sources deal with only few information sources. Few studies encompass recently available information sources (such as the Internet) in the choice set. Furthermore studies mostly link the determinants of information sources with route choice decisions and there is lack of the research that links various sources of information with travellers' mode choice decisions. Hence, the main purpose of this research is to study and predict the travellers' choice of information sources and subsequently of mode. Specifically, the objectives of this study are:

- To conduct a travel behaviour survey, to investigate the travellers' choice of information sources in different scenarios.
- To model the choice of information source and subsequent mode choice.
- To analyse and evaluate the impact of information on mode choice.

1.3 Conceptual Framework for the Study

The framework presented in Figure 1.1 is divided into four main stages consisting of preliminary investigation, survey design and data collection, model formulation and calibration, and application.

The preliminary investigation is discussed in chapter 2 and section 3.1. Data issues will be discussed in section 3.2 and the analytical issues will be discussed in section 3.3.

1.4 Thesis Structure

Following this introductory chapter, the thesis begins by taking a look at past research in the area of information search. The literature review reported in Chapter 2 is mostly focused on building an understanding of effects of different types of information on traveller's decisions. The literature review on the information need and type; age, sex, income level and education; trip purpose; information content; and the credibility of the information source suggests that they are important determinants which influence the travellers' decisions. The literature review also provides an overview of the models of response and evaluation of benefits of providing information. The chapter also details various theories of information search and use.

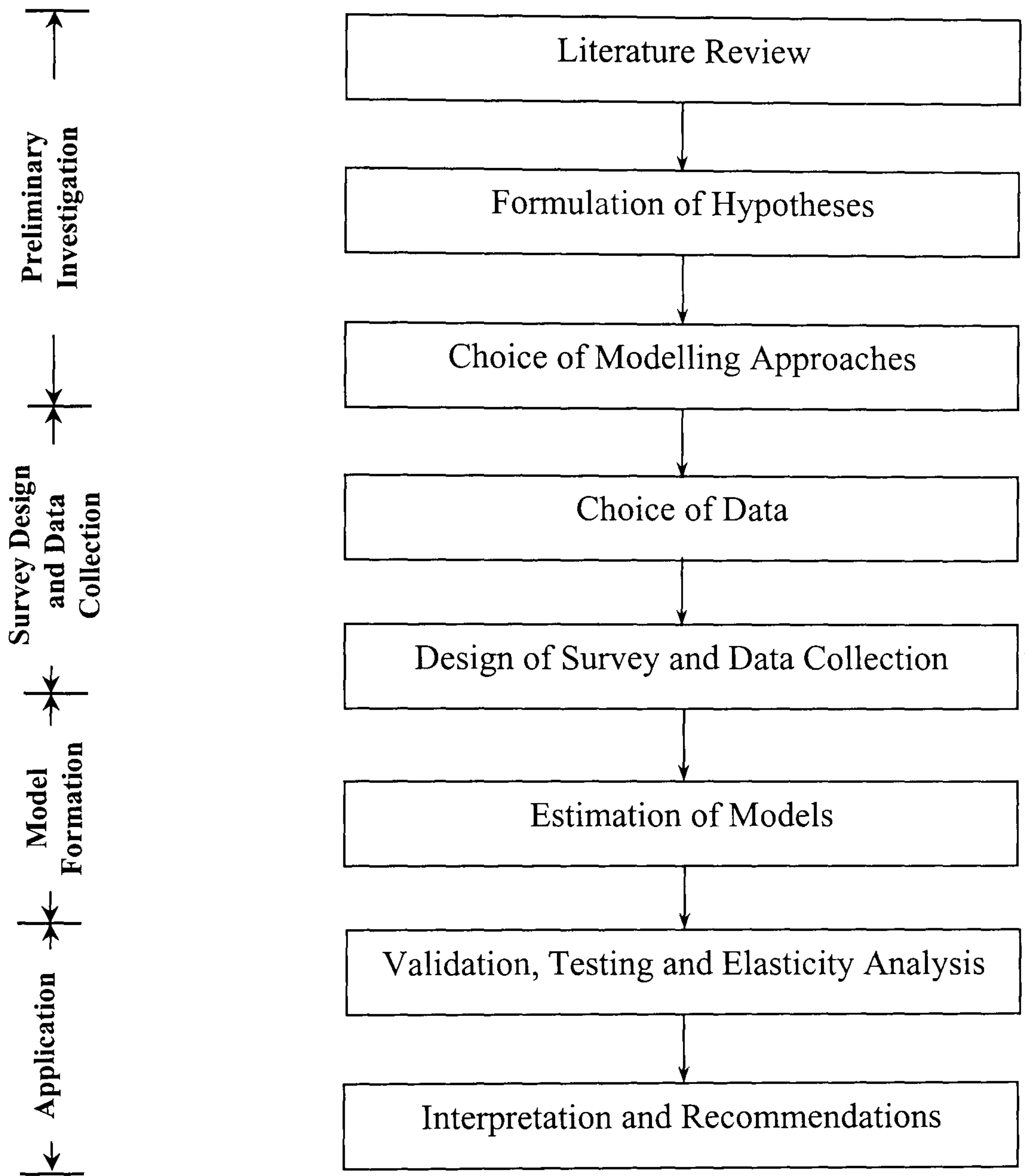

Figure 1.1: Conceptual Framework of this Research

Chapter 3 begins by discussing the issues raised by the literature review and identifying hypotheses to be tested. It discusses in detail the conceptual representation of information search and travel choice process. Based on this general framework a modelling framework is identified that simplifies the general conceptual framework into workable modelling scheme. Two models are developed namely Source Model and

Mode Model. The chapter then addresses the availability of data and the choice of data sources as well as more detailed issues such as the choice of target population, sampling strategy and sample size. Finally the chapter discusses a number of analytical issues including the choice of appropriate models to be estimated.

Chapter 4 briefly discusses the design of the questionnaire and includes the details of the SP experimental design and explains the choice of SP scenario, attributes, and levels for the selected variables for both source and mode choice experiments. It also describes the selection of the Computer Assisted Telephonic Interview (CATI) technique and the role of simulation and pilot studies for the development of the final SP exercise.

Chapter 5 details the conduct of the main survey and explains the descriptive analysis of the results. It summarises the respondents' characteristics and the cross relationship between different the socio-economic characteristics, information source selection and mode choice. It also discusses the respondents' attitudes towards different types of information sources.

Chapter 6 is the biggest chapter in the thesis. It presents the process of model development and details of the estimation techniques and analysis undertaken. The model development process includes the calibration and estimation of Revealed Preference (RP) models for mode choice, RP models for source choice, Stated Preference (SP) models for mode choice, SP models for source choice, Combined RP models for source and mode, Combined SP models for source and mode and Combined RP and SP models for source and mode. Four alternative model specifications are calibrated from the collected data. The four model specifications include the multinomial logit (Mc Fadden, 1973), the nested logit (Ben Akiva, 1974), the mixed logit (Cardel and Dunbar, 1980) and the random regret minimisation (Chorus, 2009) models. All the models were estimated to explain both mode choice and source choice decisions. The Jacknife method was applied to correct the SP Multinomial Logit (MNL) models and correction for random taste variation was applied to the mixed logit models.

In Chapter 7, the disaggregate choice models developed in Chapter 6 are applied to generate forecasts of information source and mode choice. The forecasts are made on both the estimation sample and the expanded sample taking account of the NTS 2006 long distance travellers' demographics.

Finally Chapter 8 provides the summary of the achievements of this work and identifies areas that would benefit from further research.

The Appendices include the final questionnaire of the main survey, the code of the CATI survey developed in WinMint and a glossary of variables used in the models.

1.5 Novel Features

Novel features of this work include:

- The identification of the abstract terms involved in the process of travel information acquisition and necessary to be tested in the models of information sources and subsequent mode choices.
- Development of a decision making framework for the travel information acquisition process.
- Expansion of the travellers' choice set to include different combinations of the viable sources of information.
- The inclusion of policy sensitive variables including credibility.
- Analysis of the effects of travel planning websites on travel decisions.
- The study of the influence of information sources on the mode choice decisions of the travellers under various circumstances.
- Establishment of the link between content, design, advertisements, and presentation of information on overall modal shift.
- The analysis of the travellers' treatment of low credible sources and the factors which affect the credibility of a source.
- The analysis of information source and mode choice models using Mixed Logit (MMNL) models with individual specific parameters.
- The use of the newly developed Random Regret Minimisation (RRM) models to estimate the information source choices and subsequent mode choice model.

Chapter 2
Literature Review

The literature review is mostly focused on building an understanding of effects of different types of information on the traveller's decisions. It has been observed that literature in this area has grown dramatically over the last two decades.

2.1 Travellers' Behaviour under Information

Behavioural surveys of drivers during congested traffic conditions are best suited to developing Advanced Traveller Information Systems (ATIS) behavioural models. Properly designed surveys that capture the interactions in the travel behaviour model allows for the investigation of the influence of (a) unexpected and expected congestion, (b) the various types and quality of information received about congestion and (c) drivers' experiences with congestion and related information on the whole spectrum of pre-trip and en-route decisions. In particular, these behavioural surveys allow for the relationship between a driver's response to qualitative, quantitative, predictive delay and both prescriptive and descriptive information (Dia, 1999).

In order to better understand the responses to travel information it is desirable to first determine what type of information is required, who wants traveller information, and what should be the information content. The following review of literature explores the above mentioned questions.

2.1.1 Information Need

It has been widely acknowledged that providing travellers the information about their travel choices influences their behaviour in ways that are beneficial for the efficiency and use of the transport system (Koppelman *et al.* 1980; Kanninen 1996). The literature review on the information need and type suggests that, in most of the cases the people require information about travel time and travel cost for different modes (Hague Consulting Group, 1991). Steg (2005) carried out interviews of car users and found that people consider convenience and comfort an important consideration while deciding about travel and modes. It was found that in order to influence people to use other modes the information about these less tangible characteristics could be used to influence the travellers in their mode choice decisions. These results also suggest that policy makers should not exclusively focus on instrumental motives for car use, but they should consider other social and affective motives as well.

Chorus *et al.* (2006b) carried out a web survey to study the need of travel information and the knowledge levels about different modes. It was found that travellers perceived themselves informed and resourceful (perceived awareness of alternative routes for a given mode and destination) for familiar destinations. On the other hand, level of experience with a given mode was found to be less important. Destination familiarity also was found to be substantially important factor that influences traveller's perceived reliability of their estimates about travel time and cost. The other important factor that influenced this reliability was incident occurrence. It was reported that both perceived lack of resourcefulness and particularly lack of reliable trip related knowledge trigger need for information especially under disturbances. It was also found from the study that apart from time and cost related information, a clear need exists for more advanced information e.g. trip guidance.

In recent research it has been found that latent characteristics of travel alternatives like comfort, convenience, privacy and safety are also very important considerations for travellers in travel decisions (Steg, 2001; Thogersen, 2001; Ellaway *et al.,* 2003 Bos *et al.*, 2004; Anable & Gatersleben, 2005).

Willingness to pay for the travel information has been studied by a number of researchers and they conclude that in general there is a low willingness to pay for information from the available sources. (Polydoropoulou *et al.*, 1997; Khattak *et al.*, 2003). The willingness to pay for the public transport information is also found to be low as travellers feel that they have already paid for the information by buying the ticket. (Vance & Balcombe, 1997; Molin & Chorus, 2004; Molin *et al.*, 2005).

2.1.2 Individual Characteristics

The literature on the traveller information suggests that age, sex, income level and education are the key factors that influence the use of traveller information. Literature suggests that females are risk averse and therefore are more reluctant in switching their usual travel patterns under information as compared to males (Bonsall 1992a, 1993, 1995; Khattak *et al.*, 1993a, Mannering *et al.*, 1994; Emmerink *et al.*, 1996; Petrella and Lappin, 2004). Whereas, Caplice and Mahmassani (1992) found that for radio reports, females are more likely to switch their departure times as compared to their male counterparts.

Apart from the gender, the age of the traveller also influence the response towards the information. Allen *et al.* (1991) found that older drivers are less aggressive and are less likely to change route as compared to the younger drivers. Bonsall (1992a) found that young drivers were less prone to be influenced by prescriptive information. Similarly Wardman *et al.* (1997) found that younger drivers are less likely to be influenced with the advice of VMS sign as they are less sensitive to delays as compared to older drivers.

Khattak *et al.* (1993b) reported that higher income drivers were more prone to divert under traffic information as they have higher values of time. Petrella and Lappin (2004) found that high income and well educated travellers are more at ease with the new technologies and are therefore more likely to use travel information. Emmerink *et al.* (1996) reported that professionals are more likely to use travel information as these appear to attach greater importance to make an accurate decision about their travel (Hato *et al.* 1999).

Polydoropoulou & Ben-Akiva (1998) and Yim & Khattak (2001) found that travellers who own mobile phones and are familiar with the use of internet are more likely to be using ATIS. This is because they are more likely to be experienced in handling the advanced information sources and are aware of the potential of ATIS in improving their travel related decisions. Polydoropoulou and Ben-Akiva (1998) found that people having attitude of control seeking and those who like new technology are more likely to be aware of new information sources as compared to others. Goulias *et al.* (2004) found that professionals, high income individuals, people from younger generation and car owners are more aware of new technologies and ATIS as compared to others.

2.1.3 Trip Characteristics

Trip purpose has been found to be a very important factor that drives individuals to use traveller information. It has been reported by various studies that different trip purposes have produced different responses towards traffic information. Hato *et al.* (1995) found that commuters were less likely to divert to alternate route under information as compared to travellers on shopping trips. Emmerink *et al.* (1996) and Hato *et al.* (1999) reported that on the business trips travellers are more likely to listen to traffic information as compared to drivers on other journey purposes. It was also reported that commuters were less likely to be influenced by information as compared to business travellers. Polydoropoulou & Ben-Akiva (1998) and Srinivisan *et al.* (1999) reported that travellers with arrival time sensitive trips are more likely to search and use traveller information. Petrella & Lappin (2004) reported that commuters have more inclination to use information as compared to other recreational trips.

It has also been concluded by many researchers that the longer the travel time, the more likely travellers are to use traffic information and divert to alternate route (Caplice and Mahmassani 1992; Khattak *et al.* 1993a; Khattak *et al.* 1993b; Emmerink *et al.* 1996; Lappin 2000; Yim and Khattak 2001; Targa *et al.* 2003). This is because travellers on longer journeys have more opportunities to use traffic information and to change routes. Network familiarity has also been found important factor that influences use of traffic information. Moreover it has been reported by many researchers that familiar drivers are less likely to be influenced by information (Huchingson and Dudek 1979; Khattak *et al.* 1991, Mahmassani and Chen 1991; Bonsall and Joint 1991 and Bonsall 1992a). However, in the absence of information during congested scenarios familiar drivers are more likely to divert as compared to their non familiar counterparts (Khattak *et al.*, 1993a; Bonsall and Hounsell, 1994). As familiarity with alternative routes increases the propensity to diversion under traffic information also increases (Khattak *et al.*, 1993a; Bonsall and Hounsell, 1994). On the other hand, unfamiliarity with alternative routes forces them to stay on the current route in the presence of traffic information (Wardman *et al.*, 1997). This implies that the familiar drivers are likely to make better use of information by taking their own decisions under the influence of traffic information (Bonsall and Joint, 1991; Mahmassani and Chen, 1991; Hato *et al.*, 1995). Travellers who use public transport frequently to travel to work have less need for travel information (van Wee and Dijst, 2002; Verplankenand Aarts, 1999).

2.1.4 Message Content

The following review of literature relates traveller response towards different type of information contents. Literature suggests that combination of prescriptive and quantitative information influence travellers more in their decisions as compared to only qualitative information. Following are various studies aimed at capturing traveller response towards different type of information.

a. En-route Information

Bonsall and Parry (1991) found by using interactive route choice simulator that frequent travellers preferred descriptive information over prescriptive information. This was because they believed that they could make better decisions than the real time information service and they did not wanted to be controlled by a machine.

Bonsall *et al.* (1991) and Schofer *et al.* (1993) found that unfamiliar drivers were more likely to accept prescriptive information while travellers in familiar areas were more inclined to use descriptive information in order to make their own route choices.

Khattak *et al.* (1991) carried out an extensive survey of downtown Chicago automobile commuters and found that en-route diversion behaviours influenced by source of traffic information, expected length of delay, regular travel time on the usual route, number of alternative routes used recently, anticipated congestion level on the alternate route, gender of driver, residential location, self-evaluation statements about risk behaviour and stated preferences about diverting. The key finding was that real-time traffic information influence en-route diversion behaviour. Short-term improvements in real time traffic information should focus on disseminating information about length of delay due to incidents and the congestion levels on the alternate routes surrounding the incident.

Khattak *et al.* (1993b) and Bonsall (1995) found that descriptive information that includes information without any route guidance was more influencing on travellers' route choice decisions as compared to only guidance. However Bonsall (1995) suggested that prescriptive information coupled with rationale for the advice might be more effective than if they were supplied separately. Other researchers (Mannering 1989; Mahmassani *et al.* 1990; Khattak *et al.* 1993b) found that if the combination of prescriptive and descriptive information is provided to travellers than it would influence them more in their travel decisions as compared to providing either of them separately.

Khattak *et al.* (1993a) based on analysis of SP and RP surveys, suggested that quantitative information especially if mentioning length of delay was more useful and the drivers are more likely to accept the information and make travel decisions accordingly. This was because the information was more detailed and accurate as compared to only qualitative information.

Polydoropoulou *et al.* (1994) found, from revealed preference data obtained from a survey of MIT commuters, that more reliable and more frequently updated information system than radio would stimulate the acquisition of information. They also found that drivers' own observation and accurate and precise direction that correspond to and reflect actual traffic conditions will be successful at gaining drivers' confidence in following its route choice instructions.

Khattak *et al.* (1996) conducted RP and SP survey to explore automobile commuters' pre-trip decisions and route switching response to unexpected congestion and formulated the model by combining both data types. Estimation results indicated that, given accurate quantitative delay information, commuters could overcome their behavioural inertia when faced with unexpected delay.

Polydoropoulou *et al.* (1996) studied the traveller response towards different types of ATIS information contents. Data on travellers' route switching behaviour were obtained through a survey of California Bay Area automobile commuters. It was found that there was significant heterogeneity in response to various types of ATIS messages. Travellers' propensity to take alternative route increased with prescriptive information.

The most significant increase occurred when quantitative real-time or predictive information was provided. Travel information on alternative route is critical to diversion decisions because lack of experience with alternative routes discourages diversion.

Wardman *et al.* (1997) conducted an SP survey and reported that when information about the length of delay is specified it tends to have more impact on the route choice decisions the qualitative information about the delay. They also found that when drivers received qualitative information about the delay e.g. "long delays" or "delays likely", they valued the former about 35-47 minutes and later at about 10-31 minutes depending upon the cause of delay. Whereas the effect of "all clear" message on VMS sign was different from that of a "blank" sign indicating no information about road conditions.

Khattak A. and Khattak (1998) compared behavioural responses to information across Chicago and San Francisco Bay Area and found that longer duration of residence, higher propensity of discovering new routes and characteristics tend to increase drivers' spatial knowledge. Propensity of diversion increases with higher than usual travel time on usual route plus delay and shorter alternate route travel times. It was found that information sensitive behavioural models are context dependent. It was also found that the potential benefits of ATIS must compete with the benefits already accruing from radio traffic information.

Yang *et al.* (1998) pointed out that additional ATIS information was not always better. The ideal type of ATIS information should reduce a driver's uncertainty regarding traffic conditions instead of overwhelming him or her with unneeded data. In addition, it was found that auditory delivery was adequate and effective when drivers were familiar with the traffic network owning to drivers only need short and simple information to assist them in making route diversion decision. On the other hand, for an unfamiliar network, visual display should accompanied by a corresponding auditory message to minimize driving and maximize information assimilation primarily due to lack of essential information from drivers' long-term memories.

Koo and Yim (1998) conducted a telephone survey after a major highway incident south of San Francisco. The results of their survey suggested that individual incidents do influence travel decisions to some extent if relevant information was obtained. Yet a fair number of participants did not alter their route. The likely explanation for the lack of response to information was that commuters generally do not believe that changing their route will result in shorter travel time. The key to persuade commuters lies in informing them through quantitative information.

Peeta *et al.,* (2000) investigated the effect of different message contents on driver response under Variable Message Signs (VMS). The issue was addressed through an onsite stated preference user survey in the Borman Expressway region in north-western Indiana. The analysis suggests that content in terms of level of detail of relevant information significantly affect drivers' willingness to divert. Other factors include socioeconomic characteristics, network spatial knowledge, and confidence in the displayed information.

Srinivasan *et al.* (2000) used route choice data from a dynamic interactive simulator and found that compliance and inertia mechanisms are present in route choice behaviour. The results also indicated that information quality, network loading and day to day evolution, level of service measures and trip makers' prior experiences are significant determinants of route choice through the inertial and compliance mechanisms.

b. Pre-trip Information

Adler and McNally (1994) categorized the travel decision making process into three stages i.e. pre-trip planning, en-route decision and post trip evaluation. The timings of providing information at different stages of travel have different impacts on the travellers' behaviour because at each stage the decision maker has different objectives to achieve.

Polak and Jones (1993) suggested that when information was given at a higher level of decision making process, it has more influence on various dimensions of behaviour, and the most influence being achieved at the time of acquisition. Bonsall *et al.* (1991) suggested that respondents were not interested in pre-trip information for regular trips, which requires them to get up early in the morning. On the other hand they were willing to receive accurate updated en-route information.

Liu and Mahmassani (1998) presented an empirical analysis of commuters' indifference band for joint departure time and route switching behaviour in response to real-time traffic information, based on data collected using a laboratory interactive dynamic simulator. The analysis focused on the day-to-day dynamics of commuters' joint departure time and route decision process in response to real-time traffic information. Results of this study indicated that commuters tended to switch routes both pre-trip and en-route in response to low reliability of the system perceived by the commuters. In addition, trip makers became more likely to switch routes when the system provided underestimated trip time information than when the system provided overestimated trip times. Finally, commuters tended to switch their route both pre-trip and en-route in response to higher differences between the predicted arrival time at a given decision node and their own preferred arrival times.

Mehndiratta *et al.* (1999) from a simulator based survey found that timeliness of the information and extensive coverage of routes was important to customers. Consumers had strong views on the issue of the form of guidance that they would prefer to receive. Some customers welcomed the concept of route guidance, whereas others preferred to be provided only the traffic information that they needed to make their own routing decisions. Furthermore, woman appeared to be more inclined to accept guidance than men.

Khattak and Palma (1998) conducted an extensive survey in Brussels to catch traveller behaviour under normal and unexpected travel conditions. Among automobile commuters who changed their travel pattern (about 50%) 25% reported that bad weather was important factor in changing their mode, 60% changed their departure time and 35% diverted to alternated routes, whereas 5% kept themselves informed through radio and television.

Shah *et al.* (2001) used the simulated a case study based on travel conditions in Washington D.C. to find out the time management impacts of pre-trip Advanced Traveller Information Systems. Findings from the case study indicated that benefits were significant in terms of on-time reliability but not in terms of the most frequently used measures of ATIS effectiveness, reduction of in-vehicle travel time. It was also found that commuters who did not use traveller information were three times as likely to arrival late as were counterparts who did pre-trip ATIS. Cases where pre-trip ATIS clearly benefits the user outweighed cases where pre-trip ATIS clearly disadvantages the user by 5:1. The number of late arrivals was reduced by 62.5 percent and total late schedule delay was reduced by 72 percent through pre-trip ATIS use.

Grotenhuis *et al.* (2007) carried out an online survey to study Integrated Multimodal Travel Information (IMTI) for public transport. He found that the pre-trip information is most favourite stage to collect travel information when planning multimodal travel; desired IMTI types in this stage are used to plan the part of the journey that is made by public transport. On the other hand wayside travel information is most desired when it helps the traveller to catch the right vehicle en route. Similarly on-board travellers are most concerned about timely arrival at interchanges in order to catch connecting modes.

Molin and Timmermans (2006) studied the relative importance that the travellers attach to different information types. The results indicated that real-time information is considered the most important because it has higher opportunity cost. The second most important attribute was found to be is the possibility of having different planning options available in the information system. Existing PT trip planners usually minimize travel time, but travellers value the possibility to search for options that are the cheapest or that exclude particular PT modes. The third most important attribute was found to be information about tickets.

2.1.5 Credibility

The credibility of the information source is found to be an important determinant which influences the travellers' decisions. Bonsall and Parry (1991) found that travellers tend to give less credence to received information than to evidence which they actually observe with their own eyes. Bonsall *et al.* (1991) also found that travellers prefer to test the credibility of the information randomly before considering using that source. Bonsall (1992a) reported that the influence of traffic information on route choice depended on whether the information was credible, relevant and clear. He founds that the factors like frequency of information, the degree of detail about the network, evidence it is based on, and its appropriateness. He also reported that the credibility of the information source heavily influences the extent of compliance with any guidance.

Dingus and Hulse (1993) pointed out that appropriate and timely information is an important factor to influence diversion decisions of the travellers. Shirazi *et al.* (1988) found that timely and accurate information, frequent reporting and better use of VMS signs are important determinants that improve the credibility and quality of information. Several other researchers (Polydoropulou & Ben-Akiva, 1998; Hato *et al.*, 1999; Fayish & Jovanis, 2004) also reported that reliability, timeliness and coverage of the information provided is crucial for the use of ATIS.

Zhao *et al.* (1995) found that high quality information improves the degree of compliance among the drivers. Several other studies (Bonsall and Joint 1991; Bonsall *et al.* 1994; Zhao *et al.* 1995) have found that travellers' consider the credibility of information important in their travel decisions. Hato *et al.* (1995) found that when the accuracy of the information is low, it causes negative effect on the perceived value of information, in order to influence drivers' route choice; he found that drivers required accurate and credible information. Wardman *et al.* (1997) reported that travellers were influenced more by their own observation from their windscreens as compared to the information provided by the Variable Message Signs (VMS).

2.2 Models of Response and of Network Wide Effect

The absence of information in the context of route choice models is modelled in a number of ways in the literature. These models have been developed to investigate the impacts of the information on network performance.

The absence or lack of information is often modelled by assigning informed and uninformed drivers using stochastic user equilibrium methods. This assumes that informed drivers have perfect knowledge of the link costs/time (small variance or zero variance) whereas uninformed drivers have higher link travel time (variance e.g. addition of error term). A number of studies have used this approach and have assigned different error terms for informed and uninformed drivers (Tsuiji *et al.*, 1985; Koutsopoulos and Lotan, 1989; Van Vuren and Watling, 1991). This approach has a number of limitations; the most important is the decision about the value of error term as if it is reduced to zero for the case of informed drivers it means that provision of information not only overcomes differences in knowledge but also that it overcomes differences in taste (Bonsall, 2007).

A second approach, used in some assignment models, is to represent via logit equations at each intersection (e.g. Dial, 1971). Here the value of λ could simply be decreased for drivers who have access to traffic information. This approach also has the same limitations as mentioned earlier.

The third approach could be to model the network performance on a given day by assuming informed drivers are assigned according to congested costs whereas uninformed drivers are assigned according to uncongested costs (Van Aerde *et al.*, 1989; Rakha *et al.*, 1989). This approach is far more ambitious and tends to overestimate benefits of providing information.

Another approach applied by Watling (1990), using the SATURN model assigns uninformed drivers according to a user equilibrium based on medium term conditions and informed drivers according to a user equilibrium based on actual conditions on a particular day. The approach was used to study the effects of a disturbance of a particular day in the network and demand matrix (Smith and Russam 1989; Watling 1990).

Mahmassani and Chang (1985) used a model of boundedly rational decision makers to study day to day dynamics of departure time of urban commuters. It was found that commuter behaviour could be assumed as boundedly rational and contained an indifference band of tolerable scheduled delay which could be then used to determine the acceptability of a particular departure time decision on any particular day. Mahmassani and Jayakrishnan (1989) assumed that drivers would try to keep their route same until the traffic conditions on a recommended route is perceived better than those on the original route. This reflects drivers' attitude towards the uncertainty in using new route. Jayakrishnan *et al.* (1994) developed a simulation assignment model to study effects of the information on the network called DYNASMART. In DYNASMART, uninformed drivers are assigned routes on the basis of equilibrium or they choose routes as were in their memory. The decision of informed drivers whether to change route after receiving information is based on the principle of bounded rationality. Here the travellers only switch when an alternative is perceived to provide gain in utility that exceeds a threshold level (Mahmassani and Chang, 1987).

Hounsell *et al.* (1995) used RGCONTRAM to study the driver responses to route guidance systems. The unfamiliar drivers choose routes based on perceived minimum distance or 'static' journey times if they are unguided but follow all credible route guidance received, whereas familiar drivers regularly reassess their routes, diverting if conditions justify it and following route guidance unless they perceive a better alternative.

Adler *et al.* (1993) applied another approach called conflict model. There are three primary factors in the conflict model. These are motivation improvement index, the threshold of conflict tolerance and the value of information. These directly influence route choice behaviour and real time information search and acquisition.

Prevedouros and Kasamoto (1998) studied the effects of incident management (IM) on Moanalua freeway by providing alternate route via Variable Message Sign (VMS). This research used INTEGRATION in order to simulate three scenarios: Existing or base case, Incident with no IM Program, and Incident with IM Program. This research concluded that provision of real-time driver information with an IM program contributed to a significant reduction of incident induced delays. The simulated incident increased total network travel times by 16% over base case. Real time information provided to motorists encouraged diversion onto the H-1 Freeway saving 40% over the scenario without IM.

2.3 Evaluation of the Benefits of Information Sources

From the available literature, it is evident that the benefits of ATIS are numerous. If implemented correctly, ATIS has the potential to mitigate the problems of congestion, environment and network performance by influencing travellers' route choice behaviour. Furthermore there is a significant willingness to pay for accurate and prescriptive ATIS information.

Literature on the evaluation of benefits from ATIS can be summarised as follows:

Wolinez *et al.* (2001) found from a survey in San Francisco, that overall willingness to pay for an ATIS was higher than expected from the literature. Sixty-six present of the respondents sought travel information and of these seekers 71% were willing to pay for an ATIS (average $3.34 per month, or $0.74 per call). Propensity for seeking travel information was significantly related to respondents, who were female, employed, took longer trips, faced unexpected congestion and owned a cellular phone.

Abdel-Aty (1998) used a nested model to represent incident related routing decisions. It was found that familiarity and usual use of alternate routes did not affect the decision in case of an incident. The model illustrated the significance of several socioeconomic, commuting and perceptual factors on the incident related routing decisions. A consistently significant factor was found to be traffic information acquisition. This proved that ATIS is likely to have a great impact on distributing traffic efficiently in case of incidents.

Lee (2000) estimated the benefit-cost ratio of the Washington State Department of Transportation websites that provided the current information about traffic conditions on freeways and bridges in the metropolitan Seattle area. The benefits considered in this study were time and cost savings to users as a result of informed travel choices, increased user confidence in travel choices, and reduction in congestion, pollution, and other external costs. Whereas the costs considered in this study were only the upgrade costs of website. The results obtained from this study indicated that Benefit/Cost ratio of the WSDOT website was 1.1.

Khattak *et al.* (1994) investigated the annual monetary benefits from ATIS-induced diversion in the Golden Gate Bridge corridor that range from $124 to $324 per person by using a stated preference questionnaire. In addition, this study also identified the benefits of ATIS into two groups. The first one was user benefits consisting of travel time saving, reduction of anxiety, enhancing ability to avoid congestion, improving ability to communicate during emergencies, and reduction possibility of getting lost. The other was system benefit including reductions in trip time, air pollution, and energy consumption, as well as greater safety.

Ullah *et al.* (1994) used a life-cycle cost model to evaluate a proposal for rural Advanced Traveller Information Systems (ATIS). The costs used in this model consisted of capital investment, operating, and maintenance costs. Although there were many types of benefits of ATIS such as reductions in congestion, pollution, energy consumption, and accidents, only potential reduction in the number of the occurring accidents were considered as benefits of this study to follow a conservative approach. Results indicated that both GPS and radio systems were cost effective (B/C Ratio > 1) when the accident reduction was as low as 2 percent for all types of accidents. What was more, sensitivity analysis of the cost variables also indicated that the NPV was still positive with more expensive hardware than what was used in this study.

Lappin (1996) synthesized customer satisfaction findings from ATIS research and evaluations in Seattle, San Antonio, and Phoenix dating from 1996. Evaluation findings suggested that customer demand for ATIS traffic services was based on four factors: (1) the regional traffic context (that include attributes of region, such as roadway network and capacity, level of traffic congestion, and future road expansion plan) (2) the quality of the ATIS services (that determine whether, how frequently, and with what level of confidence travellers consult traveller information) (3) the individual trip characteristics (such as the trip purpose, the time and the length of trip), and (4) the characteristics of the traveller (that include user value and attribute characteristics). Besides, it was also found that ATIS customers identified four primary benefits of service: saved time, avoided congestion, reduced stress, and avoided unsafe conditions.

Molin and Timmermans (2006) carried out an SP survey to study willingness to pay for Public Transport travel information. The results of the study indicated that even though public transport travel information is highly price sensitive; travellers are willing to pay for it if the information systems provide additional functionality such as real-time information and, to a lesser extent, additional trip planning options. It was found that travellers were more reluctant to pay for an Internet-based information service than for the telephone services that provided the same functionality. Real time information is the most valued additional feature followed by additional planning options, ticket information and timetable schedules.

Wardman et al. (2001) report that passengers value real-time information at interchange terminals as equal to 1.4 min in-vehicle-time. This measure can be recalculated into monetary willingness-to-pay (about 5 British Pence per journey).

Widlert *et al.* (1989) conducted an extensive stated preference study on Stockholmers' willingness-to-pay for public transport. It showed a significant willingness-to-pay for real-time information at bus and metro stops. Systems including real-time information signs were shown to the respondents, among other potential improvements relevant to public transport users. When choosing between, for example, certain amounts of shorter travel time, reduced ticket prices and other factors, real-time information was traded off as equal to about 12–16% lower fares or 6–8% shorter travel times

2.4 Theories of Information Search and Use

2.4.1 Maximization Concept

The microeconomic theory proposed by Samuelson (1947) has remained a dominant concept in understanding travellers' decision strategies. This theory assumes the individual is a rational decision maker who performs complete assessment of alternatives, exploring each alternative's relevant attributes and selecting an alternative that offers him maximum utility. The decision strategy serves to generate a choice from a choice set for the alternative that provides the individual with a maximum pay-off. In the presence of uncertainty the choice made from the choice set is mostly defined though the concept of maximum expected utility (Von Neumann & Morgenstern, 1947). There are some times costs involved during the process of decision making so-called transaction costs (Coase, 1937; 1960). These are costs involved in the decision making which are incurred before the actual choice.

Most of the studies in the travel demand modelling literature adopt this theory of expected utility. This framework of utility maximisation only involves the attributes of the available alternatives to the decision maker and hence the alternative not in the choice set of the decision maker are not considered at all (Chorus *et al.* 2006c). Often, it is simply assumed that the decision-maker knows all the alternatives available to him before making a decision. Microeconomic search theory has footings on the principle that there is a time investment disutility to discover acceptable opportunities of matches and therefore an individual performs a sequential search for alternatives and in the process of assessment of alternatives the new alternatives are generated (e.g. Weibull, 1978; Richardson, 1982). In such a sequential search process, apart from transaction costs, the cost of rejecting the most recently searched-for alternative can also be included. Such costs may be traded off against the expected utility to be derived from the next found alternative. After the generation of these alternatives, their attributes are than assessed under the principles of utility maximization. Several notions of microeconomic search theory can be found in travel demand studies. For examples see Williams & Ortuzar (1982), Richardson (1982), Lerman & Mahmassani (1985), Swait & Ben-Akiva (1987), Polak & Jones (1993), Ben-Akiva & Boccara (1995), Arentze & Timmermans (2005a, b). There is one important difference between the application of utility maximization principles for alternative assessment and for alternative generation: utility maximization for alternative assessment deals with choosing from alternatives, while its application on alternative generation also deals with choosing from decision strategies (should one proceed or stop searching?). Choosing a search strategy by applying utility maximization principles, the individual may well end up with an alternative having sub-optimal utility because the costs of searching are also taken into account in his decision strategy. Utility maximization principles are thus applied at different levels.

Although the application of principles of utility maximization has provided many valuable contributions to the research on individual choice (e.g. McFadden, 1974), as well as travel choice (e.g. Ben-Akiva & Lerman, 1985), researchers in general agree that its assumption of trade-off and maximization behaviour may form a less realistic representation of the actual behavioural process the individual performs (e.g. Edwards, 1954, Simon, 1955, 1978a, 1978b; Kahneman & Tversky, 1979, 1992; Hargreaves Heap *et al.*, 1992; McFadden, 1999).

2.4.2 Satisficing Concept

Herbert Simon was unhappy with the assumptions and validity of the maximization theory as a human decision making strategy, and so formulated the bounded rationality assumption (Simon, 1955, 1978a, 1978b). He assumed that human beings cannot be assumed to perform exhaustive assessment of alternatives and to thoroughly assess the alternatives found. A perspective on decision strategies that does not make such strict assumptions is that of satisficing behaviour (Simon, 1955); the individual continue a search process until he finds a satisfactory alternative that is good enough. This assumes that he has certain boundaries or aspiration levels for the relevant attributes of alternatives (which may change with time) and continue to search until he finds an alternative satisfying his standards (Olander, 1973). Bounded rationality is often called procedural rationality as in this the decision making is performed using simple 'rules of thumb' (e.g. Hey 1982, Johnson & Raab, 2003). This means a pre-decisional information search is performed to end up with an alternative of which relevant attributes meet the aspiration level set for that attribute. Notions of bounded rationality and satisficing behaviour can be found in several travel demand studies (e.g. Foerster, 1978; Mahmassani & Chang, 1987; Mahmassani & Jayakrishnan, 1991; Schofer *et al.*, 1993; Emmerink *et al.*, 1995a,b, 1996, Garling *et al.*, 2002).

2.4.3 Habit Execution

It is often argued and shown that many of the choices individuals repeatedly make are a consequence of the execution of a habit (e.g. Triandis, 1977, Hodgson, 2004). It is mere repetition of the previous decision without considering the currently available options. Since no actual decision is made in the sense of generating and assessing alternatives, habitual behaviour is often not regarded as decision-making. The decisions taken in this framework could have been optimal when first performed, but a sub-optimal situation may arise as changes concerning alternatives and situations are not observed by the individual, because he does not consciously make his decisions. Pre-decisional information acquisition is virtually non-existent in habitual behaviour (Aarts *et al.*, 1997, Verplanken *et al.*, 1997). A number of recent studies have shown the presence of this behaviour in the travel demand decisions, especially mode choices (e.g. Aarts *et al.*, 1997, 1998; Aarts & Dijksterhuis, 2000; Fujii *et al.*, 2001; Fujii & Kitamura, 2003; Fujii & Gärling, 2003; Schlich & Axhausen, 2003).

2.4.4 The Effort-Accuracy Trade-off

From the field of behavioural decision theory a perspective on decision-making originated that is often seen as an extension of bounded rationality, but that in fact incorporates several perspectives on choice-behaviour: the individual is assumed to select a decision strategy based on an effort/accuracy framework (Payne *et al.*, 1993, 1996). When choosing between alternatives, an individual first chooses a decision strategy based on (Chu & Spires, 2003) a trade-off of both the perceived effort and perceived accuracy of different decision strategies 'available' to him. Often, making decisions based on a careful trade-off of utility derived from attributes of alternatives (i.e. compensatory strategies) is not the selected decision strategy, and non-compensatory strategies such as satisficing (Simon, 1955) or some variant of lexicographic choice (Tversky, 1972) are performed instead. Only when there exists a need for and a possibility of achieving highly accurate choice-outcomes, will the costs of extensive search for and use of information be accepted by the decision-maker (see Huneke *et al.*, 2004, for a study on the effects of accountability on information search). In other cases, it is more likely that decision strategies are used that are only boundedly rational, including less extensive information search and use. Furthermore, different individuals facing the same choice-situation may perform different strategies. In recent travel demand research, explicit notions of this framework are not very widespread (for examples of the application of this framework see e.g. Gärling *et al.*, 1998, 2001, 2002; Svenson, 1998; Fujii & Gärling, 2003). Note that this effort-accuracy framework implicitly deals only with the assessment of already available or specified alternatives (Swait & Adamowicz, 2001) and not with alternative generation, while both these processes should be taken into account when studying decision-making and information acquisition (Smith, 1991; Posavac *et al.*, 2003).

2.4.5 Random Regret Minimisation

The Random Regret Minimisation (RRM) model postulates that people aim to minimise their regret with respect to the foregone alternatives (Chorus, 2009). The regret based models are based on the theory that individuals minimise anticipated regret when given a choice instead of maximising anticipated utility. The RRM estimation procedures (Chorus, 2009) assume that the regret is experienced with respect to all foregone alternatives that perform better than a chosen/intended alternative in terms of one or more alternatives unlike previous assumptions that regret is only experienced with respect to the best of foregone alternatives. In the cases tested so far, the specification produces intuitive estimation outcomes and satisfactory fit with available data (Chorus *et al.*, 2008, 2009).

2.4.6 Search Theory in Labour Economics

A key application of information search theory is in labour economics in the context of job search models in labour markets. There are some similarities of these models with the information search process of the travellers. There is large volume of literature in the field of job search models of labour economics but here we will try to discuss its relevance within information acquisition theory in transportation. Like the travel environment, the labour market is also noisy with a lot of uncertainty and incomplete information. Imperfect information requires investment of time to acquire information about opportunities and thus to unemployment (Hicks, 1932). Search theory on labour economics has footings in the principle that there is a time investment required to discover acceptable opportunities of matches in the labour market.

The simple sequential job search model can be described as follows. A jobless person looks for job opportunities each period. A job offer is a job opportunity at some fixed wage rate. The person has knowledge of the distribution of wage rates and a job offer is an independent draw from this distribution. The person has a general idea about the wage distribution in the area but he doesn't know which firm offers which wage rate. The distribution depends upon the occupation, industry and geographic location. The person's objective is to maximise his present value of income. The person adopts an optimal policy called reservation wage policy; this is a lowest acceptable wage offer below which the person rejects the offers. This reservation wage depends upon the worker's unemployment income and job search costs. Thus in the market there is a wage offer that equates the value of the continued search and the value of employment. This wage that equates the marginal costs and expected marginal benefits of continued search is the reservation wage. When the offer distribution is not known the reservation wage policy may not be optimal (Devine *et al.* 1993). An assumption can be that workers know which firms typically offer the best jobs. In this case the person will sample the firms in some order depending on the associated offer distributions. This search can be termed as systematic search, whereas, the other strategy could be random search. Search strategies may also be non-sequential; workers may choose the optimal number of applications to file each period, as well as the sequence of reservation wages.

Eckstein *et al* (2007) describes a classical job search model of labour economics. There is a measure M_i of workers of type i, who face a mortality rate η. In each period there are ηM_i newborn workers. Layoffs occur at a rate ς. When a worker of type i with skills x_i is unemployed she/he receives a flow of leisure and, possibly, unemployment benefits. In addition, there is a flow cost of search effort when the worker is unemployed. The sum of these two flows is equal to the worker utility flow in unemployment, which we denote by u_{it}^{u}. This notation highlights that all elements in this sum may change over time. When a worker of type i with skills x_i is employed she/he receives a flow of earnings of w_{it}. This minus the flow cost of search effort equals the instantaneous utility in the state of employment $u^{e}_{it}(w_{it})$. Let $d_{it} = 1$ if the individual is working and $d_{it} = 0$ if the individual is unemployed. The worker aims to maximize the expected lifetime utility. At $t = 0$, in the state of unemployment, this present value can be expressed as,

$$V_0^u = E\left[\int_0^\infty e^{-(\mu+\rho)t}(u_{it}^e(w_{it})d_{it} + u_{it}^u(1-d_{it}))dt \middle| \Omega_0\right]$$

where, the expectations are taken over all the random variables in the model. Ω_0 is the information available to the worker at time $t = 0$, and the ρ is the discount rate.

If we consider now the information search process in the travel environment the above discussed model has some similarities. Here too a traveller looks for the information about his travel options in a period of time. The person generally has some ideas about their normal travel pattern and associated parameters. The person also knows different information providers but does not know which provider provides his required information. Here too person tries to minimize his cost. There is a reservation cost which, in this case too, is less than the previous normal travel cost of the traveller. The person will continue to search for the appropriate cost minimization until the cost of search and the cost of the travel is less than the present situation. The search can be sequential as the traveller knows about some information provider being credible and accessible. On the other hand the search can be random and first optimal information stops the search process.

2.4.7 Optimal Search Theory

In optimal search theory, the relevant attributes of goods usually reduce to a single dimension i.e. price of the good sought, in our case travel information. However, prices can be viewed as hedonic and therefore summaries of more complex criteria. A probability density function (PDF) represents the searcher's imperfect information about the travel market. This dispersion of price offerings among information providers in the transport market, or at a single source, is both a measure and a manifestation of the traveller's imperfect knowledge (Stigler 1961). The parameters of the PDF(s) can be known to the searcher, or can be imperfectly known and updated through the information provided by the observed prices. In the latter case, a Bayesian updating mechanism represents the searcher's learning process. The fundamental decision mechanism in optimal search theory is the standard cost-benefit framework. Decisions involve a comparison of (expected) marginal return versus marginal cost. An early expression of this idea is the fixed *sample size* strategy in which a single decision occurs prior to search regarding how many alternatives to inspect. The search process itself simply carries out this fixed strategy, with the best of the inspected alternatives chosen for purchase (Stigler 1961). However, it is easily demonstrated that this strategy is not optimal and even nonsensical under certain conditions (see Gaswirth 1976; McCall 1965; Nelson 1970; Rothschild 1973). More attention has focused on sequential ***search*** models in which the searcher decides whether to continue search after each inspection. With sequential sampling, the cost-benefit mechanism produces a *reservation price policy* (Ratchford 1982). The reservation price is an acceptance threshold such that any price observed below this level will be accepted and search will terminate. Generally, travellers must directly inspect alternatives in order to obtain perfect knowledge of their price offerings and to acquire the information. Thus the search mode assumed by the optimal theory is most consistent with a comparison shopping.

2.5 Summary

The literature review presented above suggests that there are abundant sources of information available nowadays to a traveller and that the travel environment is information rich. The literature suggests that travellers do need information about their travel in order to plan their journeys. In addition to the travel time and travel cost for different alternatives, travellers also consider safety, reliability, security and convenience important factors for selecting different alternatives. The literature also suggests that age, sex, income level and education are key factors that influence the use of traveller information.

The studies carried out so far generally deal with only a few information sources and the effect of a rather powerful and common source, the internet, is ignored. There have been attempts to study information gain in an ICT rich environment. However, there are certain gaps in the study of all sources considered in choice set not just partial studies on the determinants of the some en-route or pre-trip sources. Furthermore most of the studies were inclined towards the linking the above mentioned determinants with the route choice decisions.

It is surprising to see that literature has few studies that comprehensively link the various sources of information with travellers' mode choice decisions. There are numerous sources of information which provide information about the different modes of travel and it is important to study the influence of these information sources on travellers' mode choice decisions under various circumstances.

Researchers in the past have done extensive research about the influence of different message contents upon travel decisions. However, the new information sources (internet, portable phones etc) have different attributes including their content, their special design, the presence advertisements, and the use of graphics which may influence the traveller and his travel decisions. If the effects of these elements can be established, it may be possible to increase the influence of these sources on overall modal shift. This issue has been rather ignored in previous research.

The issue of credibility as mentioned in the literature review section was dealt by a number of researchers. But the focus has been on the reliability of the information from the single source under study (e.g. in-vehicle guidance or VMS). There have been few attempts to consider the variation of the credibility from one source to another. Moreover, there is more to be learned about travellers' treatment of sources that have low credibility and about the factors which affect the credibility of sources.

Chapter 3
Methodological Issues

3.1 Issues to be Studied

In the light of the discussion presented in the previous chapter, it is evident that information is required by travellers to make informed decisions about their travel. Not only this, provision of information actually influences people's decisions by informing them about various less tangible characteristics of different alternatives. The previous chapter also explained in detail about the important variables which should be included in any analysis/model in order to comprehend the importance of information, travel decisions and their inter dependence. These variables could be grouped in four categories: variables associated to decision makers' characteristics, attributes associated with the information sources, characteristics of the travel options, and other external circumstances.

In order to understand the importance of these different variables, it is necessary to understand the information search and travel decision process together with its logical flow. The process of information acquisition for a traveller starts with her intention to travel. If, as an ideal case, the traveller has prior knowledge and experience, she is aware of every alternative and their attributes she may not decide to seek any further information. Similarly, if she is satisfied with the available information, she will put no value on obtaining additional information. In either case, the search process will not start at all and she decides about the travel alternatives straightaway. However, the decision to acquire information also depends on external circumstances, for example if she thinks that network is not congested at all she may not see any reason to explore available information, but if the weather conditions are poor, she may think it wise to seek further information. Secondly, it depends upon the individual traveller's preferences and attitudes. For example, if she is habitually uses a particular route, this habit will influence her to put in less efforts of obtaining information. Thirdly it also depends upon personal circumstances and individual preferences/attitudes about the traffic and driving tasks, all of which may vary with personal characteristics.

The above mentioned ideal situation is unrealistic; one deviation could be that the traveller is not aware of all available travel alternatives and hence if she decides to acquire information the aim of that information will be alternative generation, similarly if a traveller feels that she has incomplete information about the characteristics of the available alternatives, she may desire additional information about those alternatives. Here it must be noted that the randomness is not only in the environment but also there is uncertainty in the perception of the traveller about the non-stationary environment.

Once the traveller decides to acquire information and recognizes its value, she tries to search available potential sources of the information. From these sources she will decide to acquire information from a particular source depending upon awareness, its credibility, accessibility, information source attributes, and her own degree of understanding of the available sources. Awareness corresponds to the possibility of knowing the presence of a particular source and degree of success in understanding the information offered by that source. The traveller considers the information credible if it is provided by a trusted and reliable source, her experience with that particular source, her assessment of the likelihood of it meeting her needs, her trust in the motives of the information provider and the degree of its authentication by other reliable sources. The accessibility of a source depends on its cost and the ease with which it can be accessed and used. After gathering new information from a selected source the traveller tries to understand the acquired knowledge and rethinks the credibility of the new information and processes it as per her needs. She also synthesizes the new information with her beliefs and experiences prior to acquisition. If the acquired information satisfies her aspiration she will use this information in her travel decisions. On the other hand, if she is happy with the source of information but still require more information; she will try to acquire it from that source if the cost of acquisition is less than the disutility of delay and uncertainty. Alternatively, if she is not happy with the information source, she will select another source and will acquire information from it.

Finally, the decision is made on the basis of the acquired information and perceived modal and trip attributes. If user is satisfied with the decision, she forms beliefs and add/update her experiences about her travel and this particular information source. The process of information acquisition is summarised in Figure 3.1.

From the understanding of the above conceptual framework and in the light of the literature review presented in the previous chapter, it can be deduced that the decision to acquire travel information depends on the external circumstances e.g. bad weather,

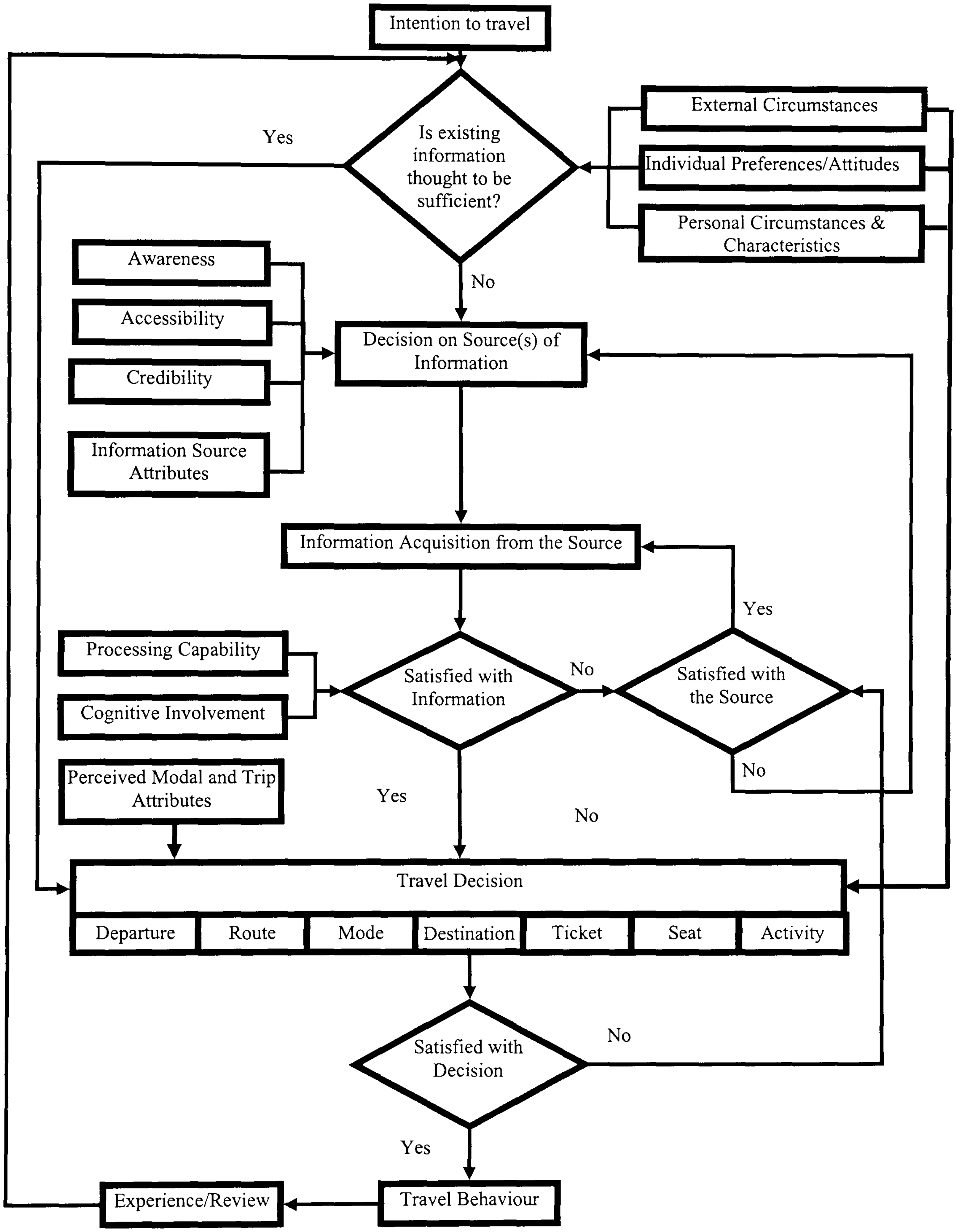

Figure 3.1 Conceptual Representations of the Information Search and Travel Choice Process

congestion, incident etc; personal attitudes and preferences; and personal circumstances. Similarly, once a decision is made to acquire information, the choice of source depends upon its accessibility and credibility, the individuals' awareness of it and information source characteristics. The attributes are thus classified in three categories: information source attributes; respondent characteristics; and scenario attributes.

The information source attributes are source owner, search time (time spent on search for required information), frequency of updating of information, type of information presented, presence of advertisements, type of information source, presentation of information (attractiveness of design of webpage), capabilities (how many probes used, CCTV camera installed, etc), price and subscription, spending on advertisements, coverage in search engines, year of start, coverage in newspapers & articles. The scenario attributes include trip purpose, bad weather, congestion, incident occurrence, and accompanying travellers. The personal characteristics include personal segmentation, socio-economic characteristics, trip frequency, travel time, frequency of using source, attitudes to optimise the journey, habit and other individual characteristics.

In addition to the above mentioned variables associated with the information acquisition process, the literature also suggests that there are characteristics which are associated with the travel alternatives. These include travel time, travel cost for different alternatives, safety, reliability, security and convenience. The details of these variables and their justification of inclusion are also discussed in detail in the next chapter.

The travel decisions as mentioned in figure 3.1 could be various i.e. choice of departure time, route, mode destinations, purchase of tickets, seat reservations and even cancellation/amendment of the activity itself. The study of all these travel decisions in one framework is quite difficult if not impossible and requires extensive data and resources. Hence, in this study, only mode choice decisions are taken into account.

In the light of the above discussion, the following research questions emerge to aid better understanding of the use of present day information sources.

- How does initial information effect subsequent search?
- Does credibility vary from source to source?

- How is information with low credibility treated?
- What affects credibility?
- Does presentation of information affect choice behaviour?
- What is the influence of information sources on the mode choice decisions?
- Does the design and content of an information source affect choice decisions?

The above research questions lead to the formulation of the following research hypotheses.

- Hypothesis 1: The presentation, design, credibility, capability and cost affect the selection of the information sources.
- Hypothesis 2: The influence of information on mode choice decisions is affected by the characteristics of the information source.

In the proposed research, an attempt will be made to develop models of the use of the predominant sources of information available in the travel related market and the influence of these sources on the selection of mode.

3.2 Data Sources

3.2.1 Choice of Data Source

The task of modelling and forecasting use of information sources and their effect on mode choice is a complicated process and requires the use of variety of data sets. These could include published data, actual observation, Revealed Preference (RP) data, Stated Preference (SP) data, and focus groups.

Published data can be used to provide background information on travel patterns and on the travelling population. As such it can be used to contextualise and scale a model built on a restricted dataset. The National Travel Survey (NTS)[1] and census data for 2001 can provide demographic information about specific populations and about long distance travellers in the UK. Data from these sources will be used in the current study. Data

[1] The National Travel Survey is a household survey of travel covering residents in Great Britain conducted by the office for National Statistics on behalf of the Department for Transport.

from the internet and from various travel planning websites can provide information about currently available information sources. This too will be used in the current study.

Observation of actual behaviour of travellers planning their journeys could form a useful part of any study. In the current context, data on travellers' search patterns while planning their journey on internet/travel planning websites could be a very good source of information. This approach would require data from website owners on the usage of their websites. The problem with this approach is that it could compromise the travellers' privacy and, moreover, website owners like to keep their data secret due to the competition from other similar websites and also due to legal privacy requirements. Thus although desirable, this data source is unlikely to be a practical proposition.

Revealed Preference (RP) data obtained via questionnaires could be very useful in exploring travel behaviour of a traveller under present conditions and his use of travel information. A detailed account of a journey could be explored by asking numerous questions regarding the previous travel decisions including frequency of travel, purpose of visits and chosen mode. In order to establish the relevant choice set, questions can be asked regarding respondents' perception of attributes of different modes available to them when they were making travel decisions. Moreover RP questions, being based on past travel decisions, are liable to avoid some of the biases inherent in the SP data. RP data, collected via a questionnaire will be used in the current study.

Stated Preference (SP) techniques are widely used to study the travellers' decisions and preferences. Usually, SP experiments offer a decision maker hypothetical scenarios in which a number of alternatives are described in terms of attributes with different attribute levels. The decision maker is then usually asked to choose which alternative they prefer. The preferences expressed indicate the relative importance of the attributes that characterise each option. A major advantage of SP data over RP data is that it allows the researcher to explore the effects of key variables in an efficient manner. An SP exercise will be used within the current study.

Focus Groups are often used to explore travel choices and preferences. In a focus group, questions are normally asked in a group setting where participants are free to talk with other group members. The main problems with this approach in the current context are that it requires specialist interview skills which the current author does not have and that it is not an efficient way to collect the quantitative data required for modelling. Also, compared to an RP or SP interview, it offers less control over the issues raised and thus a lot of time and resources may be wasted on the issues irrelevant to the study. Although focus groups could have been used to help formulate models and hypotheses, the author's lack of the necessary skills, and his intention to focus his available resources on the collection of quantitative data, resulted in the decision not to us this data source for the current study.

3.2.2 Choice of Mode of Data Collection

As noted in the previous section, it was decided to rely primarily on the data from new RP and SP surveys. By combining both approaches within one questionnaire, it was thought that the relative advantages of the two techniques could be exploited and the weaknesses of each technique could be overcome.

The use of a paper questionnaire was considered but it was concluded that this was impractical as it would become too long and could not easily use complex branching/routing for different segments of the respondents. In order to better comprehend the choice behaviour, the questionnaire needs to be interactive such that the appearance of a question sometimes depends on the previous response of the respondent. Hence in this research it is decided to use the Computer Assisted Telephonic Interview (CATI) technique to facilitate branching & routing. Another method could be online or internet based surveys but the problem with those surveys is that they only include those respondents which are already familiar with the internet or internet based sources and would result in a biased sample. Furthermore, in the online surveys, the respondents are more likely to make mistakes in the absence of any guidance particularly while dealing with more complicated and interactive questionnaires. Whereas in the CATI, the presence of an interviewer can reduce errors and the chance of the loss of a valuable respondent is minimised. There are many commercial software packages available nowadays which create CATI surveys easily. Hence in this research, a commercially available software package was used to create the CATI survey.

3.2.3 Data Sources for the Specific Issues Under Study

Table 3.1 indicates how the different data sources used in this study will supply data on issues identified in section 3.1.

3.2.4 Target Population

Use of travel information is nowadays very common for day to day journeys and, more particularly, for unknown and long distance travel. The reason for this is self explanatory as the long distance journeys are not as common as commuting and hence require information before actual travel decisions. The long distance journeys also require a decision on mode choice which, in the author's opinion, is not common in most short distance journeys. Hence in this study the use of information before a long distance journey was selected to be examined which would not only give valuable insights on the information use but also on subsequent mode choice under the influence of this acquired information. Moreover, this approach would also require a smaller sample size than would be required for short journeys, because in the latter case, the recruited respondents might not use information sources or make mode choice decisions. This means that the target population for this study is the travellers making long distance journeys (over 50 miles). Given this target population, it seems appropriate to recruit respondents at long distance transport interchanges (i.e. at train and coach terminals).

3.2.5 Sample Size and Sampling Strategy

Ortuzar and Willumsen (1996) recommend that the sample size should not be less than 250 observations. Sudam (1976) suggested that sample should be large enough so that when it is divided into groups each group will have a minimum sample size of 100.

A stratified random sampling approach, to represent the target population, is to be used in this research because, according to the literature, it offers the best means of securing sufficient respondents for all the important segments. Non-random sampling techniques offer the potential of both dramatically reducing the cost of data collection and increasing the precision of parameter estimates (everything else remaining same).

Table 3.1 Sources of Data on Specific Issues

Issues	Data Source
External Circumstances	
Bad weather	Questionnaire RP
Unexpected congestion	Questionnaire RP
Recurring congestion	Questionnaire RP
Visibility of Source	
Spending on advertisement	Original research
Coverage in search engines	Internet
Year of Start	Original research
Coverage in newspapers, articles etc	Original research
Order in the search engine	Internet
Available on Roadside/internet	Internet
Price	Internet
Marketing deals available	Internet
Subscription required or not	Internet
Attractiveness of design of webpage	Questionnaire SP
Graphical representation of attributes in real time	Internet
Characteristic of Data Source including Credibility	
Government or private service	Original research+ Questionnaire SP
Supporting data (delay, travel time)	Internet + Questionnaire SP
Real Time or static	Internet + Questionnaire SP
Type of information (Quantitative, Prescriptive)	Internet + Questionnaire SP
Capability (how many probes are used, CCTV camera)	Original research
Time taken to detect an incident	Internet
Number of irrelevant advertisement on site	Internet +Questionnaire SP
How long the source is being used or in practice	Original research
Trip Characteristics	
Out of vehicle time	Questionnaire RP, SP
In vehicle time	Questionnaire RP. SP
Cost of travel	Questionnaire RP. SP
Car ownership	Questionnaire RP
Trip purpose	Questionnaire RP
Safety of travel	Questionnaire RP
Comfort of travel	Questionnaire RP
Availability of seat on Public Transport	Questionnaire RP
Departure Time	Questionnaire RP
Frequency of travel	Questionnaire RP
Personal Characteristics	
Sex	Questionnaire RP
Age	Questionnaire RP
Income	Questionnaire RP
Occupation	Questionnaire RP
Education Level	Questionnaire RP
Personal Attitude/Preferences/Circumstances	
Habit of continuing same mode	Questionnaire RP
Attitude to optimize the journey	Questionnaire RP
Give due consideration to cost	Questionnaire RP
Frequency of long journeys	Questionnaire RP
Awareness of specific sources	Questionnaire RP
Already user of information sources	Questionnaire RP
Familiar with specific websites	Questionnaire RP
Frequency of Novel journeys	Questionnaire RP
Preparedness to spend on an information source	Questionnaire RP
Attitude for having marketing deals	Questionnaire RP
Attitude to like graphical representation of attributes	Questionnaire RP
Attitude to like more supporting information	Questionnaire RP
Attitude to like real time information	Questionnaire RP
Attitude to like resourcefulness of a source	Questionnaire RP
Attitude to irrelevant information	Questionnaire RP

However, the use of non-random sampling will require careful thought about how the resulting data should be used (Ben-Akiva and Lerman, 1985).

3.3 Analytical Issues

This aim of this section is to discuss a selection of models that will be used to explain the use of the predominant sources of information available in the travel related market and the influence of these sources on the selection of mode.

3.3.1 Random Utility Theory

In the discrete choice analysis based on random utility theory, an individual's choice is assumed to depend on 'utility' representing the satisfaction or benefits to the person from each alternative. If individuals act rationally, they are assumed to always choose the option with the highest utility to them. In this research maximization principle has been selected for analysis of the study. The reason for this selection is based on the strong footings of this theory's acceptability in wide range of disciplines. Furthermore, Utility maximization theory is easier to implement in the current context as other methods would complicate an already complicated model structure. Utility maximization provides a way by which choice probabilities can be estimated. Discrete choice models are based on choices made by individuals. They assume that the probability of an individual choosing a given alternative is the function of their socio-economic characteristics and the relative attractiveness of the options (Ortuzar and Willumsen 1994). The effect of these sources of randomness was formalized by Manski (1977) and incorporated within the utility of a given alternative i so that the *random utility* (U_i) is the sum of the observable (or systematic) component (V_i) and unobservable (or random) component (ε_i):

$$U_{in} = V_{in} + \varepsilon_{in}$$

The deterministic component, V_i is a function of observable and measureable attributes x_i. The function is commonly assumed to be linear in combining variables.

$$V_i = \sum_k \theta_{ik} x_{ik}$$

where θ_{ik} is a vector of parameters to be estimated and x_{ik} is a vector of observed data relating to alternative i.

3.3.2 Logit Models

The Multinomial Logit (MNL) and Nested Logit (NL) can be derived from Mc Fadden's (1978) Generalised Extreme Value (GEV) model, which itself can be derived from a model of Random Utility Maximisation (RUM). MNL is derived from the assumption that errors of the utility function are identically and independently Gumbel distributed. However, the main drawback of model is that it assumes that the choice options are independent and therefore fails to take account of the correlation between the alternatives. This property is known as independence from irrelevant alternatives (IIA). In MNL, the probability P_i of choosing alternative i from the choice set J, given measured utilities V_j $(j \in J)$, is given by,

$$P_i = \frac{e^{\mu V_i}}{\sum_{j \in J} e^{\mu V_j}}$$

Where J is the set of choice alternatives and μ is a strictly positive scale parameter and it is related to the variance σ^2 of the error term $\mu = \frac{\pi}{\sqrt{6}\sigma}$ and is usually normalised to be equal to one as it cannot be estimated separately from the coefficients (Ortuzar and Willumsen, 1994).

The most widely known relaxation of the MNL is the nested logit (NL) model. The model like MNL is based on the utility maximisation but allows fort the correlation between the pairs of alternatives in a common group (a nest). For a two level example the nested logit model shows the probability that alternative i is chosen as,

$$P_i = P_m P_{i|m}$$

Where P_m, the probability of choosing nest m is given by

$$P_m = \frac{\left(\sum_{j\in J_m} e^{V_j}\right)^{\theta_m}}{\sum_m \left(\sum_{j\in J_m} e^{V_j}\right)^{\theta_m}}$$

and $P_{i|m}$, the conditional probability of choosing alternative i from nest m is given by,

$$P_{(i|m)} = \frac{e^{V_i}}{\sum_{j\in J_m} e^{V_j}}$$

where N_m is the set of alternatives in nest m and θ_m is the dissimilarity parameter relating to nest m. NL is consistent with the random utility maximisation if $0 < \theta_m \leq 1$ for all θ_m.

Logit models were ultimately intended to measure the effect of information sources on the mode choice. The following is the stepwise model development process.

- Model for mode choice
- Model for source choice
- Combined model for source and mode

The Mixed Logit (MMNL) is neither GEV nor has a closed form; estimation difficulties restricted its application in the early years. More recently, development of the estimation methods, maximum simulation likelihood (Ben Akiva and Bolduc, 1991) has made this model attractive for estimations.
Brownstone *et al.* (2000) derived the utility function for alternative i as follows:

$$U_i = \beta x_i + [\eta_i + \epsilon_i]$$

Where x_i is a vector of observed variables relating to alternative i, β is a vector of structural parameters reflecting choices by the overall population, η_i is a random term with zero mean, the distribution of which varies across individuals and alternatives depends on underlying parameters and observed data relating to individuals and alternatives, and ϵ_i is a random term with zero mean that is IID across alternatives and does not depend on the underlying parameters nor data, and is normalized to set the scale of utility. Although mixed logit is not GEV, Mc Fadden and Train (2000) established that any discrete choice model from a RUM (Random Utility Maximising) model can be approximated by mixed logit.

In the context where a decision-maker makes many choices over a period of time (panel data or responses to the stated preference surveys), it is possible to accommodate the multi-period nature of the data by assuming that a respondent's tastes (η_n) do not change between choice situations. The conditional probability of the individual n's sequence of choices then becomes the product of logits:

$$S_n(\eta_n) = \prod_t F_{ni(n,t)t}(\eta_n)$$

Where i(n,t) is individual n's choice in period t.

The unconditional probability is given by:

$$P_n(\theta^*) = \int S_n(\eta_n) f(\eta_n \mid \theta^*) d\eta_n$$

Where θ^* are the parameters which describe the distribution of tastes $f(\eta_n \mid \theta^*)$.

3.3.3 Random Regret Models

The Random Regret Minimisation (RRM) are based on the theory that individuals minimise anticipated regret when given a choice instead of maximising anticipated utility. The RRM estimation procedures adopted in this section uses a specification suggested by Chorus (2009) which assumes that the regret is experienced with respect to all foregone alternatives that perform better than a chosen/intended alternative in terms of one or more alternatives unlike previous assumptions that regret is only experienced with respect to the best of foregone alternatives. The model states that a decision maker faces a set of L travel alternatives, each explained in terms of M attributes x_m that are comparable across alternatives. A decision maker would aim to minimise anticipated regret amongst the alternatives which is composed out of an IID random error (Extreme Value Type I-distributed with variance $\pi^2/6$) and a deterministic regret R. Deterministic regret is conceived to be maximum of all binary regrets associated with the comparison of the considered alternative with each of remaining alternatives (either zero or equal to the weighted difference in attribute performance. The deterministic regret associated with any alternative e.g. alternative 1 is written as,

$$R_1 = max_{l=2..L}\left\{\sum_{m=1..M} max\{0, \beta_m.(x_{lm} - x_{1m})\}\right\}$$

Assuming both terms in the max operator are stochastic and assuming that an IID random component ϵ is added, the expected maximum can then be written as:

$$E = (\max\{0 + \epsilon, \beta_m.(x_{lm} - x_{1m}) + \epsilon\})$$

$$= \ln(\exp[0] + \exp[\beta_m.(x_{lm} - x_{1m})])$$

$$= \ln(1 + \exp[\beta_m.(x_{lm} - x_{1m})])$$

Hence the deterministic regret associated with the alternative 1 can be written as

$$R = \sum_{l=2..L}\left(\sum_{m=1..M} [\ln(1 + \exp[\beta_m.(x_{lm} - x_{1m})])]\right)$$

3.3.4 Approach to be Adopted

Previous sections discussed the important issues that should be included in any study focused on the travel information acquisition and travel decisions. In particular, section 3.1 discussed the general conceptual framework of the information search process and subsequent travel decisions. Similarly, Table 3.1 discussed the determinants and their relationship for the relevant choices in general. The inclusion of all those determinants and choices in a single study is not practical and would be too ambitious. Hence in this study the choices are simplified and restricted to only two i.e. the choice of information source and the choice of mode. The selection of the determinants for these choices are also simplified and only the most relevant attributes are selected. Figure 3.2 presents the simplified modelling framework conducted in this thesis. The choices are divided in two different sections, namely choice of information source and choice of mode. The two choices are calibrated in separate models using different data sets and are combined at a later stage.

The first choice is the selection of the source of information. The determinants for this choice are classified in three categories: Personal Characteristics and Attitudes including experience of sources; Journey Characteristics; and Source attributes. This classification is consistent with table 3.1 and uses the same determinants/variables as mentioned. Figure 3.2 presents these independent variables in detail.

The second choice is the selection of mode for the travel after acquiring information from one or more sources. The determinants of this choice are also classified in three categories: Personal Characteristics and Attitudes including experience of sources; Journey Characteristics; and Modal attributes as presented by different information sources. The personal and journey characteristics are the same as in the source choice section whereas the modal attributes include the cost and time for a particular mode as described by different information sources. The restriction to only two attributes, cost and time, was carried out for econometric reasons and in order to simplify the Stated Preference (SP) exercise. The details of the SP design are discussed in the next chapter together with its justification. Various information sources were included in the questionnaire survey and only predominant sources were used in the choice set. Similarly only three modes, car coach and train, were selected for the mode choice exercise. Detailed discussion of these choices is included in Chapter 4.

In this study MNL and NL models will be calibrated from the data whereas, out of these models, only selected best performing models are estimated in the MMNL and RRM framework to check their robustness.

Variables encapsulating the issues summarised in Table 3.1 and in Figure 3.2 will be tested in the models. Details of the model development process are discussed in Chapter 6.

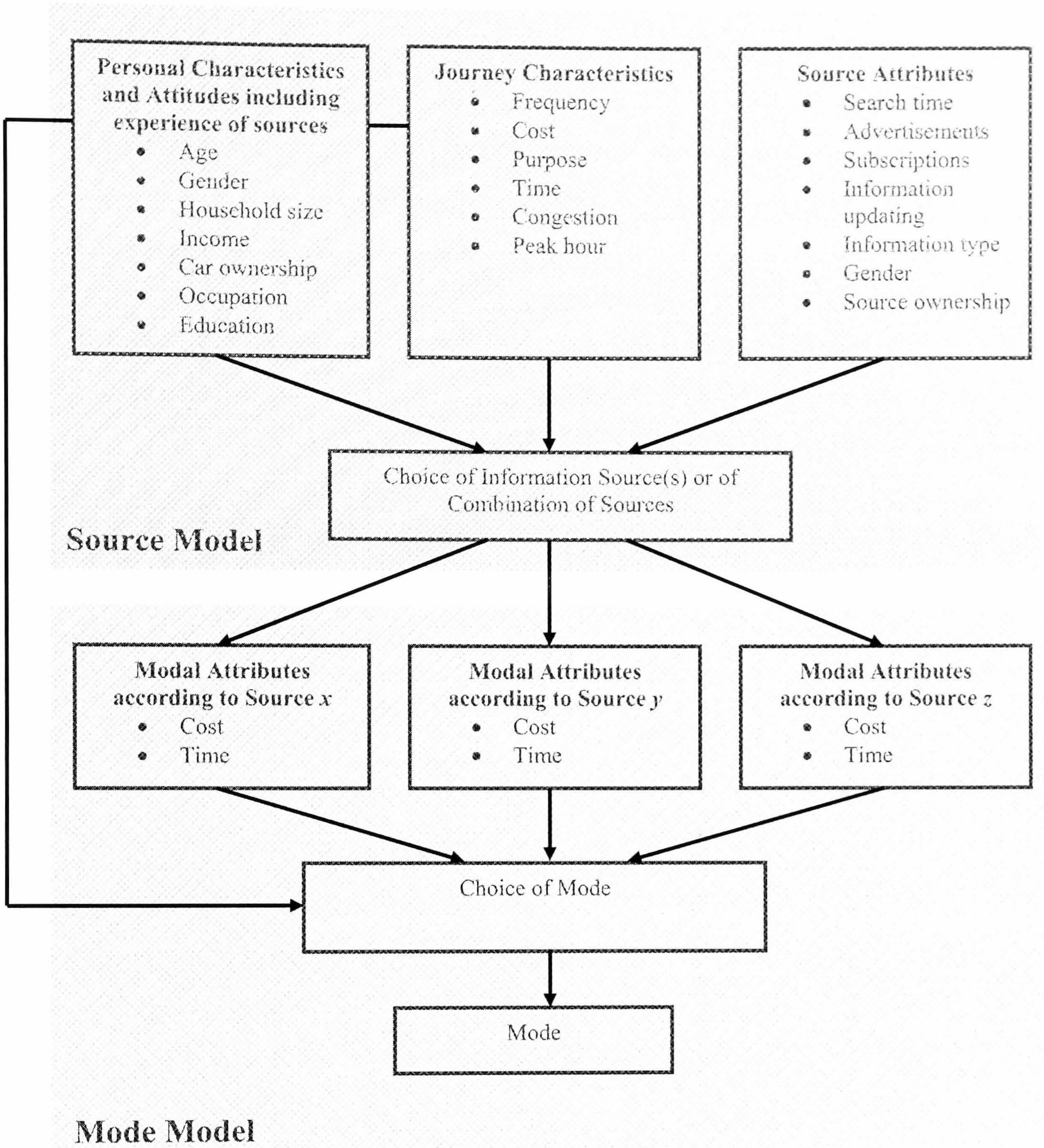

Figure 3.2 Framework for Modelling conducted in this Thesis

Chapter 4
Design of the Questionnaire

4.1 Structure of the Questionnaire

The main questionnaire developed in this study has four parts. The first part gathers Revealed Preference (RP) data for the last long journey (over 50 miles) made by the respondents. The second part includes SP survey questions to investigate travellers' choice of information sources and subsequent mode choice when making long journeys. The third part includes general questions about their attitudes towards different sources of information and on their normal search patterns. The final part contains questions about travellers' characteristics.

4.2 The RP Exercise

The Revealed Preference (RP) data consists of the exploration of the last long journey (over 50 miles) made by the respondent. A detailed account of the journey is sought by asking questions on their frequency of travel to those destinations, purpose of visits and the chosen mode. In order to establish the relevant choice set, questions are also asked regarding respondents' perception of attributes of different modes available to them when they were taking travel decisions. The RP part of the questionnaire also includes the questions about the external circumstances of the journey and about the use and effect of any information source while taking travel decisions about that journey.

4.3 The SP Exercise

4.3.1 The Principles of SP Design

An SP choice experiment consists of several choice sets, each containing two or more options (sometimes called alternatives). Respondents are shown the choice sets in turn and the response expressed indicates the relative importance of the attributes that characterise the hypothetical scenarios. The response could be a preference, rating or ranking of the alternatives. Each option is described by a set of attributes and each attribute can take one of several levels. These experiments are used in research to estimate the effect of the attributes on the attractiveness of the product under consideration.

An orthogonal method is usually applied in any SP experiment design. This assumes that there is zero correlation between the explanatory variables. The method varies the attributes, presented to respondents, independently from one another and this ensures no correlation between the coefficient estimates. Moreover, the coefficient estimates would remain the same no matter how many explanatory variables are included in the models. This makes it possible to include as many variables as possible in the models (Bates 1987; Hensher 1994) however; this would become too difficult for respondents and may overload them.

A full factorial design often has too many scenarios as the full factorial design means inclusion of all combinations of attribute levels. The total number of combinations is usually defined by the number of attributes and levels included in the SP design (Pearmain and Kroes 1990). This would make the SP survey very laborious and too long. Therefore it is useful to use a fractional factorial SP experiment design, which allows the uncorrelated estimation of all main effects under the assumption that all interactions are negligible (Addelman, 1962).

Hensher and Barnard (1990) illustrated the difficulty to retain design orthogonality when individual choice data are used to estimate discrete-choice models. Some authors have suggested that the attribute correlation problem can be circumvented by aggregating data over replications either within or across individuals, and analyzing choice frequencies (Louviere and Bunch, 1990; Van Berkum, 1987; Offen and Little, 1987), however recently some authors have used optimal orthogonal design in which every attribute level is different across alternatives. In this approach also termed as D-efficient, designs are usually constructed by algorithms that sequentially add and delete points from a potential design by using a candidate set of points spaced over the region of interest Kuhfeld *et al.* 1994, Heredia-Langner *et al.* 2003). D-efficient designs are promising, only when based on accurate prior parameter values. However, if good quality a priori information is lacking, then practitioners might be better off with shifted designs built from conventional fractional factorial designs for linear models (Ferrini and Scarpa, 2007).

4.3.2 Two Separate SP Exercises

There are two SP exercises in the second part of the questionnaire. One deals with the choice among information source with different attributes. A number of alternative scenarios varying the trip destination, purpose and accompanying individuals were tried first for the pilot survey and the results of this survey determined the selection of the most influential scenario(s). The detailed discussion about these scenarios is presented in Section 4.3.3. Fractional factorial design was be used and the fraction was selected to avoid dominance. The dominance refers to those combinations in which the respondent is better off or worse off on every dimension (Beesley & Hensher, 1987). Although it can be argued that the presence of such dominant combinations can be used to identify logical and non-logical responses, hence in this study, each SP exercise includes one dominant combination to check the issue of the consistency of the responses. It has been decided in this study to present a binary choice to the respondents in the first SP exercise regarding choice of information source. The reason for the use of binary choice here is that it is the simplest, relatively easy to implement and is easily comprehended by the respondents. The main disadvantage for the binary experiments is that they have limited use in modelling and understanding multiple choice behaviour. Their primary limitation is that they are difficult to generalize to multiple choices unless the choice process has a fairly simple structure. For example, Louviere and Woodworth (1983) showed that paired comparison experiments are significantly statistically inferior to experiments based on orthogonal fractions of 2N factorials (N = the number of choice alternatives), even in the case of simple choice models like Multinomial Logit. As described earlier, the questionnaire design, being complicated in some aspects, would in turn require simple choice exercises for the respondents to avoid overloading and hence binary choice experiments are selected here to tackle this issue despite the acknowledged drawbacks. As discussed in the previous section D-efficient designs are promising, only when based on accurate prior parameter values. As this study being novel, a good quality priori information of coefficients is lacking and the use of conventional fractional factorial designs is considered more promising.

The second SP exercise deals with the choice of mode by users and non users of the information sources. Here too a fractional factorial design is used and subjects are requested to choose between two alternatives, i.e. car and train, car and coach, or coach and train. This means each individual saw 5 choices between car and train and 4 choices between coach and train. The questionnaire also takes care of modes unavailable to a particular individual i.e. showing only those alternatives that he can use i.e. if car is unavailable to him, the questionnaire would show 9 choices between coach and train.

4.3.3 Choice of Scenarios for SP Questions

The aim of the survey is to investigate the effects of various information sources on mode choice. This research focuses on the role and use of information sources in long distance journeys. Hence in this research the scenario was selected for a long journey between the cities. The selection of the journey could be crucial for the mode selection. One option could be to present respondents with a hypothetical journey without mentioning the destination. The problem with this approach is that if the destination is not given, the use of information sources would also become unrealistic and there is a large possibility that the respondents would not truly sense the necessity of the information sources. Another option could be a journey from Leeds to London. This is a longer journey and requires acquisition of information for variety of conditions.

However, the problem with this scenario would have been that this journey is a very common journey and the choice of mode and information source would be highly influenced with the respondents' previous journey to London. In addition to this problem, the SP exercise for mode selection under influence of information would also be affected as the respondent will tend to select the choices based on previous selection. Hence in this design, a journey from Leeds to Cardiff is selected which avoids all of the above mentioned problems. The advantages of selecting Cardiff as destination are numerous, firstly in real life, there are several competing modal options available for this destination, secondly given that the interviews were carried out in Leeds, Cardiff is a relatively unknown destination which would most likely trigger the use of information sources for journey decisions, thirdly the city has leisure attraction which could be tried as a trip purpose for those respondents who are not employed or are not into any business e.g. students, housewife etc. Finally and most importantly this destination would not result in any predominated mode choice as in case of London for which the car mode is relatively unattractive.

Hence, the scenario, hence for both the SP exercises, stated that "*imagine that you need to travel from Leeds to Cardiff in 3 days time on personal business with an appointment from 1130 till 1230 in Cardiff city hall*". The scenario is not varied in the design instead the scenario attributes are explored by adding questions in the RP part of the questionnaire. The reason for keeping the scenario constant is that it would otherwise complicate an already complicated SP design and questionnaire. It would also tend to make the experiments and questionnaire lengthy and it would be difficult to keep the respondents' interest intact throughout the interview.

4.3.4 SP Exercise for Selection of Information Source

As explained in section 3.1, information acquisition is a complicated decision process that starts with the behavioural intent and stimulus which force someone to start the decision process. The decision to acquire travel information depends on external circumstances (e.g. bad weather, congestion, incident etc.); personal attitudes and preferences; and personal circumstances. Similarly, once a decision is made to acquire information, the choice of source depends upon its accessibility and credibility and the individuals' awareness of it. This theoretically requires the inclusion of more attributes in the SP exercises than would be practicable and easy for a good and comprehendible design.

The attributes are classified in three categories information source attributes; respondent characteristics; and scenario attributes. The information source attributes are source owner; search time (time spent on search for required information); frequency of updating of information; type of information presented; presence of advertisements; type of information source; presentation of information (attractiveness of design of webpage); capabilities (how many probes used, CCTV camera installed, etc); price and subscription; spending on advertisements; coverage in search engines; year of start; and coverage in newspapers & articles. The scenario attributes include trip purpose, bad weather, congestion, incident occurrence, and accompanying travellers. The personal characteristics include personal segmentation, socio-economic characteristics, trip frequency, travel time, frequency of using source, attitudes to optimise the journey, habit and other individual characteristics.

The first SP exercise deals with the attributes of the information sources and scenario attributes only. The selection of scenarios has been discussed in the previous section. From the above mentioned list of the information source attributes, the spending on advertisements, coverage in search engines, year of start, coverage in newspapers & articles are not included in the SP exercise and their values were directly included in the model from separate research and consultation of the literature. From the remaining attributes the capabilities and presentation of the information are not included in the SP experiment. The main reasons for not including the presentation of the information and attractiveness of the design (graphics) in the SP design are firstly, that the previous research has not shown this to be an important attribute, secondly that this attribute is also difficult to deal with in a SP exercise as it is difficult to present to the respondent without showing the actual designs of the prevalent information sources, and finally that such types of attributes are relative in nature and are difficult for respondent to understand. The capability of source is also not included in the SP exercise on the basis that people rarely know how the sources are getting the information and so it is slightly artificial to inform them about this. Hence the following attributes are included in the pilot survey: ownership of the source (government or private); type of information; search time; presence of advertisements; updating of information; and subscription cost. The final attributes of the SP design were decided after the execution of the pilot surveys.

A source of information could be either owned by government or private company. This may influence the credibility of the source and might be an important consideration in the selection of the sources. For type of information about travel time and delay, three levels are considered: Descriptive Quantitative real time information (e.g. specific information relevant to the journey about routes/services/prices), Descriptive Qualitative information (e.g. general information relevant to the journey about routes/services/prices), and Prescriptive information (e.g. specific information relevant to the journey based on users' criteria for routes/services/prices). For the search time three levels are considered, 5 min, 10 min and 15 min. These levels were decided on the argument that nobody would like to spend more than 15 minutes for search of the information source and will lose his interest on it. On the other hand, the minimum of 5 minutes are anticipated to be considered as a reasonable minimum search time which includes the typing and browsing time for getting the information. For the presence of advertisement there can only be two possibilities either the advertisements are there or they are not present. Hence two levels are considered, one for its presence and the other no advertisements. For frequency of information changes on the source again three levels are considered, real time, daily and weekly. Although other frequencies (e.g. hourly) could have been specified, the argument behind selecting these levels is that the difference between, say, hourly and real time, is not as great as that between the three frequencies identified. For subscription again there are two levels, no subscription and £5 already paid (it was deliberately not mentioned that whether this £5 subscription fee is for this particular trip or for some fixed time period i.e. 6 months or 1 year). The value of 5 pounds is selected here because any other value higher than this would be considered too high for an infrequent trip as mentioned in our scenario.

This means that the design must deal with 3 attributes with 2 levels and 3 attributes with 3 levels. A full factorial design would require 216 combinations of attribute levels. However, since the attributes are, arguably, not truly independent (e.g. advertising is less likely on a government site and on a subscription site than on a free commercial site), it is not necessary or appropriate to use the "full" design.

The attributes of owner of source, subscription and advertisements are therefore combined together as a one attribute with four levels:

- A government Source with no advertisements and no subscription
- A private source with advertisements and no subscription
- A private source with no advertisements but with subscription
- A private source with no advertisements and no subscription

The design thus has one attribute with 4 levels and three attributes with 3 levels. A full factorial design in this case would require 108 combinations of attribute levels. This number of combinations is still too high to be used and hence an orthogonal main effects fraction is used having 16 combinations. This means every respondent would have 8 binary choices. This allows the uncorrelated estimation of all the main effects under the assumption that all interactions are negligible.

The attributes and their levels are summarized in Table 4.1.

Table 4.1: Attributes and Levels for selection of Information Source

Attributes	Levels	
	Information Source A	**Information Source B**
Owner of Source	Private/Government	Private/Government
Subscription	No Subscription/£5 already paid	No Subscription/£5 already paid
Presence of Advertisements	No Advertisement/ Advertisements	No Advertisement/ Advertisements
Type of Information	Quantitative/Qualitative/Prescriptive	Quantitative/Qualitative/Prescriptive
Search Time	5 min/10 min/15 min	5 min/10 min/15 min
Up to date Information	Real time /Changes daily/Changes weekly	Real time /Changes daily/Changes weekly

4.3.5 SP Exercise for Selection of the Mode

The attributes of the SP design were selected considering those used and found significant in the relevant previous studies (Bonsall and Joint 1991; Khattak *et al.* 1993; Bonsall *et al.* 1994; Bonsall 1995; Bonsall and Merall 1995; Bonsall and Palmer 1995; Bonsall and Whelan 1995; Wardman *et al.* 1997). Travel times, levels of extra delay, levels of costs, and the available sources of information are selected to represent the respondent's choice situation. The sources of information were the same as those mentioned in the RP questions and their different combinations were used with each alternative.

One could have used the Adaptive Stated Preference (ASP) technique to better understand the choice of mode under the different levels of attributes as given by different sources. The initial level could have been used from the RP part of the survey and by considering the mode choice made at that point by the respondent the levels of the SP experiment could then have been altered by making the mode selected in the RP less attractive and the process goes on even in the SP experiment thus making it adaptive to the previous choice made. The problem with this approach is that it inherits in itself a number of issues and problems and requires careful design and extensive simulations to make the design robust (Bradley and Daly, 1993). Secondly this approach would have required the RP part to have similar journey as in the SP part of the survey (Leeds to Cardiff) which would not have been practicable in this case because it would then reduce the sample size considerably. Moreover artificially making chosen modes less attractive would in turn result in the less realistic levels for the respondents and make the survey artificial.

Hence in this study a traditional SP experiment is used with the careful selection of the realistic levels. In addition to this for each mode at least two information sources were used to overcome credibility bias which could affect the subsequent choices as the respondent after selecting an option on the first screen would tend to choose the same selection for the rest of screens. Each mode had at least two sources and time and cost attributes. The levels as shown in the table 4.2 are initial mean levels and are selected from the current travel time and costs for all the three modes as described by information sources during normal conditions. The other values were 20 % deviation from these average values.

Table 4.2: Attributes and initial mean levels for Selection of the Mode

Attributes	Levels		
	Car	**Coach**	**Train**
In vehicle Travel Time	5 hours including 1 hour delay	7 hours	3 hours 30 min
Return Out of Pocket Cost	£75	£60	£110

This design implies 2 variables with 3 levels for each of the mode. A full factorial design thus requires 9 combinations of attribute levels for each alternative. Table 4.3 shows the alternatives as shown to respondents. Each alternative has two columns of time and cost attributes.

Table 4.3: Choices as shown to the respondents in questionnaire

Information Source	Car		Coach	
	Friend	Multimodal Website	Recent experience	Coach Website
Time	About 4 hrs	5 hrs (including 20 min delay)	About 6 hrs (Scheduled station to station)	7 hrs (Scheduled station to station)
Out of pocket cost	About £60	£ 75	About £60	£50
Your Choice	( )		( )	

4.3.6 Simulation Tests for Attribute Levels

Before conducting the pilot survey, a simulation test was carried out to test the SP design. Fowkes and Wardman (1991) advise use of simulation to avoid inaccurate estimates of a series of relative values in the design of SP exercise. SP simulation tests involve the generation of the synthetic responses to a particular design and then estimating the utility function to check the efficiency of the design at extracting the specified parameters. The simulation test was repeated after improvement in the design following a pilot survey to check the ecological design of the survey. Logit models were estimated using the results of the pre pilot survey to check the appropriateness of the design and general understanding of the questionnaire by the respondents. The results of the pilot survey would be analysed to refine the questionnaire design. The pilot survey would also include general questions about the legibility and clarity of the questionnaire; appropriateness of the attribute levels; presentation of the SP experiments and whether it was realistic or not. Logit models would be estimated from the results of the pilot survey and the calibrated models would be checked statistically. The estimated parameters would be checked to have the correct sign and acceptable t- stats to the extent that this is possible in pilot.

As explained earlier there were three alternatives with two 2 variables on 3 levels for each of the mode. This gives us nine SP questions for each simulated respondent.

Three valuations were specified for the design. The values of the time were selected from the previous studies (WEBTAG). The selected values were -10 and -15 pence per minute. The simulations were repeated several times with the different combinations of the value of times and the attribute levels. Each run simulated the response from the 100 individuals to the nine questions making a sample size of 900 choices.

The simulation tests were repeated with and without alternate specific constants. The simulation test results are summarized in Table 4.4.

Table 4.4: Simulation Test Results for the Pilot Survey

Design	Variable	Information Source	Run	Set	Specified Values (pence/min)	Estimated Values (pence/min)	Difference	Ratio (%)
Design without Alternate Specific Constant	Car	Multimodal website	1	VOT	-15	-14.57	0.43	-2.87
		Friend			-15	-14.73	0.27	-1.80
	Coach	Coach website			-15	-25.04	10.04	66.93
		Multimodal website			-15	-15.73	0.73	4.87
	Train	Train website			-15	-3.07	11.93	-79.53
		Recent Experience			-15	-8.37	6.63	-44.20
Design with alternate Specific Constant	Car	Multimodal website	2	VOT	-15	-23.69	8.69	57.95
		Friend			-15	-41.13	26.13	174.23
	Coach	Coach website			-15	-34.16	19.15	127.72
		Multimodal website			-15	-6.03	8.97	-59.81
	Train	Train website			-15	-19.76	4.75	31.72
		Recent Experience			-15	-37.38	22.37	149.18

The 6th column indicates the artificial utility values specified by the designer and the 7th column shows the estimated values using the simulated responses. To show the relative deviations between the specified values and estimated values, the last column summarises the calculated ratio between them. The relative attribute valuation can be obtained by taking ratios of the estimated and specified values which will cancel out the scale parameter and be compared with the design values. The ratio was calculated for each statistic given by the general expression

$$\text{Ratio} = (\text{Estimated Value} - \text{Specified Values}) / \text{Specified Value}$$

The results showed that the simulation for the design without alternative specific constants was slightly better which indicate that the design can recover a reasonable range of relative valuations. Therefore the designs of the SP questions were considered capable of supporting the intended analysis.

4.4 Pilot Surveys

4.4.1 Pilot Data Collection

A simple pilot survey was conducted to test the feasibility of the SP design before the actual survey. The survey was intended to help improve the presentation of the questionnaire; to improve the design of different levels of the attributes to avoid dominance in the choices of different alternatives; and to help in representing realistic values of different levels and rates.

The survey was conducted as a CAPI exercise among research students at the University of Leeds in early February 2008. A total of 30 respondents were recruited and interviewed. They were shown both the SP exercises on the screen of the computer by using WinMint (a survey software developed by HCG to conduct CATI or CAPI). Although almost all of the respondents were male, had their income levels well below the national average and were regular, it is not thought that this mattered given the purpose of the pilot survey.

The purpose of the pilot survey was firstly, to identify potential improvements in the presentation of the questionnaire. Secondly, to improve the design of different levels of the attributes to avoid dominance in the choices of different alternatives. Lastly, to help establish realistic attribute levels and durations for the survey. In addition to this, a few alternative scenarios (including destination, trip purpose and accompanying passengers) were tested in the pilot survey and the most appropriate was selected in the final survey.

4.4.2 Results from First Pilot Survey

The first (CAPI) pilot survey revealed that the questionnaire was reasonably understandable and the design of the SP experiments was satisfactory. However, there were some problems in the questionnaire which are discussed in the following paragraphs.

The pilot survey revealed that the questionnaire was reasonably long and had taken in average of about 25 min to complete. This was a very important factor in the survey as longer questionnaire would in turn reduce the chance of interviewing a respondent at the site (train/bus station). It was thus decided to carry out the entire survey as CATI (Computer Assisted Telephone Interview) instead of CAPI. The other reason for this decision was that carrying out CAPI survey would otherwise require interviewing the respondents at their homes or workplaces which is quite difficult in this research because this would require a lot of interviewers and travel costs for them to reach the home/workplace of a respondent recruited for the interview at bus/train station.

It was noted that the original interview script (as used in the pilot) was rather long and complicated. Hence it was decided to rephrase and simplify the questionnaire. It was also decided to shorten the questions to reduce the overall duration of the interview.

One of the consequences of the decision to use CATI was that the SP exercises had to be distributed beforehand to the respondents in the form of the survey cards

4.4.3 Second Pilot Survey - Data Collection

After the first CAPI pilot survey a simple CATI pilot survey was conducted to test the feasibility of the CATI questionnaire and of the SP design before the actual survey. The survey was intended to help improve the verbal presentation of the questionnaire; to improve the design of different levels of the attributes to avoid dominance in the choices of different alternatives; to help in representing realistic values of different levels and rates; and to check the presentation of the survey cards (containing SP exercises and personal segmentation questions).

The second pilot survey was conducted at the University of Leeds in mid February 2008 using the survey software WinMint as a CATI tool. A total of 5 respondents were recruited and interviewed. As with the first pilot, their characteristics were not representative of the intended sample for the main survey (all were research students, almost were males, and all had below average incomes and almost all the respondents were users of the internet and traveller information services) but this is not thought to have mattered given the purpose of the pilot.

4.4.4 Results of Second Pilot Survey

The CATI pilot survey revealed that the questionnaire was reasonably understandable and the design of the SP experiments was satisfactory. However, there were some problems in the questionnaire which are discussed in the following paragraphs.

The CATI pilot survey indicated that the presentation of the cards was reasonably understandable and the respondents required very little or no explanation. It was found that the software presented the SP questions randomly whereas in the case of CATI in which respondents were using pre-printed cards this feature was undesirable. Hence in order to keep the randomness, it was required to print different sets of cards for different respondents. It was also found that some longer sentences should be further shortened and that the scenario description was too long to read out and so should be given to the respondents as a card.

4.5 Final Design

The questionnaire was again rephrased and simplified and the questions/sentences were further shortened to reduce the overall duration of the interview. The final version of the questionnaire is included as Appendix A and the WinMint software code is included as Appendix B.

It was decided to distribute the set of cards to represent the SP exercises at the time of recruitment. The cards included the sets of SP questions for the respondent to consult while answering the questions on the telephone. Every respondent was given four sets of SP cards from which he would use only one. In these four cards one was for those respondents who can use all the three modes (i.e. car. coach and train) whereas the other was for respondents whose choice of modes was limited for some reason. Figure 4.1 shows an example of cards distributed to the respondents during recruitment.

It was noted that the target respondents would include wide variety of public with different vocations (students, housewives, working class etc) hence, to understand the influence of the information on these different type of travellers, it was decided to use the SP experiment with different trip purposes of personal business and leisure trip. As noted in section 3.2.5 (about the minimum sample size), it was decided to conduct survey of at least 150 respondents for each trip purpose. This means interviewing in all 300 respondents.

The CATI interview calls were to be made to respondents while they were at home. Recruitment for this interview was to be via a simple recruitment interview conducted at the rail station and at the coach station (150 to be recruited at each location). The recruitment interviews were kept very simple and only required basic contact information from a potential respondent. The recruitment questions devised to recruit a typical respondent are as follows:

I am a PhD student at Leeds University; can I ask you a couple of questions as part of the data collection for my PhD thesis?

- Do you live in Leeds?
- Do you ever make journeys over 50 miles for leisure or personal business?
- Would you be prepared to help me in my research by taking part in a short phone interview about your travel choices?
- What phone number should I ring?
- Who should I ask for?
- When would you be available for the telephone interview?

101

CARD 4

Which website would you prefer?

1

WEBSITE A	WEBSITE B
Commercial Source Advertisements on the webpage No subscription required	Commercial Source Advertisements on the webpage Subscription £5 that you have already paid
Specific information, relevant to your journey, about routes/services/prices	General information, relevant to your journey, about routes/services/prices
Search time 15 minutes	Search time 10 minutes
Information Changes/updates in real time	Information Changes/updates daily

2

WEBSITE C	WEBSITE D
Government Source No advertisements on the webpage No subscription required	Commercial Source No advertisements on the webpage No subscription required
Specific information, based on your own criteria, for routes/services/prices	Specific information, based on your own criteria, for routes/services/prices
Search time 10 minutes	Search time 10 minutes
Information Changes/updates daily time	Information Changes/updates in real time

3

WEBSITE E	WEBSITE F
Government Source No advertisements on the webpage No subscription required	Commercial Source No advertisements on the webpage No subscription required
General information, relevant to your journey, about routes/services/prices	General information, relevant to your journey, about routes/services/prices
Search time 5 minutes	Search time 15 minutes
Information Changes/updates in real time	Information Changes/updates daily

4

WEBSITE G	WEBSITE H
Government Source No advertisements on the webpage No subscription required	Commercial Source Advertisements on the webpage Subscription £5 that you have already paid
Specific information, relevant to your journey, about routes/services/prices	Specific information, relevant to your journey, about routes/services/prices
Search time 15 minutes	Search time 5 minutes
Information Changes/updates daily	Information Changes/updates daily

201

CARD 5

Which mode would you prefer?

Choice 1

	Car		Train	
	Information Source 1	Information Source 2	Information Source 1	Information Source 2
	Friend	Multimodal website	Train website	Recent experience
Time	About 5 hrs 10 min (including 30 min delay)	5 hrs (including 20 min delay)	3:50 hrs (Scheduled station to station)	4: 05 min (Scheduled station to station)
Out of pocket cost	About £50	£ 75	£90	£115
Choice	(A)		(B)	

Choice 2

	Car		Train	
	Information Source 1	Information Source 2	Information Source 1	Information Source 2
	Friend	Multimodal website	Train website	Recent experience
Time	About 4 hrs 40 min	4 hrs 30 min	3 hrs 30 min (Scheduled station to station)	3 hrs 40 min (Scheduled station to station)
Out of pocket cost	About £70	£ 75	£110	£115
Choice	(C)		(D)	

Choice 3

	Car		Train	
	Information Source 1	Information Source 2	Information Source 1	Information Source 2
	Friend	Multimodal website	Train website	Recent experience
Time	About 5 hrs 10 min (including 30 min delay)	4 hrs 30 min	3 hrs 50 min (Scheduled station to station)	4 hrs (Scheduled station to station)
Out of pocket cost	About £60	£ 90	£110	£95
Choice	(E)		(F)	

Choice 4

	Car		Train	
	Information Source 1	Information Source 2	Information Source 1	Information Source 2
	Friend	Multimodal website	Train website	Recent experience
Time	About 4 hrs 20 min	4 hrs 30 min	3 hrs 30 min (Scheduled station to station)	4 hrs (Scheduled station to station)
Out of pocket cost	About £50	£ 60	£90	£125
Choice	(G)		(H)	

Figure 4.1 Sample cards distributed to the respondents during recruitments

In order to reduce any embarrassment in revealing personal information, and thereby maximise response, the potential answers were printed to cards distributed at the time of recruitment by showing the income and age categories on the cards in random order instead of showing them in typical ascending/descending order an impression of anonymity was created.

Chapter 5
Analysis of Descriptive Data from the Main Survey

5.1 Introduction

This chapter aims to report the descriptive analysis of the results of the survey. As described in Chapter 4 and shown in Appendix A, the main questionnaire had four parts. The first part gathered Revealed Preference (RP) data for the respondents' most recent long journey (over 50 miles) and included questions on frequency of travel to that destination, purpose of visit, the chosen mode, other available modes, the external circumstances of the journey, and the use and effect of any information source used while planning that journey. The second part included SP survey questions designed to investigate the traveller's choice of information sources and subsequent mode choice when making long journeys. The third part included general questions about their attitudes towards different sources of information and their normal search patterns. The final part contained questions about the traveller's socio-economic characteristics. This chapter will discuss the results from the first, third and fourth parts of the questionnaire.

Section 5.2 reports about the recruitment process. Section 5.3 summarizes the respondents' characteristics and travel patterns. Finally, section 5.4 reports respondents' use of the information sources and their stated effect on their travel decisions.

5.2 Recruitment

As described in Chapter 4, respondents were recruited at the main long distance transport interchanges in Leeds i.e. the coach Station and the train station. Recruits were interviewed by CATI. The response rate in the first few weeks was lower than expected and so other PhD students were employed to help with recruitment. About 950 members of the public were recruited to achieve the target sample of 300.

5.3 Respondents' Characteristics and Travel Patterns

Table 5.1 shows the distribution of socio–economic characteristics of the samples and compares them with the Census data (HMSO 2001) for Leeds residents and National Travel Survey (NTS) 2006 data of long distance[2] travellers. About 61 percent of respondents were males. This compares with a figure of 48 percent as per 2001 Census data (HMSO 2001) for Leeds residents. This difference is not unexpected because the target population of the research was long distance travellers not residents. On the other hand, as expected the figure matches with a figure of 61 percent as per NTS data of long distance travellers. About 30 percent of the respondents were under 30 (compared to 40 percent for Leeds residents as per 2001 Census of city of Leeds) and about 63 percent of the respondents had ages between 30 and 50 (compared to 40 percent for Leeds residents). The apparent under representation of people as per NTS data of long distance travellers under 20 and over 50 is again probably related to the fact that frequent intercity travellers are generally between 20 and 50.

The income distribution summarised in Figure 5.1 shows that 54% of the respondents said that their household income was between £20,000 and £40,000 and around 34% percent said it was under £20,000. This could be because there are a lot of students living in Leeds and they travel frequently to different cities.

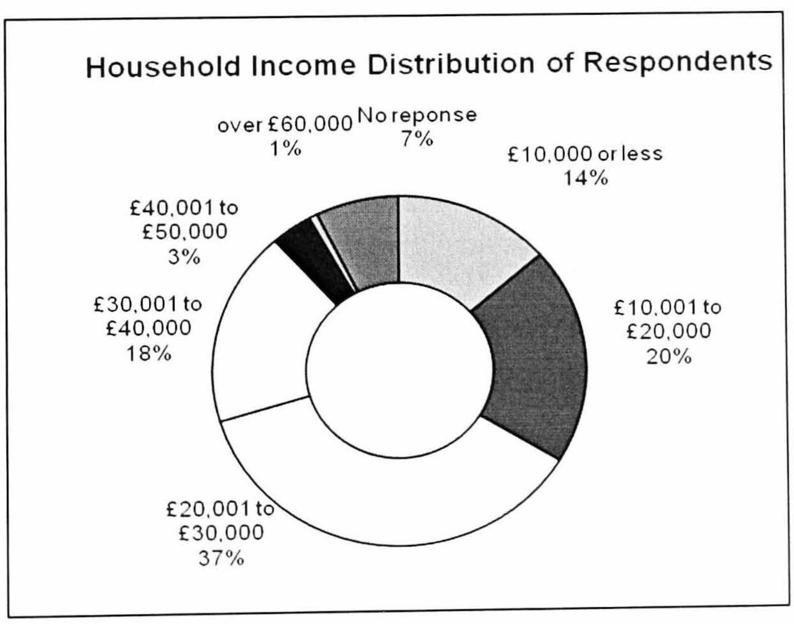

Figure 5.1 Distribution of Respondents' income

[2] Long distance travel is defined as trips of 50 miles or more one-way in the NTS as is used in this study.

Figure 5.2 shows that 67% percent of the respondents were employed (compared to 59 percent for Leeds residents) and about 28% were students.

Table 5.1 Distribution of Socio-Economics

Characteristics	Main Survey (by author)		Census 2001 for Leeds	NTS 2006
	Frequency	Percent		
Gender				
Female	108	39	52	39
Male	170	61	48	61
Total	278	100	100	100
Age				
less than 20	6	2	25	-
20 to 29	78	29	15	-
30 to 39	116	43	15	-
40 to 49	52	19	14	-
50 to 64	8	3	16	-
65 and over	4	1	15	-
Total	270	100	100	-
Occupation				
Employed part time	12	4	-	12
Retired	2	1	12	13
Student	76	28	11	3
looking after home/housewife	6	2	5	4
without work	2	1	13	1
Full time employed	170	63	59	66
Total	268	100	100	100
Car Ownership				
0	63	24	34	8
1	100	38	42	34
2 or more	102	38	24	58
Total	268	100	100	100
Household Size				
1	12	5	32	-
2	88	33	33	-
More than 2	164	62	35	-
Total	264	100	100	-
Income				
£20,000 or less	94	34	-	18
£20,001 to £40,000	152	54	-	30
over £40,000	12	5	-	52
No response	20	7	-	-
Total	278	100	-	100

About 76% of the respondents owned a car and 50% of them had more than one car available to their household. According to the 2001 Census data 65.5 % of the population of Leeds owned at least one car – this difference is again not surprising since the respondents were recruited at public transport hubs.

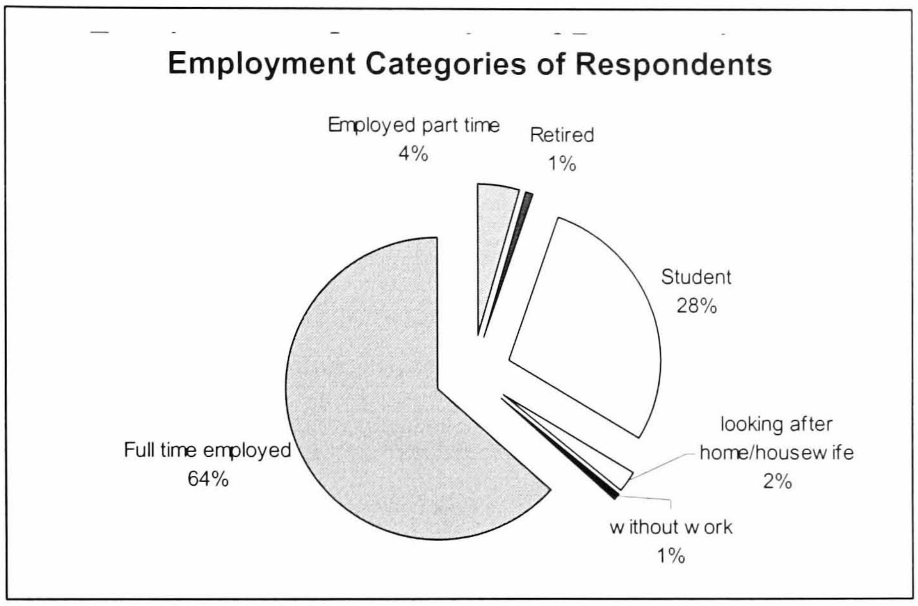

Figure 5.2 Employment distribution of Respondents

The respondents' household size was fairly evenly distributed; as can be seen in Figure 5.3 about 36% had 4 or more family members whereas 38% had 2 or less family members.

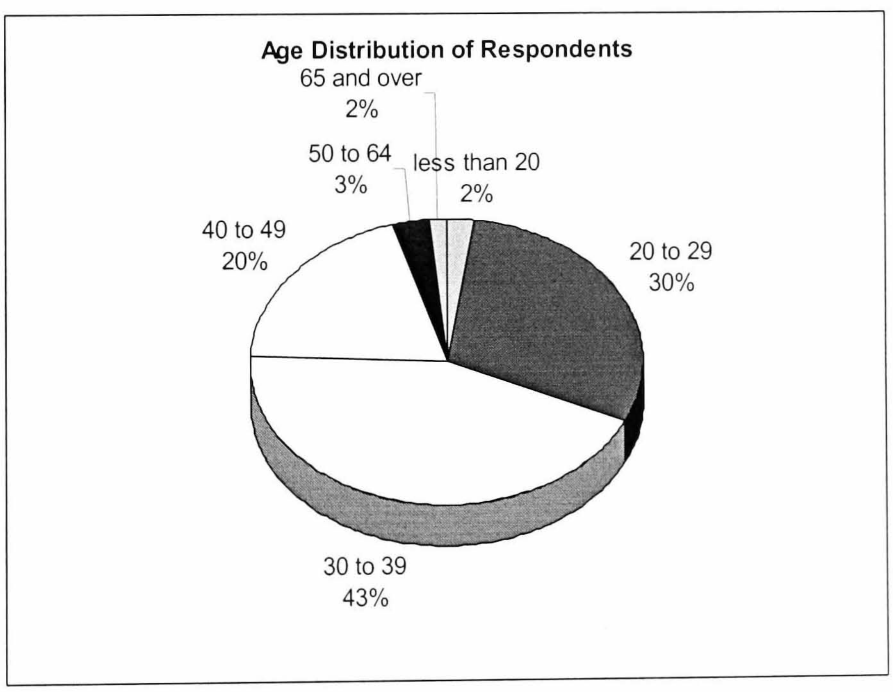

Figure 5.3 Age Distribution of Respondents

Table 5.2a shows the use of information sources by respondents before the journey reported in the RP part of the questionnaire. The table shows that the 87% of the respondents used websites as an information source before their previous intercity journey and that 44% of them used websites as their only source of information. Websites are clearly the predominant information sources nowadays.

Table 5.2b illustrates respondents' use of different types of websites and clearly shows that car and train websites are used more commonly than coach or multimodal websites.

Table 5.2 Use of Information Sources

a. Types of Source

Source	Frequency	Percent
Car-web exclusively	66	24
Coach-web exclusively	26	9
Train-web exclusively	58	21
Multimodal-web exclusively	6	2
Car-web with other	116	42
Coach-web with other	62	22
Train-web with other	116	42
Multimodal with other	58	21
Car and multimodal	20	7
Train and coach	20	7
Multimodal, train and coach	8	3

b. Types of Websites

Website	Frequency	Percent
Website Used exclusively	124	45
Website with other	242	87
Website with friend	64	23
Website with map	30	11
Website with map and friend	16	6
Friend exclusive	6	2
Friend with other	92	33
Friend and map	2	1

5.4 Cross Relationship of the Use of Website and Mode on the Socio-Economic Characteristics.

The use of information sources and subsequent mode choice depends on a variety of factors and characteristics. These include socio-economic characteristics (gender, education level, and monthly income); trip characteristics (trip purpose and frequency); and other behavioural factors. Figure 5.4 shows the cross relationship between use of website and mode choice. It is clear that the choice of car decreases with increase in the use of websites. This is intuitively understandable as public transport journeys are nowadays planned/tickets booked via websites more commonly than car journeys.

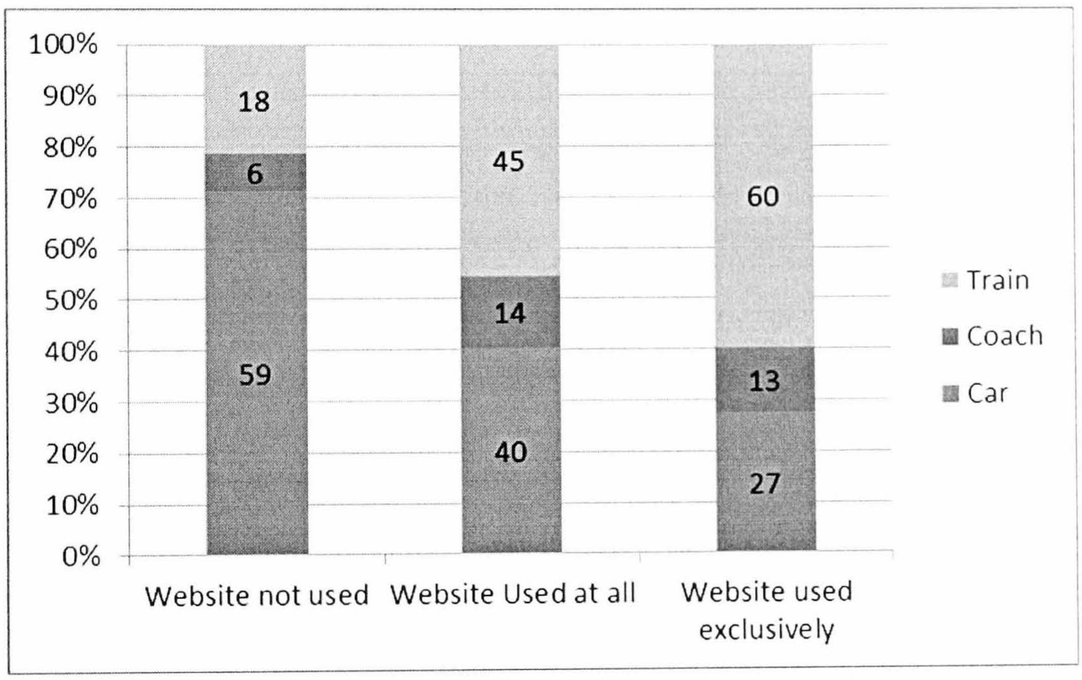

Figure 5.4 Cross relationship between use of website and mode choice

We will now examine the effects of gender, age, income, trip frequency and trip purpose on information and mode use.

a. Gender

Figure 5.5 reveals that, among those respondents who used websites as a source of information, there is not much difference in mode choice between male and females. Males are slightly more prone to driving after they got travel information, where as females more likely to use public transport slightly more than the males. The Figure 5.5 also shows that among those respondents who didn't use websites as source of travel information, there is again not much difference in mode choice between male and females. Respondents, who only used websites tended to use car less than public transport. This shows that people using public transport use websites more as compared to the car users. Table 5.3 shows the overall counts and percentages.

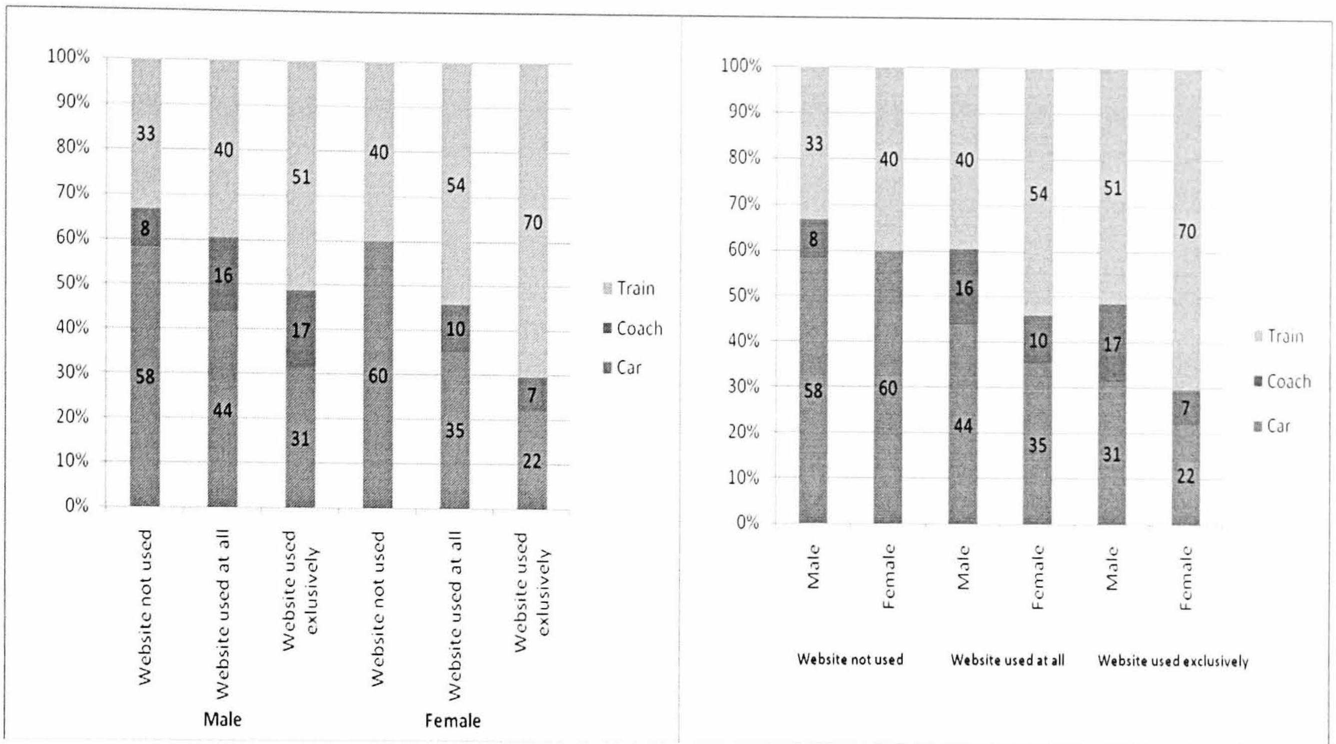

Figure 5.5 Choice of Mode by Sex (among Website Users)

Table 5.3 Use of website exclusively and related mode choice by sex

Relationship		Chosen Mode			Total
		Car	Coach	Train	
Male	Observations	22	12	36	70
	% within Sex[3]	31	17	51	100
	% within Chosen Mode[4]	65	75	49	56
	% of Total[5]	18	10	29	56
Female	Observations	12	4	38	54
	% within Sex	22	7	70	100
	% within Chosen Mode	35	25	51	44
	% of Total	10	3	31	44
Total	Observations	34	16	74	124
	% within Sex	27	13	60	100
	% within Chosen Mode	100	100	100	100
	% of Total	27	13	60	100

[3] % within Sex = Percentage of gender by mode e.g. 31 % of males choose car, 12% Coach and so on
[4] % within Chosen Mode = Percentage of mode by gender e.g. 65% of Car users are males and so on
[5] % of Total = Percentage of total observations e.g. 18 % of all population using car are males and so on

b. Age

Respondents were divided into three age groups; those under 29, those who are between 30 and 39 and those over 40. Figure 5.6 reveals that among respondents who used websites, as a source of information, younger respondents and seniors were particularly likely to use train for their intercity travel. Choice of car decreases as age increases and train is more preferred than the coach by the intercity travellers. Figure 5.6 also reveals that among those respondents who only used websites, younger respondents and seniors were again particularly likely to use train for their intercity travel. There is a mixed trend for the choice of car, while train is more preferred than the coach by the intercity travellers. Table 5.5 shows all the counts and percentages.

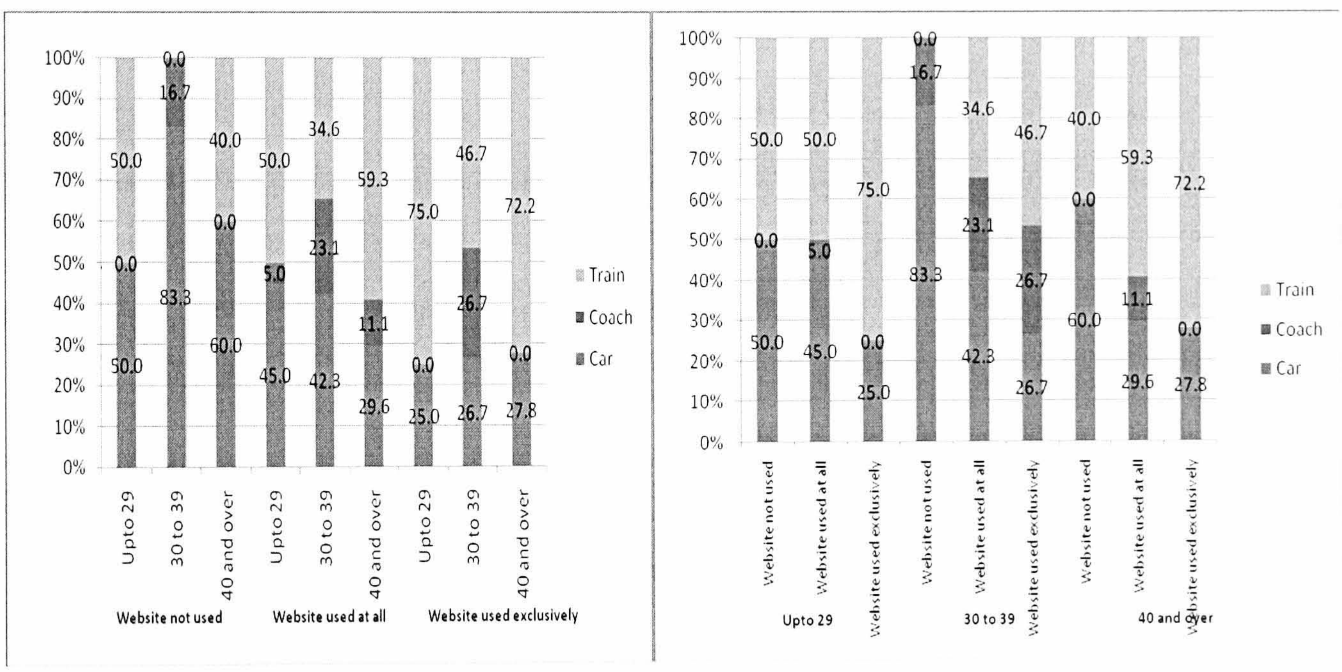

Figure 5.6 Choice of mode by Age (Among Website Users)

c. Income

Respondents were categorized into four annual income levels, under £10,000, between £10,000 and £20,000, between £20,000 and £30,000, and over £30,000. Figure 5.7 reveals that among those respondents who used websites, as a source of information, the choice of coach decreases as income increases whereas the choice of train and car remains fairly constant. This could be because coach is unattractive as a mode in long journeys because it takes more time than trains or car but, since it is relatively cheap, it is attractive to people on low incomes.

Table 5.4 Use of website exclusively and related mode choice by Age

Relationship				Chosen Mode: Car	Coach	Train	Total
Website not used	Age	Up to 29	Count	2	0	2	4
			% within Age	50	0	50	100
			% within Chosen Mode	11	0	33	13
			% of Total	6	0	6	13
		30 to 39	Count	10	2	0	12
			% within Age	83	17	0	100
			% within Chosen Mode	56	100	0	38
			% of Total	31	6	0	38
		40 and over	Count	6	0	4	10
			% within Age	60	0	40	100
			% within Chosen Mode	33	0	67	31
			% of Total	19	0	13	31
		Total	Count	18	2	6	32
			% within Age	56	6	19	100
			% within Chosen Mode	100	100	100	100
			% of Total	56	6	19	100
Website used at all	Age	Up to 29	Count	36	4	40	80
			% within Age	45	5	50	100
			% within Chosen Mode	38	12	37	34
			% of Total	15	2	17	34
		30 to 39	Count	44	24	36	104
			% within Age	42	23	35	100
			% within Chosen Mode	46	71	33	44
			% of Total	18	10	15	44
		40 and over	Count	16	6	32	54
			% within Age	30	11	59	100
			% within Chosen Mode	17	18	30	23
			% of Total	7	3	13	23
		Total	Count	96	34	108	238
			% within Age	40	14	45	100
			% within Chosen Mode	100	100	100	100
			% of Total	40	14	45	100
Website used exclusively	Age	Up to 29	Count	6	0	18	24
			% within Age	25	0	75	100
			% within Chosen Mode	19	0	25	20
			% of Total	5	0	15	20
		30 to 39	Count	16	16	28	60
			% within Age	27	27	47	100
			% within Chosen Mode	50	100	39	50
			% of Total	13	13	23	50
		40 and over	Count	10	0	26	36
			% within Age	28	0	72	100
			% within Chosen Mode	31	0	36	30
			% of Total	8	0	22	30

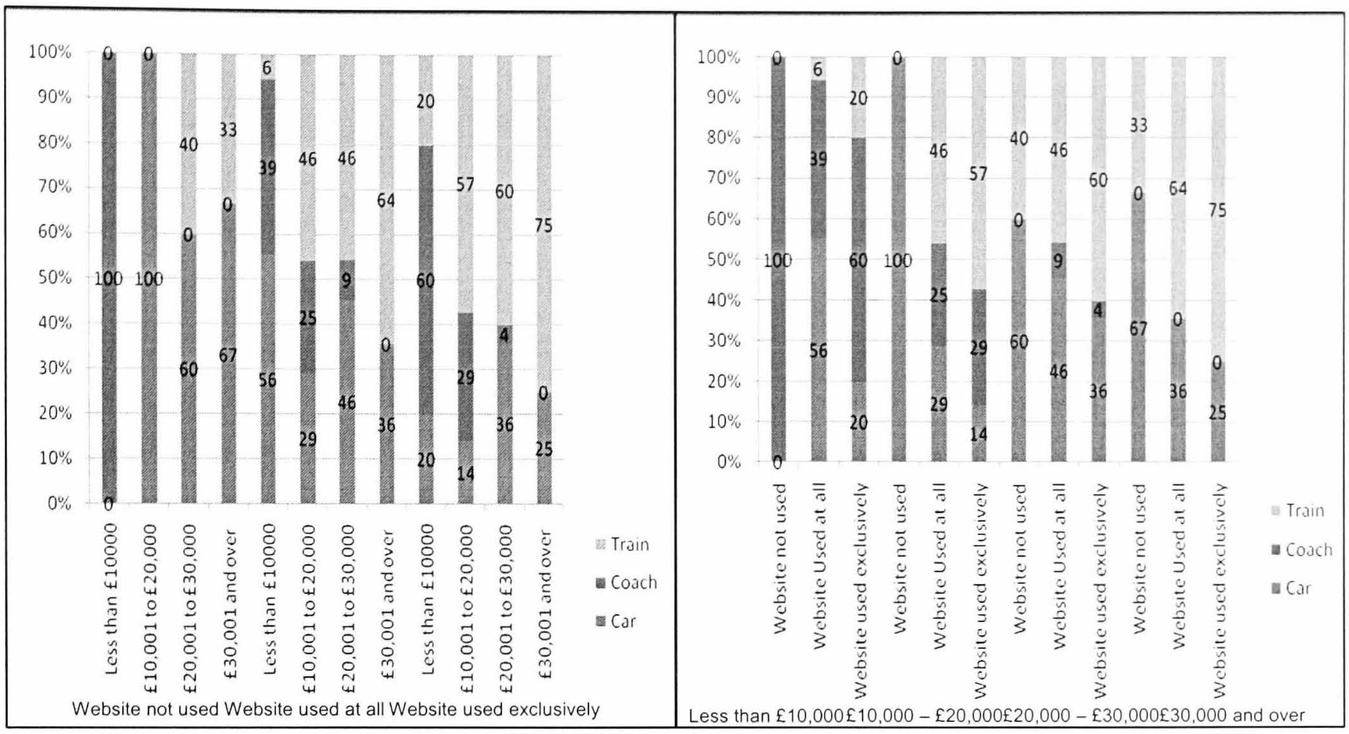

Figure 5.7 Choice of mode by Income (Among Website Users)

The Figure 5.7 also reveals that among those respondents, who only used websites, choice of train increases with increased income whereas choice of coach decreases as income increases. There is a mixed pattern for the choice of car. Table 5.6 shows all the counts and percentages.

d. Frequency

Respondents were categorised into four levels of trip frequency; less than 5 times per year, between 6 and 12 times per year, between 13 and 40 times per year, and over 40 times per year. Figure 5.8 reveals that among those respondents who only used websites, the choice of public transport increases with the use of website with all travel frequencies. This could be because the websites are more often used to plan public transport journeys rather than car journeys.

Table 5.5 Use of website exclusively and related mode choice by Income

Relationship		Chosen Mode			Total
		Car	Coach	Train	
Less than £10000	Count	2.0	6.0	2.0	10.0
	% within Income	20.0	60.0	20.0	100.0
	% within Chosen Mode	5.9	37.5	2.7	8.1
	% of Total	1.6	4.8	1.6	8.1
£10,001 to £20,000	Count	4.0	8.0	16.0	28.0
	% within Income	14.3	28.6	57.1	100.0
	% within Chosen Mode	11.8	50.0	21.6	22.6
	% of Total	3.2	6.5	12.9	22.6
£20,001 to £30,000	Count	18.0	2.0	30.0	50.0
	% within Income	36.0	4.0	60.0	100.0
	% within Chosen Mode	52.9	12.5	40.5	40.3
	% of Total	14.5	1.6	24.2	40.3
£30,001 and over	Count	8.0	0.0	24.0	32.0
	% within Income	25.0	0.0	75.0	100.0
	% within Chosen Mode	23.5	0.0	27.0	22.6
	% of Total	6.5	0.0	16.1	22.6
Total	Count	34.0	16.0	74.0	124.0
	% within Income	27.4	12.9	59.7	100.0
	% within Chosen Mode	100.0	100.0	100.0	100.0
	% of Total	27.4	12.9	59.7	100.0

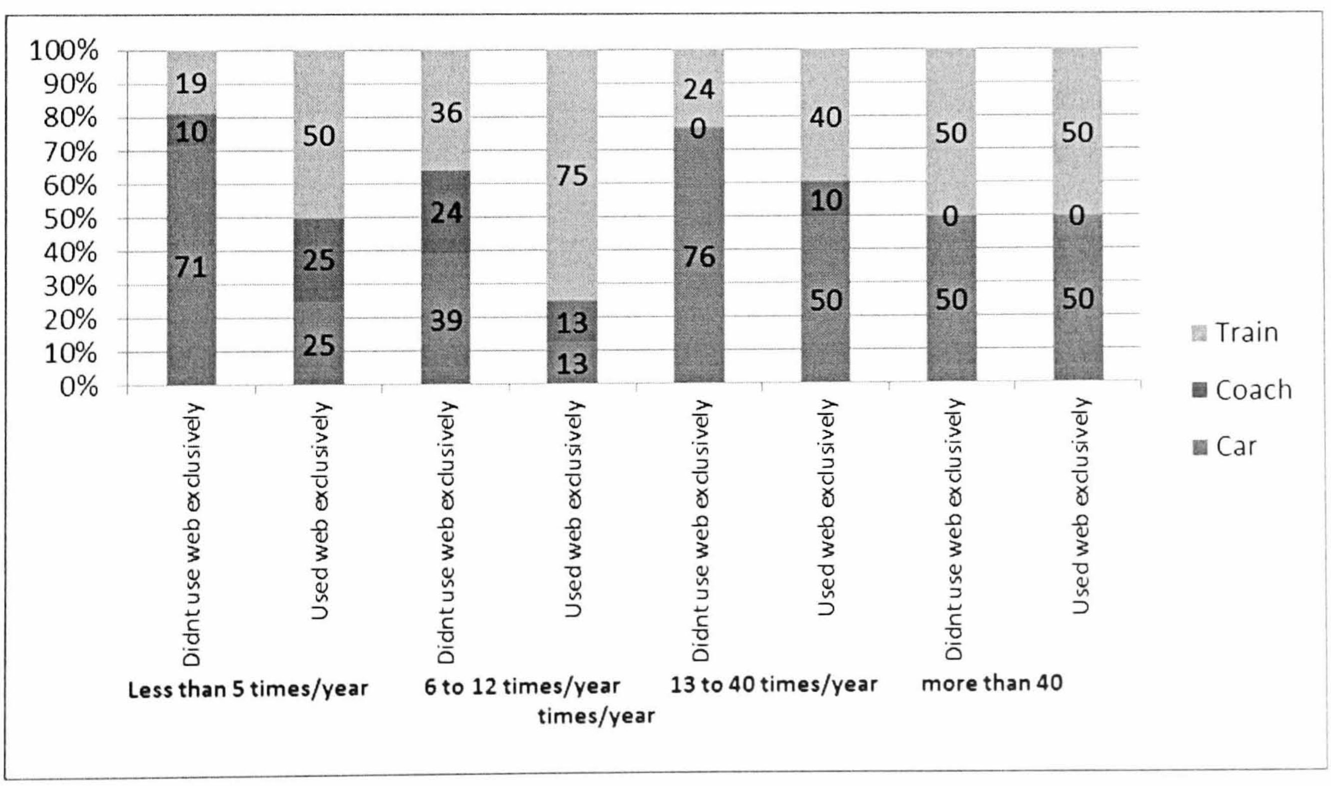

Figure 5.8 Choice of mode by Travel Frequency (Among Website Users)

e. Trip Purpose

Respondents were categorized into seven trip purposes as shown in figure 5.9.

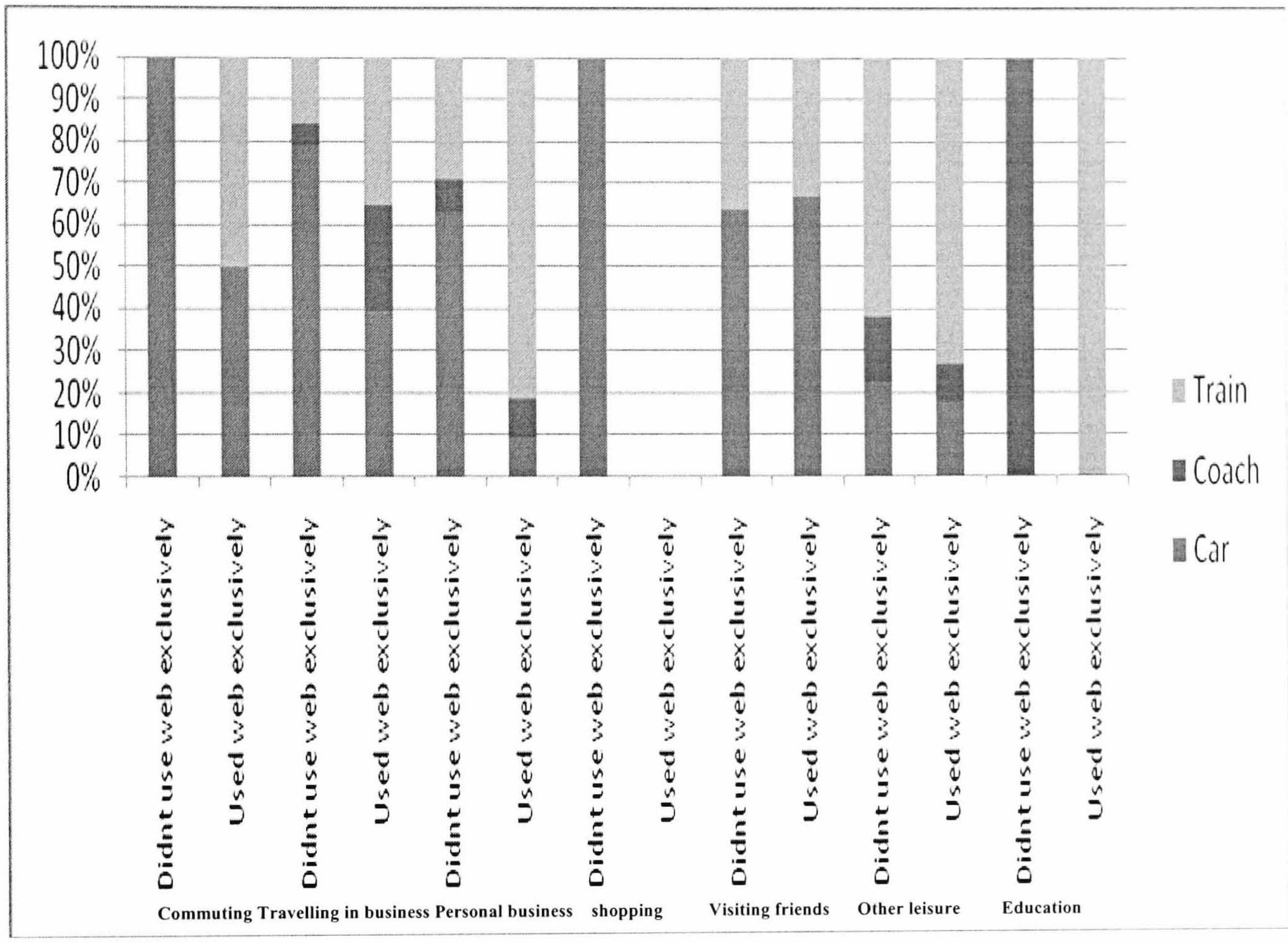

Figure 5.9 Choice of mode by Travel Purpose (Among Website Users)

Table 5.6 Use of website exclusively and related mode choice by Travel Frequency

Relationship			Chosen Mode			
			Car	Coach	Train	Total
less than 5 times/year	Didn't use web exclusively	Count	30	4	8	42
		% within only web	71	10	19	100
		% within Chosen Mode	88	50	50	72
		% of Total	52	7	14	72
	Used web exclusively	Count	4	4	8	16
		% within only web	25	25	50	100
		% within Chosen Mode	12	50	50	28
		% of Total	7	7	14	28
	Total	Count	34	8	16	58
		% within only web	59	14	28	100
		% within Chosen Mode	100	100	100	100
		% of Total	59	14	28	100
6 to 12 times per year	Didn't use web exclusively	Count	26	16	24	66
		% within only web	39	24	36	100
		% within Chosen Mode	76	67	33	51
		% of Total	20	12	18	51
	Used web exclusively	Count	8	8	48	64
		% within only web	13	13	75	100
		% within Chosen Mode	24	33	67	49
		% of Total	6	6	37	49
	Total	Count	34	24	72	130
		% within only web	26	18	55	100
		% within Chosen Mode	100	100	100	100
		% of Total	26	18	55	100
13 to 40 times per year	Didn't use web exclusively	Count	26	0	8	34
		% within only web	76	0	24	100
		% within Chosen Mode	57	0	33	46
		% of Total	35	0	11	46
	Used web exclusively	Count	20	4	16	40
		% within only web	50	10	40	100
		% within Chosen Mode	43	100	67	54
		% of Total	27	5	22	54
	Total	Count	46	4	24	74
		% within only web	62	5	32	100
		% within Chosen Mode	100	100	100	100
		% of Total	62	5	32	100
more than 40 times per year	Didn't use web exclusively	Count	2	0	2	4
		% within only web	50	0	50	100
		% within Chosen Mode	50	0	50	50
		% of Total	25	0	25	50
	Used web exclusively	Count	2	0	2	4
		% within only web	50	0	50	100
		% within Chosen Mode	50	0	50	50
		% of Total	25	0	25	50
	Total	Count	4	0	4	8
		% within only web	50	0	50	100
		% within Chosen Mode	100	0	100	100
		% of Total	50	0	50	100

The Figure 5.9 reveals that among those who only used websites, the commuting, business travel and education groups used both car and public transport whereas those travelling for leisure used car predominantly. This could be because the flexibility offered by the car is particularly important for the leisure travellers.

Chapter 6
Model Development

6.1 General

An approach combining Revealed Preference (RP) and Stated Preference (SP) is adopted in this study to examine the impact of traveller information on the mode choice decisions of travellers. First this chapter develops a series of models in order to find the best fitting model for the survey results. Several models have been developed using different variables and different combination of variables. From these models a best model is selected. Secondly the selected model is analyzed and discussed in detail. Multinomial and nested logit models are used to analyze the behaviour. In this chapter, the model development is discussed in detail. Section 6.2 briefly summarises the RP and SP data which is analysed. Section 6.3 deals with the modelling approach used. Section 6.4 reports the process of model development. Section 6.5 outlines the issue of repeated measurement. Section 6.6 shows the estimations of mixed logit models. Section 6.7 shows the calibration of random regret minimisation framework. Finally Section 6.8 presents a summary of the work in this chapter and conclusions.

6.2 Brief Data Description

The Revealed Preference (RP) data concerns the past behaviour of the respondents for their last long journey (over 50 miles). A detailed account of the journey was explored by asking numerous questions including their frequency of travel to that destination, purpose of visits and the chosen mode. In order to make a choice set, it also contains questions regarding respondents' perception of attributes of different modes available to them when they were taking the travel decisions. Questions were also asked regarding the external circumstances of the journey together with questions about the use and effect of any information source while taking travel decisions about that journey.

The SP design was discussed in chapter 4 but it is summarized here for convenience. The first SP exercise (SP1) offered the respondents two choices for the selection of the information source under a hypothetical travel situation. The attributes included ownership of the source; type of information; search time; presence of advertisements;

frequency with which the information is updated; and any subscription cost. *Ownership of the information source* could be either "government" or "private company". This may influence the credibility of the source and might be an important consideration in the selection of the sources. For *type of information* about travel time and delay, three levels were considered: "Descriptive Quantitative real time information" (e.g. specific information relevant to the journey about routes/services/prices), "Descriptive Qualitative information" (e.g. general information about routes/services/prices relevant to the journey), and "Prescriptive information" (e.g. specific advice relevant to the journey based on users' criteria for routes/services/prices). For the *search time* three levels were considered, "5 min", "10 min" and "15 min". For the *presence of advertisement* there were only two possibilities: advertisements are either "there" or they are "not". For *frequency of information* changes on the source again three levels were considered, "real time", "daily" and "weekly". For *subscription* again there were two levels, "no subscription" and "£5 already paid". The selection of these attributes was discussed in detail in section 4.4 of the previous chapter. A copy of the questionnaire is included in appendix A.

The second SP exercise (SP2) required the respondent to choose between three modes. Each mode had at least two sources and time and cost attributes. The base levels were selected from the current travel time and costs as described by information sources during normal conditions in spring 2008. The other values were +20% and -20% deviations from these average values. In all about 950 members of the public were recruited to achieve the target sample size of 300.

6.3 Model Estimation Methods

A computer software package, Alogit, was used to estimate the models for both RP and SP data separately. The estimation process is based on maximum likelihood method.

6.3.1 Data Structure for Estimation

As per requirement of the Alogit Software, the collected data is formatted in a single data sheet to support mode choice model estimation. For the combined SP and combined RP-SP models, more than one file is used as an input in Alogit. Hence the data collected from respondents are merged by using Merge software and then assembled by using MS Excel. Then the data are saved as a csv file required as input for the software.

6.3.2 Random Utility Theory

Discrete choice models are based on choices made by individuals. They assume that the probability of an individual choosing a given alternative is the function of their socio-economic characteristics and the relative attractiveness of the options (Ortuzar and Willumsen 1994). Random Utility Theory is the most common theoretical base for discrete choice models. Manski (1973) identified four sources of randomness;

- Unobserved attributes
- Unobserved taste variations
- Measurement error and imperfect information; and
- Instrumental or proxy variables.

The effect of these sources of randomness was formalized by Manski (1977) and incorporated within the utility of a given alternative so that the *random utility* (U_i) is the sum of the observable (or systematic) component (V_i) and unobservable (or random) component (ε_i):

$$U_{in} = V_{in} + \varepsilon_{in}$$

The deterministic component, V_i is a function of observable and measureable attributes x_i. The function is commonly assumed to be linear in combining variables.

$$V_i = \sum_k \theta_{ik} x_{ik}$$

where θ_{ik} is a vector of parameters to be estimated and x_{ik} is a vector of observed data relating to alternative *i.*

And the probability of choosing an alternative from a choice set is:

$$P(i \mid M_n) = P[V_{in} + \varepsilon_{in} \geq V_{in} + \varepsilon_{in}\ ,\ all\ j \in M_n]$$

6.3.3 The Multinomial Logit

One of the main advantages of MNL is that it is relatively easy to estimate, the coefficients are easy to interpret and the forecast are, most of the time, robust. MNL is derived from the assumption that errors of the utility function are identically and independently Gumbel distributed. However, the main drawback of model is that it assumes that the choice options are independent and therefore fails to take account of the correlation between the alternatives. This property is known as independence from irrelevant alternatives (IIA).

The model was first introduced in the context of binary choice models. McFadden (1973) can be credited with its derivation from random utility theory. In MNL, the probability P_i of choosing alternative i from the choice set J, given measured utilities V_j $(j \in J)$, is given by,

$$P_i = \frac{e^{\mu V_i}}{\sum_{j \in J} e^{\mu V_j}}$$

Where J is the set of choice alternatives and μ is a strictly positive scale parameter and it is related to the variance σ^2 of the error term $\mu = \frac{\pi}{\sqrt{6}\sigma}$ and is usually normalised to be equal to one as it cannot be estimated separately from the coefficients (Ortuzar and Willumsen, 1994).

6.3.4 The Nested Logit

The most widely known relaxation of the MNL is the nested logit model. The model like MNL is based on the utility maximisation but allows for the correlation between the pairs of alternatives in a common group (a nest). The nested logit (NL) can be written as the series of MNL models at each level of the tree structure. By allowing the correlation among the subsets of alternatives the model reduces some of the problems associated with the MNL.

According to Ortuzar (2001), the application of the nested logit preceded its theoretical derivation. Ortuzar attributed the application to Ben-Akiva (1974) and the theoretical derivation to Williams (1977) and Daly and Zachary (1978). For a two level example the nested logit model shows the probability that alternative i is chosen as,

$$P_i = P_m P_{i|m}$$

Where P_m, the probability of choosing nest m is given by

$$P_m = \frac{\left(\sum_{j \in J_m} e^{V_j}\right)^{\theta_m}}{\sum_m \left(\sum_{j \in J_m} e^{V_j}\right)^{\theta_m}}$$

and $P_{i|m}$, the conditional probability of choosing alternative i from nest m is given by,

$$P_{(i|m)} = \frac{e^{V_i}}{\sum_{j \in J_m} e^{V_j}}$$

where N_m is the set of alternatives in nest m and θ_m is the dissimilarity parameter relating to nest m. NL is consistent with the random utility maximisation if $0 < \theta_m \leq 1$ for all θ_m.

6.3.5 Maximum Likelihood

Discrete choice models are mostly estimated using Maximum Likelihood. The likelihood function for a general multinomial model is expressed as,

$$L = \prod_{n=1}^{N} \prod_{i \in C_n} P_n(i)^{y_{in}}$$

where N is the sample size i represents alternative and C_n is Chice set,

$$y_{in} \begin{cases} 1 \; if\ observation\ n\ chose\ alternative\ i \\ 0 \; otherwise \end{cases}$$

The objective of the estimation is to set the coefficients of the chosen model to maximise the likelihood of function. The function is linearised by taking the natural logarithm where the log-likelihood is maximised.

6.4 Model Development

Logit models were ultimately intended to measure the effect of information sources on the mode choice. The process of model development was incremental starting with multinomial models for both SP and RP models. Models for RP and SP data were first estimated separately. The following is the stepwise model development process.

- RP model for mode choice
- RP model for source choice
- SP model for mode choice
- SP model for source choice
- Combined RP model for source and mode
- Combined SP model for source and mode
- Combined RP and SP model for source and mode

6.4.1 RP Model for Mode Choice

Multinomial models were calibrated for the choice of mode in the RP data. As mentioned earlier the data consists of respondents' past behaviour for the last long journey (over 50 miles). Variables used in the model were selected as discussed in Chapter 3 and were added to the model incrementally (from simpler to more complex). A Multinomial Logit Model (MNL) is constructed with the dependent variable being the choice among car, coach and train. The base case for this model is travelling by car. The choice set C_n of each individual thus consists of three alternatives. The utility functions are given by,

1. $U_{(Car)} = \mathrm{Time} * Time_1 + \mathrm{Cost} * Cost_1 + \mathrm{Car}_1 * DCar_1$

2. $U_{(Coach)}$ = Time*$Time_2$ + Cost*$Cost_2$ + BPur_2*$DBPur_2$ + LFreq_2*$DLFreq_2$ + GWeather_2*$DGWeather_2$+ PeakP_2*$DPeakP_2$ + TraAlone_2*$DTraAlone_2$ + ImpSfty_2*$DImpSfty_2$ + ImpCmfrt_2*$DImpCmfrt_2$ + ImpSeats_2*$DImpSeats_2$+ Male_2*$DMale_2$ + Educ_2*$DEduc_2$ + FEmpl_2*$DFEmpl_2$ + Income_2*$DIncome_2$ + Age_2*$DAge_2$+ASC_2

3. $U_{(Train)}$ = Time*$Time_3$ +Cost*$Cost_3$ + BPur_3*$DBPur_3$ + LFreq_3*$DLFreq_3$ + GWeather_3*$DGWeather_3$+ PeakP_3*$DPeakP_3$ + TraAlone_3*$DTraAlone_3$ + ImpSfty_3*$DImpSfty_3$ + ImpCmfrt_3*$DImpCmfrt_3$ + ImpSeats_3*$DImpSeats_3$+ Male_3*$DMale_3$ + Educ_3*$DEduc_3$ + FEmpl_3*$DFEmpl_3$ + Income_3*$DIncome_3$ + Age_3*$DAge_3$+ ASC_3

Model estimates are shown in table 6.1. The variables and associated coefficients are defined below.

Variables

Where: (with subscript n indicating modes)

$Time_n$	***(Generic, in Minutes)***
$Cost_n$	***(Generic, in Pence)***
$DCar_n$	***(Dummy, if car is available in household = 1, otherwise = 0)***
$DBPur_n$	***(Dummy, Business Purpose = 1, otherwise = 0)***
$DLFreq_n$	***(Dummy, Trip Frequency less than 13/year =1, otherwise = 0)***
$DGWeather_n$	***(Dummy, Good weather =1, otherwise = 0)***
$DIncident_n$	***(Dummy, Incident occurred =1, otherwise = 0)***
$DPeakP_n$	***(Dummy, Travelled in the Peak Period = 1, otherwise = 0)***
$DReasMode_n$	***(Dummy, Reason for selecting mode is Time/Cost = 1, otherwise = 0)***
$DTravAlone_n$	***(Dummy, Travelling alone = 1, otherwise = 0)***
$DImpSfty_n$	***(Dummy, Safety important = 1, otherwise = 0)***
$DImpCmfrt_n$	***(Dummy, Comfort important = 1, otherwise = 0)***
$DImpSeats_n$	***(Dummy, Seat availability important = 1, otherwise = 0)***
$DMale_n$	***(Dummy, If male =1, otherwise = 0)***
$DEduc_n$	***(Dummy, Left full time education at or after 20 = 1, otherwise = 0)***
$DEduc25_n$	***(Dummy, Left full time education at or after 25 = 1, otherwise = 0)***
$DFEmpl_n$	***(Dummy, Full time employed =1, otherwise = 0)***
$DIncome_n$	***(Dummy, If income over £30,000 =1, otherwise = 0)***
$DAge_n$	***(Dummy, If Age less than 50 =1, otherwise = 0)***
$DAge40_n$	***(Dummy, If Age less than 40 =1, otherwise = 0)***

Coefficients

Where: (with subscript n indicating modes)

Time	**(is a parameter vector to be estimated for Time, Generic, in Minutes)**
Cost	**(is a parameter vector to be estimated for Cost, Generic, in Pence)**
Car_n	**(is a parameter vector to be estimated for Dummy, if car is available in household = 1, otherwise = 0)**
$BPur_n$	**(is a parameter vector to be estimated for Dummy, Business Purpose = 1, otherwise = 0)**
$LFreq_n$	**(is a parameter vector to be estimated for Dummy, Trip Frequency less than 13/year =1, otherwise = 0)**
$GWeather_n$	**(is a parameter vector to be estimated for Dummy, Good weather =1, otherwise = 0)**
$Incident_n$	**(is a parameter vector to be estimated for Dummy, Incident occurred =1, otherwise = 0)**
$PeakP_n$	**(is a parameter vector to be estimated for Dummy, Travelled in the Peak Period = 1, otherwise = 0)**
$ReasMode_n$	**(is a parameter vector to be estimated for Dummy, Reason for selecting mode is Time/Cost = 1, otherwise = 0)**
$TravAlone_n$	**(is a parameter vector to be estimated for Dummy, Travelling alone = 1, otherwise = 0)**
$ImpSfty_n$	**(is a parameter vector to be estimated for Dummy, Safety important = 1, otherwise = 0)**
$ImpCmfrt_n$	**(is a parameter vector to be estimated for Dummy, Comfort important = 1, otherwise = 0)**
$ImpSeats_n$	**(is a parameter vector to be estimated for Dummy, Seat availability important = 1, otherwise = 0)**
$Male_n$	**(is a parameter vector to be estimated for Dummy, If male =1, otherwise = 0)**
$Educ_n$	**(is a parameter vector to be estimated for Dummy, Left full time education at or after 20 = 1, otherwise = 0)**
$Educ25_n$	**(is a parameter vector to be estimated for Dummy, Left full time education at or after 25 = 1, otherwise = 0)**
$FEmpl_n$	**(is a parameter vector to be estimated for Dummy, Full time employed =1, otherwise = 0)**
$Income_n$	**(is a parameter vector to be estimated for Dummy, If income over £30,000 =1, otherwise = 0)**
Age_n	**(is a parameter vector to be estimated for Dummy, If Age less than 50 =1, otherwise = 0)**
$Age40_n$	**(is a parameter vector to be estimated for Dummy, If Age less than 40 =1, otherwise = 0)**
ASC_n	**(is a parameter vector to be estimated for Alternative Specific Constant)**

Table 6.1 RP Mode Model

	Model A1	ModelA2	ModelA3	ModelA4	ModelA5	ModelA6	ModelA7	ModelA8
Observations	121	121	121	121	121	121	121	121
Final log (L)	-107.4	-97.7	-87.0	-80.0	-82.7	-82.8	-80.1	-80.7
$Rho^2(c)$	0.114	0.194	0.282	0.340	0.317	0.317	0.339	0.334
VOT (£/h)	36	34	68	66	59	58	41	34
Time	**-0.0143 (-3.6)**	**-0.0140 (-3.3)**	**-0.0149 (-3.1)**	**-0.0154 (-2.8)**	**-0.0138 (-2.7)**	**-0.0137 (-2.7)**	**-0.0143 (-2.8)**	**-0.0136 (-2.7)**
Cost	**-2.4e-4 (-1.9)**	**-2.5e-4 (-1.9)**	-1.3e-4 (-0.9)	-1.4e-4 (-0.9)	-1.4e-4 (-0.9)	-1.4e-4 (-0.9)	-2.1e-4 (-1.3)	-2.4e-4 (-1.5)
BPur2	1.13 (1.5)	1.18 (1.6)	1.08 (1.3)	1.14 (1.2)	1.06 (1.2)	1.07 (1.2)	0.703 (0.8)	0.723 (0.8)
LFreq2	**1.97 (2.4)**	**1.91 (2.3)**	**2.43 (2.8)**	1.66 (1.5)	1.67 (1.5)	1.70 (1.6)	1.43 (1.3)	1.49 (1.4)
BPur3	0.184 (0.4)	-0.111 (-0.2)	-0.448 (-0.7)	-0.412 (-0.6)	-0.288 (-0.5)	-0.264 (-0.5)	-0.513 (-0.9)	-0.467 (-0.8)
LFreq3	**1.32 (2.7)**	**1.12 (2.1)**	**1.42 (2.5)**	**1.31 (2.0)**	**1.16 (1.9)**	1.15 (1.8)	1.04 (1.6)	0.925 (1.4)
ASC2	**-2.68 (-2.6)**	**-2.92 (-1.8)**	-16.6 (-0.1)	-15.8 (-0.1)	-15.3 (-0.1)	-15.4 (-0.1)	-14.3 (-0.1)	-14.5 (-0.1)
ASC3	-0.968 (-1.7)	1.04 (1.1)	-0.391 (-0.2)	1.30 (0.5)	0.921 (0.4)	0.954 (0.5)	2.77 (1.4)	2.17 (1.1)
GWeather2		0.373 (0.3)	0.962 (0.7)	0.705 (0.5)	0.527 (0.4)	0.553 (0.4)	0.288 (0.2)	0.160 (0.1)
PeakP2		-0.260 (-0.4)	-0.197 (-0.3)	-0.112 (-0.1)	-0.0693 (-0.1)			
GWeather3		**-2.20 (-3.2)**	**-2.39 (-2.9)**	**-2.45 (-3.0)**	**-2.12 (-2.8)**	**-2.11 (-2.8)**	**-2.45 (-3.1)**	**-2.31 (-3.0)**
PeakP3		0.339 (0.7)	0.310 (0.6)	0.233 (0.4)	0.0953 (0.2)			
TraAlone2			12.0 (0.1)	12.4 (0.1)	14.0 (0.1)	14.0 (0.1)	14.3 (0.1)	14.2 (0.1)
ImpSfty2			1.12 (0.0)	1.26 (0.0)				
ImpCmfrt2			0.126 (0.2)	0.797 (0.9)				
impSeats2			0.126 (0.1)	0.0328 (0.0)				
TraAlone3			**2.14 (2.2)**	**2.07 (2.1)**	**1.73 (2.2)**	**1.75 (2.2)**	**1.86 (2.3)**	**1.91 (2.4)**

ImpSfty3			-1.79 (-0.9)	-1.84 (-0.9)				
ImpCmfrt3			0.0019 (0.0)	-0.0332 (-0.1)				
impSeats3			**1.45 (1.8)**	**1.43 (1.8)**				
Male2				-0.0393 (-0.0)	-0.0553 (-0.1)	-0.0617 (-0.1)	-0.295 (-0.3)	-0.184 (-0.2)
Male3				0.172 (0.3)	0.112 (0.2)	0.118 (0.2)	-0.103 (-0.2)	-0.0266 (-0.0)
Educ2				0.357 (0.3)	0.327 (0.3)	0.316 (0.3)	0.0264 (0.0)	
FEmpl2				0.263 (0.3)	0.132 (0.1)	0.155 (0.2)	0.751 (0.8)	0.827 (0.8)
Income2				**-2.60 (-2.5)**	**-2.36 (-2.5)**	**-2.38 (-2.5)**	**-2.16 (-2.3)**	**-2.17 (-2.3)**
Age2				-0.0662 (-0.0)	-0.218 (-0.1)	-0.262 (-0.1)		
Educ3				-0.772 (-0.9)	-0.773 (-0.9)	-0.780 (-0.9)	-1.08 (-1.3)	
FEmpl3				-0.349 (-0.6)	-0.427 (-0.7)	-0.438 (-0.7)	-0.154 (-0.2)	-0.289 (-0.4)
Income3				0.0719 (0.1)	0.0785 (0.1)	0.0881 (0.1)	0.182 (0.2)	-0.0977 (-0.1)
Age3				-0.633 (-0.5)	-0.651 (-0.5)	-0.659 (-0.5)		
Car1							1.51 (1.8)	1.34 (1.7)
Age402							0.296 (0.3)	0.203 (0.2)
Age403							-0.739 (-1.2)	-0.885 (-1.4)
Educ252								0.291 (0.4)
Educ253								-0.380 (-0.7)

As described earlier the models were developed with a minimal specification, which includes those variables considered as essential to any reasonable model (e.g. time and cost). An incremental approach was then used to incorporate other important variables. Statistical tests were used to verify the model and to reformulate the utility function. The focus was to find a more precise model that performs better statistically and consistent with priori expectations such as signs and magnitudes. The coefficients of the utility function can be estimated by Alogit. The *t* values of all variables, ρ^2, log likelihood with zero coefficients, and log likelihood at convergence can also be obtained by using the same package.

All the models were compared and tested by likelihood ratio tests. More complex models had lower t statistics less than simpler models. Table 6.1 presents specification of each of these models. The Value of time for the models 1,2,4 and 8 are similar to the long distance traveller's value of time as calculated by similar studies (33 Pound/hour for long distance travel for business, Dargay, 2010). This shows the values of time are within acceptable limits. The *t*-statistic values of some variables were not satisfactory.

The following results are worthy of note:

The alternative specific constants in all the models reflect the average effects of variables which are omitted in the model. The negative sign for coach implies natural aversion to go by coach (all else being equal). This suggests that users exhibit an inclination to travel by car in long distance journeys this looks reasonable as in long intercity journeys coach has longer travel time and is inconvenient for most of the travellers. Whereas, the alternate specific constant for train was not significant.

Travel Related Coefficients: As expected, *travel time* is negative in all models, which shows travellers prefer to select the alternative which offers the lowest travel time. The cost is also negative but becomes insignificant with more complicated models.
The Peak period travel by both coach and train is not statistically significant whereas travelling in good weather was significant for train, this shows if all being equal the travellers in good weather try to travel by car as in good weather driving is more enjoyable. If the frequency of the travel is lower, the people are more likely to travel by

train and coach. This seems plausible as in high frequency travel people know all the routes and try to travel by car.

Traveller Related Coefficients: Income is a socioeconomic characteristic that significantly influences the rejection of coach as the travel mode. This means that higher income travellers don't like to travel by coach in long journeys, this is again is intuitively reasonable since coach being more uncomfortable in long journeys, is unattractive (all else being equal). Travellers travelling alone in long journeys prefer train as compared to other modes. This is reasonable as in long journeys travelling by car becomes cheaper if travelling with friends or family.

6.4.2 RP Model for Source Choice

Multinomial models were estimated for the choice of mode in RP data. As mentioned earlier the data consists of respondents' past behaviour for the last long journey (over 50 miles). Variables used in the model were selected as discussed in the previous chapter and were added in the model incrementally. A Multinomial Logit Model (MNL) is constructed with the dependent variable being the choice among different important information sources as gathered from the data i.e. onlyweb, friend+web, map+web, and map+friend+web. The choice set C_n of each individual thus consists of four alternatives. The utility functions are given by,

1.

$$U_{(onlyweb)} = \mathrm{LFreq}_1 * DLFreq_1 + \mathrm{GWeather}_1 * DGWeather_1 + \mathrm{PeakP}_1 * DPeakP_1 + \mathrm{Subs}_1 * DSubs_1 + \mathrm{ImpAd}_1 * DImpAd_1 + \mathrm{ImpRealt}_1 * DImpRealt_1 + \mathrm{ImpGInfo}_1 * DImpGInfo_1 + \mathrm{Male}_1 * DMale_1 + \mathrm{Educ}_1 * DEduc_1 + \mathrm{FEmpl}_1 * DFEmpl_1 + \mathrm{Income}_1 * DIncome_1 + \mathrm{Age}_1 * DAge_1$$

2.

$$U_{(friweb)} = \mathrm{LFreq}_2 * DLFreq_2 + \mathrm{GWeather}_2 * DGWeather_2 + \mathrm{PeakP}_2 * DPeakP_2 + \mathrm{Subs}_2 * DSubs_2 + \mathrm{ImpAd}_2 * DImpAd_2 + \mathrm{ImpRealt}_2 * DImpRealt_2 + \mathrm{ImpGInfo}_2 * DImpGInfo_2 + \mathrm{Male}_2 * DMale_2 + \mathrm{Educ}_2 * DEduc_2 + \mathrm{FEmpl}_2 * DFEmpl_2 + \mathrm{Income}_2 * DIncome_2 + \mathrm{Age}_2 * DAge_2 + \mathrm{ASC}_2$$

3. $U_{(mapweb)}$ = $LFreq_3*DLFreq_3$ + $GWeather_3*DGWeather_3$ + $PeakP_3*DPeakP_3$ + $Subs_3*DSubs_3$ + $ImpAd_3*DImpAd_3$ + $ImpRealt_3*DImpRealt_3$ + $ImpGInfo_3*DImpGInfo_3$ + $Male_3*DMale_3$ + $Educ_3*DEduc_3$ + $FEmpl_3*DFEmpl_3$ + $Income_3*DIncome_3$ + Age_3*DAge_3 + ASC_3

4. $U_{(mpfrweb)}$ = $LFreq_4*DLFreq_4$ + $GWeather_4*DGWeather_4$ + $PeakP_4*DPeakP_4$ + $Subs_4*DSubs_4$ + $ImpAd_4*DImpAd_4$ + $ImpRealt_4*DImpRealt_4$ + $ImpGInfo_4*DImpGInfo_4$ + $Male_4*DMale_4$ + $Educ_4*DEduc_4$ + $FEmpl_4*DFEmpl_4$ + $Income_4*DIncome_4$ + Age_4*DAge_4 + ASC_4

5. $U_{(onlymap)}$ = ASC_5

Variables

Where: (with subscript n indicating information sources)

$DCar_n$	***(Dummy, if car is available in household = 1, otherwise = 0)***
$DBPur_n$	***(Dummy, Business Purpose = 1, otherwise = 0)***
$DLFreq_n$	***(Dummy, Trip Frequency less than 13/year =1, otherwise = 0)***
$DGWeather_n$	***(Dummy, Good weather =1, otherwise = 0)***
$DIncident_n$	***(Dummy, Incident occurred =1, otherwise = 0)***
$DPeakP_n$	***(Dummy, Travelled in the Peak Period = 1, otherwise = 0)***
$DSubs_n$	***(Dummy, Subscribed to a website = 1, otherwise = 0)***
$DImpAd_n$	***(Dummy, free from Advertisements important = 1, otherwise = 0)***
$DImpRealt_n$	***(Dummy, Real time information important = 1, otherwise = 0)***
$DImpGInfo_n$	***(Dummy, General information important = 1, otherwise = 0)***
$DMale_n$	***(Dummy, If male =1, otherwise = 0)***
$DEduc_n$	***(Dummy, Left full time education at or after 20 = 1, otherwise = 0)***
$DFEmpl_n$	***(Dummy, Full time employed =1, otherwise = 0)***
$DIncome_n$	***(Dummy, If income over £30,000 =1, otherwise = 0)***
$DAge_n$	***(Dummy, If Age less than 50 =1, otherwise = 0)***

Coefficients

Where: (with subscript n indicating information sources)

Car_n **(is a parameter vector to be estimated for Dummy, if car is available in household = 1, otherwise = 0)**

$BPur_n$	**(is a parameter vector to be estimated for Dummy, Business Purpose = 1, otherwise = 0)**
$LFreq_n$	**(is a parameter vector to be estimated for Dummy, Trip Frequency less than 13/year =1, otherwise = 0)**
$GWeather_n$	**(is a parameter vector to be estimated for Dummy, Good weather =1, otherwise = 0)**
$Incident_n$	**(is a parameter vector to be estimated for Dummy, Incident occurred =1, otherwise = 0)**
$PeakP_n$	**(is a parameter vector to be estimated for Dummy, Travelled in the Peak Period = 1, otherwise = 0)**
$Subs_n$	**(is a parameter vector to be estimated for Dummy, Subscribed to a website = 1, otherwise = 0)**
$ImpAd_n$	**(is a parameter vector to be estimated for Dummy, free from Advertisements important = 1, otherwise = 0)**
$ImpRealt_n$	**(is a parameter vector to be estimated for Dummy, Real time information important = 1, otherwise = 0)**
$ImpGInfo_n$	**(is a parameter vector to be estimated for Dummy, General information important = 1, otherwise = 0)**
$Male_n$	**(is a parameter vector to be estimated for Dummy, If male =1, otherwise = 0)**
$Educ_n$	**(is a parameter vector to be estimated for Dummy, Left full time education at or after 20 = 1, otherwise = 0)**
$FEmpl_n$	**(is a parameter vector to be estimated for Dummy, Full time employed =1, otherwise = 0)**
$Income_n$	**(is a parameter vector to be estimated for Dummy, If income over £30,000 =1, otherwise = 0)**
Age_n	**(is a parameter vector to be estimated for Dummy, If Age less than 50 =1, otherwise = 0)**
ASC_n	**(is a parameter vector to be estimated for Alternative Specific Constant)**

Model estimates are shown in table 6.2. The *t*-statistic values of some variables were not satisfactory. The following results are worthy of note:

The alternative specific constants in the models reflect the average effects of variables which are omitted in the model. The negative sign for Alternative specific constant in all models for friend + web implies that if all else being equal, travellers that are website users believe more in the information provided by the website alone as compared to the information provided by friends. This suggests that travellers consider websites as reliable source of travel information. On the other hand travellers that use

Table 6.2 RP Source Model

File	**Soumodel1**	**Soumodel2**	**Soumodel3**	**Soumodel4**	**Soumodel5**
Observations	124	124	124	124	124
Final log (L)	-143.6	-145.9	-145.1	-140.2	-120.9
D.O.F	20	16	20	32	52
Rho²(0)	0.280	0.269	0.273	0.298	0.394
Rho²(c)	0.103	0.088	0.093	0.124	0.245
BPur1	-12.2 (-0.0)				
LFreq1	0.386 (0.4)	0.561 (0.7)	0.366 (0.4)	0.510 (0.5)	0.933 (0.6)
GWeather1	-0.629 (-0.5)	-0.465 (-0.4)	-0.625 (-0.5)	-0.258 (-0.2)	0.803 (0.5)
PeakP1	-1.03 (-1.1)	-1.27 (-1.4)	-1.23 (-1.4)	-1.65 (-1.7)	**-2.91 (-1.9)**
BPur2	-12.4 (-0.0)				
LFreq2	**2.69 (2.4)**	**2.89 (2.7)**	**2.75 (2.5)**	**2.83 (2.5)**	2.79 (1.6)
GWeather2	0.136 (0.1)	0.316 (0.3)	0.216 (0.2)	0.459 (0.4)	1.17 (0.7)
PeakP2	-0.470 (-0.5)	-0.732 (-0.8)	-0.665 (-0.7)	-1.08 (-1.1)	-2.48 (-1.6)
BPur3	-12.5 (-0.0)				
LFreq3	-0.0446 (-0.0)	0.157 (0.2)	-0.0683 (-0.1)	-0.0241 (-0.0)	1.35 (0.8)
GWeather3	0.591 (-0.5)	-0.390 (-0.3)	-0.603 (-0.5	-0.204 (-0.2)	1.16 (0.6)
PeakP3	-0.733 (-0.7)	-1.02 (-1.0)	-0.982 (-1.0)	-1.55 (-1.4)	**-2.92 (-1.8)**
BPur4	-12.7 (-0.0)				
LFreq4	1.69 (1.2)	1.87 (1.4)	1.69 (1.2)	1.55 (1.1)	1.76 (0.9)
GWeather4	12.5 (0.0)	10.7 (0.1)	10.5 (0.1)	10.9 (0.1)	13.5 (0.0)
PeakP4	-2.13 (-1.5)	**-2.44 (-1.8)**	-2.39 (-1.7)	**-3.13 (-2.1)**	**-3.76 (-1.9)**
ASC2	**-3.30 (-3.1)**	**-3.42 (-3.4)**	**-3.61 (-3.4)**	**-3.44 (-2.1)**	-3.47 (-1.5)
ASC3	-1.16 (-1.2)	-1.35 (-1.6)	-1.24 (-1.3)	-12.9 (-0.0)	-18.9 (-0.0)
ASC4	-15.3 (-0.0)	-13.7 (-0.1)	-13.6 (-0.1)	-24.5 (-0.1)	-40.6 (-0.0)
ASC5	-15.0 (-0.1)	**-2.96 (-2.1)**	**-3.45 (-2.3)**	-1.09 (-0.6)	-20.4 (-0.0)
Subs1			-0.887 (-1.0)	-1.37 (-1.4)	-1.57 (-1.2)
Subs2			-0.542 (-0.6)	-0.802 (-0.7)	-1.05 (-0.7)
Subs3			-1.10 (-1.1)	-1.65 (-1.4)	-2.14 (-1.5)
Subs4			-1.10 (-0.8)	-1.18 (-0.8)	-2.30 (-1.1)
ImpAd1				1.23 (1.3)	1.81 (1.4)
ImpRealt1				1.53 (1.1)	1.83 (0.9)
ImpGInfo1				0.468 (0.5)	0.893 (0.7)
ImpAd2				0.439 (0.4)	1.09 (0.8)
ImpRealt2				1.99 (1.2)	2.35 (1.0)
ImpGInfo2				0.514 (0.5)	0.762 (0.6)
ImpAd3				1.13 (1.0)	1.47 (1.0)
ImpRealt3				13.4 (0.0)	16.1 (0.0)
ImpGInfo3				0.573 (0.6)	0.881 (0.6)
ImpAd4				0.218 (0.2)	0.883 (0.6)
ImpRealt4				13.0 (0.0)	14.9 (0.0)
ImpGInfo4				1.11 (0.9)	1.34 (0.9)
Male1					0.0753 (0.1)
Male2					0.0304 (0.0)
Male3					2.49 (1.6)
Male4					-1.21 (-0.7)
Educ1					**3.45 (2.4)**
FEmpl1					-11.8 (-0.0)
Income1					0.0337 (0.0)
Age1					-11.3 (-0.0)
Educ2					**3.56 (2.3)**
FEmpl2					-12.6 (-0.0)
Income2					-0.275 (-0.2)
Age2					-9.97 (-0.0)
Educ3					**3.84 (2.2)**
FEmpl3					-11.2 (-0.0)
Income3					1.45 (1.0)
Age3					-11.8 (-0.0)
Educ4					**4.83 (2.3)**
FEmpl4					-10.4 (-0.0)
Income4					-2.19 (-1.2)
Age4					1.45 (0.0)

websites are more satisfied with the information provided to them as compared to the information provided by the maps only. This suggests that users exhibit an inclination to gather information from the websites. Whereas, the alternative specific constants for map+web and map+friend+web were not found significant.

Information Source Related Coefficients: The analysis suggests that the frequent travellers require, in addition to website, the information from a person who has travelled before. The sign of PeakP in all models is negative which suggests that Peak period travellers do like to stay with the maps only. This implies that in peak periods, travellers are more likely to take diversions from the usual route and thus require maps while they travel.

Traveller Related Coefficients: Education is a socioeconomic characteristic that significantly influence the choice of information sources. It appears that people who left their education after the age of 20 like to use all the information sources including websites. This result is as per priori expectation that higher educated people are at ease with technology and are more frequent users of websites.

6.4.3 SP Model for Mode Choice

Multinomial models were calibrated for the choice of mode in the SP data. As mentioned earlier the data consists of respondent's stated choice between the three mode alternatives under the influence of information. Variables used in the model were selected as discussed in the previous chapter and were added in the model incrementally. A Multinomial Logit Model (MNL) is constructed with the dependent variable being the choice among car, coach and train. The base case for this model travelling by coach. The choice set C_n of each individual thus consists of three alternatives. The utility functions are given by,

1. $U_{(car)} = \mathrm{Carfrit}_1 * Carfrit_1 + \mathrm{Carmult}_1 * Carmult_1 + \mathrm{Carfric}_1 * Carfric_1 + \mathrm{Carmulc}_1 * Carmulc_1 + \mathrm{Male}_1 * DMale_1 + \mathrm{FEmp}_1 * DFEmp_1 + \mathrm{DInc}_1 * DInc_1 + \mathrm{Age}_1 * DAge_1$

2. $U_{(coach)} = \text{Coawt}_2 * Coawt_2 + \text{Coamult}_2 * Coamult_2 + \text{Coawc}_2 * Coawc_2 + \text{Coamulc}_2 * Coamulc_2$

3. $U_{(train)} = \text{Trainwt}_3 * Trainwt_3 + \text{Trainext}_3 * Trainext_3 + \text{Trainwc}_3 * Trainwc_3 + \text{Trainexc}_3 * Trainexc_3 + \text{Male}_3 * DMale_3 + \text{FEmp}_3 * DFEmp_3 + \text{Inc}_3 * DInc_3 + \text{Age}_3 * DAge_3$

Variables

Where: (with subscript n indicating modes)

$Carfrit_n$	***(Time by car information by friend, in Minutes)***
$Carmult_n$	***(Time by car information by multimodal website, in Pence)***
$Carfric_n$	***(Cost by car information by friend, in Minutes)***
$Carmultc_n$	***(Cost by car information by multimodal website, in Pence)***
$DMale_n$	***(Dummy, If male =1, otherwise = 0)***
$DFEmp_n$	***(Dummy, if full employment = 1, otherwise = 0)***
$DInc_n$	***(Dummy, if income over 30,000 pounds =1, otherwise = 0)***
$DAge_n$	***(Dummy, if age less than 40 =1, otherwise = 0)***
$Trainwt_n$	***(Time by train information by train website, in Minutes)***
$Trainext_n$	***(Time by train information by past experience, in Minutes)***
$Trainwc_n$	***(Cost by train information by train website, in Pence)***
$Trainexc_n$	***(Cost by train information by past experience, in Pence)***
$Coawt_n$	***(Time by coach information by coach website, in Minutes)***
$Coawc_n$	***(Cost by coach information by coach website, in Pence)***
$Coamult_n$	***(Time by coach information by multimodal website, in Minutes)***
$Coamulc_n$	***(Cost by coach information by multimodal website, in Pence)***

Coefficients

Where: (with subscript n indicating modes)

Carfrit_n	**(is a parameter vector to be estimated for Time by car information by friend, in Minutes)**
Carmult_n	**(is a parameter vector to be estimated for Time by car information by multimodal website, in Pence)**
Carfric_n	**(is a parameter vector to be estimated for Cost by car information by friend, in Minutes)**
Carmultc_n	**(is a parameter vector to be estimated for Cost by car information by multimodal website, in Pence)**
Male_n	**(is a parameter vector to be estimated for Dummy, If male =1, otherwise = 0)**
FEmp_n	**(is a parameter vector to be estimated for Dummy, if full employment = 1, otherwise = 0)**

Inc_n	**(is a parameter vector to be estimated for Dummy, if income over 30,000 pounds =1, otherwise = 0)**
Age_n	**(is a parameter vector to be estimated for Dummy, if age less than 40 =1, otherwise = 0)**
$Trainwt_n$	**(is a parameter vector to be estimated for Time by train information by train website, in Minutes)**
$Trainext_n$	**(is a parameter vector to be estimated for Time by train information by past experience, in Minutes)**
$Trainwc_n$	**(is a parameter vector to be estimated for Cost by train information by train website, in Pence)**
$Trainexc_n$	**(is a parameter vector to be estimated for Cost by train information by past experience, in Pence)**
$Coawt_n$	**(is a parameter vector to be estimated for Time by coach information by coach website, in Minutes)**
$Coawc_n$	**(is a parameter vector to be estimated for Cost by coach information by coach website, in Pence)**
$Coamult_n$	**(is a parameter vector to be estimated for Time by coach information by multimodal website, in Minutes)**
$Coamulc_n$	**(is a parameter vector to be estimated for Cost by coach information by multimodal website, in Pence)**

Model estimates are shown in Table 6.3.

All the models were compared and tested by likelihood ratio tests. More complex models had lower t-statistics whereas simpler models were performing better. Table 6.3 presents specification of each of these models. The *t*-statistic values of some variables were not satisfying the requirements this could be due to the presence of repeated measurement problem in SP data. This issue will be dealt in the next chapter.

The Following results are worthy of note:

Travel Related Coefficients: As expected, travel time, and cost by different information sources for all the models were negative, which shows travellers will prefer any alternative which offers the lowest expected travel time and cost. The t stats for the time and cost for the initial models were not significant but when the coefficients for each information sources was constrained to be same for all modes, the models t stats improved. The value of time for the Model 5are 24, 18 and 10 £/hour for Car, train and coach respectively. The values are similar to the corresponding long distance journey studies of 24, 28 and 10 respectively for the long distance journey of more than 150 miles. (Dargay, 2010).

Table 6.3 SP Mode Models

File	ModelSPM1	ModelSPM2	ModelSMP3	ModelSMP4	ModelSMP5
Converged	True	True	True	True	True
Observations	1143	1143	1143	1143	1143
Final log (L)	-613.7	-613.1	-612.9	-613.1	-534.8
D.O.F.	8	10	12	11	17
$Rho^2(0)$	0.225	0.226	0.226	0.226	0.325
$Rho^2(c)$	0.080	0.081	0.081	0.081	0.198
Carfrit1	-0.0044 (-1.4)	-0.0060 (-1.7)	-0.0075 (-1.6)	-0.0059 (-1.6)	-0.0070 (-1.7)
Carmult1			-0.0012 (-0.3)		
CarCoamult	**-0.0040 (-2.5)**	-0.0033 (-1.7)		-0.0033 (-1.7)	**-0.0047 (-2.5)**
Carfric1	**-2.4e-4 (-1.8)**	**-2.3e-4 (-1.8)**	-2.0e-4 (-1.4)	-2.3e-4 (-1.7)	**-2.8e-4 (-2.0)**
Carmulc1			-7.3e-5 (-0.9)	-9.2e-5 (-1.2)	
CarCoamulc	**-1.2e-4 (-1.9)**	-9.0e-5 (-1.3)			**-1.2e-4 (-1.8)**
Trainwt3		-0.0043 (-0.8)	2.4e-5 (0.0)	-0.0044 (-0.8)	
CoaTrawt	**-0.0073 (-3.5)**				**-0.0056 (-2.3)**
Trainext3	-3.2e-4 (-0.1)	-0.0042 (-0.8)	-0.0078 (-0.9)	-0.0041 (-0.7)	-0.0049 (-0.8)
Trainwc3		**-1.6e-4 (-3.4)**	**-1.5e-4 (-2.7)**	**-1.6e-4 (-3.3)**	**-1.9e-4 (-3.5)**
CoaTrawc	**-1.8e-4 (-4.1)**				
Trainexc3		-1.1e-4 (-1.7)	-9.2e-5 (-1.2)	-1.1e-4 (-1.7)	**-1.2e-4 (1.8)**
Coawt2	**-1.1e-4 (-1.8)**	**-0.0065 (-2.9)**	-0.0050 (-1.3)	**-0.0066 (-2.6)**	
Coamult2			-0.0047 (-1.5)		
Coawc2		**-3.3e-4 (-2.2)**	**-3.0e-4 (-1.8)**	**-3.3e-4 (-2.1)**	**-3.5e-4 (-2.1)**
Coamulc2			-3.3e-5 (-0.2)	-8.0e-5 (-0.4)	
Male1					**1.38 (3.8)**
FEmp1					**1.22 (3.5)**
Inc1					**1.04 (2.6)**
Age1					-1.13 (-1.5)
Male3					0.192 (0.6)
FEmp3					**0.584 (2.1)**
Inc3					**2.10 (6.2)**
Age3					-1.06 (-1.5)

Traveller Related Coefficients: Being male, in full time employment, and with higher income increased the propensity to travel by car and train all else being equal. This is in line with the expectation that coach is inconvenient for longer journeys. Age on the other hand was negative which means that the younger people like to travel by coach, this is again as expected because coach, being relatively cheap, is attractive to younger people and students.

6.4.4 SP Model for Source Choice

For the Source choice, respondents were given two websites with different attributes as mentioned in section 6.2. Multinomial models were calibrated for the choice of website as source to analyse different important attributes and their influence on travellers. Variables used in the model were selected as discussed in the previous chapter and were added in the model incrementally. A Multinomial Logit Model (MNL) is constructed with the dependent variable being the choice among website one and website two. The base case for this model is website 1. The choice set C_n of each individual thus consists of two alternatives. The utility functions are given by,

1. $U_1 = \text{Seatime}_1 * Seatime_1 + \text{MaleStim}_1 * DMale_1 * Seatime_1 + \text{EducStim}_1 * DEducation_1 * Seatime_1 + \text{FEmpStim}_1 * DFEmpl_1 * Seatime_1 + \text{IncStim}_1 * DIncome_1 * Seatime_1 + \text{YgStim}_1 * DYoung_1 * Seatime_1 + \text{ComAd}_1 * DComAd_1 + \text{ComAdsub}_1 * DComAdsub_1 + \text{Com}_1 * DCom_1 + \text{InfoSpec}_1 * DInfoSpec_1 + \text{InfoPersc}_1 * DInfoPersc_1 + \text{Updaily}_1 * DUpdaily_1 + \text{Upweekly}_1 * DUpweekly_1$

2. $U_2 = \text{Seatime}_2 * Seatime_2 + \text{MaleStim}_2 * DMale_2 * Seatime_2 + \text{EducStim}_2 * DEducation_2 * Seatime_2 + \text{FEmpStim}_2 * DFEmpl_2 * Seatime_2 + \text{IncStim}_2 * DIncome_2 * Seatime_2 + \text{YgStim}_2 * DYoung_2 * Seatime_2 + \text{ComAd}_2 * DComAd_2 + \text{ComAdsub}_2 * DComAdsub_2 + \text{Com}_2 * DCom_2 + \text{InfoSpec}_2 * DInfoSpec_2 + \text{InfoPersc}_2 * DInfoPersc_2 + \text{Updaily}_2 * DUpdaily_2 + \text{Upweekly}_2 * DUpweekly_2$

Variables

Where: (with subscript n websites)

$DComAd_n$:	***(dummy that equals 1 if Commercial Ads No Sub; otherwise 0)***
$DComAdSub_n$:	***(dummy that equals 1 if Commercial Ads Subs; otherwise 0)***
$DCom_n$:	***(dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)***
$Seatime_n$	***(5 min, 10 min, 15 min)***
$DUpdaily_n$:	***(dummy that equals 1 if Website updates daily; otherwise 0)***
$DUpweekly_n$:	***(dummy that equals 1 if Website updates weekly; otherwise 0)***
$DInfoSpec_n$:	***(dummy that equals 1 if Specific info available; otherwise 0)***
$DInfoPerc_n$:	***(dummy that equals 1 if Info w.r.t own criteria; otherwise 0)***
$DMale_n$	***(Dummy, If male =1, otherwise = 0)***
$DEducation_n$	***(Dummy, Left full time education at 20 = 1, otherwise = 0)***
$DFEmpl_n$	***(Dummy, Full time employed =1, otherwise = 0)***
$DIncome_n$	***(Dummy, If income over £30,000 =1, otherwise = 0)***
$DYoung_n$	***(Dummy, If Age less than 40 =1, otherwise = 0)***

Coefficients

Where: (with subscript n indicating modes)

$ComAd_n$:	**(dummy that equals 1 if Commercial Ads No Sub; otherwise 0)**
$ComAdSub_n$:	**(dummy that equals 1 if Commercial Ads Subs; otherwise 0)**
Com_n:	**(dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)**
$Seatime_n$	**(5 min, 10 min, 15 min)**
$Updaily_n$:	**(dummy that equals 1 if Website updates daily; otherwise 0)**
$Upweekly_n$:	**(dummy that equals 1 if Website updates weekly; otherwise 0)**
$InfoSpec_n$:	**(dummy that equals 1 if Specific info available; otherwise 0)**
$InfoPerc_n$:	**(dummy that equals 1 if Info w.r.t own criteria; otherwise 0)**
$MaleStim_n$	**(Dummy, If male =1, otherwise = 0)**
$EducStim_n$	**(Dummy, Left full time education at 20 = 1, otherwise = 0)**
$FEmpStim_n$	**(Dummy, Full time employed =1, otherwise = 0)**
$IncStim_n$	**(Dummy, If income over £30,000 =1, otherwise = 0)**
$YgStim_n$	**(Dummy, If Age less than 40 =1, otherwise = 0)**

Model estimates are shown in Table 6.4.

Table 6.4 SP Source Models

File	ModelSPS1	ModelSPS2	ModelSPS3	ModelSPS4	ModelSPS5	ModelSPS6	ModelSPS7	ModelSPS8	ModelSPS9	ModelSPS10
Converged	True	True	True	True	True	True	True	True	True	True
Observations	1056	1056	1056	1056	1056	1056	1056	1056	1056	1056
Final log (L)	-685.0	-593.5	-413.8	-577.3	-436.8	-436.7	-436.7	-436.0	-435.0	-554.6
D.O.F.	2	6	10	8	12	13	13	14	17	11
Rho²(0)	0.064	0.189	0.435	0.211	0.403	0.403	0.403	0.404	0.406	0.242
Rho²(c)	0.064	0.189	0.434	0.211	0.403	0.403	0.403	0.404	0.406	0.242
Scaling	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Seatime1	**-0.104 (-9.0)**	**-0.0756 (-4.0)**	**0.308 (4.4)**	**-0.185 (-6.0)**	**-0.971 (-11.3)**	**-0.9674 (-11.3)**	**-0.971 (-11.3)**	**-0.973 (-11.3)**	**-0.975 (-11.3)**	**-0.165 (-11.1)**
Seatime2	**-0.101 (-8.7)**	**-0.0650 (-5.2)**	**-0.0399 (-2.0)**	**-0.0373 (-2.6)**	**0.629 (7.4)**	**0.630 (7.4)**	**0.630 (7.4)**	**0.630 (7.4)**	**0.631 (7.4)**	
InfoSpec1		0.203 (1.0)	**-4.15 (-5.1)**	**1.26 (4.6)**	**1.88 (4.3)**	**1.86 (4.3)**	**1.88 (4.3)**	**1.88 (4.3)**	**1.89 (4.3)**	**0.237 (1.8)**
InfoPersc1		**3.46 (7.7)**	**-5.61 (-3.3)**	**5.34 (9.3)**	-0.0232 (-0.0)		-0.0313 (-0.0)	-0.0235 (-0.0)	-0.0237 (-0.0)	**1.01 (2.1)**
infoSpec2		**0.609 (4.0)**	**6.18 (6.2)**	-0.342 (-1.5)	**4.05 (5.3)**	**4.08 (5.3)**	**4.06 (5.3)**	**4.06 (5.3)**	**4.07 (5.3)**	
infoPersc2		**2.54 (5.9)**	**3.95 (6.9)**	**1.81 (3.7)**	-0.0787 (-0.1)		-0.0755 (-0.1)	-0.0795 (-0.1)	-0.0833 (-0.1)	
Updaily1		**5.09 (4.4)**	**-1.53 (-5.4)**	-0.570 (-1.2)	-0.567 (-1.2)	-0.549 (-1.2)	-0.572 (-1.2)	-0.577 (-1.2)	**-0.305 (-2.7)**	**-0.305 (-2.7)**
Upweekly1		0.0156 (0.0)								
Updaily2		-0.0751 (-0.2)	**-0.847 (-3.6)**	**-1.83 (-5.5)**	**-1.83 (-5.5)**	**-1.81 (-5.5)**	**-1.83 (-5.5)**	**-1.84 (-5.5)**		
Upweekly2		**-7.00 (-9.5)**								
ComAd1				0.256 (1.0)	0.258 (1.0)	0.265 (1.0)	0.257 (1.0)	0.257 (1.0)	**-1.47 (-7.3)**	**-1.47 (-7.3)**
Com1				**-1.26 (-1.8)**	**-1.28 (-1.8)**	**-1.267 (-1.8)**	**-1.27 (-1.8)**	**-1.27 (-1.8)**	**0.691 (3.7)**	**0.691 (3.7)**
ComAdSub2					**-17.5 (-8.8)**	**-17.5 (-8.8)**	**-17.5 (-8.8)**	**-17.8 (-8.8)**	**-18.1 (-8.9)**	
Com2				**-15.2 (-7.6)**	**-15.2 (-7.6)**	**-15.2 (-7.6)**	**-15.4 (-7.7)**	**-15.8 (-7.7)**		
MaleStim11					0.0741 (0.4)		0.0629 (0.4)	0.0951 (0.5)	0.189 (1.3)	
Male11										
EducStim1							-0.279 (-1.2)	-0.241 (-0.9)	-0.0065 (-0.0)	
FEmpStim1							-0.120 (-0.6)	-0.0883 (-0.5)		
IncStim1							0.0992 (0.5)	0.134 (0.8)		
YgStim1							-0.377 (-1.1)	0.142 (0.6)		
MInfoSpec1										0.189 (1.3)
EdcInfoSpec1										-0.0065 (-0.0)
FEpinfoSpec1										-0.0883 (-0.5)
IncinfoSpec1										0.134 (0.8)
YginfoSpec1										0.142 (0.6)
MInfoPersc1										0.189 (1.3)
EdcinfoPerc1										-0.0065 (-0.0)
FEpInfoPerc1										-0.0883 (-0.5)
IncInfoPerc1										0.134 (0.8)
YginfoPerc1										0.142 (0.6)

As the alternatives in the model are only "website one" and "website two", different variables were interacted with each other to reveal the full effect of the different attributes. Search time, information update and type of information were, in turn, used as interactive variables and multiplied with the other dummy variables. Table 6.4 presents the specification of each of these models. The *t*-statistic values of some variables were not significant.

Key properties of the estimated coefficients are discussed below:

Website Related Coefficients: As expected, search time for the website was found significant and negative which shows that, other things being equal, travellers will prefer an alternative which offers the lowest expected search time. Information type i.e. specific information available and information w.r.t own criteria were also found significant and positive which means more information increases the utility of a website. Updating of information was also as per priori expectation and negative which tells that if in comparison to real time information updating, daily and weekly have negative effect on the utility. Commercial Ads No Sub; Commercial Ads Subs and Commercial No Ads No Sub were also found negative which means the base case i.e. government with No advertisements and no subscription has higher attraction and credibility within the respondents.

Traveller Related Coefficients: Gender, full time employment and higher income and younger travellers were found insignificant – as were variables interacting with these dummies. This shows that the importance of information does not depend on the socioeconomic characteristics of the respondents.

6.4.5 RP MNL Model for Source and Mode Choice

Multinomial models were calibrated for the choice of mode and source together using the RP data. Variables used in the model were selected as discussed in the previous chapter and were added in the model incrementally. A Multinomial Logit Model (MNL) is constructed with the dependent variable being the combined choice of mode and source. The modes were car, coach and train whereas the sources used by respondents in their past journey were only web; friend with web; map with web; map with friend and web; and only map. Hence by combining these alternatives, the resulting combined alternatives should be fifteen but as map was not used by the coach and train travellers, the alternatives were reduced to nine. The base case for the dummy variables was travelling by car and using website as the source of information. Another base case of travelling by car and using map is also tested. The choice set C_n of each individual thus consists of nine alternatives. The utility functions for the nine alternatives are given by,

1. $U_{(carweb)} = \mathrm{Time}_1 * Time_1 + \mathrm{Cost}_1 * Cost_1$

2. $U_{(carfrweb)} = \mathrm{Time}_2 * Time_2 + \mathrm{Cost}_2 * Cost_2 + \mathrm{Male2} * DMale_2 + \mathrm{Educ}_2 * DEduc_2 + \mathrm{FEmpl}_2 * DFEmpl_2 + \mathrm{Income}_2 * DIncome_2 + \mathrm{Age40}_2 * DAge40_2$

3. $U_{(carmpweb)} = \mathrm{Time}_3 * Time_3 + \mathrm{Cost}_3 * Cost_3 + \mathrm{Male}_3 * DMale_3 + \mathrm{Educ}_3 * DEduc_3 + \mathrm{FEmpl}_3 * DFEmpl_3 + \mathrm{Income}_3 * DIncome_3 + \mathrm{Age40}_3 * DAge40_3$

4. $U_{(carmfweb)} = \mathrm{Time}_4 * Time_4 + \mathrm{Cost}_4 * Cost_4 + \mathrm{Male}_4 * DMale_4 + \mathrm{Educ}_4 * DEduc_4 + \mathrm{FEmpl}_4 * DFEmpl_4 + \mathrm{Income}_4 * DIncome_4 + \mathrm{Age40}_4 * DAge40_4$

5. $U_{(carmap)} = \mathrm{Time}_5 * Time_5 + \mathrm{Cost}_5 * Cost_5 + \mathrm{Male}_5 * DMale_5 + \mathrm{Educ}_5 * DEduc_5 + \mathrm{FEmpl}_5 * DFEmpl_5 + \mathrm{Income}_5 * DIncome_5 + \mathrm{Age40}_5 * DAge40_5$

6. $U_{(busweb)} = \mathrm{Time}_6 * Time_6 + \mathrm{Cost}_6 * Cost_6 + \mathrm{Male}_6 * DMale_6 + \mathrm{Educ}_6 * DEduc_6 + \mathrm{FEmpl}_6 * DFEmpl_6 + \mathrm{Income}_6 * DIncome_6 + \mathrm{Age40}_6 * DAge40_6$

7. $U_{(busfrweb)} = \mathrm{Time}_7 * Time_7 + \mathrm{Cost}_7 * Cost_7 + \mathrm{Male}_7 * DMale_7 + \mathrm{Educ}_7 * DEduc_7 + \mathrm{FEmpl}_7 * DFEmpl_7 + \mathrm{Income}_7 * DIncome_7 + \mathrm{Age40}_7 * DAge40_7$

8. $U_{(traweb)} = \mathrm{Time}_8 * Time_8 + \mathrm{Cost}_8 * Cost_8 + \mathrm{Male}_8 * DMale_8 + \mathrm{Educ}_8 * DEduc_8 + \mathrm{FEmpl}_8 * DFEmpl_8 + \mathrm{Income}_8 * DIncome_8 + \mathrm{Age40}_8 * DAge40_8$

9. $U_{(trafrweb)} = \mathrm{Time}_9 * Time_9 + \mathrm{Cost}_9 * Cost_9 + \mathrm{Male}_9 * DMale_9 + \mathrm{Educ}_9 * DEduc_9 + \mathrm{FEmpl}_9 * DFEmpl_9 + \mathrm{Income}_9 * DIncome_9 + \mathrm{Age40}_9 * DAge40_9$

Variables

Where: (with subscript n indicating alternatives)

Time	***(Generic, in Minutes)***
Cost	***(Generic, in Pence)***
$DCar_n$	***(Dummy, if car is available in household = 1, otherwise = 0)***
$DMale_n$	***(Dummy, If male =1, otherwise = 0)***
$DEduc_n$	***(Dummy, Left full time education at 20 = 1, otherwise = 0)***
$DEduc25_n$	***(Dummy, Left full time education at 25 = 1, otherwise = 0)***
$DFEmpl_n$	***(Dummy, Full time employed =1, otherwise = 0)***
$DIncome_n$	***(Dummy, If income over £30,000 =1, otherwise = 0)***
$DAge_n$	***(Dummy, If Age less than 50 =1, otherwise = 0)***
$DAge40_n$	***(Dummy, If Age less than 40 =1, otherwise = 0)***

Coefficients

Where: (with subscript n indicating alternatives)

Time	**(is a parameter vector to be estimated for Time, Generic, in Minutes)**
Cost	**(is a parameter vector to be estimated for Cost, Generic, in Pence)**
Car_n	**(is a parameter vector to be estimated for Dummy, if car is available in household = 1, otherwise = 0)**
$Male_n$	**(is a parameter vector to be estimated for Dummy, If male =1, otherwise = 0)**
$Educ_n$	**(is a parameter vector to be estimated for Dummy, Left full time education at 20 = 1, otherwise = 0)**
$Educ25_n$	**(is a parameter vector to be estimated for Dummy, Left full time education at 25 = 1, otherwise = 0)**
$FEmpl_n$	**(is a parameter vector to be estimated for Dummy, Full time employed =1, otherwise = 0)**
$Income_n$	**(is a parameter vector to be estimated for Dummy, If income over £30,000 =1, otherwise = 0)**
Age_n	**(is a parameter vector to be estimated for Dummy, If Age less than 50 =1, otherwise = 0)**
$Age40_n$	**(is a parameter vector to be estimated for Dummy, If Age less than 40 =1, otherwise = 0)**

Model estimates are shown in table 6.5. In table 6.5, models 5, 4, and 3 have car + map as base case for dummy variables where as model 8 has car + web as base case for dummy variables.

Table 6.5 RP Mode Source MNL Model

	ModelRPSM1	ModelRPSM2	ModelRPSM3	ModelRPSM4
Converged	True	True	True	True
Observations	226	226	226	226
Final log (L)	-476.5	-387.1	-488.1	-387.1
D.O.F	10	42	2	42
$Rho^2(0)$	0.040	0.220	0.017	0.220
$Rho^2(c)$	-0.035	0.159	-0.060	0.159
VOT	11	20	7.5	20
Time	-0.0020 (-0.8)	-0.0020 (-0.6)	-0.0021 (-1.0)	-0.0020 (-0.6)
Cost	**-1.8e-4 (-2.0**	-9.9e-5 (-0.9	**-2.8e-4 (-3.7)**	-9.9e-5 (-0.9)
Male1	0.405 (1.1)	0.830 (1.1)		
Male2	5e-16 (0.0)	-0.126 (-0.1)		-0.956 (-1.3)
Male3	**1.04 (3.1)**	0.433 (0.6)		-0.396 (-0.8)
Male4	0.154 (0.4)	-0.974 (-1.1)		**-1.80 (-2.4)**
Male6	0.108 (0.3)	-0.0318 (-0.0)		-0.862 (-1.4)
Male7	0.618 (1.6)	1.27 (1.7)		0.439 (0.8)
Male8	-0.299 (-0.6)	-1.38 (-1.6		**-2.21 (-2.8)**
Male9	-0.299 (-0.6)	-0.761 (-0.7)		-1.59 (-1.6)
Educ1		-1.25 (-1.2)		
FEmpl1		-0.910 (-1.2)		
Income1		**3.08 (3.3)**		
Age401		-0.113 (-0.2)		
Educ2		**-1.82 (-1.8)**		-0.563 (-0.6)
FEmpl2		-0.0336 (-0.0)		0.876 (1.2)
Income2		-0.870 (-0.8		**-3.95 (-3.7)**
Age402		**2.13 (2.5)**		**2.25 (2.8)**
Educ3		-0.236 (-0.3)		1.02 (1.3)
FEmpl3		-0.525 (-0.8)		0.385 (0.7)
Income3		**2.64 (3.3)**		-0.440 (-0.6)
Age403		0.0374 (0.1)		0.151 (0.3)
Educ4		-0.0862 (-0.1)		1.17 (1.1)
FEmpl4		**-1.65 (-2.2)**		-0.737 (-1.0)
Income4		0.387 (0.4)		**-2.69 (-3.0)**
Age404		**1.83 (2.0)**		**1.95 (2.3)**
Educ6		**-2.67 (-2.6)**		-1.42 (-1.6)
FEmpl6		**-1.69 (-2.2)**		-0.782 (-1.2)
Income6		**2.62 (2.8)**		-0.465 (-0.6)
Age406		**2.65 (3.1)**		**2.76 (3.6)**
Educ7		**-2.09 (-2.1)**		-0.840 (-1.0)
FEmpl7		-0.671 (-0.9)		0.239 (0.4)
Income7		**2.95 (3.2)**		-0.131 (-0.2)
Age407		0.524 (0.7)		0.638 (1.1)
Educ8		-0.337 (-0.3)		0.918 (0.9)
FEmpl8		0.582 (0.7)		**1.49 (1.9)**
Income8		-0.550 (-0.6)		**-3.63 (-3.8)**
Age408		1.20 (1.4)		1.31 (1.7)
Educ9		**-7.35 (-3.9)**		**-6.09 (-3.4)**
FEmpl9		**2.22 (2.0)**		**3.13 (2.9)**
Income9		**5.36 (3.1)**		2.28 (1.4)
Age409		-0.288 (-0.3)		-0.175 (-0.2)
Male5				-0.830 (-1.1)
Educ5				1.25 (1.2)
FEmpl5				0.910 (1.2)
Income5				**-3.08 (-3.3)**
Age405				0.113 (0.2)

Table 6.5 presents specification of each of these models. The *t*-statistic values of some variables were not satisfying the requirement.

The important properties of the estimated coefficients are discussed below.

Travel Related Coefficients: As expected, *travel time* is negative, but it is insignificant. The cost is also negative but becomes insignificant with more complicated models. This shows that when many taste variations are at stake the travel time and travel cost becomes less important to the travellers. The values of time are lower as compared to other similar studies; this can be due to the fact that the time coefficient is not significant in all models.

Traveller Related Coefficients: Model 1 shows that when travelling by car the males tend to choose to consult websites as well as maps (rather than only maps). Similarly in Model 4 in which car and website are base case, males are less like to travel by train if they use website as source of travel information.

Model 2 shows that income is also an important socioeconomic characteristic that significantly predisposes people to use website as a source of travel information. Similarly, travellers who left their education by the age of 20 are less likely to use website as source of travel information.

Model 4 shows that high income people tend to use website more as compared to other sources similarly if they used website as a source of information, they tend to travel by car. Model 4 also suggests that younger travellers care more for other sources in addition to websites. They also are prone to travel by train and bus if they got travel information from a website. Moreover travellers in full time employment are more likely to travel by train as compared to car.

6.4.6 RP NL Model for Source and Mode Choice

Nested Logit Models were calibrated for the combined choice of mode and source using the RP data. Variables used in the model were selected as discussed in the previous chapter and were added in the model incrementally. Like the Multinomial Logit Model (MNL), Nested Logit Models (NL) are also constructed with the dependent variable being the combined choice of mode and source. The modes were car, coach and train where as the sources used by respondents in their past journey were only web; friend with web; map with web; map with friend and web; and only map. Hence by combining these alternatives, the resulting combined alternatives should be fifteen but as map was not used by the coach and train travellers the alternatives were reduced to nine. The base case for this model was travelling by car and using website as source of information.

First the following nesting structure was tried. In this structure there is more substitution between modes than between information sources.

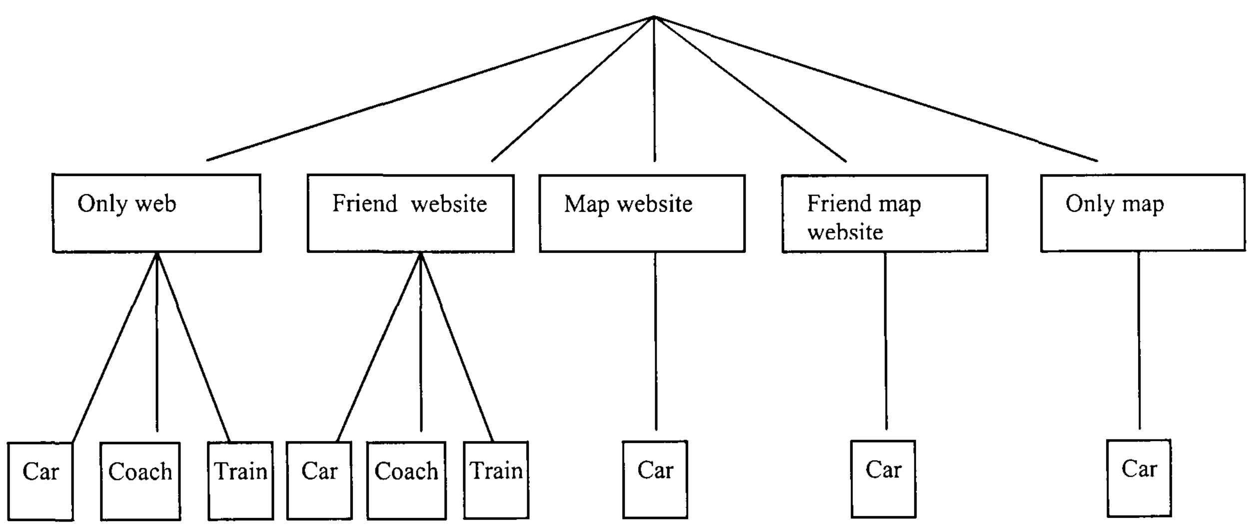

Figure 6.1 Nested Structures

There is also an alternative nest structure considered in order to better comprehend the combined effect of all variables. The alternative nest structure, in which there is more substitution between sources than between modes, is shown in figure 6.2.

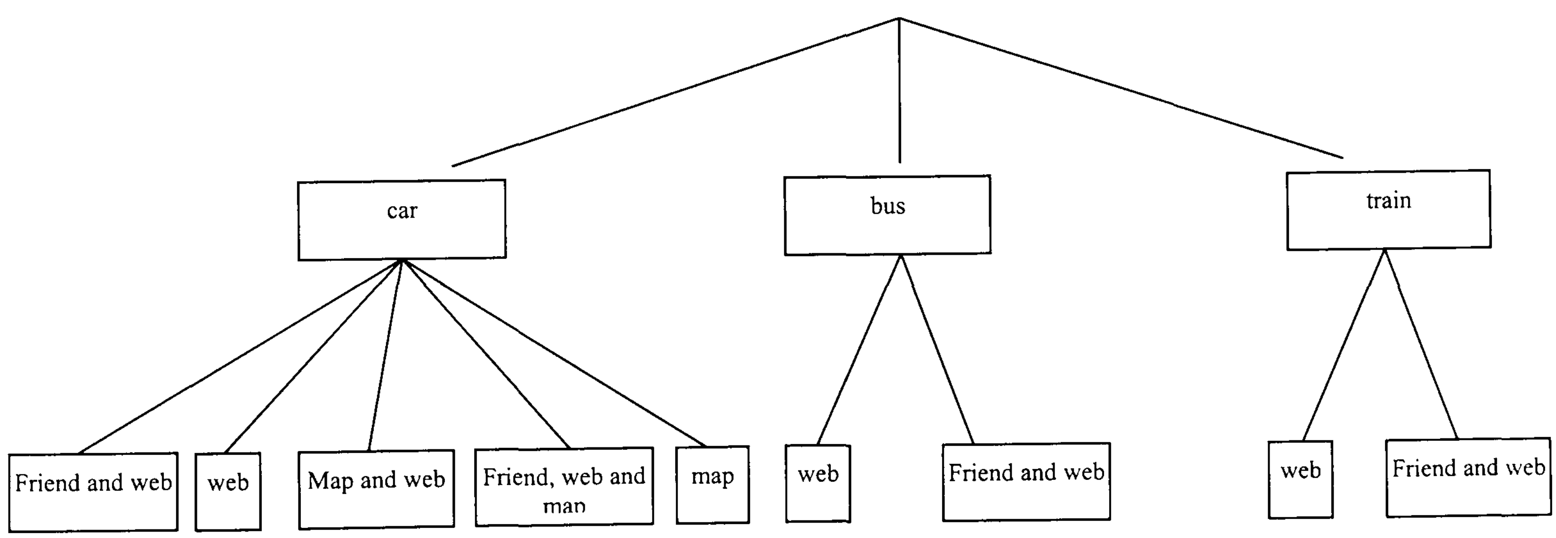

Figure 6.2 Alternative Nested Structure

Various models with the abovementioned structures were tried with different combination of variables and nesting coefficients. The initial models did not converge due to correlation between the variables. Finally the Structure in Figure 6.1 was adopted (information sources on top and modes under the nest). The nesting coefficient was kept constant for all the nests and was denoted as *thetamode*. The choice set C_n of each individual thus consists of nine alternatives. The utility functions for the nine alternatives are given by,

1. $U_{(carweb)} = \text{Time}_1 * Time_1 + \text{Cost}_1 * Cost_1$

2. $U_{(carfrweb)} = \text{Time}_2 * Time_2 + \text{Cost}_2 * Cost_2 + \text{Male2} * DMale_2 + \text{Educ}_2 * DEduc_2 + \text{FEmpl}_2 * DFEmpl_2 + \text{Income}_2 * DIncome_2 + \text{Age40}_2 * DAge40_2$

3. $U_{(carmpweb)} = \text{Time}_3 * Time_3 + \text{Cost}_3 * Cost_3 + \text{Male}_3 * DMale_3 + \text{Educ}_3 * DEduc_3 + \text{FEmpl}_3 * DFEmpl_3 + \text{Income}_3 * DIncome_3 + \text{Age40}_3 * DAge40_3$

4. $U_{(carmfweb)} = \text{Time}_4 * Time_4 + \text{Cost}_4 * Cost_4 + \text{Male}_4 * DMale_4 + \text{Educ}_4 * DEduc_4 + \text{FEmpl}_4 * DFEmpl_4 + \text{Income}_4 * DIncome_4 + \text{Age40}_4 * DAge40_4$

5. $U_{(carmap)} = \text{Time}_5 * Time_5 + \text{Cost}_5 * Cost_5 + \text{Male}_5 * DMale_5 + \text{Educ}_5 * DEduc_5 + \text{FEmpl}_5 * DFEmpl_5 + \text{Income}_5 * DIncome_5 + \text{Age40}_5 * DAge40_5$

6. $U_{(busweb)} = \text{Time}_6 * Time_6 + \text{Cost}_6 * Cost_6 + \text{Male}_6 * DMale_6 + \text{Educ}_6 * DEduc_6 + \text{FEmpl}_6 * DFEmpl_6 + \text{Income}_6 * DIncome_6 + \text{Age40}_6 * DAge40_6$

7. $U_{(busfrweb)} = \text{Time}_7 * Time_7 + \text{Cost}_7 * Cost_7 + \text{Male}_7 * DMale_7 + \text{Educ}_7 * DEduc_7 + \text{FEmpl}_7 * DFEmpl_7 + \text{Income}_7 * DIncome_7 + \text{Age40}_7 * DAge40_7$

8. $U_{(traweb)} = \text{Time}_8 * Time_8 + \text{Cost}_8 * Cost_8 + \text{Male}_8 * DMale_8 + \text{Educ}_8 * DEduc_8 + \text{FEmpl}_8 * DFEmpl_8 + \text{Income}_8 * DIncome_8 + \text{Age40}_8 * DAge40_8$

9. $U_{(trafrweb)} = \text{Time}_9 * Time_9 + \text{Cost}_9 * Cost_9 + \text{Male}_9 * DMale_9 + \text{Educ}_9 * DEduc_9 + \text{FEmpl}_9 * DFEmpl_9 + \text{Income}_9 * DIncome_9 + \text{Age40}_9 * DAge40_9$

Variables

Where: (with subscript n indicating alternatives)

Time	*(Generic, in Minutes)*
Cost	*(Generic, in Pence)*
$DCar_n$	*(Dummy, if car is available in household = 1, otherwise = 0)*
$DMale_n$	*(Dummy, If male =1, otherwise = 0)*
$DEduc_n$	*(Dummy, Left full time education at 20 = 1, otherwise = 0)*
$DEduc25_n$	*(Dummy, Left full time education at 25 = 1, otherwise = 0)*
$DFEmpl_n$	*(Dummy, Full time employed =1, otherwise = 0)*
$DIncome_n$	*(Dummy, If income over £30,000 =1, otherwise = 0)*
$DAge_n$	*(Dummy, If Age less than 50 =1, otherwise = 0)*
$DAge40_n$	*(Dummy, If Age less than 40 =1, otherwise = 0)*

Coefficients

Where: (with subscript n indicating alternatives)

Time	**(is a parameter vector to be estimated for Time, Generic, in Minutes)**
Cost	**(is a parameter vector to be estimated for Cost, Generic, in Pence)**
Car_n	**(is a parameter vector to be estimated for Dummy, if car is available in household = 1, otherwise = 0)**
$Male_n$	**(is a parameter vector to be estimated for Dummy, If male =1, otherwise = 0)**
$Educ_n$	**(is a parameter vector to be estimated for Dummy, Left full time education at 20 = 1, otherwise = 0)**
$Educ25_n$	**(is a parameter vector to be estimated for Dummy, Left full time education at 25 = 1, otherwise = 0)**
$FEmpl_n$	**(is a parameter vector to be estimated for Dummy, Full time employed =1, otherwise = 0)**
$Income_n$	**(is a parameter vector to be estimated for Dummy, If income over £30,000 =1, otherwise = 0)**
Age_n	**(is a parameter vector to be estimated for Dummy, If Age less than 50 =1, otherwise = 0)**
$Age40_n$	**(is a parameter vector to be estimated for Dummy, If Age less than 40 =1, otherwise = 0)**

Model estimates of some of the selected are shown in table 6.6.

Table 6.6 NL RP Model for Model and Source

File	Soumorpnn1	Soumorpnn2	Soumorpnn3
Converged	True	True	True
Observations	226	226	226
Final log (L)	-471.8	483.9	-449.5
D.O.F	11	3	19
$Rho^2(0)$	0.050	0.026	0.095
$Rho^2(c)$	-0.025	-0.051	0.024
Time	-6.4e-4 (-0.2)	5.1e-5 (0.0)	-0.0017 (-0.5)
Cost	1.3e-5 (0.1)	**-1.7e-4 (-1.8)**	-5.8e-5 (-0.3)
Male2	-0.259 (-0.3)		0.688 (0.6)
Male3	0.482 (0.5)		-0.232 (-0.5)
Male4	-2.82 (-0.9)		-0.0747 (-0.1)
Male5	-3.40 (-0.9)		0.178 (0.3)
Male6	-0.368 (-0.9)		0.0077 (0.0)
Male7	0.291 (0.3)		0.630 (0.7)
Male8	**-1.14 (-2.1)**		-0.767 (-1.2)
Male9	-0.999 (-1.0)		0.723 (0.6)
Thetamode	0.268 (1.2)	**0.560 (3.9)**	0.970 (1.0)
Educ2			-1.31 (-1.0)
Educ3			**0.924 (2.6)**
Educ4			-0.228 (-0.2)
Educ5			-0.706 (-0.5)
Educ6			-0.357 (-0.8)
Educ7			-0.486 (-0.5)
Educ8			-0.257 (-0.4)
Educ9			-2.35 (-1.3)

Table 6.6 presents specification of each of these models. The *t*-statistic values of some variables were not satisfying the requirement. Due to high correlation between the variables and the modelling structure more complicated models were not converging even after several trials of different combinations of variables. The values of the nesting coefficients for models including soumorpnn61 and soumorpnn63 also were not

significant even after many trials of different combinations. This confirms that the Multinomial model reflects the choice behaviour of the sample better than the Nested logit model. The multinomial logit structure was therefore adopted for the models where the RP and SP data sets were combined.

The Important properties of the estimated coefficients are discussed below.

Travel Related Coefficients: As expected, *travel time* is negative, but it is insignificant. The cost is also negative but becomes insignificant with more complicated models. This shows that when many taste variations are at stake the travel time and travel cost becomes less important to the travellers.

Traveller Related Coefficients: Model 1 shows that when travelling by car the males tend to choose to consult websites. This means that males are less likely to travel by train if they use websites only as source of travel information. Model 3 shows that Education significantly predisposes people to use website as a source of travel information. Similarly travellers who left their education by the age of 20 are less likely to use website only as source of travel information. The thetamode (nesting coefficient for modes) was not significant in model 1 and 3 where it was significant in model 2 which is very simple and is not reflecting any effect.

6.4.7 Combined RP and SP Model for Mode Choice

As discussed in Section 6.2, there were three data sets in this study. The first was RP data which explored previous behaviours and choices of the travellers when choosing information sources in travelling. The second data set included an SP exercise (SP1) which explored the choice of website as a information source prior to travel. And finally, the third data set included another SP exercise (SP2) which explored the choice of modes under the influence of different information sources. The RP model developed in section 6.4.5 offers advantages that it is based on the actual choices. The SP models on the other hand offer an advantage in that they contain detailed information on the sensitivity of the choice to changes in a range of attributes of information sources in particular subscription costs and reliability. A combination of all three datasets was considered in order to capture their respective advantages.

Although both SP experiments were hypothetical in nature and both involved similar tasks for the respondent under similar conditions, it is important to examine the data to test for any systematic differences in the scale of the utility functions. The methodology proposed by Ben-Akiva and Morikawa (1990) examines difference in scale between revealed and stated preference data. Bradley and Daly (1991) propose simultaneous estimation procedures in which the data are modelled jointly to improve the efficiency of the estimations.

Ben-Akiva and Morikawa postulate that the differences in the error terms between any two data sets can be represented as a function of the variance of the error term of each of the data set. The potential differences in error between the datasets can be removed by multiplying the parameters of SP2 by the scale parameter. In the current study there is no single continuous coefficient in SP1 and RP and thus a combined model for SP1 is not possible. The analysis reported in Section 6.4.4 showed search time to be a very important factor in the choice of information sources and thus it cannot be compromised in the combined model. On the other hand dummies cannot be used as the basis for the combined model. Hence it is not possible to combine the SP1 data with the RP or SP2 data.

In theory, information from question 20 of the questionnaire (which asks for names of websites used) could have been used to construct variables describing the characteristics of websites used. This RP information might then have provided a "bridge" to the website characteristics included in the SP exercises. However, not only would this have been an onerous task but it would have been difficult to develop continuous variables from the RP part of the data. Hence, regrettably, it is not possible to combine data from SP 1 (the source choice experiment) with the remaining of the data sets.

In order to develop a single choice model, data from the RP questions and from the SP 2 (the mode choice experiment), an artificial tree structure proposed by Bradley and Daly (1991) was used. Table 6.7 presents the results of the combined RP-SP model

Table 6.7 RP-SP Model for Mode

File	Sprpjointn1	Sprpjointn2	sprpjointn3
Converged	True	True	True
Observations	1369	1369	1369
Final log (L)	-1113.6	-1105.6	-1105.3
D.O.F.	3	4	5
$Rho^2(0)$	0.136	0.142	0.142
$Rho^2(c)$	0.012	0.019	0.020
VOT	32		
VOTweb		31	25
Time	-0.0024 (-1.3)[6]	-0.0030 (-1.5)	-0.0026 (-1.3)
Cost	-4.5e-5 (-1.2)[7]	**-2.8e-4 (-3.8)**	**-2.8e-4 (-3.8)**
scale2	3.00 (1.3)	**0.296 (3.2)**	**0.228 (2.2)**
Timeweb[8]		**-0.0145 (-3.6)**	**-0.0207 (-1.9)**
Costweb[9]			-4.9e-4 (-1.4)

The utility functions of the RP model are given as

1. $U_{(carweb)} = \mathrm{Time}_1 * Time_1 + \mathrm{Cost}_1 * Cost_1$

2. $U_{(carfrweb)} = \mathrm{Time}_1 * Time_2 + \mathrm{Cost}_1 * Cost_2$

3. $U_{(carmpweb)} = \mathrm{Time}_1 * Time_3 + \mathrm{Cost}_1 * Cost_3$

4. $U_{(carmfweb)} = \mathrm{Time}_1 * Time_4 + \mathrm{Cost}_1 * Cost_4$

5. $U_{(carmap)} = \mathrm{Time}_1 * Time_5 + \mathrm{Cost}_1 * Cost_5$

6. $U_{(busweb)} = \mathrm{Time}_1 * Time_6 + \mathrm{Cost}_1 * Cost_6$

7. $U_{(busfrweb)} = \mathrm{Time}_1 * Time_7 + \mathrm{Cost}_1 * Cost_7$

8. $U_{(traweb)} = \mathrm{Time}_1 * Time_8 + \mathrm{Cost}_1 * Cost_8$

[6] Time is generic and is kept same for all time based variables
[7] Cost is generic and is kept same for all cost based variables
[8] Timeweb is coefficient for time variables with web or multimodal website as information source
[9] Costweb is coefficient for cost variables with web or multimodal website as information source

9. $U_{(trafrweb)} = \mathrm{Time}_1 * Time_9 + \mathrm{Cost}_1 * Cost_9$

Variables

Time$_n$	***(Generic, in Minutes) for alternative n***
Cost$_n$	***(Generic, in Pence) for alternative n***

Coefficients

Time$_n$	**(is a parameter vector to be estimated for Time, Generic, in Minutes) for alternative *n***
Cost$_n$	**(is a parameter vector to be estimated for Cost, Generic, in Pence) for alternative *n***

The second SP experiment had following utility functions,

1. $U(car) = \mathrm{Time}_1 * Carfrit_1 + \mathrm{Timeweb}_1 * Carmult_1 + \mathrm{Cost}_1 * Carfric_1 + \mathrm{Costweb}_1 * Carmulc_1$

2. $U(coach) = \mathrm{Timeweb}_1 * Coawt_2 + \mathrm{Timeweb}_1 * Coamult_2 + \mathrm{Cost}_1 * Coawc_2 + \mathrm{Cost}_1 * Coamulc_2$

3. $U(train) = \mathrm{Timeweb}_1 * Trainwt_3 + \mathrm{Time}_1 * Trainext_3 + \mathrm{Costweb}_1 * Trainwc_3 + \mathrm{Cost}_1 * Trainexc_3$

Variables

Carfrit$_n$	***(Time by car information by friend, in Minutes)for mode n***
Carmult$_n$	***(Time by car information by multimodal website, in Pence) for mode n***
Carfric$_n$	***(Cost by car information by friend, in Minutes) for mode n***
Carmultc$_n$	***(Cost by car information by multimodal website, in Pence) for mode n***
Trainwt$_n$	***(Time by train information by train website, in Minutes) for mode n***
Trainext$_n$	***(Time by train information by past experience, in Minutes) for mode n***
Trainwc$_n$	***(Cost by train information by train website, in Pence) for mode n***
Trainexc$_n$	***(Cost by train information by past experience, in Pence) for mode n***
Coawt$_n$	***(Time by coach information by coach website, in Minutes) for mode n***
Coawc$_n$	***(Cost by coach information by coach website, in Pence) for mode n***

$Coamult_n$	***(Time by coach information by multimodal website, in Minutes) for mode n***
$Coamulc_n$	***(Cost by coach information by multimodal website, in Pence) for mode n***

Coefficients

$Time_n$	**(a parameter vector to be estimated for Time, Generic, in Minutes) for alternative n**
$Cost_n$	**(a parameter vector to be estimated for Cost, Generic, in Pence) for alternative n**
$Timeweb_n$	**(a parameter vector to be estimated for web/multimodal webTime, Generic, in Minutes) for alternative n**
$Costweb_n$	**(a parameter vector to be estimated for web/multimodal webCost, Generic, in Pence) for alternative n**

Table 6.7 provides the specification of each of these models. The *t*-statistic values of some variables did not satisfy the requirement. Due to high correlation between the variables and the modelling structure more complicated models did not converge even after several trials of different combinations.

The important properties of the estimated coefficients are discussed below.

Travel Related Coefficients: The models show that, as expected, travel time and cost are negative and significant. The values of time are also similar to comparable studies as discussed earlier.

Information Source Related Coefficients: Model 2 and Model 3 suggest that journey time described by websites have higher values for the travellers. It also suggests that there is a higher opportunity cost for the information given by the websites and travellers consider the values of pre-trip time information, more important than time and cost actually spent on travelling. Models 2 and 3 also show that value of time is significantly reduced with the introduction of web time as a variable. This provides evidence of differences in travellers' perception of different kinds of time while taking travel decisions. Model 3 also suggests that travellers consider information from websites four times more important than the normal travel time and this is higher than the waiting times on stops in the normal mode choice models.

6.5 Accounting for Repeated Measurement Problem

All standard methods for analyzing SP choice data assume that each observation is independent. This assumption is not strictly valid when several repeated choices are made by each respondent (Bates, 1997, Ortuzar *et al.* 1997). The fact that individuals make series of repeated choices, implies that the resulting data is nested within individuals and the informational content of the data is reduced. Although it is generally assumed that the coefficients estimated on the assumption of independence will not be biased, the associated t-stats will be upward biased implying increased significance of explanatory variables (Ortuzar *et al.* 1997; Bates, 1997).

A number of correction procedures have been suggested in the literature. The simplest method involves dividing the t-stats of the uncorrected method by the square root of the number of observations per individual assuming the perfect correlation of the errors across the choices of each individual (Kocur *et al.* 1982; MVA *et al.* 1987; Khattak *et al.* 1993a). The coefficients by this method are the same as those from the uncorrected models, but the values of the t-stats are reduced to account for the influence of the repeated measurements on the significance of the estimates. Although this method is easy to implement, it has a tendency to be more conservative and is considered to be the other extreme to the uncorrected method.

A less extreme but computationally more difficult method is to assume that repeated observations introduce an additional component in the error term. Estimation of this model is relatively straightforward with mixed logit modelling approach in which the error terms are assumed to be both Gumbel or any other distribution to deal with the individual specific effect. The estimation process requires only one additional parameter. Another approach, alternative to mixed logit, is to just add an individual specific constant but this requires estimation of many coefficients.

Ouwersloot and Rietveld (1996) used another method which treated each observation separately, estimating separate models based on each subgroup one by one and combining these estimates to produce an overall parameter estimate using a 'minimum distance' method. Since only one observation per individual is used in each model, there is no longer a correlation problem due to repeated observations. The only disadvantage of this approach is that it is very complicated to perform and requires a lot of computations. A similar approach was adopted by Abdel-Aty *et al.* (1995) in which a binary logit model was used rather than a probit model as was used by Ouwersloot and Rietveld to estimate the influence of repeated measurements.

Another approach recommended by Cirillo, Daly and Lindveld (2000) is based on re-sampling. The purpose of re-sampling is to find out the true variance of the estimates affected by the repeated measurement problem and to observe the way coefficients change as the number of sub-samples changes. The difference between the estimates obtained from all the small samples gives a more reliable estimate of overall variance. Selecting the particular form of sample reduction gives the most efficient means of variance reduction. Therefore it is necessary to observe the difference between small sample estimates that are also affected by the repeated measurement problem. The most popular of these techniques are known as 'Jack-knifing' and 'Bootstrapping'. Both techniques estimate a series of models on different sub-samples of the data and compute the required coefficient estimates and standard errors from these models. The difference between the two approaches is that jack-knifing systematically omits a small fraction of the data whereas bootstrapping creates a sub-sample by drawing randomly, with replacement. Hence the Jack-knife requires less computational work than bootstrapping (Wonnacott and Wonnacott 1990; Cirillo *et al.* 1996). It was also reported by Shao and Tu (1995) that the bootstrap variance was down-biased and was not as efficient as the Jack-knife variance estimator.

In this study the Jack-knife method is selected because the other methods are too complicated and require a lot of computational resources (Jack-knife is available within a software program) and because other studies, as mentioned earlier, also recommended Jack-knife method for logit models.

To undertake a Jack-knife, the sample is divided into n groups of individuals and form n Jack-knife sub-samples, where each sub-sample is formed by deleting one of the groups from the sample.

The Jack-knife variance estimator is formed as follows:

$$v_j = \frac{n-1}{n}\sum_{r=1}^{n}(t_j - t)^2$$

where t_j is the estimate produced from the replicate with the j^{th} group deleted, and t is the estimate produced from the full sample.

Jack-knife often reduces variance slightly, especially if a large number of sub-groups *r* are used. Even if the variance is actually reduced, the reduced biased is usually sufficient to effect an improvement in terms of mean squared error.

It is desirable to make *r* as large as possible because it improves the power of significance tests and reduces the expected length of confidence intervals, as well as make variance standard much more stable. A large *r* also tends to reduce the bias in standard errors which often seems to slightly overestimate variance where *r* is very small (Bissell and Ferguson, 1975).

6.5.1 Application of Jack-knife on SP Model of Source Choice

As discussed before, the number of sub-samples is important in Jack-knife implementation because it improves the power of significance tests and make variance stable. The ideal number of sub-samples is the number of samples (i.e. r = n). The program Alogit only allows the number of subsamples between 2 and 99. In this study, six models were estimated each with a different number of randomly selected sub-samples (5, 10, 20, 40, 60 and 90).

The final model i.e. ModelSPS6 was taken from section 6.4.4. The utility function is as follows,

1. U_1 = $\text{Seatime}_1 * Seatime_1 + \text{ComAd}_1 * DComAd_1 + \text{Com}_1 * DCom_1 + \text{InfoSpec}_1 * DInfoSpec_1 + \text{InfoPersc}_1 * DInfoPersc_1 + \text{Updaily}_1 * DUpdaily_1$

2. U_2 = $\text{Seatime}_2 * Seatime_2 + \text{ComAdsub}_2 * DComAdsub_2 + \text{Com}_2 * DCom_2 + \text{InfoSpec}_2 * DInfoSpec_2 + \text{InfoPersc}_2 * DInfoPersc_2 + \text{Updaily}_2 * DUpdaily_2$

Variables

Where: (with subscript n websites)

$DComAd_n$:	***(dummy that equals 1 if Commercial Ads No Sub; otherwise 0)***
$DComAdSub_n$:	***(dummy that equals 1 if Commercial Ads Subs; otherwise 0)***
$DCom_n$:	***(dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)***
$Seatime_n$	***(5 min, 10 min, 15 min)***
$DUpdaily_n$:	***(dummy that equals 1 if Website updates daily; otherwise 0)***
$DInfoSpec_n$:	***(dummy that equals 1 if Specific info available; otherwise 0)***
$DInfoPerc_n$:	***(dummy that equals 1 if Info w.r.t own criteria; otherwise 0)***

Coefficients

Where: (with subscript n indicating modes)

ComAd_n:	**(is a parameter vector to be estimated for dummy that equals 1 if Commercial Ads No Sub; otherwise 0)**
ComAdSub_n:	**(is a parameter vector to be estimated for dummy that equals 1 if Commercial Ads Subs; otherwise 0)**
Com_n:	**(is a parameter vector to be estimated for dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)**
Seatime_n	**(is a parameter vector to be estimated for search time, 5 min, 10 min, 15 min)**
Updaily_n:	**(is a parameter vector to be estimated for dummy that equals 1 if Website updates daily; otherwise 0)**
InfoSpec_n:	**(is a parameter vector to be estimated for dummy that equals 1 if Specific info available; otherwise 0)**
InfoPerc_n:	**(is a parameter vector to be estimated for dummy that equals 1 if Info w.r.t own criteria; otherwise 0)**

Model estimates of the uncorrected model and Jack-knifed estimates with 5, 10, 20, 30, 40, 60 and 90 are presented in Table 6.8.

Table 6.8 Comparison between Uncorrected Method and Jack-knife Source Choice Models

	Uncorrected method		Jack-knife 5 sub-samples		Jack-knife 10 sub-samples		Jack-knife 20 sub-samples		Jack-knife 30 sub-samples		Jack-knife 40 sub-samples		Jack-knife 60 sub-samples		Jack-knife 90 sub-samples	
	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio
Seatime1	-0.9674	-12.91	-0.9514	-7.57	-0.9469	-7.82	-0.9301	-9.07	-0.9459	-11.27	-0.9234	-9.04	-0.9345	-12.03	-0.9306	-12.30
DinfoSpec1	1.8617	6.87	1.8198	6.86	1.8236	6.27	1.8050	6.24	1.8258	6.65	1.8099	6.25	1.8188	6.98	1.8089	6.09
DComAd1	0.2653	1.03	0.2754	1.05	0.3483	0.81	0.2671	0.64	0.2852	0.87	0.2854	0.71	0.2627	0.84	0.2841	0.99
DCom1	-1.2661	-2.24	-1.2483	-1.83	-1.2289	-1.85	-1.1663	-1.95	-1.2584	-2.01	-1.1336	-1.69	-1.2079	-1.97	-1.1699	-2.01
DUpdaily1	-0.5487	-2.67	-0.5360	-2.58	-0.5243	-3.17	-0.5200	-2.40	-0.5381	-3.17	-0.5310	-1.96	-0.5425	-2.96	-0.5473	-2.68
Seatime2	0.6298	9.92	0.6214	6.42	0.6287	5.90	0.6147	6.48	0.6214	7.83	0.6122	6.10	0.6160	8.39	0.6152	8.68
DinfoSpec2	4.0865	11.66	4.0376	10.25	4.0572	8.36	3.9449	7.26	4.0399	10.75	3.9413	7.73	4.0030	10.21	4.0028	11.08
DComAdSub2	-17.5278	-13.15	-17.2672	-8.67	-17.2308	-8.24	-16.9350	-9.39	-17.1989	-11.50	-16.8037	-8.75	-17.0242	-12.41	-16.9579	-12.66
DCom2	-15.2016	-11.96	-14.9925	-7.51	-15.0042	-7.41	-14.6739	-7.99	-14.9380	-10.28	-14.5924	-7.66	-14.7694	-10.80	-14.7322	-11.20
DUpdaily2	-1.8044	-7.26	-1.7682	-5.54	-1.7753	-5.09	-1.7667	-7.99	-1.7542	-7.16	-1.7454	-4.36	-1.7372	-7.35	-1.7306	-7.18

The Jack-knife estimates show that, regardless of the number of sub-samples most coefficients on the Jack-knife method are very close to those of the uncorrected model estimates. These results show that the coefficients of the uncorrected model estimate were quite accurate despite of the repeated measurement problem. However the t ratios are slightly reduced in case of the Jack-knife method which indicates that the uncorrected model slightly overestimated the significance of the parameters.

6.5.2 Application of Jack-knife on SP Model of Mode Choice

In this study, total six models were estimated each with a different number of randomly chosen sub-samples; 5, 10, 20, 30, 40, 60 and 90.

The final model i.e. ModelSPM16 was taken from section 6.4.3. The utility function is as follows,

1. $U_{(car)} = \mathrm{Carfrit_1}*Carfrit_1 + \mathrm{Carmult_1}*Carmult_1 + \mathrm{Carfric_1}*Carfric_1 + \mathrm{Carmulc_1}*Carmulc_1 + \mathrm{Male_1}*DMale_1 + \mathrm{FEmp_1}*DFEmp_1 + \mathrm{DInc_1}*DInc_1 + \mathrm{Age_1}*DAge_1$

2. $U_{(coach)} = \mathrm{Coawt_2}*Coawt_2 + \mathrm{Carmult_1}*Coamult_2 + \mathrm{Coawc_2}*Coawc_2 + \mathrm{Carmulc_1}*Coamulc_2$

3. $U_{(train)} = \mathrm{Coawt_2}*Trainwt_3 + \mathrm{Trainext_3}*Trainext_3 + \mathrm{Trainwc_3}*Trainwc_3 + \mathrm{Trainexc_3}*Trainexc_3 + \mathrm{Male_3}*DMale_3 + \mathrm{FEmp_3}*DFEmp_3 + \mathrm{Inc_3}*DInc_3 + \mathrm{Age_3}*DAge_3$

Variables

Where: (with subscript n indicating modes)

$Carfrit_n$	***(Time by car information by friend, in Minutes)***
$Carmult_n$	***(Time by car information by multimodal website, in Pence)***
$Carfric_n$	***(Cost by car information by friend, in Minutes)***
$Carmultc_n$	***(Cost by car information by multimodal website, in Pence)***
$DMale_n$	***(Dummy, If male =1, otherwise = 0)***
$DFEmp_n$	***(Dummy, if full employment = 1, otherwise = 0)***
$DInc_n$	***(Dummy, if income over 30,000 pounds =1, otherwise = 0)***
$DAge_n$	***(Dummy, if age less than 40 =1, otherwise = 0)***
$Trainwt_n$	***(Time by train information by train website, in Minutes)***
$Trainext_n$	***(Time by train information by past experience, in Minutes)***
$Trainwc_n$	***(Cost by train information by train website, in Pence)***
$Trainexc_n$	***(Cost by train information by past experience, in Pence)***
$Coawt_n$	***(Time by coach information by coach website, in Minutes)***

$Coawc_n$ *(Cost by coach information by coach website, in Pence)*
$Coamult_n$ *(Time by coach information by multimodal website, in Minutes)*
$Coamulc_n$ *(Cost by coach information by multimodal website, in Pence)*

Coefficients

Where: (with subscript n indicating modes)

$Carfrit_n$ **(is a parameter vector to be estimated for Time by car information by friend, in Minutes)**
$Carmult_n$ **(is a parameter vector to be estimated for Time by car information by multimodal website, in Pence)**
$Carfric_n$ **(is a parameter vector to be estimated for Cost by car information by friend, in Minutes)**
$Carmultc_n$ **(is a parameter vector to be estimated for Cost by car information by multimodal website, in Pence)**
$Male_n$ **(is a parameter vector to be estimated for Dummy, If male =1, otherwise = 0)**
$FEmp_n$ **(is a parameter vector to be estimated for Dummy, if full employment = 1, otherwise = 0)**
Inc_n **(is a parameter vector to be estimated for Dummy, if income over 30,000 pounds =1, otherwise = 0)**
Age_n **(is a parameter vector to be estimated for Dummy, if age less than 40 =1, otherwise = 0)**
$Trainwt_n$ **(is a parameter vector to be estimated for Time by train information by train website, in Minutes)**
$Trainext_n$ **(is a parameter vector to be estimated for Time by train information by past experience, in Minutes)**
$Trainwc_n$ **(is a parameter vector to be estimated for Cost by train information by train website, in Pence)**
$Trainexc_n$ **(is a parameter vector to be estimated for Cost by train information by past experience, in Pence)**
$Coawt_n$ **(is a parameter vector to be estimated for Time by coach information by coach website, in Minutes)**
$Coawc_n$ **(is a parameter vector to be estimated for Cost by coach information by coach website, in Pence)**
$Coamult_n$ **(is a parameter vector to be estimated for Time by coach information by multimodal website, in Minutes)**
$Coamulc_n$ **(Cost by coach information by multimodal website, in Pence)**

Model estimates of the uncorrected model and Jack-knifed estimates with 5, 10, 20,30,40,60, and 90 samples are presented in the Table 6.9.

Here too the Jack-knife estimates show that the estimates are close to the MNL model, regardless of the number of sub-sample. These results show that the coefficients of the uncorrected model estimate were quite accurate despite of the repeated measurement problem.

Table 6.9 Comparison between Uncorrected Method and Jack-knife Mode Choice Models

	Uncorrected method		Jack-knife 5 sub-samples		Jack-knife 10 sub-samples		Jack-knife 20 sub-samples		Jack-knife 30 sub-samples		Jack-knife 40 sub-samples		Jack-knife 60 sub-samples		Jack-knife 90 sub-samples	
	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio
carfrit1	-0.0070	-1.74	-0.0054	-1.69	0.0069	-1.79	-0.0072	-1.7	**-0.0072**	**-1.98**	**-0.0066**	**-2.27**	**-0.0071**	**-1.95**	**-0.0072**	**-1.80**
carmult1	**-0.0047**	**-2.54**	**-0.0042**	**-2.25**	**-0.0046**	**-2.44**	**-0.0046**	**-2.56**	**-0.0046**	**-2.66**	**-0.0045**	**-2.22**	**-0.0046**	**-2.62**	**-0.0046**	**-2.59**
carfric1	**-0.0003**	**-1.97**	**-0.0003**	**-2.64**	-0.0003	-1.68	**-0.0003**	**-1.97**	-0.0003	-1.63	**-0.0003**	**-2.08**	-0.0003	-1.66	**-0.0003**	**-1.81**
carmulc1	**-0.0001**	**-1.80**	**-0.0001**	**-3.12**	**-0.0001**	**-2.01**	**-0.0001**	**-2.02**	**-0.0001**	**-2.09**	**-0.0001**	**-2.06**	**-0.0001**	**-2.12**	-0.0001	-1.51
Dmale1	**1.3801**	**3.76**	**1.3329**	**6.32**	**1.3805**	**3.50**	**1.3884**	**4.26**	**1.4003**	**3.58**	**1.3661**	**7.07**	**1.4082**	**5.68**	**1.3865**	**3.89**
Dfemp1	**1.2216**	**3.47**	**1.3088**	**6.22**	**1.1932**	**2.70**	**1.2176**	**3.21**	**1.2118**	**3.13**	**1.2377**	**2.90**	**1.2117**	**3.49**	**1.2133**	**3.14**
Dinc1	**1.0415**	**2.64**	**0.9819**	**3.44**	1.0234	1.49	**1.0241**	**2.50**	**1.0371**	**1.82**	**1.0105**	**3.15**	**1.0425**	**3.01**	**1.0232**	**2.00**
Dage1	-1.1312	-1.45	-0.9786	-1.54	-0.9060	-1.15	-0.9372	-0.93	-0.9420	-1.03	-0.9231	-1.16	-0.9225	-0.86	-0.8996	-1.08
Trainext3	-0.0049	-0.80	-0.0031	-1.29	-0.0050	-0.84	-0.0053	-0.94	-0.0052	-0.95	-0.0045	-0.91	-0.0052	-1.06	-0.0055	-0.81
Trainwc3	**-0.0002**	**-3.54**	**-0.0002**	**-2.65**	**-0.0002**	**-3.57**	**-0.0002**	**-3.72**	**-0.0002**	**-4.13**	**-0.0002**	**-3.36**	**-0.0002**	**-3.77**	**-0.0002**	**-3.80**
Trainexc3	-0.0001	-1.76	**-0.0001**	**-3.01**	**-0.0001**	**-2.41**	-0.0001	-1.82	-0.0001	-1.75	**-0.0001**	**-4.26**	-0.0001	-1.48	**-0.0001**	**-1.80**
Dmale3	0.1920	0.62	0.1806	1.01	0.2005	0.62	0.2008	0.71	0.2125	0.69	0.1999	1.10	0.2235	1.15	0.1977	0.57
Dfemp3	**0.5836**	**2.14**	**0.6354**	**2.56**	**0.5632**	**2.21**	**0.5796**	**2.40**	**0.5744**	**2.42**	**0.5965**	**2.16**	**0.5805**	**2.63**	**0.5747**	**2.00**
Dinc3	**2.0101**	**6.21**	**1.9470**	**7.46**	**1.9781**	**4.28**	**1.9894**	**5.81**	**1.9973**	**4.64**	**1.9729**	**8.87**	**1.9849**	**6.12**	**1.9961**	**4.90**
Dage3	-1.0583	-1.54	-0.8985	-1.32	-0.8125	-1.06	-0.8469	-1.01	-0.8665	-1.05	-0.8439	-1.28	-0.8700	-1.00	-0.8160	-1.07
Coawt2	**-0.0056**	**-2.25**	**-0.0057**	**-1.97**	**-0.0054**	**-2.03**	**-0.0055**	**-2.35**	**-0.0055**	**-2.17**	-0.0056	-1.28	**-0.0056**	**-2.37**	**-0.0054**	**-2.07**
Coawc2	**-0.0003**	**-2.13**	-0.0003	-1.16	-0.0003	-1.61	**-0.0003**	**-1.84**	**-0.0003**	**-1.95**	-0.0003	-1.28	-0.0003	-1.46	**-0.0003**	**-1.89**

6.6 Mixed Logit Estimations

The origin of the mixed logit can be traced to Cardell and Dunbar (1980). Since the mixed logit is neither Generalised Extreme Value (GEV) nor has a closed form, estimation difficulties restricted its application in the early years. More recently, development of the estimation methods, maximum simulation likelihood (Ben Akiva and Bolduc, 1991) has made this model attractive for estimations.

Brownstone *et al.* (2000) derived the utility function for alternative *i* as follows:

$$U_i = \beta x_i + [\eta_i + \epsilon_i]$$

Where x_i is a vector of observed variables relating to alternative i, β is a vector of structural parameters reflecting choices by the overall population, η_i is a random term with zero mean, the distribution of which varies across individuals and alternatives depends on underlying parameters and observed data relating to individuals and alternatives, and ϵ_i is a random term with zero mean that is IID across alternatives and does not depend on the underlying parameters nor data, and is normalized to set the scale of utility. The model was derived by assuming a general distribution for η and an IID extreme value for ϵ. For given η, the remaining error is IID extreme value and hence the conditional choice probability is given by

$$F_i(\eta) = \frac{e^{(\beta v_i + \eta_i)}}{\sum_{j \epsilon J} e^{(\beta v_j + \eta_j)}}$$

However, since in practice η is not given, the unconditional choice probability is MNL integrated over all values of η weighted by the density of η:

$$P_i = \int F_i(\eta) f(\eta \,|\, \Omega) d\eta$$

Where, Ω is a vector of parameters which describe the distribution of η.

Although mixed logit is not GEV, Mc Fadden and Train (2000) established that any discrete choice model from a RUM (Random Utility Maximising) model can be approximated by mixed logit.

Unlike the estimation of standard logit models exact maximum likelihood estimation is not possible in mixed logit models. On the contrary, a simulated likelihood function is specified in which the probability is approximated by summation over randomly chosen values of η_n. The process is repeated for R random draws of η_n and the simulated probability of the individuals' sequence of choices is:

$$SP_n(\theta) = \left(1/R\right)\sum_{r=1}^{R} S_n(\eta_n^{r|\theta})$$

Where $(\eta_n^{r|\theta})$ is the r^{th} draw from $f(\eta_n\,|\,\theta)$. So long the numbers of random draws is sufficiently large, the simulated probability is an unbiased estimate of the true probability and the simulated likelihood function is constructed as

$$SLL(\theta) = \sum_n \ln(SP_n(\theta))$$

In the context where a decision-maker makes many choices over a period of time (panel data or responses to the stated preference surveys), it is possible to accommodate the multi-period nature of the data by assuming that a respondent's tastes (η_n) do not change between choice situations. The conditional probability of the individual n's sequence of choices then becomes the product of logits:

$$S_n(\eta_n) = \prod_t F_{ni(n,t)t}(\eta_n)$$

Where $i(n,t)$ is individual n's choice in period t.

The unconditional probability is given by:

$$P_n(\theta^*) = \int S_n(\eta_n) f(\eta_n\,|\,\theta^*) d\eta_n$$

Where θ^* are the parameters which describe the distribution of tastes $f(\eta_n\,|\,\theta^*)$.

In this study only selected models would be calibrated in the mixed logit framework to check their robustness.

6.6.1 Mixed Logit Model on SP Model of Source Choice

The final model i.e. ModelSPS6 was taken from section 6.4.4 for the calibration of Mixed Logit Model (MMNL) for SP mode choice model. This model has similar specification to the multinomial logit (MNL) but following notable differences.

Firstly, normally distributed coefficients were specified for search time to explain additional variation in the model (MMNL in Table 6.10). Secondly, in order to accommodate the repeated measurement problem, the search time was also identified to vary across individuals and not across observations. (Panel MMNL in Table 6.10) This was done by using Biogeme's panel data specifications (for details refer Biogeme manual).

The utility function is as follows,

1. $U_1 = \mathrm{Seatime}_1 * Seatime_1\,[\mathrm{Sigma}_1] + \mathrm{ComAd}_1 * DComAd_1 + \mathrm{Com}_1 * DCom_1 + \mathrm{InfoSpec}_1 * DInfoSpec_1 + \mathrm{InfoPersc}_1 * DInfoPersc_1 + \mathrm{Updaily}_1 * DUpdaily_1$

2. $U_2 = \mathrm{Seatime}_1 * Seatime_2\,[\mathrm{Sigma}_1] + \mathrm{ComAdsub}_2 * DComAdsub_2 + \mathrm{Com}_2 * DCom_2 + \mathrm{InfoSpec}_2 * DInfoSpec_2 + \mathrm{InfoPersc}_2 * DInfoPersc_2 + \mathrm{Updaily}_2 * DUpdaily_2$

Variables

Where: (with subscript n websites)

$DComAd_n$:	***(dummy that equals 1 if Commercial Ads No Sub; otherwise 0)***
$DComAdSub_n$:	***(dummy that equals 1 if Commercial Ads Subs; otherwise 0)***
$DCom_n$:	***(dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)***
$Seatime_n$	***(5 min, 10 min, 15 min)***
$DUpdaily_n$:	***(dummy that equals 1 if Website updates daily; otherwise 0)***
$DInfoSpec_n$:	***(dummy that equals 1 if Specific info available; otherwise 0)***
$DInfoPerc_n$:	***(dummy that equals 1 if Info w.r.t own criteria; otherwise 0)***

Coefficients

Where: (with subscript n indicating websites)

ComAd_n:	**(is a parameter vector to be estimated for dummy that equals 1 if Commercial Ads No Sub; otherwise 0)**

$ComAdSub_n$:	**(is a parameter to be estimated for dummy that equals 1 if Commercial Ads Subs; otherwise 0)**
Com_n:	**(is a parameter to be estimated for dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)**
$Seatime_n$	**(is a parameter to be estimated for Searchtime)**
$Updaily_n$:	**(is a parameter to be estimated for dummy that equals 1 if Website updates daily; otherwise 0)**
$InfoSpec_n$:	**(is a parameter to be estimated for dummy that equals 1 if Specific info available; otherwise 0)**
$InfoPerc_n$:	**(is a parameter to be estimated for dummy that equals 1 if Info w.r.t own criteria; otherwise 0)**
$Sigma_n$:	**(is a normally distributed error component)**

The models were estimated in Biogeme V1.8 and the estimated coefficients are shown in Table 6.10.

Table 6.10 Comparison between MNL, Jack-knifed Method and Mixed Logit Source Choice Model

File	MNLSPS6	Jacknifed	MMNL	Panel MMNL
Converged	True	True	True	True
Observations	1056	1056	1056	1056
Final log (L)	-503.8	-503.8	-503.7	-503.8
D.O.F.	11	11	12	12
$Rho^2(0)$	0.312	0.312	0.312	0.312
$Rho^2(c)$	0.312	0.312	0.312	0.312
Seatime1	**-0.194 (-7.5)**	**-0.194 (-7.9)**		
Mean of Seatime1			**-0.192 (-8.0)**	**-0.202 (-10.7)**
S.D. of Seatime1			-0.0061 (-0.3)	0.0116 (0.7)
DinfoSpec1	**1.63 (4.9)**	**1.64 (4.2)**	**1.60 (4.4)**	**1.70 (5.4)**
DinfoPers1	**3.91 (5.8)**	**3.60 (4.9)**	**3.85 (5.9)**	**1.92 (4.7)**
DComAd1	**-1.27 (-5.4)**	**-1.26 (-5.0)**	**-1.26 (-4.9)**	**-1.31 (-6.5)**
DCom1	**3.55 (8.1)**	**3.52 (7.4)**	**3.53 (8.2)**	**3.58 (8.5)**
DUpdaily1	**-1.79 (-4.7)**	**-1.79 (-4.7)**	**-1.75 (-4.4)**	**-1.88 (-6.3)**
DinfoSpec2	-0.619 (-1.4)	-0.623 (-1.6)	-0.565 (-1.4)	**-0.758 (-2.6)**
DinfoPers2	0.288 (0.5)	-0.0304 (-0.1)	0.322 (0.5)	**-1.92 (-6.4)**
DComAdSub2	0.227 (0.5)	0.255 (0.5)	0.170 (0.4)	0.326 (0.8)
DCom2	**2.30 (4.2)**	**2.32 (4.9)**	**2.24 (4.9)**	**2.47 (7.1)**
DUpdaily2	**-1.05 (-3.7)**	**-1.06 (-4.2)**	**-1.02 (-3.9)**	**-1.13 (-5.1)**

The results above show that the mixed logit calibration of the model increased the efficiency of the model with increased t stats. The inclusion of the error component terms improves the fit to the data, but only marginally. The results confirm the findings of the MNL model. The standard deviations of the normally distributed components are not significant in both MMNL and Panel MMNL this suggests that the MNL model specification represents the choice of the respondents adequately.

6.6.2 Mixed Logit Model on SP Model of Mode Choice

The final model i.e. ModelSPM16 was taken from section 6.4.3 for the calibration of mixed logit model (MMNL) for SP mode choice. This model has similar specification to the multinomial logit (MNL) but with following differences.

Firstly (MMNL in Table 6.11), four normally distributed coefficients were specified for multimodal website time, multimodal website cost, mode specific website time and train website cost. The other normally distributed coefficients for coach website cost, friend time, friend cost, train experience cost and previous experience time were also tried but were found insignificant were thus were subsequently dropped. Secondly (Panel MMNL in Table 6.11), in another model, the above mentioned coefficients were identified to vary only across individuals to cope with repeated measurement problem.

The utility function is as follows,

1. $U(car) = \text{Carfrit}_1 * Carfrit_1 + \text{Carmult}_1 * Carmult_1\ [\text{Sigma}_1] + \text{Carfric}_1 * Carfric_1 + \text{Carmulc}_1 * Carmulc_1\ [\text{Sigma}_2] + \text{Male}_1 * DMale_1 + \text{FEmp}_1 * DFEmp_1 + \text{DInc}_1 * DInc_1 + \text{Age}_1 * DAge_1$

2. $U(coach) = \text{Coawt}_2 * Coawt_2\ [\text{Sigma}_3] + \text{Carmult}_1 * Coamult_2\ [\text{Sigma}_1] + \text{Coawc}_2 * Coawc_2 + \text{Carmulc}_1 * Coamulc_2\ [\text{Sigma}_2]$

3. $U(train) = \text{Coawt}_2 * Trainwt_3\ [\text{Sigma}_3] + \text{Trainext}_3 * Trainext_3 + \text{Trainwc}_3 * Trainwc_3\ [\text{Sigma}_4] + \text{Trainexc}_3 * Trainexc_3 + \text{Male}_3 * DMale_3 + \text{FEmp}_3 * DFEmp_3 + \text{Inc}_3 * DInc_3 + \text{Age}_3 * DAge_3$

Variables

Where: (with subscript n indicating modes)

$Carfrit_n$	***(Time by car information by friend, in Minutes)***
$Carmult_n$	***(Time by car information by multimodal website, in Pence)***

$Carfric_n$	*(Cost by car information by friend, in Minutes)*
$Carmultc_n$	*(Cost by car information by multimodal website, in Pence)*
$DMale_n$	*(Dummy, If male =1, otherwise = 0)*
$DFEmp_n$	*(Dummy, if full employment = 1, otherwise = 0)*
$DInc_n$	*(Dummy, if income over 30,000 pounds =1, otherwise = 0)*
$DAge_n$	*(Dummy, if age less than 40 =1, otherwise = 0)*
$Trainwt_n$	*(Time by train information by train website, in Minutes)*
$Trainext_n$	*(Time by train information by past experience, in Minutes)*
$Trainwc_n$	*(Cost by train information by train website, in Pence)*
$Trainexc_n$	*(Cost by train information by past experience, in Pence)*
$Coawt_n$	*(Time by coach information by coach website, in Minutes)*
$Coawc_n$	*(Cost by coach information by coach website, in Pence)*
$Coamult_n$	*(Time by coach information by multimodal website, in Minutes)*
$Coamulc_n$	*(Cost by coach information by multimodal website, in Pence)*

Coefficients

Where: (with subscript n indicating modes)

$Carfrit_n$	**(is a parameter vector to be estimated for Time by car information by friend, in Minutes)**
$Carmult_n$	**(is a parameter vector to be estimated for Time by car information by multimodal website, in Pence)**
$Carfric_n$	**(is a parameter vector to be estimated for Cost by car information by friend, in Minutes)**
$Carmulc_n$	**(is a parameter vector to be estimated for Cost by car information by multimodal website, in Pence)**
$Male_n$	**(is a parameter vector to be estimated for Dummy, If male =1, otherwise = 0)**
$FEmp_n$	**(is a parameter vector to be estimated for Dummy, if full employment = 1, otherwise = 0)**
Inc_n	**(is a parameter vector to be estimated for Dummy, if income over 30,000 pounds =1, otherwise = 0)**
Age_n	**(Dummy, if age less than 40 =1, otherwise = 0)**
$Trainext_n$	**(is a parameter vector to be estimated for Time by train information by past experience, in Minutes)**
$Trainwc_n$	**(is a parameter vector to be estimated for Cost by train information by train website, in Pence)**
$Trainexc_n$	**(is a parameter vector to be estimated for Cost by train information by past experience, in Pence)**
$Coawt_n$	**(is a parameter vector to be estimated for Time by coach information by coach website, in Minutes)**
$Coawc_n$	**(is a parameter vector to be estimated for Cost by coach information by coach website, in Pence)**
$Sigma_n$:	**(is a normally distributed error component)**

The results of the model suggest that there is no significant difference in the models. The estimates of the standard deviations of the normally distributed terms were not significant (or if they are significant, as happens sometimes in the panel specification, this takes place at the expense of the precision of coefficients of other attributes, without significant gains in the overall Loglikelihood value). This all means that the MNL model reflects the characteristics of the population satisfactorily.

Table 6.11 Comparison between MNL, Jack-knife Method and Mixed Logit Mode Choice Model

File	ModelSPM16	Jacknifed	MMNL	Panel MMNL
Converged	True	True	True	True
Observations	1143	1143	1143	1143
Final log (L)	-534.8	-534.8	-533.1	-534.3
Rho²(0)	0.325	0.325	0.439	0.495
Rho²(c)	0.198	0.198	0.327	0.161
Carfrit1	-0.0070 (-1.7)	**-0.0072 (-1.8)**	**-0.0074 (-1.8)**	-0.0030 (-0.5)
Carmult1	**-0.0047 (-2.5)**	**-0.0046 (-2.6)**		
Mean of Carmult1			**-0.0050 (-2.6)**	-0.0022 (-0.4)
S.D. of Carmult1			1.9e-4 (0.4)	-0.0012 (-0.0)
Carfric1	**-2.8e-4 (-2.0)**	**-2.9e-4 (-1.8)**	**-2.9e-4 (-1.9)**	**-3.9e-4 (-2.5)**
Carmulc1	**-1.2e-4 (-1.8)**	-1.2e-4 (-1.5)		
Mean of Carmulc1			**-1.2e-4 (-1.8)**	-6.5e-5 (-0.6)
S.D. of Carmulc1			-3.1e-6 (-0.1)	**-1.0e-4 (-2.5)**
DMale1	**1.38 (3.8)**	**1.39 (3.9)**	**1.38 (3.5)**	1.8e-4 (0.0)
DFEmp1	**1.22 (3.5)**	**1.21 (3.1)**	**1.28 (3.6)**	8.8e-5 (0.0)
DInc1	**1.04 (2.6)**	**1.02 (2.0)**	**1.00 (2.3)**	-9.1e-5 (-0.0)
DAge1	-1.13 (-1.5)	-0.900 (-1.1)	-1.13 (-1.5)	-5.2e-6 (-0.0)
Trainext3	-0.0049 (-0.8)	-0.0055 (-0.8)	-0.0053 (-0.9)	0.0014 (0.1)
Trainwc3	**-1.9e-4 (-3.5)**	**-1.9e-4 (-3.8)**		
Mean of Trainwc3			**-1.9e-4 (-3.5)**	**-3.0e-4 (-4.5)**
S.D. of Trainwc3			-2.8e-5 (-1.4)	**1.4e-4 (7.4)**
Trainexc3	**-1.2e-4 (-1.8)**	**-1.2e-4 (-1.8)**	**-1.3e-4 (-1.8)**	-2.2e-4 (-1.7)
DMale3	0.192 (0.6)	0.198 (0.6)	0.177 (0.5)	-1.8e-4 (-0.0)
DFEmp3	**0.584 (2.1)**	**0.575 (2.0)**	**0.603 (2.2)**	-8.8e-5 (-0.0)
DInc3	**2.01 (6.2)**	**2.00 (4.9)**	**2.02 (5.7)**	9.1e-5 (0.0)
DAge3	-1.06 (-1.5)	-0.816 (-1.1)	-1.10 (-1.6)	5.2e-6 (0.0)
Coawt2	**-0.0056 (-2.3)**	**-0.0054 (-2.1)**		
Mean of Coawt2			**-0.0059 (-2.3)**	0.0016 (0.1)
S.D of Coawt2			-5.6e-4 (-0.7)	-5.6e-4 (-0.2)
Coawc2	**-3.5e-4 (-2.1)**	**-3.4e-4 (-1.9)**	**-3.5e-4 (-2.1)**	0 (0.0)

6.7 Random Regret Minimisation Estimations

The Random Regret Minimisation (RRM) model postulates that people aim to minimise their regret with respect to the foregone alternatives (Chorus, 2009). The regret based models are based on the theory that individuals minimise anticipated regret when given a choice instead of maximising anticipated utility. The RRM estimation procedures adopted in this section uses a specification suggested by Chorus (2009) which assumes that the regret is experienced with respect to all foregone alternatives that perform better than a chosen/intended alternative in terms of one or more alternatives unlike previous assumptions that regret is only experienced with respect to the best of foregone alternatives. The specification produces intuitive estimation outcomes and satisfactory fit with available data (Chorus *et al.,* 2008, 2009). The model states that a decision maker faces a set of L travel alternatives, each explained in terms of M attributes x_m that are comparable across alternatives. A decision maker would aim to minimise anticipated regret amongst the alternatives which is composed out of an iid random error (Extreme Value Type I-distributed with variance $\pi^2/6$) and a deterministic regret R. Deterministic regret is conceived to be maximum of all binary regrets associated with the comparison of the considered alternative with each of remaining alternatives (either zero or equal to the weighted difference in attribute performance. The deterministic regret associated with any alternative e.g. alternative 1 is written as,

$$R_1 = max_{l=2..L}\left\{\sum_{m=1..M} max\{0, \beta_m.(x_{lm} - x_{1m})\}\right\}$$

Assuming both terms in the max operator are stochastic and assuming that an iid random component ϵ is added, the expected maximum can then be written as:

$$\text{E} = (\max\ \{0 + \epsilon,\ \beta_m.(x_{lm} - x_{1m}) + \epsilon\})$$

$$= \ln\ (\ \exp[0] + \exp[\ \beta_m.(x_{lm} - x_{1m})])$$

$$= \ln\ (1 + \exp[\beta_m.(x_{lm} - x_{1m})])$$

Hence the deterministic regret associated with the alternative 1 can be written as

$$R = \sum_{l=2..L}\left(\sum_{m=1..M}[\ln(1 + \exp[\beta_m.(x_{lm} - x_{1m})])]\right)$$

6.7.1 Application of Regret Minimization on SP Model of Source Choice

The final model i.e. ModelSPS6 was thus formulised in the RRM paradigm. Although when the choice sets are binary, RRM reduces to Random Utility Maximisation's (RUM's) linear additive binary logit model, the model was still calibrated to validate the applicability and performance of the RRM formulation.

The utility functions are as follows,

1. $$R_1 = ln\,(1 + e^{\text{Seatime1}*(Seatime2-Seatime1)}) + ln\,(1 + e^{\text{ComAd1}*(DComAd2-DComAd1)}) + ln\,(1 + e^{\text{Com1}*(DCom2-DCom1)}) + ln\,(1 + e^{\text{InfoSpec1}*(DInfoSpec2-DInfoSpec1)}) + ln\,(1 + e^{\text{InfoPersc1}*(DInfoPersc2-DInfoPersc1)}) + ln\,(1 + e^{\text{Updaily1}*(DUpdaily2-DUpdaily1)})$$

2. $$R_2 = ln\,(1 + e^{\text{Seatime1}*(Seatime1-Seatime2)}) + ln\,(1 + e^{\text{ComAdSub2}*(DComAdSub1-DComAdSub2)}) + ln\,(1 + e^{\text{Com2}*(DCom1-DCom2)}) + ln\,(1 + e^{\text{InfoSpec2}*(DInfoSpec1-DinfoSpec2)}) + ln\,(1 + e^{\text{InfoPersc2}*(DInfoPersc1-DInfoPersc2)}) + ln\,(1 + e^{\text{Updaily2}*(DUpdaily1-DUpdaily2)})$$

Variables

Where: (with subscript n websites)

$DComAd_n$:	***(dummy that equals 1 if Commercial Ads No Sub; otherwise 0)***
$DComAdSub_n$:	***(dummy that equals 1 if Commercial Ads Subs; otherwise 0)***
$DCom_n$:	***(dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)***
$Seatime_n$	***(5 min, 10 min, 15 min)***
$DUpdaily_n$:	***(dummy that equals 1 if Website updates daily; otherwise 0)***
$DInfoSpec_n$:	***(dummy that equals 1 if Specific info available; otherwise 0)***
$DInfoPerc_n$:	***(dummy that equals 1 if Info w.r.t own criteria; otherwise 0)***

Coefficients

Where: (with subscript n indicating websites)

$ComAd_n$:	**(is a parameter vector to be estimated for dummy that equals 1 if Commercial Ads No Sub; otherwise 0)**
$ComAdSub_n$:	**(is a parameter vector to be estimated for dummy that equals 1 if Commercial Ads Subs; otherwise 0)**
Com_n:	**(is a parameter vector to be estimated for dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)**

Seatime$_n$ (is a parameter vector to be estimated for Search time i.e. 5 min, 10 min, 15 min)

Updaily$_n$: (is a parameter vector to be estimated for dummy that equals 1 if Website updates daily; otherwise 0)

InfoSpec$_n$: (is a parameter vector to be estimated for dummy that equals 1 if Specific info available; otherwise 0)

InfoPerc$_n$: (is a parameter vector to be estimated for dummy that equals 1 if Info w.r.t own criteria; otherwise 0)

Table 6.12 Comparison between Uncorrected Method, Jack-knifed and Regret Source Choice Model

File	ModelSPS6	Jacknifed	RRM
Converged	True	True	True
Observations	1056	1056	1056
Final log (L)	-503.8	-503.8	-503.8
D.O.F.	11	11	11
$Rho^2(0)$	0.312	0.312	0.312
$Rho^2(c)$	0.312	0.312	0.312
Seatime1	**-0.194 (-7.5)**	**-0.194 (-7.9)**	**- 0.194 (-8.3)**
DinfoSpec1	**1.63 (4.9)**	**1.64 (4.2)**	**-1.63 (-4.5)**
DinfoPersc1	**3.91 (5.8)**	**3.60 (4.9)**	**-3.91 (-5.9)**
DComAdSub1	**-1.27 (-5.4)**	**-1.26 (-5.0)**	**1.27 (4.9)**
DCom1	**3.55 (8.1)**	**3.52 (7.4)**	**- 3.55 (-8.2)**
DUpdily1	**-1.79 (-4.7)**	**-1.79 (-4.7)**	**1.79 (4.5)**
DinfoSpec2	-0.619 (-1.4)	-0.623 (-1.6)	0.618 (1.5)
DinfoPersc2	0.288 (0.5)	-0.0304 (-0.1)	- 0.291 (-0.5)
DComAdSub2	0.227 (0.5)	0.255 (0.5)	- 0.227 (-0.5)
DCom2	**2.30 (4.2)**	**2.32 (4.9)**	**- 2.30 (-4.9)**
DUpdaily2	**-1.05 (-3.7)**	**-1.06 (-4.2)**	**1.05 (3.9)**

The results from the above table show that the RRM model is almost equal to the MNL model which confirms the appropriateness of the RRM model. Although all the variables were first tried inside the Regret function, the resulting model lost its significance even on the important explanatory variables; hence only search time was tried in the regret function (presented in Table 6.12). The above table also shows that the signs of the dummy variables (outside the regret function) are reverse as compared to the MNL model, this confirms that the model shows the regret of an individual not a utility as in the other models.

6.7.2 Application of Regret Minimization on SP Model of Mode Choice

The final model i.e. ModelSPM16 was now taken for the calibration of RRM model for SP mode choice model. Two models were estimated, one with all the cost and time variables inside the regret function. In the other model information from friend and previous experience were kept outside the regret function. The dummies were outside the regret function for both the models. The utility function is as follows,

1. $$R(car) = ln\,(1+ e^{\text{Carfrit1}*(Coawt2 - Carfrit1)}) + ln\,(1+ e^{\text{Carfrit1}*(Coamult2 - Carfrit1)}) + ln\,(1+ e^{\text{Carfrit1}*(Trainwt3 - Carfrit1)}) + ln\,(1+ e^{\text{Carfrit1}*(Trainext3 - Carfrit1)}) + ln\,(1 + e^{\text{Carmult1}*(Coawt2 - Carmult1)}) + ln\,(1 + e^{\text{Carmult1}*(Coamult2 - Carmult1)}) + ln\,(1 + e^{\text{Carmult1}*(Trainwt3 - Carmult1)}) + ln\,(1 + e^{\text{Carmult1}*(Trainext3 - Carmult1)}) + ln\,(1+ e^{\text{Carfric1}*(Coawc2 - Carfric1)}) + ln\,(1+ e^{\text{Carfric1}*(Coamulc2 - Carfric1)}) + ln\,(1+ e^{\text{Carfric1}*(Trainwc3 - Carfric1)}) + ln\,(1+ e^{\text{Carfric1}*(Trainexc3 - Carfric1)}) + ln\,(1 + e^{\text{Carmulc1}*(Coawc2 - Carmulc1)}) + ln\,(1 + e^{\text{Carmulc1}*(Coamulc2 - Carmulc1)}) + ln\,(1 + e^{\text{Carmulc1}*(Trainwc3 - Carmulc1)}) + ln\,(1 + e^{\text{Carmulc1}*(Trainexc3 - Carmulc1)}) + \text{Male}_1 * DMale_1 + \text{FEmp}_1 * DFEmp_1 + \text{DInc}_1 * DInc_1 + \text{Age}_1 * DAge_1$$

Similarly other Regret functions for coach and train also were formulised in similar way.

Variables

Where: (with subscript n indicating modes)

$Carfrit_n$	***(Time by car information by friend, in Minutes)***
$Carmult_n$	***(Time by car information by multimodal website, in Pence)***
$Carfric_n$	***(Cost by car information by friend, in Minutes)***
$Carmultc_n$	***(Cost by car information by multimodal website, in Pence)***
$DMale_n$	***(Dummy, If male =1, otherwise = 0)***
$DFEmp_n$	***(Dummy, if full employment = 1, otherwise = 0)***
$DInc_n$	***(Dummy, if income over 30,000 pounds =1, otherwise = 0)***
$DAge_n$	***(Dummy, if age less than 40 =1, otherwise = 0)***
$Trainwt_n$	***(Time by train information by train website, in Minutes)***
$Trainext_n$	***(Time by train information by past experience, in Minutes)***
$Trainwc_n$	***(Cost by train information by train website, in Pence)***
$Trainexc_n$	***(Cost by train information by past experience, in Pence)***

$Coawt_n$	*(Time by coach information by coach website, in Minutes)*
$Coawc_n$	*(Cost by coach information by coach website, in Pence)*
$Coamult_n$	*(Time by coach information by multimodal website, in Minutes)*
$Coamulc_n$	*(Cost by coach information by multimodal website, in Pence)*

Coefficients

Where: (with subscript n indicating modes)

$Carfrit_n$	**(is a parameter vector to be estimated for Time by car information by friend, in Minutes)**
$Carmult_n$	**(is a parameter vector to be estimated for Time by car information by multimodal website, in Pence)**
$Carfric_n$	**(is a parameter vector to be estimated for Cost by car information by friend, in Minutes)**
$Carmultc_n$	**(is a parameter vector to be estimated for Cost by car information by multimodal website, in Pence)**
$Male_n$	**(is a parameter vector to be estimated for Dummy, If male =1, otherwise = 0)**
$FEmp_n$	**(is a parameter vector to be estimated for Dummy, if full employment = 1, otherwise = 0)**
Inc_n	**(is a parameter vector to be estimated for Dummy, if income over 30,000 pounds =1, otherwise = 0)**
Age_n	**(Dummy, if age less than 40 =1, otherwise = 0)**
$Trainext_n$	**(is a parameter vector to be estimated for Time by train information by past experience, in Minutes)**
$Trainwc_n$	**(is a parameter vector to be estimated for Cost by train information by train website, in Pence)**
$Trainexc_n$	**(is a parameter vector to be estimated for Cost by train information by past experience, in Pence)**
$Coawt_n$	**(is a parameter vector to be estimated for Time by coach information by coach website, in Minutes)**
$Coawc_n$	**(is a parameter vector to be estimated for Cost by coach information by coach website, in Pence)**

Table 6.13 Comparison between Uncorrected Method, Jack-knifed and Regret Mode Choice Models

File	ModelSPM16	Jacknifed	Regret1(All)	Regret (without fri & exp)
Converged	True	True	True	True
Observations	1143	1143	1143	1143
Final log (L)	-534.8	-534.8	-534.8	-534.7
D.O.F.	17	17	17	17
$Rho^2(0)$	0.325	0.325	0.437	0.434
$Rho^2(c)$	0.198	0.198	0.325	0.325
Carfrit1	-0.0070 (-1.7)	**-0.0072 (-1.8)**	- 0.0026 (-1.2)	**- 0.0032 (-2.3)**
Mult1	**-0.0047 (-2.5)**	**-0.0046 (-2.6)**	- 4.4e-4 (-0.3)	4.6e-4 (0.4)
Carfric1	**-2.8e-4 (-2.0)**	**-2.9e-4 (-1.8)**	- 1.4e-4 (-1.6)	1.8e-5 (1.5)
Mulc1	**-1.2e-4 (-1.8)**	-1.2e-4 (-1.5)	- 6.5e-5 (-1.6)	**- 6.6e-5 (-1.9)**
DMale1	**1.38 (3.8)**	**1.39 (3.9)**	**- 1.39 (-3.6)**	**- 1.40 (-3.9)**
DFEmp1	**1.22 (3.5)**	**1.21 (3.1)**	**- 1.23 (-3.6)**	**- 1.21 (-3.5)**
DInc1	**1.04 (2.6)**	**1.02 (2.0)**	**- 1.05 (-2.5)**	**- 1.06 (-3.0)**
DAge1	-1.13 (-1.5)	-0.900 (-1.1)	0.914 (1.2)	**0.900 (1.8)**
Trainext1	-0.0049 (-0.8)	-0.0055 (-0.8)	- 0.0045 (-1.1)	- 0.059 (-1.2)
Trainwc1	**-1.9e-4 (-3.5)**	**-1.9e-4 (-3.8)**	- 2.1e-5 (-0.5)	**- 2.2e-4 (-2.0)**
Trainexc1	**-1.2e-4 (-1.8)**	**-1.2e-4 (-1.8)**	2.0e-5 (0.6)	- 2.6e-5 (-0.9)
DMale2	0.192 (0.6)	0.198 (0.6)	- 0.206 (-0.6)	- 0.219 (-0.9)
DFEmp2	**0.584 (2.1)**	**0.575 (2.0)**	**- 0.595 (-2.2)**	**- 0.591 (-2.2)**
DInc2	**2.01 (6.2)**	**2.00 (4.9)**	**- 2.02 (-5.8)**	**- 2.16 (-6.0)**
DAge2	-1.06 (-1.5)	-0.816 (-1.1)	0.847 (1.3)	0.943 (1.1)
WT1	**-0.0056 (-2.3)**	**-0.0054 (-2.1)**	- 0.0019 (-0.8)	**- 0.0027 (-2.8)**
Coawc1	**-3.5e-4 (-2.1)**	**-3.4e-4 (-1.9)**	**- 2.0e-4 (-2.1)**	**- 2.4e-5 (-1.9)**

The results from Table 6.13 suggest that all the significant parameters in both the RRM models have the expected signs which indicate the regret paradigms of the models. Some of the variables had reverse signs but they were not significant. In terms of comparison between RRM and MNL models, it appears that the MNL model fits the data slightly better than its RRM counterpart. The significance levels of some of the parameter were increased in RRM. RRM estimates of the parameters were about half of their MNL counterparts. This suggests that the respondents had lower anticipated regret of the foregone alternatives.

6.8 Summary

This chapter has described the development of a range of models to explain and predict the choice of information sources as a function of a range of explanatory variables and their influence on the subsequent mode choice. The models are calibrated to data from CATI questionnaire survey conducted by the author at the main transport interchanges in Leeds. The range of models developed includes multinomial logit, nested logit, mixed logit to account for the correlations between choice alternatives and random regret minimisation model to study the concept of newly developed regret minimisation framework.

The chapter has also performed a useful comparison of traditional GEV modelling techniques with mixed logit modelling and random regret minimisation framework. The selected methodology was found to be capable of addressing various interesting complications, including merging of data from different sources, incorporating for the heterogeneity of tastes, accounting for repeated measurement problems and encompassing minimisation of regret by the travellers while choosing between alternatives. In order to combine the relative strengths of the RP and SP data sets and to permit for the development of the forecasts for both source and mode choice models, the two data sets are combined and joint RP-SP models of source and mode are estimated.

The source models explained source choice as a function of search time, specific information for the journey, prescriptive information, updating of the information, presence of advertisements, subscription of the source, effect of ownership of the source and other socioeconomic characteristics. The mode models were estimated by calibrating mode choice as a function of time and cost of the modes as provided by different information sources and other socioeconomic characteristics.

The process of model development was incremental starting with multinomial models for both SP and RP models. The process included estimating RP model for mode choice, followed by RP model for source choice, SP model for mode choice, SP model for source choice, combined RP model for source and mode, combined SP model for source and mode, combined RP and SP model for source and mode, mixed logit model for source and mode and finally random regret models were estimated.

The results from the RP (MNL) mode choice model indicate that generally the travellers exhibit an inclination to travel by car in long distance journeys and all else being equal; they prefer to select the alternative which offers the lowest travel time. Moreover in the good weather conditions they prefer to travel by car as in good weather driving is more enjoyable and less tiring. However, if the frequency of the travel is lower, the people are more likely to travel by public transport. Coach is not a popular mode in long distance journeys for high income people. On the other hand travellers travelling alone prefer train as compared to other modes. The RP (MNL) source choice model suggests that travellers consider websites as reliable source of travel information and they are more satisfied with the information provided to them by websites as compared to the information provided by the maps only. Moreover frequent travellers require, in addition to website, the information from a person who has travelled before. Higher educated people are at ease with technology and are more frequent users of websites.

The SP (MNL) mode choice model also establishes the fact that travellers will prefer any alternative which offers the lowest expected travel time and cost. It also suggests that males, in full time employment, and with higher income use car and train whereas coach is inconvenient for longer journeys. Moreover, younger people like to travel by coach. The SP (MNL) source choice model, on the other hand, suggests that travellers prefer an information source which offers the lowest expected search time. A website offering specific information on users' own criteria increases the utility of that website. Similarly, in comparison to real time information updating, daily and weekly updating have negative effect on the utility. Another important finding is that government websites with no advertisements and no subscription have higher attraction and credibility within the respondents.

The RP models are then calibrated for the choice of mode and source together in MNL and NL frameworks. The model RP (MNL) source+mode mode suggests that when many taste variations are at stake the travel time and travel cost becomes less important to the travellers. It also indicates that males, when travelling by car, tend to consult websites together with maps (rather than only maps). Similarly, males are less like to travel by train if they use website as a source of travel information. Income significantly predisposes people to use of website as a source of travel information and if they use website they are more likely to travel by car. Travellers who left their education by the age of 20 are less likely to use websites as source of travel information. Younger people are prone to travel by train and bus if they got travel information from a website whereas travellers in full time employment are more likely to travel by train as compared to car. The RP (NL) source+mode model confirms these results.

The Combined RP-SP model is calibrated next to use the advantages of both data sets. The model indicated that information from websites time is a very important attribute of and travellers consider it more important than time and cost actually spent on travelling. The value of time is reduced significantly with the introduction of web time as a variable. This suggests that the information from websites influence travel behaviour significantly and travellers consider it 4 times more important than the normal travel time and this is higher than the multiplier on waiting times on stops in the normal mode choice models.

The Jack-knife estimates of the SP model ModelSPS6 from section 6.4.4 for the source choice show that, regardless of the number of sub-samples most coefficients on the Jack-knife method are very close to those of the uncorrected model estimates. However the t ratios are slightly reduced in case of the Jack-knife method which indicates that the uncorrected model slightly overestimated the significance of the parameters. In the SP Mode choice model ModelSPM16 from section 6.4.3, the Jack-knife estimates show that the estimates are close to the MNL model, regardless of the number of sub-sample. These results show that the coefficients of the uncorrected model estimate were quite accurate despite of the repeated measurement problem.

The preferred model i.e. ModelSPS6 is taken from section 6.4.4 for the calibration of Mixed Logit Model (MMNL) for SP source choice model. In order to accommodate the repeated measurement problem, the search time was also identified to vary across individuals and not across observations. The results show that the mixed logit calibration of the model increased the efficiency of the model with increased t stats. The inclusion of the error component terms improves the fit to the data, but only marginally. The results confirm the findings of the MNL model. The standard deviations of the normally distributed components are not significant in both MMNL and Panel MMNL. This suggests that the MNL model specification represents the choice of the respondents adequately. Similarly model i.e. ModelSPM16 is taken from section 6.4.3 for the calibration of mixed logit model for SP mode choice. The results of the model also suggest that there is no significant difference in the models. The estimates of the standard deviations of the normally distributed terms were not significant (or if they are significant, as happens sometimes in the panel specification, this takes place at the expense of the precision of coefficients of other attributes, without significant gains in the overall Log likelihood value). This all means that the MNL model reflects the characteristics of the population satisfactorily.

The final model i.e. ModelSPS6 was also formulised in the RRM paradigm. Although when the choice sets are binary, RRM reduces to Random Utility Maximisation's (RUM's) linear additive binary logit model, the model was still calibrated to validate the applicability and performance of the RRM formulation. The results from the above table show that the RRM model is almost equal to the MNL model which confirms the appropriateness of the RRM model. The results show that the signs of the dummy variables (outside the regret function) are reverse as compared to the MNL model; this confirms that the model shows the regret of an individual not a utility as in the other models. The model i.e. ModelSPM16 is then used to estimate RRM model for SP mode choice model. Two models are estimated, one with all the cost and time variables inside the regret function. In the other model information from friend and previous experience are kept outside the regret function. The results suggest that all the significant parameters in both the RRM models have the expected signs which indicate the regret paradigms of the models. Some of the variables had reverse signs but they were not significant. In terms of comparison between RRM and MNL models, it appears that the MNL model fits the data slightly better than its RRM counterpart. The significance levels of some of the parameter were increased in RRM. RRM estimates of the parameters were about half of their MNL counterparts. This suggests that the respondents had lower anticipated regret of the foregone alternatives.

The above mentioned results and conclusions from the selected models give detailed insights for the choice of sources in order to make travel decisions. The results would help to understand travellers' behaviour in selecting the information sources and modes.

The research questions raised in section 3.1 are explored and the results are discussed below one by one respectively.

- The previous experience or initial information significantly affects the utility of the selected mode. The selected Model ModelSPM16 suggests that travellers having previous experience with a mode give more importance to it as compared to other modes. The model also suggests that the travellers anticipate more credence with their previous experience as compared to any other source.

- Credibility does vary from source to source. It was found that travellers give varying degrees of importance to different information sources. It was found that travellers give credence to government owned sources. Similarly travellers give more importance to their own previous experiences followed by multimodal websites, train websites, friends and coach websites respectively. Moreover the credibility of an information source increases significantly when two information sources give the same information.

- The travellers give less value to the information sources with low credibility. It was found that commercial websites with advertisements have reduced utility as compared to other sources. Furthermore, travellers try to gather more information from another source if they find that the information provided by a particular source is less credible.

- A number of factors affect credibility of a source, they include past experience of the traveller with that source/mode, ownership of the source, and presence of irrelevant information. It was also found that multimodal websites are considered more credible as compared to mono modal website.

- A website offering specific information on users' own criteria increases the utility of that website. Prescriptive information also increases the utility of choice of an information source. Similarly, the presentation of real time and updated information has increased influence on the travellers' choice.

- The study found that the effect of information provided by multimodal website is more as compared to the information provided by friends or mode specific websites. The study also found that if a person gets information from both sources, the market share increases even more. Moreover, the market share of the modes increases when information sources show decreased travel time and cost values and the maximum results are achieved when different information sources give the same information to the travellers. These results show that information from a information sources could be used to influence the mode choice of the travellers.

- A website with comprehensive information about the intended travel and offering specific information on users' own criteria increases the utility of that website. On the other hand presence of irrelevant information and advertisement significantly reduce the choice of that website and hence decrease its market share. Another important factor is that the government websites with no advertisements and no subscription have higher attraction and credibility within the respondents.

The models developed in the current chapter are applied to generate forecasts under different policy assumptions in the next chapter.

Chapter 7
Model Applications

7.1 General

The previous chapter reported the development of a range of mode choice and information source choice models. The objective of this chapter is to develop a suitable way to apply the models so as to forecast uptake of variant information sources and to identify its desirable properties. The remainder of this chapter is structured as follows. Section 7.2 sets out a forecasting methodology and provides a description of the forecasting assumptions. Sections 7.3 documents the forecasting of usage of information sources with desirable properties as estimated from the models on estimation sample. Section 7.4 shows the range of mode choice forecasts on estimation sample. Section 7.5 reports the forecasting of information source choice for expanded sample to represent population and section 7.6 does the same for mode choice forecasts. Finally section 7.7 discusses the policy implications of this research.

7.2 Forecasting Methodology

The aim of this section is to elaborate the methodology in which the source choice and mode choice models are applied to generate forecasts of aggregate demand for information sources with different hypothetical characteristics.

In this study sample enumeration method is used to generate forecasts. This approach rests on the assumption that the sample (on which the choice model is calibrated) can be made representative of the population through expansion. The forecast share for each alternative is estimated as

$$Q(i) = \sum_{n=1}^{N_s} w_n . P(i\text{I}X_g)$$

$$W(i) = \frac{1}{N_T}\sum_{n=1}^{N_s} w_n . P(i\text{I}X_g)$$

Where w_n, is a weight or expansion factor attached to decision maker *n* in order to make its sum representative to the population. If the sample is representative, the weight for each alternative is simply equal to N_T/N_S. Using this approach both the total demand Q(i) and the market share W(i) of a given choice alternative *i* are estimated.

The following assumptions have been made to the forecasting process

- The sample can be expanded to be representative of the population.
- Everything else remaining constant
- Hypothetical values of the desirable characteristics of information source can be used to understand the attractiveness of the sources.

7.3 Source Choice Forecasts on Estimation Sample

Using the methodology and assumptions outlined in section 7.2, the Jacknifed SP model for source choice developed in section 6.6 of the previous chapter was used to predict the choice probabilities for the source forecasts based on the estimation sample. A model based on SP data alone cannot be used to forecast market shares. This requires additional support from observed (RP) data. The effects, calculated using the SP only model, are based on un-calibrated models and as such the numerical magnitudes are only valid in comparisons across models. These models cannot be used to forecast market shares without calibration using revealed preference shares. There was no common variable between the SP source models and other RP and SP models and hence it was not possible to combine both data sets. However, the forecasting has been carried out on the SP model on the understanding that model was calibrated on the same sample and search time had the same value of time as that of other RP models. The selected Jacknifed SP Source Model has following utility function,

1. $U_1 = \text{Seatime}_1 * Seatime_1 + \text{ComAdsub}_1 * DComAdsub_1 + \text{Com}_1 * DCom_1 + \text{InfoSpec}_1 * DInfoSpec_1 + \text{InfoPerc}_1 * DInfoPerc_1 + \text{Updaily}_1 * DUpdaily_1$

2. $U_2 = \text{Seatime}_1 * Seatime_2 + \text{ComAdsub}_2 * DComAdsub_2 + \text{Com}_1 * DCom_2 + \text{InfoSpec}_2 * DInfoSpec_2 + \text{InfoPerc}_2 * DInfoPerc_2 + \text{Updaily}_2 * DUpdaily_2$

Variables

Where: (with subscript n websites e.g. A or B)

$DComAdSub_n$:	***(dummy that equals 1 if Commercial Ads Subs; otherwise 0)***
$DCom_n$:	***(dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)***
$Seatime_n$	***(Search time equals to 5 min, 10 min, 15 min)***
$DUpdaily_n$:	***(dummy that equals 1 if Website updates daily; otherwise 0)***
$DInfoSpec_n$:	***(dummy that equals 1 if Specific info available; otherwise 0)***
$DInfoPerc_n$:	***(dummy that equals 1 if Info w.r.t own criteria; otherwise 0)***

Coefficients

Where: (with subscript n indicating websites)

$ComAdSub_n$:	**(is a parameter vector to be estimated for dummy that equals 1 if Commercial Ads Subs; otherwise 0)**
Com_n:	**(is a parameter vector to be estimated for dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)**
$Seatime_n$	**(5 min, 10 min, 15 min)**
$Updaily_n$:	**(is a parameter vector to be estimated for dummy that equals 1 if Website updates daily; otherwise 0)**
$InfoSpec_n$:	**(is a parameter vector to be estimated for dummy that equals 1 if Specific info available; otherwise 0)**
$InfoPerc_n$:	**(is a parameter vector to be estimated for dummy that equals 1 if Info w.r.t own criteria; otherwise 0)**

The tests carried out on this model included the analysis of search time, type of the information provided, and whether the sources are commercial requiring subscriptions and with advertisements. These policy variables were selected because they were found significant in the model estimations and needed further analysis. Table 7.1 shows the forecasts for different values of the search time compared to the original models. The methodology employed in this exercise involves forecasting of use of information source (website) with different values of the search time and resulting market share, everything else remaining same. For each respondent in the sample, a forecast is made for the probability that it will choose website A or website B as presented to them in the SP exercise. The choice probabilities for each respondent are then aggregated to generate forecasts.

Table 7.1 shows that the search time is very important characteristic of a website. Reduced search time causes more people to use the website.

Table 7.1 Effect of Search Time on Market Share

Search time in minutes for LHS Website	Market share % of LHSWebsite	Market share % of LHSWebsite (Search time of RHS website 30 min)
combination of search time as presented in the SP exercise	51	95
30	5	51
20	21	99
15	35	100
10	51	100
5	67	100
3	73	100
2	76	100
1	79	100

Figure 7.1 shows the trend for this relationship. The first curve represents various search times for LHS website and corresponding market share keeping the RHS website with the levels as presented in the SP exercise whereas second curve shows search times for LHS website and corresponding market share by restricting the levels of search time for RHS website as 30 minutes. Both curves show slightly nonlinear trends. The slope of the first curve increases as the search time decreases, the market shares increase from 5 % at 30 minutes of search time to 79 % at search time of 1 minute. This shows the increase in the importance of search time with every minute saved in the search. The trend is somehow reverse for the curve B where search time for website B is restricted to be 30 minutes for every case. Here the curve is concave in contrast to the first curve which suggests that as search time of website A approaches that of website B, the effect of unit change increases.

Another important variable is the presentation of the information about the journey. Table 7.2 explains the importance of the information and its effect on market share of the website. It reveals that the choice of type of information is also related to the search time of the websites. If the website provides information as per the traveller's criteria, the market share of the website raises from 41% when no information is provided to 63%. But when the search time of the competing website is increased the importance of specific information reduces. Similarly prescriptive information becomes also has the same tendency. This could be due to the complete reliance of the decision maker on the website as there is no competition between the websites and the selected website is much better and credible. Similarly, when a commercial website has advertisements on it, its market share decreases from 60% to 41%. On the other hand, if a commercial website has no advertisements and no subscription its share increases to 92%. Updating of the information also plays an important part in this phenomenon. And the real time information updating increase the share of a website from 38% to 63% when the information was updated only once daily. Table 7.3 describes this relationship in detail.

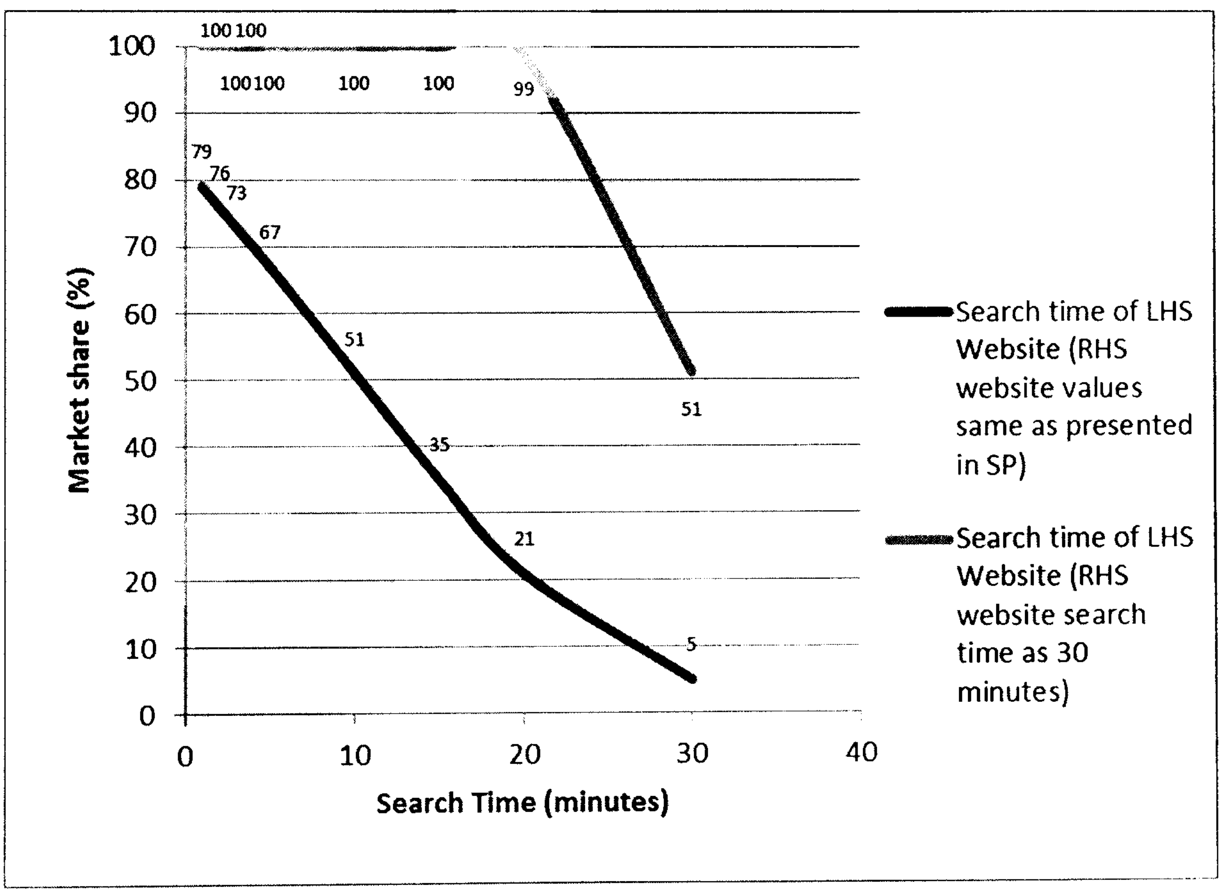

Figure 7.1 Search Time and Corresponding Market Share

Table 7.2 Effect of Type of Information on Market Share

Attribute Level	Market share of website A (%)		
	Search time as presented in SP	Website B Search time 15 minutes	Website B Search time 30 minutes
Information as per criteria of the traveller (InfoSpec)			
No information provided	41	54	88
As presented in SP	51	67	95
Information provided	63	76	96
Prescriptive information (InfoPerc)			
No information provided	32	47	89
As presented in SP	51	67	95
Information provided	86	93	99

Table 7.3 Effect of Adverts and Subscription on Market Share

Attribute Level	Market share for website A (%)		
	Search time as presented in SP	Website B Search time 15 minutes	Website B Search time 30 minutes
Commercial website with adverts and subscription			
Commercial website no adverts but with subscription	60	77	98
As presented in SP	51	67	95
Government with adverts with subscription	41	56	93
Commercial website no adverts no subscription			
Government with adverts and subscription	46	64	94
As presented in SP	51	67	95
Commercial no adverts no subscription	92	97	100
Daily updating of Information			
Only daily updates	38	52	87
As presented in SP	51	67	95
Real-time updating	63	76	96

7.4 Mode Choice Forecasts on Estimation Sample

Here too, the combined RPSP model developed in section 6.4.7 of the previous chapter was used to predict the choice probabilities for the mode forecasts. The reason for this choice was again that the model had been developed on the aggregation of both data sets (RP and two SP). The model has following utility function,

1. $U_{(car)} = \mathrm{Time}_1 * Carfrit_1 + \mathrm{Time}_1 * Carmult_1 + \mathrm{Cost}_1 * Carfric_1 + \mathrm{Cost}_{1\,1} * Carmulc_1$

2. $U_{(coach)} = \mathrm{Time}_1 * Coawt_2 + \mathrm{Time}_1 * Coamult_2 + \mathrm{Cost}_1 * Coawc_2 + \mathrm{Cost}_1 * Coamulc_2$

3. $U_{(train)} = \mathrm{Time}_1 * Trainwt_3 + \mathrm{Time}_1 * Trainext_3 + \mathrm{Cost}_1 * Trainwc_3 + \mathrm{Cost}_1 * Trainexc_3$

Variables

Where: (with subscript n indicating modes)

$Carfrit_n$	***(Time by car information by friend, in Minutes)***
$Carmult_n$	***(Time by car information by multimodal website, in Pence)***
$Carfric_n$	***(Cost by car information by friend, in Minutes)***
$Carmultc_n$	***(Cost by car information by multimodal website, in Pence)***
$Trainwt_n$	***(Time by train information by train website, in Minutes)***
$Trainext_n$	***(Time by train information by past experience, in Minutes)***
$Trainwc_n$	***(Cost by train information by train website, in Pence)***
$Trainexc_n$	***(Cost by train information by past experience, in Pence)***
$Coawt_n$	***(Time by coach information by coach website, in Minutes)***
$Coawc_n$	***(Cost by coach information by coach website, in Pence)***
$Coamult_n$	***(Time by coach information by multimodal website, in Minutes)***
$Coamulc_n$	***(Cost by coach information by multimodal website, in Pence)***

Coefficients

Where: (with subscript n indicating alternatives)

Time_n	**(is a parameter vector to be estimated for Time, Generic, in Minutes)**
Cost_n	**(is a parameter vector to be estimated for Cost, Generic, in Pence)**

The methodology involves forecasting using different values for time and cost from different information sources and subsequent mode choice. For each respondent in the sample, a forecast is made of the probability that it will choose car, coach or train. Table 7.4 shows the resulting market share of the modes under different time and cost values provided by the information sources. First the lowest levels of the SP design were tested for each information source. Then the values of time and cost were reduced by 20% from the SP design lowest level for each source and mode. Later both time and cost were reduced together for each of the source. And finally the time and cost were reduced for one source and it was made worse for other modes and sources i.e. highest SP design value. The table reveals that the effect of information provided by multimodal website is more as compared to the information provided by friends for car users. On the other hand previous experience of a travel with a mode influences the choice more if it is also validated by other sources. Table 7.4 also reveals that if a person gets information from both sources, the market share increases even more.

Table 7.4 Effect on the Market Share of the Modes w.r.t Time and Cost as Presented by Different Sources (RPSP Model)

Mode / Time and Cost Values	Car		Train		Coach	
	Friend	**Multimodal Website**	**Train Website**	**Previous Experience**	**Coach Website**	**Multimodal Website**
Average of SP	52	52	30	30	17	17
Lowest value of SP	54	56	34	34	19	19
Time (-20%)[10]	57	59	35	35	21	22
Cost (-20%)[11]	54	56	37	37	19	19
Both (-20%)	59	62	41	42	22	23
From both sources	69		52		28	

7.5 Source Choice Forecasts on Expanded Sample

As was shown in table 5.1 the estimation sample is almost representative of the NTS data for long distance travellers. However, it can be observed that the income levels of the base sample are not representative of the population (NTS 2006). Hence, it was decided to adjust the weights (w_n) so that the sample can be considered to be representative. The process involves weighing a base sample so that the sum of the weighted observations show similar aggregate characteristics of the relevant planning

[10] The 20% reduced time for car, train and coach was calculated to be 200, 170 and 300 minutes respectively.
[11] The 20% reduced cost for car, train and coach was calculated to be 4000, 6000 and 3500 pence respectively.

authorities. The base sample is the sample obtained from the questionnaire survey conducted by the author at the main transport interchanges in Leeds. Whereas, the target population is the National Travel Survey (NTS) 2006 population of long distance travellers in the United Kingdom. Using the methodology and assumptions outlined in section 7.2, the combined RPSP model developed in section 6.4.7 of the previous chapter was used to predict the choice probabilities for the mode choice RP and SP model forecasts. Similarly the Jacknifed SP source model developed in section 6.6 was also used to predict the choice probabilities for the source forecasts on the assumption that it has been drawn from the same sample and has the same value of time. The selected SPRP model and Jacknifed SP source model have the same utility functions and variables as described in section 7.3.

The tests carried out included the analysis of search time, type of the information provided, and whether the sources are commercial requiring subscriptions and with advertisements. Table 7.5 shows the forecasts for different values of the search time compared to the original models. The methodology employed was same as described in section 7.3 which involves forecasting of use of information source (website) with different values of the search time and resulting market share, everything else remaining same. For each respondent in the sample, a forecast is made for the probability that it will choose website A or website B as presented to them in the SP exercise. The choice probabilities of each respondent are then weighted and aggregated to generate forecasts.

Table 7.5 shows, as expected, that the search time is very important characteristic of a website. Reduced search time attracts more people to use the website as an information source. Figure 7.2 shows the trend for this relationship. The first curve represents various search times for LHS website and corresponding market share keeping the RHS website with the levels as presented in the SP exercise whereas second curve shows search times for LHS website and corresponding market share by restricting the levels of search time for RHS website as 30 minutes. Table 7.5 reveals that the adjusted sample reflects even more people to use websites as compared to the previous forecast of the estimation sample and the market shares increase from 5 % at 30 minutes of search time to 97 % at search time of 1 minute. This shows additional 18% percent users as compared to the previous forecast. This confirms that the adjustment of the sample with the higher income people increases the use of website as higher educated people are at ease with the websites as an information source which was also concluded from the analysis in chapter 6. This also gives an interesting insight to the website owners and transport planning authorities to attract more users to implement their transport policy by putting more efforts in reducing the search time of the information.

Table 7.5 Effect of Search Time on Market Share (Expanded Sample)

Search time in minutes for LHS Website	Market share % of LHS Website
combination of search time as presented in the SP exercise	62
30	5
20	25
15	43
10	63
5	82
3	90
2	93
1	97

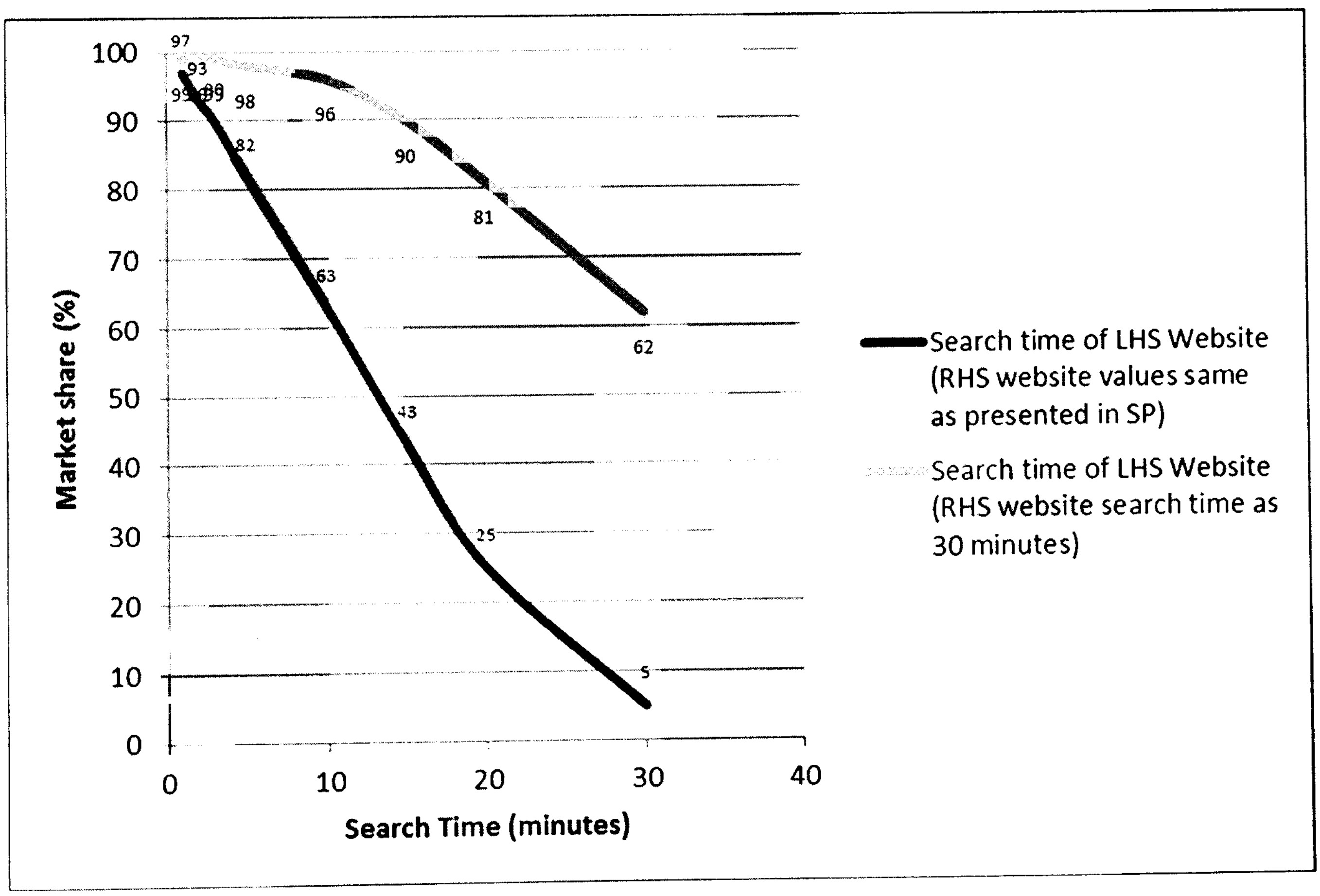

Figure 7.2 Search Time and Corresponding Market Share (Expanded Sample)

Again the important variable to be checked is the presentation of the information about the journey. Table 7.6 explains the importance of the information and its effect on market share of the website. It reveals that if the website provides information as per criteria of the traveller, the market share of the website raises from 51 % when no information provided to 77% and even more people are attracted if the search time of the competitive website is increased and the market share reaches 94%. On the other hand, prescriptive information also has a large effect on the market share of a website and it even goes to 100% of the market . This shows that with the increase of the higher income share in the population the credibility of websites increases and more travellers believe in the information prescribed by the website. This again is a very interesting result for the policy makers as they can prescribe different modes to distribute the travellers to obtain maximum performance of the available resources.

Table 7.6 Effect of Type of Information on Market Share (Expanded Sample)

Attribute Level	Market share of website A (%)	
	Search time as presented in SP	**Website B Search time 15 minutes**
Information as per criteria of the traveller (InfoSpec)		
No information provided	51	66
As presented in SP	62	82
Information provided	77	94
Prescriptive information (InfoPerc)		
No information provided	39	58
As presented in SP	62	82
Information provided	100	100

Similarly, again when any commercial websites has advertisements on it, its market share decreases from original 74% to 48%. On the other hand, if a commercial website has no advertisements and no subscription its share increases to 100%. Table 7.7 describes this relationship in detail. This shows that the irrelevant information or subscription costs are, by large, not popular in the website users.

Table 7.7 Effect of Adverts and Subscription on Market Share (Expanded Sample)

Attribute Level	Market share for website A (%)	
	Search time as presented in SP	Website B Search time 15 minutes
Commercial website with adverts and subscription		
Website with either adverts or subscription	74	95
As presented in SP	62	82
Commercial with adverts with subscription	48	69
Commercial website no adverts no subscription		
Website with adverts and/or subscription	57	79
As presented in SP	62	82
Commercial no adverts no subscription	100	100
Daily updating of Information		
Only daily updates	46	64
As presented in SP	62	82
Real-time updating	78	93

7.6 Mode Choice Forecasts on Expanded Sample

The combined RPSP model developed in section 6.4.7 of the previous chapter was again used to predict the choice probabilities for the mode forecasts. The reason for this choice was again that the model had been developed on the aggregation of both data sets (RP and two SP). The model has the same utility as mention earlier in section 7.4.

The methodology as described earlier involves forecasting using different values for time and cost from different information sources and subsequent mode choice. For each respondent in the sample, a forecast is made of the probability that it will choose car, coach or train. Table 7.8 reveals these resulting market shares in details. The table reveals that the effect of information provided by multimodal website is more as compared to the information provided by friends for car users. On the other hand previous experience of a travel with a mode influences the choice more if it is also validated by other sources. Table 7.8 also reveals that if a person gets information from both sources, the market share increases even more.

Table 7.8 Effect on the Market Share of the Modes w.r.t Time and Cost as Presented by Different Sources (RPSP Model) on Expanded Sample

Mode / Time and Cost Values	Car		Train		Coach	
	Friend	Multimodal Website	Train Website	Previous Experience	Coach Website	Multimodal Website
Average of SP	64	64	20	20	16	16
Lowest value of SP	67	70	24	23	18	18
Time (-20%)[12]	70	73	24	24	19	20
Cost (-20%)[13]	67	69	26	26	17	17
Both (-20%)	73	77	31	31	21	21
From both sources	85		43		26	

Table 7.8 shows that market share of the modes increase when information sources show decreased travel time and cost values. This shows that information sources could be used as a policy measure to distribute the travel miles among the modes. And the maximum results are achieved when different information sources give the same information to the travellers. This shows the effect of credibility on the sources as travellers tend to believe more when they observe same information about the journey from multiple sources.

7.7 Policy Implications of This Study

7.7.1 Search time

The study found that the search time of an information source is a very important factor to attract the travellers to that source. In addition to that, the search time of a website can be used to influence other travel decisions of the travellers including mode choice. The study found that the travellers consider search time, 4 times more important to the time spent on travelling. The study showed the increase in the market shares of the information sources with varying search times and the resulting market share can even go up to 97%. This means a website can attract many users with simply improving its

[12] The 20% reduced time for car, train and coach was calculated to be 200, 170 and 300 minutes respectively.
[13] The 20% reduced cost for car, train and coach was calculated to be 4000, 6000 and 3500 pence respectively.

performance. These results could be used by the website owners to increase their market share and influence in the travel market. These results can be adopted by the public transport operators and policy makers to simply attract more travellers on their websites by improving them and distribute the overall modal shift.

7.7.2 Provision of Information to Users' Criteria

The study also found this very important factor to attract the users to an information source. The provisions of information as desired by a traveller can increase the use of information source up to 94%. This means that the information source providers can, from the search patterns of the users, provide specific information that is useful to their registered users. The type of information and subsequent details as desired by the users, increases the credibility of the website and attract more users.

7.7.3 Prescriptive Information

Prescriptive information also has a large effect on the market share of a website and it even goes to 100% of the market. This is also a very important finding of this study and this can be viewed as a measure of credibility of any website. Travellers do like prescriptive information and thus they can be influenced to improve the overall transport network.

7.7.4 Advertisements and Subscriptions

This study revealed that the advertisements and subscription requirements of a website has a big impact on its use. When a commercial websites has advertisements on it, its market share decreases from 74% to 48% all else being equal. On the other hand, if a commercial website has no advertisements and no subscription, its share can go up to 100%. This means that the design and the content of an information source could have huge impacts on their use. If a website has irrelevant information and adverts on it, its use is largely reduced due to these factors. The study tried a subscription rate of £5 as a variable to understand the issue of subscription decisions. It was revealed that people generally don't like the idea of subscription and are reluctant to pay any subscription fee. Hence the government websites with no subscription fee could be used to influence the travellers more conveniently.

7.7.5 Effect of Information Sources on Mode Choice

The study gave detailed insights on the effect of different types of information sources on the mode choice decisions of the travellers. It was found that there is a lot of difference between the impact of a multimodal website and a mono-modal website on mode choice. It was found that the effect of information, provided by multimodal websites, is even more important than the advice from friends. This is a very interesting finding as this shows the effect of websites on the decisions of common travellers. It was found that even the previous experience of a traveller has the same impact on her travel decisions as of any website. This is also a very interesting result. This also shows the usability of websites by the common people nowadays and its huge impact on them. On the other hand previous experience of a travel with a mode influences the choice more if it is also validated by other sources. It also reveals that if a person gets information from both sources, the market share increases even more. It was found that the market share of the modes increase when information sources show decreased travel time and cost values. This shows that information sources could be used as a policy measure to distribute the travel miles among the modes. And the maximum results are achieved when different information sources give the same information to the travellers. This shows the effect of credibility on the sources as travellers tend to believe more when they observe same information about the journey from multiple sources.

7.8 Methodological Implications of This Study

7.8.1 Repeated Measurement Problem

This study tested the repeated measurement problem is the data analysis. The main SP survey results were analysed using both uncorrected and Jacknife methods. The results showed that the data of this study were not much influenced by the repeated measurement problem; however, the t-stats of some variables were much improved. The study also used panel data formulations in the analysis of mixed logit models to test the repeated measurement issues. It is recommended that such procedures should be adopted in any study and if the data is influenced with this problem it should be corrected.

7.8.2 Tests on SP design

It is recommended to test the SP design with the pilot surveys in addition to the simulation tests. As discussed in the Chapter 5, this study tested Sp survey design through pilot surveys and simulations. As a result of those steps the initial design required significant changes and modifications. It becomes more important while testing new concepts as were designed in this study. Simulations alone cannot establish whether the SP design is capable of using correct values for the levels and whether the respondents are able to understand what is being asked to them. This necessitates the use of pilot survey to allow the design to be more user friendly and efficient.

7.8.3 CATI Survey

The survey was conducted at the main transport centres in Leeds and two pilot surveys were conducted in the University of Leeds. The response rate initially was very low and it was very difficult to recruit and survey the respondent at the same time. Hence in this study, CATI survey was adopted which not only gave the satisfactory response rate but also incorporated complex branching and looping of the questions which were otherwise very difficult to carry out. Hence it is recommended to use CATI surveys if the study requires branching and looping of the questions. A problem which sometimes occurred in this study was difficult to track the respondents at the agreed time for the surveys. But the subsequent calls could reduce this difficulty.

7.8.4 Survey Software

In this study commercially available software was used to carry out the CATI survey. The software was easy to use and was better as compared to other online software which requires respondents to be users of the internet. In this research it was found that the use of software is quite easy and could reduce the enormous paperwork of the normal paper based questionnaires. It is therefore concluded that this software that this software was a good tool for designing the SP surveys and implementing them easily.

Chapter 8
Conclusions

8.1 Introduction

Traveller information is now not a new concept and is long been used to improve the traffic congestion and many types of information services and products are already in the market and research is continuing to introduce the second generation of these systems. This study takes into account the properties of available information sources in the market, their attractiveness and use.

The motivation behind this research was to investigate the factors that people consider in selecting different sources of information. It is important to explicitly model the abstract terms involved in the process of information acquisition and subsequent travel decisions. The understanding of these factors not only benefits information service providers, manufacturers and suppliers of these products to understand the impacts of their ATIS and predict their profitability but also helps government agencies and policy makers to stimulate changes in the travel decisions of travellers and influence their mode choice, and also benefit public transport providers to use these sources in attracting and retaining customers.

The goal of this study was to develop a comprehensive choice model that can capture the information acquisition process by predicting the choice of information sources together with its effects on mode choice of the travellers.

The main purpose of this research was to study and predict the travellers' choice of information sources and subsequently of mode. Specifically, the objectives of this study were:

- To conduct a travel behaviour survey, to investigate the travellers' choice of information sources in different scenarios.
- To model the choice of information source and subsequent mode choice.
- To analyse and evaluate the impact of information on mode choice.

Additional sub-objectives of the thesis were:

- To identify the abstract terms involved in the process of travel information acquisition and necessary to be tested in the models of information sources and subsequent mode choices.
- To develop a decision making framework for the travel information acquisition process.
- To expand the travellers' choice set to include different combinations of the viable sources of information.
- To include policy sensitive variables including credibility.
- To analyse the effects of travel planning websites on travel decisions.
- To study the influence of information sources on the mode choice decisions of the travellers under various circumstances.
- To establish the link between content, design, advertisements, and presentation of information on overall modal shift.
- To analyse the travellers' treatment of low credible sources and the factors which affect the credibility of a source.
- To analyse information source and mode choice models using Mixed Logit (MMNL) models with individual specific parameters.
- To use the newly developed Random Regret Minimisation (RRM) models to estimate the information source choices and subsequent mode choice model.

To achieve these objectives, the study employs a wide range of modelling methodologies and draws on the relative strengths of data sets including a revealed and stated preference questionnaire survey. A summary of the achievements of each chapter is discussed in section 8.2 which is followed by suggestions for future work in section 8.3 and finally, section 8.4 presents a summary of the thesis as a whole.

8.2 Summary of Achievements

8.2.1 Literature Review

A review of relevant literature is presented in the second chapter. The chapter is mostly focused on building an understanding of effects of different types of information on the traveller's decisions.

a. Traveller Behaviour under Information

The first part of chapter 2 provides a review of the travellers' behaviour under the influence of information. The literature review on the information need and type suggests that, in most of the cases the people require information about travel time and travel cost for different modes. The literature also suggests that age, sex, income level and education are the key factors that influence the use of traveller. Trip purpose has been found to be very important factor that drives individuals to use traveller information. It has been reported by various studies that different trip purposes have produced different responses towards traffic information. It was found that commuters were less likely to divert to alternate route under information as compared to other trip purposes. Literature suggests that combination of prescriptive and quantitative information influence travellers more in their decisions as compared to only qualitative information.

The credibility of the information source is found to be an important determinant which influences the travellers'' decisions. It has been found that travellers tend to give less credence to the information in comparison to what they actually observe with their own eyes. Moreover they prefer to test the credibility of the information randomly before considering using that source and the influence of traffic information on route choice depended on whether the information was credible, relevant and clear with credibility heavily influencing its compliance. This part of chapter 2 later discusses the models of response and their effect on the network. The absence of information in the context of route choice models with various methodologies was discussed to investigate the impacts of the information on network performance. This is followed by the detailed discussion on evaluation of benefits of information sources. It was found that if implemented correctly, ATIS has the potential to mitigate the problems of congestion, environment and network performance by influencing travellers' route choice behaviour. Furthermore willingness to pay for ATIS is also significant for accurate and prescriptive information.

b. Theories of Information Search and Use

Because the traveller information is effective in mitigating the problem of congestion and influencing the travellers, the second part of chapter 2 extends the literature review to provide a summary of methodologies aimed at modelling information search and use. The review indentifies following theories of consumer choice.

- Maximisation Concept
- Satisficing Concept
- Habit Execution
- Effort-Accuracy Trade-off
- Search Theory in Labour Economics
- Random Regret Minimisation

A detailed discussion of each of the above mentioned theories is provided. But in summary, it is assumed, in almost all the models developed in this study, that decisions are based on the utility maximisation concept. Since there have been significant advances in the choice modelling over last 10 yeas, a new concept of random regret minimisation was also tested in some of the models in this thesis to understand the effects of information sources more comprehensively.

8.2.2 Development of Methodology

In Chapter 3, a detailed discussion is presented to develop the methodology of this study. This methodology was developed by the author to better comprehend the importance of information, travel decisions and their inter dependence. The factors to be studied were indentified and grouped in four categories: variables associated to decision makers' characteristics, attributes associated with the information sources, characteristics of the travel modes, and other external circumstances.

A conceptual framework of the information acquisition process was developed. It was deduced that the decision to acquire travel information depends on the external circumstances e.g. bad weather, congestion, incident etc.; personal attitudes and preferences; and personal circumstances. Similarly once a decision is made to acquire information, the choice of source depends upon its accessibility and credibility and the individuals' awareness of it. The attributes are thus classified in three categories, information source attributes; respondent characteristics; and scenario attributes. The information source attributes are source owner, search time, frequency of updating of information, type of information presented, presence of advertisements, type of information source, presentation of information, capabilities, price and subscription, spending on advertisements, coverage in search engines, year of start, coverage in newspapers & articles. The scenario attributes include trip purpose, bad weather, congestion, incident occurrence, and accompanying travellers. The personal characteristics include personal segmentation, trip frequency, travel time, frequency of using source, attitudes to optimise the journey, habit and other individual characteristics. The rest of the chapter includes the decisions made on the selection of data sources, target population, sampling strategy, sample size and choice of models to be estimated. Finally the chapter links the developed conceptual framework with a more precise and practical modelling framework and identifies the above discussed determinants precisely and links them with the developed modelling choices.

The novel features for this work include:

- The identification of the abstract terms (variables) necessary to be tested in the models of information sources and subsequent mode choice.
- Development of a decision making framework for the travel information acquisition process.

8.2.3 Data Collection and Analysis

This study employed Computer Assisted Telephone Interview (CATI) technique to collect data. The main questionnaire had four parts. The first part gathered Revealed Preference (RP) data for the respondents' most recent long journey (over 50 miles) and included questions on frequency of travel to that destination, purpose of visit, the chosen mode, other available modes, the external circumstances of the journey, and the use and effect of any information source used while planning that journey. The second part included Stated Preference (SP) survey questions designed to investigate the traveller's choice of information sources and subsequent mode choice when making long journeys. The third part included general questions about their attitudes towards different sources of information and their normal search patterns. The final part contained questions about the traveller's socio-economic characteristics. The details of the questionnaire and design of SP survey is discussed in detail in Chapter 4.

The first SP exercise (SP1) offered the respondents two choices for the selection of the information source under a hypothetical travel situation. The attributes included ownership of the source; type of information; search time; presence of advertisements; frequency with which the information is updated; and any subscription cost. *Ownership of the information source* could be either "government" or "private". For *type of information* about travel time and delay, three levels were considered: "Descriptive Quantitative real time information", "Descriptive Qualitative information", and "Prescriptive information". For the *search time* three levels were considered, "5 min", "10 min" and "15 min". For the *presence of advertisement* there were only two possibilities: advertisements are either "there" or they are "not". For *frequency of information* changes on the source again three levels were considered, "real time", "daily" and "weekly". For *subscription* again there were two levels, "no subscription" and "£5 already paid".

The second SP exercise (SP2) required the respondent to choose between three modes. Each mode had at least two sources and time and cost attributes. The base levels were selected from the current travel time and costs as described by information sources during normal conditions in spring 2008. The other values were +20% and -20% deviations from these average values. A survey software WinMint was used as a CATI tool. The process of questionnaire development included simulation test and two pilot surveys before the main survey. In all about 950 members of the public were recruited at the main long distance transport interchanges in Leeds i.e. the coach station and the train station to achieve the target sample size of 300. See chapter 5 for descriptive analysis of the sample.

The novel features for this work include:

- The development of an interactive CATI questionnaire with complicated branching and looping.
- Design of an SP exercise with attributes and levels of information sources and modes together.
- Levels of perceived cost and time as offered by different information sources.

8.2.4 RP MNL Model

Multinomial logit models were estimated for the choice of mode and source from the RP data. The data consists of respondents' past behaviour for the last long journey (over 50 miles) from the survey developed for this thesis. Variables used in the model were selected as discussed in Chapter 3 and were added to the model incrementally (from simpler to more complex).

a. Mode Choice Model

A Multinomial Logit Model (MNL) was constructed with the dependent variable being the choice among car, coach and train. The base case for this model is travelling by car.

Significant findings of the research indicate:

- Users exhibit an inclination to travel by car in long distance journeys, all else being equal.
- Travellers prefer to select the alternative which offers the lowest travel time.
- Travellers in good weather try to travel by car as in good weather driving is more enjoyable/less tiring.
- If the frequency of the travel is lower, the people are more likely to travel by train and coach.
- Higher income travellers don't like to travel by coach in long journeys
- Travellers travelling alone in long journeys prefer train as compared to other modes.

b. Source Choice Model

The model framework incorporates MNL models showing the probability that a given traveller will choose among different important information sources as gathered from the data i.e. onlyweb, friend+web, map+web, and map+friend+web.

The novel features for this work include

- The expansion of the travellers' choice set to include different combination of viable sources of information.
- The inclusion of policy sensitive variables including credibility.

Significant findings of the research indicate:

- Travellers that are website users believe more in the information provided by the website alone as compared to the information provided by friends.
- Travellers consider websites as reliable source of travel information.
- Travellers that use websites are more satisfied with the information provided to them as compared to the information provided by the maps only.
- Peak period travellers do like to stay with the maps only.
- Higher educated people are at ease with technology and are more frequent users of websites.

8.2.5 SP MNL Model

a. Mode Choice Model

Multinomial models were estimated for the choice of mode on the SP data developed in this study. As mentioned earlier the data consists of respondent's stated choice between the three mode alternatives (car, train and coach) under the influence of information. A Multinomial Logit Model (MNL) is constructed with the dependent variable being the choice among car, coach and train. The base case for this model being travelling by coach.

The following results are worthy of note:

- Travellers will prefer any alternative which offers the lowest expected travel time and cost.
- Being male, in full time employment, and with higher income increased the propensity to travel by car and train.
- Coach is inconvenient for longer journeys.
- Younger people like to travel by coach.

b. Source Choice Model

Multinomial models were estimated for the choice of website as source to analyse different important attributes and their influence on travellers. A Multinomial Logit Model (MNL) is constructed with the dependent variable being the choice among website one and website two. The base case for this model is website 1. As the alternatives in the model are only "website one" and "website two", different variables were interacted with each other to reveal the full effect of the different attributes. Search time, information update and type of information were, in turn, used as interactive variables and multiplied with the other dummy variables.

Key properties of the estimated coefficients are summarised below:

- Travellers prefer an alternative which offers the lowest expected search time.
- Specific information under the own criteria of the respondents increases the utility of a website.
- In comparison to real time information updating, daily and weekly updating have a negative effect on the utility.
- Government websites with no advertisements and no subscription have higher attraction and credibility within the respondents.

8.2.6 RP MNL Model for Source and Mode Choice

Multinomial models were estimated for the choice of mode and source together using the RP data with the dependent variable being the combined choice of mode and source. The modes were car, coach and train whereas the sources used by respondents in their past journey were only web; friend with web; map with web; map with friend and web; and only map. Hence by combining these alternatives, the resulting combined alternatives should be fifteen but as map was not used by the coach and train travellers, the alternatives were reduced to nine. The base case for the dummy variables was travelling by car and using website as the source of information. Another base case of travelling by car and using map is also tested.

The important properties of the estimated coefficients are summarised below.

- When many taste variations are at stake the travel time and travel cost becomes less important to the travellers.
- When travelling by car the males tend to choose to consult websites as well as maps (rather than only maps).
- Males are less like to travel by train if they use website as source of travel information.
- Travellers who left their education by the age of 20 are less likely to use website as source of travel information.
- High income people tend to use website more as compared to other sources similarly if they used website as a source of information, they tend to travel by car.
- Younger travellers care more for other sources in addition to websites.
- Younger people are prone to travel by train and bus if they got travel information from a website.
- Travellers in full time employment are more likely to travel by train as compared to car.

8.2.7 RP NL Model for Source and Mode Choice

Nested Logit (NL) Models were estimated for the combined choice of mode and source using the RP data. Nested Logit Models (NL) were also constructed with the dependent variable being the combined choice of mode and source. The modes were car, coach and train where as the sources used by respondents in their past journey were only web; friend with web; map with web; map with friend and web; and only map. Hence by combining these alternatives, the resulting combined alternatives should be fifteen but as map was not used by the coach and train travellers the alternatives were reduced to nine. The base case for this model was travelling by car and using website as source of information. Various models with the different nesting structures were tried with different combination of variables and nesting coefficients. The initial models did not converge due to correlation between the variables. Then a model with information sources on top and modes below, with a nested structure and the nesting coefficient was kept constant for all the nests and was denoted as *thetamode*.

The important properties of the estimated coefficients are summarised below.

- When many taste variations are at stake the travel time and travel cost becomes less important to the travellers.
- When travelling by car the males tend to consult websites.
- Males are less likely to travel by train if they use websites only as source of travel information.
- Travellers who left their education by the age of 20 are less likely to use website only as source of travel information.

8.2.8 RP and SP Combined model for Source and Mode

As discussed in Section 6.2, there were three data sets in this study. The first was RP data which explored previous behaviours and choices of the travellers when choosing information sources and modes in travelling. The second data set included SP game (SP1) which explored the choice of website as an information source prior to travel. And finally, the third data set included another SP exercise (SP2) which explored the choice of modes under the influence of different information sources. The RP model developed in section 6.4.5 offers advantages that it is based on the actual choices. The SP models on the other hand offer an advantage in that they contain detailed information on the sensitivity of the choice to changes in a range of attributes of information sources in particular subscription costs and reliability. The combined RP and SP model thus required combination of all these three datasets to encompass the relative advantages of both datasets. Ben-Akiva and Morikawa postulate that the differences in the error terms between any two data sets can be represented as a function of the variance of the error term of each of the data set. The potential differences in error between the datasets can be removed by multiplying the parameters of SP2 by a scale parameter. There was not any common coefficient in SP1 and RP and thus the common model in this case was not possible. In order to develop a single mode choice model, data from the RP questions was merged with data from the second SP experiment using an artificial tree structure as proposed by Bradley and Daly (1991).

The important properties of the estimated coefficients in this combined model are:

- Travel time information provided by websites has a higher value for the travellers in their travel decisions.

- There is a higher opportunity cost for the information given by the websites and travellers consider the values of pre-trip time information, more important than time and cost actually spent on travelling.
- The value of time is significantly reduced with the introduction of web time as a variable.
- Travellers consider information from websites four times more important than the normal travel time.

8.2.9 Mixed Logit Model on SP Model

a. Source Choice Model

The final model (ModelSPS6) was taken from section 6.4.4 for the estimation of Mixed Logit Model (MMNL) for SP mode choice model. This model has a similar specification as the multinomial logit (MNL) but the following notable differences.

Firstly, normally distributed coefficients were specified for search time to explain additional variation is the model (MMNL in Table 6.14). Secondly, in order to accommodate the repeated measurement problem, the search time was also identified to vary across individuals and not across observations. (Panel MMNL in Table 6.14) This was done by using Biogeme's panel data specifications (for details refer Biogeme manual).

The results show that the mixed logit calibration of the model increased the efficiency of the model with increased t stats. The inclusion of the error component terms improves the fit to the data, but only marginally. The results confirm the findings of the MNL model. The standard deviations of the normally distributed components are not significant in both MMNL and Panel MMNL this suggests that the MNL model specification represents the choice of the respondents adequately.

b. Mode Choice Model

The final model (ModelSPM16) was taken from section 6.4.3 for the estimation of Mixed logit model for SP mode choice. This model has a similar specification to the multinomial logit (MNL) but with the following differences.

Firstly (MMNL in Table 6.15), four normally distributed coefficients were specified for multimodal website time, multimodal website cost, mode specific website time and train website cost. The other normally distributed coefficients for coach website cost, friend time, friend cost, train experience cost and previous experience time were also tried but were found insignificant and thus were subsequently dropped. Secondly (Panel MMNL in Table 6.15), in another model, the above mentioned coefficients were identified to vary only across individuals to cope with repeated measurement problem.

The results of the model suggest that there is no significant difference in the models. The estimates of the standard deviations of the normally distributed terms were not significant (or if they are significant, as happens sometimes in the panel specification, this takes place at the expense of the precision of coefficients of other attributes, without significant gains in the overall Loglikelihood value). This all means that the MNL model reflects the characteristics of the population satisfactorily.

8.2.10 Application of Regret Minimization on SP Model

a. Source Choice Model

The final model (ModelSPS6) was also formulised in the RRM paradigm. Although when the choice sets are binary, RRM reduces to Random Utility Maximisation's (RUM's) linear additive binary logit model, the model was still calibrated to validate the applicability and performance of the RRM formulation.

The results show that the RRM model is almost equal to the MNL model which confirms the appropriateness of the RRM model. Although all the variables were first tried inside the Regret function, the resulting model lost its significance even on the important explanatory variables; hence only search time was tried in the regret function.

The estimates show that the signs of the dummy variables (outside the regret function) are reverse as compared to the MNL model; this confirms that the model shows the regret of an individual not a utility as in the other models.

b. Mode Choice Model

The final model (ModelSPM16) was taken for the calibration of RRM model for SP mode choice model. Two models were estimated, one with all the cost and time variables inside the regret function. In the other model the dummies for information from friend and previous experience were kept outside the regret function. The results suggest that all the significant parameters in both the RRM models have the expected signs which indicate the regret paradigms of the models. Some of the variables had reverse signs but they were not significant. In terms of comparison between RRM and MNL models, it appears that the MNL model fits the data slightly better than its RRM counterpart. The significance levels of some of the parameter were increased in RRM. RRM estimates of the parameters were about half of their MNL counterparts. This suggests that the respondents had lower anticipated regret of the foregone alternatives.

8.2.11 Application to the Models

The overall objective of the Chapter 7 was to apply the disaggregate models of information sources and modes so as to generate forecast uptake of variant information sources and to identify its desirable properties. More specifically this work involved:

- The application of the joint RP-SP model of mode choice and Jacknifed SP source model, developed in Chapter 6, to generate forecasts on estimation sample.
- The application of the joint RP-SP model of mode choice and Jacknifed SP source model, developed in Chapter 6, to generate forecasts on expanded sample as per NTS 2006 population of long distance travellers.
- The application of a range of sensitivity tests under different policies.

The main findings of the analysis are:

- The market share of a website increases from 5 % at 30 minutes of search time to 97 % at search time of 1 minute this indicates search time is a controlling factor for any website owner/policy makers to attract the users and influence their decisions..

- The market share of a website rises from 51 % when no information was provided to 77% and even more people are attracted if the search time of the competitive website is increased and eventually the market share reaches 94%.
- Prescriptive information also has a large effect on the market share of a website and it even goes to 100% of the market if the search time of competitive websites is 15 minutes.
- When any commercial websites has advertisements on it, its market share decreases from original 74% to 48%.
- If a commercial website has no advertisements and no subscription its share increases from 57% to even 100%.
- The effect of information provided by multimodal website is more as compared to the information provided by friends for car users.
- Previous experience of a travel with a mode influences the choice more if it is also validated by other sources.

The analysis gave following valuable policy implications:

- The study found that the travellers consider information of travel time from a website 4 times more important to the time spent on travelling. These results could be used by the website owners, public transport providers and policy makers to attract more travellers on their websites by improving them and distribute the overall modal shift.
- The provisions of information as desired by a traveller can increase the use of information source up to 94%. The type of information and subsequent details as desired by the users, increases the credibility of the website and attract more users.
- Prescriptive information also has a large effect on the market share of a website and it even goes to 100% of the market. Thus travellers can be influenced to improve the overall transport network.
- This study revealed that the advertisements and subscription requirements of a website has a big impact on its use. Hence the government websites with no subscription fee could be used to influence the travellers more conveniently.

- It was found that there is a lot of difference between the impact of a multimodal website and a mono-modal website on mode choice. It was found that the effect of information, provided by multimodal websites, is even more important than the advice from friends. It was found that even the previous experience of a traveller has the same impact on her travel decisions as of any website.
- It was found that the market share of the modes increase when information sources show decreased travel time and cost values. This shows that information sources could be used as a policy measure to distribute the travel miles among the modes. And the maximum results are achieved when different information sources give the same information to the travellers. This shows the effect of credibility on the sources as travellers tend to believe more when they observe same information about the journey from multiple sources.

8.3 Conclusions

The above mentioned results and conclusions from the selected models give detailed insights for the choice of sources in order to make travel decisions. The results would help to understand travellers' behaviour in selecting the information sources and modes. The research questions raised in section 3.1 are explored and the results are discussed below.

- *How does initial information effect subsequent search?*

  The study found out that the previous experience or initial information significantly affects the utility of the selected mode. The selected Model ModelSPM16 suggests that travellers with a previous experience with a mode give more importance to it as compared to other sources. The model suggests that the traveller anticipate more credence with the previous experience as compared to any other source.

- *Does credibility vary from source to source?*

  Credibility does vary from source to source. It was found that travellers give varying degrees of importance to different information sources. It was found that travellers give credence to government owned sources. Similarly travellers give more importance to their own previous experiences followed by multimodal websites, train websites, friends and coach websites respectively. Moreover the credibility of an information source increases significantly when two information sources give the same information.

- *How is information with low credibility treated?*

  The travellers give less value to the information sources with low credibility. It was found that commercial websites with advertisements have reduced utility as compared to other sources. Furthermore, travellers try to gather more information from another source if they find that the information provided by a particular source is less credible.

- *What affects credibility?*

  A number of factors affect credibility of a source, they include past experience of the traveller with that source/mode, ownership of the source, and presence of irrelevant information. It was also found that multimodal websites are considered more credible as compared to mono modal website.

- *Does presentation of information affect choice behaviour?*

  A website offering specific information on users' own criteria increases the utility of that website. Prescriptive information also increases the utility of choice of an information source. Similarly, the presentation of real time and updated information has increased influence on the travellers' choice.

- *What is the influence of information sources on the mode choice decisions?*

  The study found that the effect of information provided by multimodal website is more as compared to the information provided by friends or than mode specific websites. The study also found that if a person gets information from both sources, the market share increases even more. Moreover, the market share of the modes increases when information sources show decreased travel time and cost values and the maximum results are achieved when different information sources give the same information to the travellers. These results show that information from the information sources could be used to influence the mode choice of the travellers.

- *Does the design and content of an information source affect choice decisions?*

  A website with comprehensive information about the intended travel and offering specific information on users' own criteria increases the utility of that website. On the other had presence of irrelevant information and advertisement significantly reduce the choice of that website and hence decrease its market share. Another important factor is that the government websites with no advertisements and no subscription have higher attraction and credibility within the respondents.

8.4 Recommendations for Further Research

Utility theory has been the basis of most analyses of response to traffic information and modes. This research also employed Random Regret Minimisation Theory to understand the regret associated with each alternative and its effect on the choice of decision maker. This research has confirmed that Random Regret Minimisation Theory can also be used in this context and can provide very important insights in the behaviour of the traveller. Since traffic information is often treated as uncertain, the wider application of the Regret concept is recommended. This is expected to give deeper understanding of travel choices and of information acquisition. Another approach to study RRM framework would involve collecting data in a minimisation framework. For instance, in an SP exercise, there could be questions that require the minimisation of regret in the choices instead of selecting the best alternative. A future study could employ an SP exercise specially designed to study the effect of an RRM modelling framework.

Another interesting approach to collect the data could be to use the travellers' search patterns while planning their journey on the internet/travel planning websites. This approach would require data from the website owners on the usage of their websites. This approach although desirable, was not used in this study for reasons mentioned in section 3.3. It is however recommended that such data, if obtainable in future, would give very interesting insights on the travellers' search behaviour and has valuable research prospects.

Although considerable progress has been made in understanding the information acquisition process and search behaviour, there is scope for further work. Travel information acquisition and online travel behaviour research is far from perfect and further improvements are needed. Firstly, there is a need to more fully understand the link between travel information search, subsequent travel choices and the underlying complex dynamics of decision making. Secondly, as society becomes increasingly dependent on the Internet, it is likely that the websites will be designed and used for specific travel tasks. Their use could be very important in the future of transport planning and policy. Finally, dynamic Internet-based search behaviours may require a new style of modelling and may open up the possibility of new methods of data collection. There may be the opportunity to gather data on travellers' online search behaviour and, if this can be gathered over a significant period of time, emerging search patterns and information effects may be assessed. More immediately there is scope to examine the performance of models using data from travel diaries of long distance as well as short distance travellers. This will provide a good validation test for the existing models.

References

Aarts, H., Verplanken, B., Knippenberg, A. van, 1997. Habit and information use in travel mode choices. Acta Psychologica 96, 1-14.

Aarts, H., Verplanken, B., Knippenberg, A. van, 1998. Predicting behaviour from actions in the past: repeated decision making or a matter of habit? Journal of Applied Social Psychology 28, 1355-1374.

Aarts, H., Dijksterhuis, A., 2000. The automatic activation of goal directed behaviour: the case of travel habit. Journal of Environmental Psychology 20, 75-82.

Abdel-Aty, M.A. (1998). Modelling Incident-Related Routing Decisions by Using a Nested Logit Structure. In Transportation Research Record 1645, TRB, National Research Council, Washington, D.C., pp. 103-110.

Abdel-Aty, M., and Abdulla, M. F. 2004. Modelling Drivers Diversion from Normal Routes under ATIS using Generalised Estimating Equations and Bionomial Probit Link Function. Transportation , 31(3), 327-348.

Abdel-Aty, M., R. Kitamura, P. P. Jovanis 1995. Exploring Route Choice Behaviour using Geographic Information System- Based Alternate Routes and Hypothetical Travel Time Information Input. In Transportation Research Record 1493, TRB, National Research Council, Washington, D.C., pp. 74-80.

Adler, J.L., Recker, W.W., and McNally, M.G., (1993). A conflict model and interactive simulator (FASTCARS) for predicting en-route driver behaviour in response to real-time traffic condition information. Transportation 20, 83-106.

Adler, J.L., and McNally, M.G., 1994. In laboratory experiments to investigate driver behaviour under advanced traveller information systems. Transportation Research Vol.2,No.3, 149-164.

Al-Deek, H., Khattak, A.J., Thaanjevan, P., 1998. A combined traveller behaviour and system performance model with advanced traveller information systems. Transportation Research Part A 32 (7), 479-493.

Allen, R., Ziedman, D., Rosenthal, T.J., Stein, A.C., Torres, J.F., and Halati, A.B., 1991. Laboratory assessment of driver route diversion in response to in vehicle navigation and motorists information systems. Transportation Research Record 1306, 82-91

Anable, J., Gatersleben, B., 2005. All work and no play? The role of instrumental and affective factors in work and leisure journeys by different travel modes. Transportation Research 39A, 163-181.

Arasan, V.T., Rengaraju, V.R., and Rao, K.V.K 1996. Trip Characteristics of Travellers Without Vehicles", In Journal of Transportation Engineering, Vol. 122, No. 1, p.p. 76-81

Arentze, T.A., Timmermans, H.J.P., 2005a. Representing mental maps and cognitive learning in microsimulation models of activity-travel choice dynamics. Transportation 32, 321-341.

Arentze, T.A., Timmermans, H.J.P., 2005b. Information gain, novelty seeking and travel: a model of dynamic activity-travel behaviour under conditions on uncertainty. Transportation Research 39A, 125-145.

Arnott, R., de Palma, A., Lindsey, R., 1991. Does providing information to drivers reduce traffic congestion? Transportation Research Part A 25, 309-318.

Ben-Akiva, M., and Lerman, S., 1985. Discrete Choice Analysis. The MIT Press, Cambridge, Mass.

Ben-Akiva, M., A. D. Palma, and I. Kaysi 1991. Dynamic Network Models and Driver Information Systems. In Transportation Research Part A, Vol. 25A, No. 5, pp. 251-266.

Ben-Akiva, M., Boccara, B., 1995. Discrete choice models with latent choice sets. International Journal of Marketing 12, 9-24.

Ben-Akiva, M., H. N. Koutsopoulos, R. Mishalani and Q. Yang 1997. Simulation laboratory for evaluating dynamic traffic management systems. In ASCE Journal of Transportation Engineering, Vol. 123, No. 4, pp. 283-289.

Bogers, E.A.I, Viti, F., Hoogendoorn, S.P., 2005. Joint modelling of ATIS, habit and learning impacts on route choice by laboratory simulator experiments. Transportation Research Records 1926. 189-197.

Bonsall, P.W., Pickup, L. and Stathopoulos, A., 1991. Measuring behavioural responses to road transport informatics. Proc DRIVE Conference on advanced telematics in road transport, Brussels VOL 2, 1457-1487.

Bonsall, P.W., Parry, T., 1991. Using an interactive route-choice simulator to investigate drivers' compliance with route guidance advice. In Transportation Research Record 1306, TRB, National Research Council, Washington, D.C., pp. 59- 68.

Bonsall, P.W., and Joint, M., 1991. Driver compliance with route guidance advice: the evidence and its implications, Vehicle Navigation and Information Systems Conference Proceedings (VNIS), Deororm, Michighan, USA October, 47-59

Bonsall, P.W., 1992a. The influence of route guidance advice on route choice in urban networks. Transportation 19(1), 1-23.

Bonsall, P.W., 1992b. Drivers' acceptance of route choice advice; an international comparison. Proceedings of International Conference on Vehicle Navigation and Information systems Oslo, 617-625.

Bonsall, P.W., 1993. Assessing the impacts and benefits of in car route guidance advice via field trials, IEEE-IEE Vehicle navigation and information systems conference Ottawa-VNIS, 93.

Bonsall, P.W., and Hounsell, N., 1994. Modelling drivers' route choice behaviour in the context on route guidance and information systems, towards an Intelligent Transportation Systems. Presented at proceedings of the first world congress on applications of transport telematics and intelligent vehicle-highway systems, Palais de congress de pares, Vol 2, 508-515.

Bonsall, P.W., 1995. Analysing and modelling the influence of roadside variable message displays on drivers' route choice. 7th World Conference on Transportation Research Sydney, 16-21.

Bonsall, P.W., 2007. Chapter in Handbook of Transport Modelling. Hensher and Buttons 2007 edition.

Bos, D.M., van der Hiejden, R.E.C.M., Molin, E.J.E., Timmermans, H.J.P., 2004. The choice of park and ride facilities: an analysis using a context dependent hierarchical choice experiment. Environment and Planning 36A, 1673-1686

Caplice, C., Mahmassani, H.S., 1992. Aspects of commuting behaviour: preferred arrival time, use of information, and switching propensity. Transportation Research 26A, 409-418.

Chatterjee, K., McDonald, M., 2004. Effectiveness of using variable message signs to disseminate dynamic traffic information: evidence from field trials in European cities. Transport Reviews, 24, 559-585.

Chen, P. S., and H. S. Mahmassani 1993. A Dynamic Interactive Simulator for Studying Commuter Behaviour Under Real-time Traffic Information Supply Strategies, In Transportation Research Record 1413, TRB, National Research Council, Washington, D.C., pp. 12-21.

Chorus, C.G., Molin, E.J.E., van Wee, G.P., 2006a. Use and effects of Advanced Traveller Information Services (ATIS): a review of the literature. Transport Reviews, 26(2), pp 127-149

Chorus, C.G., Molin, E.J.E., Arentze, T.A., Timmermans, H.J.P., Van Wee, G.P., 2006b. Travellers' need for information: an empirical study into the role of knowledge. Paper presented at the 85th annual meeting of the Transportation Research Board, Washington, D.C.

Chorus, C.G., Walker, J.L., Ben-Akiva, M.E., 2006c. Travel information use and effects. Paper to be presented at the 11th International Conference on Travel Behaviour Research, Kyoto, August 2006

Cowell, M.P.H., James, H., and Silcock, D.T., 1988. Real time passenger information at Heworth Metro/Bus interchange. Research Report 75, Transport Operation Research Group, University of Newcastle upon Tyne.

Chu, L., H. X. Liu, W. Recker, and H. M. Zhang 2001. Development of a Simulation Laboratory for Evaluating Ramp Metering Algorithm, Institute of Transportation Studies, University of California, Irvine.

Chu, P.C., Spires, E.E., 2003. Perception of accuracy and effort of decision strategies. Organisational Behaviour and Human Decision Processes 91, 203-214.

Coase, R.H., 1937. The Nature of the Firm. Economica 4, 386-405.

Coase, R.H., 1960. The Problem of social Cost. Journal of Law & Economics 3.

Da Penha Sanches, S., and F. Serra de Arruda 2002. Incorporating Non-motorized Modes in a Mode Choice Model. In Transportation Research Record 1818, TRB, National Research Council, Washington, D.C., pp. 89-93.

Dia, Hussein 2000. A Conceptual Framework for Modelling the Environmental Impacts of Intelligent Transportation Systems, Proceedings of the XI Pan America Conference on Traffic and Transportation Engineering.

Dia, Hussein 1999. Modelling the impacts of Advanced Traveller Information Systems using Intelligent Agents, Department of Civil Engineering, the University of Queensland.

Dial RB 1971. A probabilistic multipath traffic assignment model which obviates path enumeration. Transportation Research 5(2).

Eckstein, Z., van den Berg, G.J., 2007. Empirical labor search: A survey. Journal of Econometrics136, 531-564.

Edwards, W., 1954. The theory of decision making. Psychological Bulletin 51, 380-517.

Ellaway, A., Macintyre, S., Hiscock, R., Kearns, A., 2003. In the driving seat: psychological benefits from private motor vehicle transport compared to public transport. Transportation Research 6F, 217-231.

Emmerink, R., Axhausen, K.W., Nijkamp, P., Rietveld, P., 1995a. Effects of information on road transport networks with recurrent congestion. Transportation 22, 21-53.

Emmerink, R., Axhausen, K.W., Nijkamp, P., Rietveld, P., 1995b. The potential of information provision in a simulated road transport network with non-recurrent congestion. Transportation Research Part C 3, 293-309.

Emmerink, R., Nijkamp, P., Rietveld, P., Van Ommeren, J.N., 1996. Variable message signs and radio traffic information: an integrated empirical analysis of drivers' route choice behaviour. Transportation Research 30 A, 135-153.

Fowkes, A.S., and Wardman, M., 1991. Disaggregate Methods, Chapter 4, in Fowkes, A.S., and Nash, C.A.,, Analysing Demand for Rail Travel, Avebury.

Foerster, J.F., 1978. Mode choice decision process models: a comparison of compensatory and non compensatory structures. Transportation Research 13A, 17-28.

Forsyth, F., and Silcock, D.T., An evaluation of the provision of real time information for passengers in the Northern Line of the London Underground. Research Report 59 Transport Operations Research Group, University of Newcastle upon Tyne.

FTA 2000. A Self Instructing Course in Disaggregate Mode Choice Modelling, Federal Transportation Administration, 2000.

Fujii, S., Garling, T., Kitamura, R., 2001. Changes in drivers' perceptions and use of public transport during a freeway closure. Environment and Behaviour 33, 796-808.

Fujii, S., Garling, T., 2003. Development of script based travel mode choice after forced change. Transportation Research 6F, 117-124.

Fujii, S., Kitamura, R., 2003. What does a one month free bus ticket do to habitual drivers? Transportation 30, 81-95.

Garling, T., Laitila, T., Westin, K., 1998. Theoretical foundations of travel choice modelling: an introduction. In Garling, T., Laitila, T., Westin, K. (eds.). Theoretical foundations of travel choice modelling. Elsevier, Amsterdam.

Garling, T., Fujii, S., Boe, O., 2001. Empirical tests of a model of determinants of script based driving choice. Transportation Research 4F, 89-102.

Garling, T., Eek, D., Loukopoulos, P., Fujii, S., Johansson-Stenman, O., Kitamura, R., Pendyala, R., Vilhelmson. B., 2002. A conceptual analysis of the impact of travel demand management on private car use. Transport Policy 9, 59-70.

Goalies, K.G., Kim, T., Pribyl, O., 2004. A longitudinal analysis of awareness and use for advanced traveller information systems. Paper presented at 83rd meeting of Transportation Research Board, Washington D.C.

Hicks, J.R., 1932. The theory of wages (second Edition, 1964), Macmillan London.

Hague Consulting Group, 1991 Factors influencing mode choice. Final Report.

Hargreaves Heap, S., Hollis, M., Lyons, B., Sugden, R., Weale, A., 1992. The theory of choice: a critical guide. Blackwell, Oxford.

Jayakrishnan R, Mahmassani HS and Hu T-Y 1994. An evaluation tool for advanced traffic information and management systems in urban networks. Transportation Research 2C(3), 129-147.

Jeffery, D., 1990. Driver route guidance systems: state of art. Soekkha, H.M. (ed). In Telematics-Transportation and Spatial Development. VSP, The Netherlands.

Johnson, J.G., Raab, M., 2003. Take the first: option-generating and resulting choices. Organisational Behaviour and Human Decision Processes 91, 215-229.

Jou, R. C., T. Y. Hu, and C. W. Lin 1997. Empirical Results from Taiwan and their Implications for Advanced Traveller Pre-trip Information Systems. In Transportation Research Record 1607, TRB, National Research Council, Washington, D.C., pp. 126-133.

Hato, E., Taniguchi, M., and Sugie, Y., 1995. Influence of traffic information on drivers; route choice. Paper presented at WCTR.

Hato, E., Taniguchi, M., Sugie, Y., Kuwahara, M., Morita, H., 1999. Incorporating an information acquisition process into a route choice model with multiple information sources. Transportation Research 7C, 109-129.

Hey, J.D., 1982. Search for rules of search. Journal of economic behaviour and organisation 3, 65-81.

Hodgson, G.M., 2004. Reclaiming habit for institutional economics. Journal of Economic Psychology 25, 651-660.

Hounsell NB, McDonald M and Njoze SR 1995. Strategies for route guidance systems taking account of driver response. Proc Vehicle Navigation and Information Systems and Pacific Rim TransTech Conference, Seattle, 328-33, ISBN 0-7803-2587-7.

Huchingson, R.D., and Dudek, C., 1979. Delay, time saved, and travel time information for freeway traffic management. Transportation Research Record 722, 36-39.

Huneke, M.E., Cole, C., Levin, I.P., 2004. How varying levels of knowledge and motivation affects search and confidence during consideration and choice. Marketing Letters 15, 67-79.

Kahneman, D., Tversky, A., 1979. Prospect Theory: an analysis of decision under risk. Econometrica 47, 263-291.

Kahneman, D., Tversky, A., 1992. Advances in prospect theory: cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 297-323.

Kanninen, B., J., 1996. Intelligent transportation systems: an economic and environmental policy assessment. Transportation Research 30A, pp 1-10

Kenyon, S., Lyons, G., 2003. The value of integrated multimodal traveller information and its potential contribution to modal change. Transportation Research 6F, 1-21.

Khattak, A., J. Schofer, and F. Koppelman 1991. Factors Influencing Commuters' En Route Diversion Behaviour in Response to Delay. In Transportation Research Record 1318, TRB, National Research Council, Washington, D.C., pp. 125-136.

Khattak, A.J., Koppelman, F.S., Schofer, J.L., 1993a. Stated preference for investigating commuters' diversion propensities. Transportation 20, 107-127.

Khattak, A.J., Schofer, J.L., Koppelman, F.S., 1993b. Commuters' en-route diversion and return decisions: Analysis and implications for advances traveller information systems. Transportation Research 27A, No.2.

Khattak, A., Kanafani, A., and Le Colletter, E., 1994. Stated and Reported Route Diversion Behaviour: Implications of Benefits of Advanced Traveller Information System, In Transportation Research Record 1464, TRB, National Research Council, Washington, D.C., pp. 55-62.

Khattak, A., A. Polydoropoulou, and M. Ben Akiva 1996. Modelling Revealed and Stated Pre-trip Travel Response to Advanced Traveller Information Systems. In Transportation Research Record 1537, TRB, National Research Council, Washington, D.C., pp. 46-54.

Khattak, A. J., and A. Khattak 1998. Comparative Analysis of Spatial Knowledge and En-route Diversion behaviour in Chicago and San Francisco: Implication for Advanced Traveller Information Systems. In Transportation Research Record 1621, TRB, National Research Council, Washington, D.C., pp. 27-35.

Khattak, A., and Palma, A.D., 1998. The Impact of Adverse weather conditions on the Propensity to Change Travel Decision: A Survey of Brussels Commuters. In Transportation Research Part A, Vol. 31 No.3, pp. 181-203.

Khattak, A.J., Yim, Y., Prokopy, L.S. 2003. Willingness to pay for travel information. Transportation Research, 11C. 137-159.

Koo, R., and Yim, Y., 1998. Commuter Response to Traffic Information on an Incident. In Transportation Research Record 1621, TRB, National Research Council, Washington, D.C., pp. 36-42.

Koppelman, F.S., Pas, E.I., 1980. Travel choice behaviour: models of perceptions, feelings, preference, and choice. Transportation Research Record 765, pp 26-33.

Koutsopoulos, H., Lotan, T., 1989. Motorist information Systems and recurrent traffic congestion: sensitivity analysis of expected results. Transportation Research Record 1281, 148-157.

Lappin, J., 2000. Advanced traveller information service (ATIS): who are ATIS customers? Working paper prepared for Volpe National Transportation Systems Centre. Cambridge, MA.

Lee, D.B. Jr. 2000. Benefit-cost Evaluation of Traveller Information (Seattle's Washington State department of Transportation Website), Transportation Research Record 1739, TRB, National Research Council, Washington, D.C., pp.25-34.

Lerman, S.R., Mahmassani, H.S., 1985. The econometrics of search. Environmental and Planning 17A, 1009-1024.

Liu, H. X., Jun-Seok Oh, and W. Recker 2002. Adaptive signal Control System with On-Line Performance Measure, Paper for the 81st Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C.

Liu, Y. H., and Mahmassani, H., 1998. Dynamic Aspects of Commuter Decisions under Advanced Traveller Information Systems: Modelling Framework and Experimental Results. In Transportation Research Record 1645, TRB, National Research Council, Washington, D.C., pp. 111-119.

Mahmassani, H.C., and Chang, G.L., 1985. Dynamics aspects of departure time choice behaviour in a commuting system: theoretical framework and experimental analysis. Transportation Research Record 1037.

Mahmassani, H.C., and Chang, G.L., 1987. On boundedly rational user equilibrium in transportation systems, Transportation Science 21(2), 89-99

Mahmassani, H.C., and Jayakrishnan, K., 1989. System performance and user response under real time information in a congested traffic corridor. Presented at seminar on Urban Traffic Networks, Capri, June.

Mahmassani, H.S., Caplice, C., and Walton, C., 1990. Characteristics of urban commuter behaviour switching propensity and use of information. Transportation Research Record 1285, 57-69.

Mahmassani, H.S., and Chen, P.S., 1991. Comparative assessment of origin-based and en-route real-time information under alternative user behaviour rules. Transportation Research Record 1306, 69-81.

Mahmassani, H., Jayakrishnan, R., 1991. System performance and user response under real time information in a congested traffic corridor. Transportation research Part A 25(5), 293-307.

Mahmassani, H., Liu, Y.S., 1999 Dynamics of commuting decision behaviour under advanced traveller information systems. Transportation Research Part C 7, 91-107.

Mannering, F., 1989. Poisson analysis of commuter flexibility in changing routes and departure times. Transportation Research 23B, 53-60.

Mannering, F., Kim, S.G., Barfield, W., and Ng, L., 1994. Statistical analysis of commuters' route, mode, and departure time flexibility. Transportation Research 2C, 35-47.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behaviour. In Zarembka, P., eds. Frontiers in Econometrics, Academic Press, New York, 105-142.

McFadden, D., 1999. Rationality for economists? Journal of Risk and Uncertainty 19, 73-105.

Mckay, B. 2000. Transportation Modellers Meet a Simulated Real World, AITPM international Conference: The Travel Bug, 2000.

Mehndiratta, S.R., Kemp, M.A., Lappin, J.E., and Brand, D., 1999. What Advanced Traveller System Information Do Users Want? (Evidence from In-vehicle Navigation Device Users), Transportation Research Record 1679, TRB, National Research Council, Washington, D.C., pp. 41-49.

Meyer, M. D., and E. J. Miller 2001. Urban Transportation Planning: A Decision Oriented Approach, Mc Graw-Hill, Network.

Molin, E.J.E., Chorus, C.G., 2004. Willingness to pay for personalized public transport information services. Paper presented at the 83rd meeting of the Transportation Research Board, Washington D.C.

Molin, E.J.E., Chorus, C.G., van Sloten, 2005. The effect of un reliable travel conditions on the need for advanced public transport information services. Working Paper.

Olander, F., 1973. Search behaviour in non-simultaneous choice-situations: satisficing or maximising? In selected proceedings of an interdisciplinary research conference, Rome.

Ortuzar, J. D., and Willumsen, L.G. 1994. modelling Transport, 2nd Ed., John Wiley & Sons Ltd.

Payne, J.W., Bettman, J.R., Johnson, E.J., 1993. The adaptive decision maker. Cambridge University Press, Cambridge.

Payne, J.W., Bettman, J.R., Luce, M.F., 1996. When time is money: decision behaviour under opportunity-cost time-pressure. Organisational Behaviour and Human Decision Processes 66, 131-152.

Petrella, M., Lappin, J., 2004. Los Angeles and Seattle: A comparative analysis of customer response to online traffic information. Paper presented at the 83rd meeting of the transportation Research Board, Washington D.C.

Peeta, S., Ramos, J.L., and Pasupathy, R., 2000. Content of Variable Message Signs and On-Line Driver Behaviour, In Transportation Research Record 1725, TRB, National Research Council, Washington, D.C., pp. 102-108.

Polak, J., Jones, P., 1993. The acquisition of pre-trip information: a stated preference approach. Transportation 20, 179-198.

Polydoropoulou, A., M. Ben Akiva, and I. Kaysi 1994. Influence of Traffic Information on Drivers' Route choice Behaviour. In Transportation Research Record 1453, TRB, National Research Council, Washington, D.C., pp. 56-65.

Polydoropoulou, A., M. Ben Akiva, A. Khattak, and G. Lauprete 1996. Modelling Revealed and Stated En-Route Travel Responses to Advanced Traveller Information Systems. In Transportation Research Record 1537, TRB, National Research Council, Washington, D.C., pp. 38-45.

Polydoropoulou, A., Gopinath, D.A., Ben-Akiva, M., 1997. Willingness to pay for traveller information systems: smartraveller case study. Transportation Research Record 1588, 1-9.

Polydoropoulou, A., Ben-Akiva, M., 1998. The effect of advanced traveller information systems on travellers' behaviour. In Emmerink, A., & Nijkamp, P., 1998. Behavioural and network impacts of driver information systems. Aldershot, Ashgate.

Posavac, S.S., Herzenstein, M., Sanbonmatsu, D.M., 2003. The role of decision importance and the salience of alternatives in determining the consistency between consumers' attitudes and decisions. Marketing Letters 14, 47-57.

Prevedouros, P.D., and Kasamoto, K.D. 1998. Incident Management Simulation on a Two Freeway Corridor in Honolulu, Department of Civil Engineering, University of Hawaii at Manoa, Hawaii.

Rakha, H., Van Aerde, M., Case, E., Ugge, A., 1989. Evaluating the benefits and interactions of route guidance and traffic control strategies using simulations. Presented in proceedings of first VNIS conference, Toronto, IEEE.

Richardson, A., 1982. Search models and choice set generation. Transportation Research 16A, 403-419.

Richardson, A. J. 2002. Simulation Study of Estimation of Individual Specific Values of Time Using Adaptive Stated-Preference Survey. Transportation Research Record 1804, TRB, National Research Council, Washington, D.C., pp. 117-125.

Samuelson, P.A., 1947. Foundations of economic analysis. Harvard University Press, Cambridge.

Sawaya, O. B., D. L. Doan, and A. K. Ziliaskopoulos 2000. Predictive Time-Based Feedback Control Approach for Managing Freeway Incidents, In Transportation Research Record 1710, TRB, National Research Council, Washington, D.C., pp. 79-84.

Schlich, R., Axhausen, K.W., 2003. Habitual travel behaviour: evidence from a six week travel diary. Transportation 30, 13-36.

Schofer, J.L., Khattak, A., and Koppelman, F.S., 1993. Behavioural issues in the design and evaluation of advanced traveller information systems. Transportation Research 1C, 107-117.

Simon, H.A., 1955. A behavioural model of rational choice. Quarterly Journal of Economics 69, 99-118.

Simon, H.A., 1978a. Rationality as process and as product of thought. Richard T. Ely Lecture before the American Economic Association.

Simon, H.A., 1978b. Rational decision making in business organisation. Nobel Memorial Lecture, Stockholm.

Shah, V.P., Wunderlich, K., and Larkin, J., 2001. Time Management Impacts of Pre-trip Advanced Traveller Information Systems (Finding from a Washington, D.C., Case study), Transportation Research Record 1774, TRB, National Research Council, Washington, D.C., pp. 36-43.

Smith, V., 1991. Rational Choice: the contrast between economics and psychology. Journal of Political Economy 99, 877-897.

Spasovic L. N., M. P. Boile, and A. K. Bladikas 1995. Framework for Assessing Benefits of Highway Traveller Information Services. In Transportation Research Record 1485, TRB, National Research Council, Washington, D.C., pp. 27-34.

Srinivasan, K. Chen, I., Reddy, P., Jovanis, P.P., 1999. Pre-trip information systems: an investigation into users' information acquisition process. Paper presented at the 78th meeting of the Transportation Research Board, Washington D.C.

Srinivasan, K., and Mahmassani, H., 2000. Modelling Inertia and Compliance Mechanisms in Route Choice Behaviour under Real-Time Information. In Transportation Research Record 1725, TRB, National Research Council, Washington, D.C., pp. 45-53.

Steg, L., Vlek, C., Slotegraaf, G., 2001. Instrumental reasons and symbolic-affective motives using a motor car. Transportation Research 4F, 151-169.

Steg, L., 2005. Car use: lust and must Instrumental, symbolic and affective motives for car use. Transportation Research 39A, 147-162.

Stough, R.R., 2001. Intelligent Transport Systems: Cases and Policies. Edward Elgar, Cheltenham.

Suen, L., and Geehan, T., 1987. Information for public transport users. In Bonsall, P.W., and Bell, M., (eds.). Information Technology Applications in Transport. VNU Science Press, 287-318.

Svenson, O., 1998. The perspective from behavioural decision theory on modelling travel choice. In Garling, T., Laitila, T., Westin, K. (eds.). Theoretical foundations of travel choice modelling. Elsevier, Amsterdam.

Swait, J., Ben-Akiva, M., 1987. Incorporating random constraints in discrete models of choice set generation. Transportation Research 21B, 91-102.

Swait, J., Adamowicz, W., 2001. Choice environment, market complexity, and consumer behaviour: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice. Organisational Behaviour and Human Decision Processes 86, 141-167.

Targa, F., Khattak, A.J., Yim, Y., 2003. Understanding access and use of dynamic travel information. Paper presents at the 82nd meeting of the Transportation Research Board, Washington D.C.

Technical Innovation 1992. Passenger information systems. Technical innovation supplement. Technical Innovation, Winter 91-92

Thogerse, J., 2001. Structural and psychological determinants of the use of PT. Paper presented at the TRIP-colloquium at Horsholm.

Tversky, H.C., 1972. Elimination by Aspects: a theory of choice. Psychological Review 79, 281-299.

Triandis, H.C., 1977. Interpersonal behaviour. Brooks/Cole Publishing Company, Monterey.

Ullah, K., O'Neill,, W.A., and Bishop, A.B., 1994. Cost-Benefit Analysis of a Rural Advanced Traveller Information System, Transportation Research Record 1450, TRB, National Research Council, Washington, D.C., pp. 25-32.

Van Aerde, M.W., Voss, J., and Blum, Y., 1989. Modelling the Burlington skyway during recurring and non recurring congestion. Traffic Engineering and Control Vol.30, No.5, 228-240.

Vance, C., Balcombe, R., 1997. How to tell bus passengers what they need to know. Working paper of Transportation Research Laboratory.

Van Vuren T and Watling D 1991. A multiple user class assignment model for route guidance. Transportation Research Record 1306, 22-31.

Vaughn, K., Abdel-Aty, M., Kitamura, R., Jovanis, P.P., and Yang, H., (1993). Experimental Analysis and Modelling of Sequential Route Choice under an Advanced Traveller Information System in a Simplistic Traffic Network. In Transportation Research Record 1408, TRB, National Research Council, Washington, D.C., pp. 75-82.

Verplanken, B., Aarts, H., Knippenberg, A. van, 1997. Habit, information acquisition and the process of making modal choices. European Journal of Social Psychology 27, 539-560.

Von Nueman, J., Morgenstern, O., 1947. Theory of games and economic behaviour. Princeton University Press.

Wardman, M., Bonsall, P.W., and Shires, J.D., 1997. Driver response to variable message signs: A stated preference investigation. Transportation Research 5C, 389-405.

Weibull, J.W., 1978. A search model for microeconomic analysis with spatial application in Karlqvist, A., (eds.). Spatial interaction theory and planning models, North Holland Publishing Company, Amsterdam.

Williams, H.C.W.L., Ortuzar, J.D., 1982. Behavioural theories of dispersion and the miss-specification of travel demand models. Transportation Research 16B, 167-219.

Wolinetz, L.D., Khattak, A., and Yim, Y., 2001. Why Will Some Individuals Pay for Travel Information When It Can Be free. In Transportation Research Record 1759, TRB, National Research Council, Washington, D.C., pp. 9-18.

Yang, C.Y., and Fricker, J.D., 2001. Using Human Information Processing Principles to Design Advanced Traveller Information Systems. Transportation 1759, TRB, National Research Council, Washington, D.C., pp.1-8.

Yang, C.Y., Fricker, J. D. and Kuczek, T., 1998. Design Advanced Traveller Information Systems from a Driver's Perspective: Results of Driving Simulation Study, Transportation Research Record 1464, TRB, National Research Council, Washington, D.C., pp. 20-26.

Yang, Q., Koutsopouloulos, H. N., and Ben-Akiva, M., 2000. Simulation Laboratory for Evaluating Dynamic Traffic Management Systems. In Transportation Research Record 1710, TRB, National Research Council, Washington, D.C., pp. 122-130.

Yim, Y., Khattak, A.J., 2001. Traveller response to new dynamic information sources: analyzing corridor and area-wide behavioural surveys. Paper presented at the 81st meeting of the Transportation Research Record, Washington D.C.

Appendix A
CATI Questionnaire
(Final Version)

Part 1 (questions 1-34)	**relates to a recent long distance journey**
Part 2 (questions 35a and 35b)	**relates to a hypothetical choices of information source and mode**
Part 3 (questions 36-54)	**seeks background information**
Part 4 (questions 55-62)	**seeks personal information**

Note that the options shown for the SP questions (35a and 35b) are single examples drawn from the full experimental design. Each respondent was asked 8 variants on questions 35a and 9 variants on questions 35b (see section 4.3 and for details)

QUESTIONNAIRE

Code........................

This survey is being undertaken as part of a PhD project of Institute for Transport Studies, University of Leeds to study travellers' choice of information sources while selecting their modes and departure times. Information will be used for academic purposes only. Thank You.

Please read all the questions carefully, answering each as accurately as possible.

Part 1: General

1. When did u last make a journey to somewhere at least 50 miles away?
   Date:.....................................

2. Please provide the following details for your above mentioned journey
   Origin:.............................   Destination:.................................
   Distance:...........................

3. What were the main purposes of your journey to [*destination from Q.No.2*]? *Please Tick (√) all that apply and Double Tick (√√) the most important.*
   ( ) Travelling to/from work
   ( ) Travelling in the course of business
   ( ) Personal business
   ( ) Shopping
   ( ) Visiting friends/relatives
   ( ) Sports/entertainment
   ( ) Other leisure
   ( ) Education
   ( ) Other please specify ....

4. What mode did you use for your journey to [*destination from Q.No.2*]?
   ( ) Car   ( ) Coach   ( ) Train   ( ) Airplane   ( ) Other............

5. What other modes you think you <u>could</u> have used for your journey to [*destination from Q.No.2*]? *Please Tick (√) all that apply and Double Tick (√√) the most likely.*
   ( ) Car   ( ) Coach   ( ) Train   ( ) Airplane   ( ) Other............
   ( ) None other than [*mode selected in Q.No.4*]

6. How long did it take to get to
   a. the coach station? .... (minutes). [*if included in Q. No.4*]
   b. the train station? .... (minutes). [*if included in Q. No.4*]
   c. the airport? .... (minutes). [*if included in Q. No.4*]

7. How long do you think it would have taken to get to
   a. the coach station? .... (minutes). [*if included in Q. No.5*]
   b. the train station? .... (minutes). [*if included in Q. No.5*]
   c. the airport? .... (minutes). [*if included in Q. No.4*]

8. How long before the [*mode selected in Q.No.4*] [left/departed] did you reach the [*mode selected in Q.No.4*] station
   .... (minutes).

9. How long before the [*mode(s) selected in Q.No.5*] scheduled departure time would you have tried to reach the [*mode(s) selected in Q.No.5*] station
   .... (minutes).

10. Once onboard, how long did it take to get to the
    a. [*destination from Q.No.2*] by car .... (minutes). [*if included in Q. No.4*]
    b. Coach station at [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.4*]
    c. Train station at [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.4*]
    d. Airport at [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.4*]

11. Once onboard, how long do you think it would have taken to get to the
    a. [*destination from Q.No.2*] by car .... (minutes). [*if included in Q. No.5*]
    b. Coach station at [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.5*]
    c. Train station at [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.5*]
    d. Airport at [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.5*]

12. How long did it take to get to your final destination after reaching
    a. Coach station of [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.4*]
    b. Train station of [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.4*]
    c. Airport of [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.4*]

13. How long do you think it would have taken to get to your final destination after reaching
    a. Coach station of [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.5*]
    b. Train station of [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.5*]
    c. Airport of [*destination from Q.No.2*]? .... (minutes). [*if included in Q. No.5*]

14. What was your out of pocket cost(s) (petrol, parking, any tolls, return fare, taxi costs) for the return journey if you travel
    a. by car? £ .... [*if included in Q. No.4*]
    b. by coach? £ .... [*if included in Q. No.4*]
    c. by train? £ .... [*if included in Q. No.4*]
    d. by air? £ .... [*if included in Q. No.4*]

15. What do you think you out of pocket cost(s) (petrol, parking, any tolls, return fare, taxi costs) would have been for the return journey if you travel
    a. by car? £ .... [*if included in Q. No.5*]
    b. by coach? £ .... [*if included in Q. No.5*]
    c. by train? £ .... [*if included in Q. No.5*]
    d. by air? £ .... [*if included in Q. No.5*]

16. Was bad weather expected on the day of your journey?
( ) Yes ( ) No

17. Did you have advance knowledge about any incident/congestion along the route before you started your journey?
( ) Yes ( ) No

18. Did you make the journey in peak hours?
( ) Yes ( ) No

19. What information source(s) did you consult? *Please Tick (√) all that apply and Double Tick (√√) the most important.*
( ) Friend(s) [*go to Q.No.23*]
( ) Map(s) [*go to Q.No.23*]
( ) A paper timetable [*go to Q.No.23*]
( ) A telephone enquiry line [*go to Q.No.23*]
( ) website providing information about public transport <u>and</u> car journeys
( ) website providing information about car journeys <u>and</u> road conditions
( ) website providing information about Coach journeys/services <u>only</u>
( ) website providing information about Train journeys/services <u>only</u>
( ) SMS alerts [*go to Q.No.23*]
( ) Radio bulletin or TV Tele-text [*go to Q.No.23*]
( ) Other please specify .... [*go to Q.No.23*]
( ) I didn't use any information source [*go to Q.No.27*]

20. Which website(s)/journey planner(s) did you consult? (e.g. Transport Direct, Multi Map, AA, Traffic Master, Traffic-I, traveline, nationalrailenquiries etc)?
....

21. How many days before the actual departure date did consult the website? .... days.

22. In what reasons have you used websites for your journey to [*destination from Q.No.2*]? *Please Tick (√) all that apply and Double Tick (√√) the most important.*
( ) plan a public transport journey
( ) plan a car journey
( ) selecting a route for car journey
( ) compare a car journey with public transport
( ) confirm/check travel distances
( ) confirm/check travel costs
( ) check parking facilities
( ) book public transport ticket(s)
( ) reserve seats on train or coach
( ) enquire public transport timetables
( ) to seek public transport departure/arrival times
( ) book a car park
( ) look for live traffic news
( ) Any other ....

23. Did you have a particular mode in mind for your journey to [*destination from Q.No.2*] before looking at the website?
( ) Yes ( ) No [*go to Q.No.26*]

24. What mode was it?
( ) Car ( ) Coach ( ) Train ( ) Airplane ( ) Other............

{If [*mode selected in Q.No.24*] = [*mode selected in Q.No.4*] goto *Q.No.27* Else If [*Q.No.23 is Yes*] go to *Q.No.25* Otherwise go to *Q.No.26*}

25. What was the reason for changing the mode?
( ) Travel Cost [*go to Q.No.28*]
( ) Travel time [*go to Q.No.28*]
( ) Bad weather [*go to Q.No.28*]
( ) Congestion on the route [*go to Q.No.28*]
( ) Safety [*go to Q.No.28*]
( ) Comfort and convenience [*go to Q.No.28*]
( ) Travelling with friends [*go to Q.No.28*]
( ) Other please specify .... [*go to Q.No.28*]

26. What was the reason for selecting the mode?
( ) Travel Cost [*go to Q.No.28*]
( ) Travel time [*go to Q.No.28*]
( ) Bad weather [*go to Q.No.28*]
( ) Congestion on the route [*go to Q.No.28*]
( ) Safety [*go to Q.No.28*]
( ) Comfort and convenience [*go to Q.No.28*]
( ) Travelling with friends [*go to Q.No.28*]
( ) Other please specify .... [*go to Q.No.28*]

27. Please provide the reasons for not using the information source(s)? *Please Tick (√) all that apply and Double Tick (√√) the most important.*
( ) I didn't know about any of the sources
( ) I know my journey thoroughly so had no need for extra information
( ) I used them before but didn't like it (confusing)
( ) I used them before but didn't like it (time consuming)
( ) Information provided was ambiguous and inconsistent

( ) My journey was organised for me (e.g. by tour operator/office)
( ) Any other, please mention
……………………………………………….

28. How many times have you travelled to [*destination from Q.No.2*] in the last year?
( ) Only once ( ) Less than 5 times ( ) 6 – 12 times ( ) 13 – 40 times ( ) More than 40 times

Part 2:

We would like you to imagine that you need to travel from Leeds to Cardiff in 3 days time on personal business with an appointment from 1130 till 1230 in Cardiff city hall.

29. Have you ever travelled to Cardiff from Leeds?
( ) Yes ( ) No

30. What modes would you consider? *Please Tick (√) all that apply.*
( ) Car ( ) Coach ( ) Train ( ) Airplane

31. Which mode do you think you would use?
( ) <u>Certainly</u> car
( ) <u>Certainly</u> coach
( ) <u>Certainly</u> train
( ) <u>Certainly</u> airplane
( ) Not sure (I would need more information)

32. If by car, which route will you use?
( ) <u>Certainly</u> Motorway
( ) <u>Certainly</u> non Motorway
( ) Not sure (I would need more information)

33. If by car, where will you try to park? (different for different cities)

( ) <u>Certainly</u> Castle Mews
( ) <u>Certainly</u> North Road
( ) <u>Certainly</u> Greyfriars Road
( ) <u>Certainly</u> Dumfries Place
( ) <u>Certainly</u> Capitol
( ) Other please specify ……….
( ) Not sure (I would need more information)

34. [*If needed information in Q.No.31 or 32 or 33*], what information Source(s) you would not use? *Please Tick (√) all that apply.*

( ) Friend(s) [*go to Q.No.35(b)*]
( ) Map(s) [*go to Q.No.35(b)*]
( ) A paper timetable [*go to Q.No.35(b)*]

( ) A telephone enquiry line *[go to Q.No.35(b)]*
( ) website providing information about public transport <u>and</u> car journeys
( ) website providing information about car journeys <u>and</u> road conditions
( ) website providing information about Coach journeys/services <u>only</u>
( ) website providing information about Train journeys/services <u>only</u>
( ) SMS alerts *[go to Q.No.35(b)]*
( ) Radio bulletin or TV Tele-text *[go to Q.No.35(b)]*

Part 2: SP Questions

35. ***(a) Only if chosen information source is internet in 30 answer the following***

Choice 1

Consider you are trying to collect information from the website for your journey to Cardiff and the website provides you following series of information about your journey. Which of the following websites would you prefer?

Information Source A	**Information Source B**
Private Source Subscription £5 that you have already paid No advertisements on the webpage	Government Source No subscription required No Advertisements on the webpage
Accurate travel Time and delay information of the journey	Qualitative information about travel time and delay
Search time 15 min	Search time 10 min
Information Changes in real time	Information Changes daily

( ) Prefer A ( ) Prefer B

35. ***(b) Again assuming that you need to travel from Leeds to Cardiff in 3 days time on personal business with an appointment from 1130 till 1230 in Cardiff city hall.***

Given that you have chosen [*unclick answers fromQ.No.34*] ***as data*** **source**[*s*]***, what will be your mode choice if the source***[*s*] ***provide***[*s*] ***you the following sets of information about the different modes?***

	Car		**Coach**	
Information Source [*only those will be showed which are unclick in Question No.30*]	**Friend**	**Multimodal Website**	**Personal experience**	**Coach Website**
Time	About 5 hrs	5:30 hrs	About 8 hrs	7:20 hrs
Out of pocket cost	About £40	£ 30	About £25	£20
Your Choice	( )		( )	

Part 3: General Background

The following questions are about your attitudes towards choice of mode and information sources in your day to day travel decisions

36. Which of the following sources have you ever used? *Please Tick (√) all that apply and Double Tick (√√) the most important*
   ( ) Friend(s)
   ( ) Map(s)
   ( ) A paper timetable
   ( ) A telephone enquiry line
   ( ) website providing information about public transport <u>and</u> car journeys
   ( ) website providing information about car journeys <u>and</u> road conditions
   ( ) website providing information about Coach journeys/services <u>only</u>
   ( ) website providing information about Train journeys/services <u>only</u>
   ( ) SMS alerts
   ( ) Radio bulletin or TV Tele-text
   ( ) Other please specify ....

37. Which sources would you use if bad weather is expected? *Please Tick (√) all that apply and Double Tick (√√) the most important*
   ( ) Friend(s)
   ( ) Map(s)
   ( ) A paper timetable
   ( ) A telephone enquiry line
   ( ) website providing information about public transport <u>and</u> car journeys
   ( ) website providing information about car journeys <u>and</u> road conditions
   ( ) website providing information about Coach journeys/services <u>only</u>
   ( ) website providing information about Train journeys/services <u>only</u>
   ( ) SMS alerts
   ( ) Radio bulletin or TV Tele-text
   ( ) Other please specify ....

38. Which source would you use if you learn that there has been an incident on your intended route? *Please Tick (√) all that apply and Double Tick (√√) the most important*
   ( ) Friend(s)
   ( ) Map(s)
   ( ) A paper timetable
   ( ) A telephone enquiry line
   ( ) website providing information about public transport <u>and</u> car journeys

( ) website providing information about car journeys <u>and</u> road conditions
( ) website providing information about Coach journeys/services <u>only</u>
( ) website providing information about Train journeys/services <u>only</u>
( ) SMS alerts
( ) Radio bulletin or TV Tele-text
( ) Other please specify ....

39. Have you subscribed to any of the following sources
( ) Telephone enquiry
( ) website which provides information about public transport <u>and</u> car journeys
( ) website providing information about car journeys <u>and</u> road conditions
( ) website which provides information about Coach journeys/services <u>only</u>
( ) website which provides information about Train journeys/services <u>only</u>
( ) SMS

40. [*if yes to any in Q.No.39*] Please name them?
....

41. How frequently you make long journeys (over 50 miles)?
( ) Never
( ) Less than 5 times per year
( ) 6 – 12 times per year
( ) 13 – 40 times per year
( ) More than 40 times per year

42. How often do you use websites as a source of travel information prior to making this kind of journey?

( ) Always	[*go to Q.No.44*]
( ) very often	[*go to Q.No.44*]
( ) Sometimes	[*go to Q.No.44*]
( ) Rarely	[*go to Q.No.43*]
( ) Never	[*go to Q.No.43*]

43. [*if answered Rarely or Never in Q.No.42*] Please select the reasons for not using websites as your travel information source from the following. *Please Tick (√) all that apply and Double Tick (√√) the most important.*
( ) I didn't know about any of the sources [*go to Q.No.46*]
( ) I know my journey thoroughly so had no need for extra information [*go to Q.No.46*]
( ) I used them before but didn't like it (confusing) [*go to Q.No.46*]
( ) I used them before but didn't like it (time consuming) [*go to Q.No.46*]
( ) Information provided was ambiguous and inconsistent [*go to Q.No.46*]
( ) My journey was organised for me (e.g. by tour operator/office) [*go to Q.No.46*]
( ) Any other, please mention... [*go to Q.No.46*]

44. If you use websites for travel information, what method(s) do you use to find it/them? *Please Tick (√) all that apply and Double Tick (√√) the most important.*
( ) I have it/them in my 'favourites' *[go to Q.No.46]*
( ) I type in the website address *[go to Q.No.46]*
( ) I type a source name into a search engine (e.g. Google) *[go to Q.No.46]*
( ) I type a generic phrase (e.g. travel information, time table, national express etc)

45. *[if answered using generic phrase in Q.No.44]* What generic phrase would you use for the journey to Cardiff mentioned earlier? ....

46. How much a month would you be prepared to spend on subscribing to a website which provides you with accurate and up-to-date information about your travel time, delay information, speeds on the network, cost of different modes of travel and departure times for public transport?..............................a month.

47. How important is it that a website is free from irrelevant advertisements and pop ups?
( ) Essential ( ) Important ( ) Not important

48. How important is real time information (<u>current</u> congestion levels, <u>current</u> travel times, incidents) in selecting a source
( ) Essential ( ) Important ( ) Not important

49. How important is it for you to know where the information source gets its information?
( ) Essential ( ) Important ( ) Not important

50. How important are safety aspects for you when selecting a mode
( ) Essential ( ) Important ( ) Not important

51. How important is comfort for you when selecting a mode
( ) Essential ( ) Important ( ) Not important

52. How important is the availability of seats for you when selecting a mode
( ) Essential ( ) Important ( ) Not important

53. Do you like to consider changing/discovering new modes?
( ) Always ( ) very often ( ) Sometimes ( ) Rarely ( ) Never

54. Do you try to optimize your journey in terms of travel time?
( ) Always ( ) very often ( ) Sometimes ( ) Rarely ( ) Never

55. Do you always try to minimize the cost of your journey?

( ) Always ( ) very often ( ) Sometimes ( ) Rarely ( ) Never

Part 4: Personal Information

56. Are you ( ) Male ( ) Female [*Omit if face to face interview*]

57. How old are you?
( ) Less than 20 ( ) 20-29 ( ) 30-39 ( ) 40-49 ( ) 50-64 ( ) 65 and over

58. At what age did you leave full time education? .... years

59. What is your occupation *Please Tick (√) all that apply and Double Tick (√√) the most important.*
( ) Employed Full Time (18 hours or more per week)
( ) Employed Part Time (less than 18 hours per week)
( ) Retired
( ) Self-employed
( ) Full time student
( ) Unemployed
( ) Looking after home
( ) Others (please specify).................

60. How many vehicles (cars, company vehicles, motorcycles, vans or trucks) are there in your household?
( ) None ( ) one ( ) two or more

61. How many people live in your household?

Aged 17 or older	( ) 0	( ) 1	( ) 2	( ) 3	( ) 4	( ) more
Aged 5 to 16	( ) 0	( ) 1	( ) 2	( ) 3	( ) 4	( ) more
Aged less than 5	( ) 0	( ) 1	( ) 2	( ) 3	( ) 4	( ) more

62. What is your household's annual income before tax?
( ) £10,000 or less
( ) £10,001-£20,000
( ) £20,001-£30,000
( ) £30,001-£40,000
( ) £40,001-£50,000
( ) £50,001-£60,000
( ) over £60,000

Appendix B

WinMint Code for Questionnaire

```
****************************************
* Travellers' Choice of Information Sources and
* of modes    *
* November 2007 Agha Faisal Habib Pathan       *
* SURVEY QUESTIONNAIRE           *
* Institute for Transport Studies          *
* University of Leeds            *
*                        *
****************************************
*P
*S INTRODUCTION
*Q 0 COLOR
*T 00 ^C00 Colortype 00^C00
*T 01 ^C01 Colortype 01^C00
*T 02 ^C02 Colortype 02^C00
*T 03 ^C03 Colortype 03^C00
*T 04 ^C04 Colortype 04^C00
*T 05 ^C05 Colortype 05^C00
*T 06 ^C06 Colortype 06^C00
*T 07 ^C07 Colortype 07^C00
*T 08 ^C08 Colortype 08^C00
*T 09 ^C09 Colortype 09^C00
*T 10 ^C10 Colortype 10^C00
*T 11 ^C11 Colortype 11^C00
*T 12 ^C12 Colortype 12^C00
*T 13 ^C13 Colortype 13^C00
*T 14 ^C14 Colortype 14^C00
*T 15 ^C15 Colortype 15^C00
*>
*
P
S Travellers' Choice of Information Sources
U C
Q 2 DATE
T ^B+SURVEY QUESTIONNAIRE^B-
T
U L
T Please enter the date of the interview
T
T
T (The interviewer must fill this by himself)
>
*
Q 3 HHNUM
U L
T Please enter Household Number
T
T
T (The interviewer must fill this by himself)
L 1
H 1000
>
*
Q 3 PERSNUM
T Please enter Person Number
T
T
T (The interviewer must fill this by himself)
L 1
H 5000
>
*
Q 1 TITLE
T Please select respondent's title
T
T
T (The interviewer must fill this by himself)
A Mr
A Ms
>
*
Q 1 TYPE
T Please select type of the questionnaire to be asked?
T
T
T (The interviewer must fill this by himself)
A 1
A 2
A 3
A 4
>
*
Q 7 TIME
T Please enter the time of recruitment?
T
T
T (The interviewer must fill this by himself)
>
*
Q 1 PLACE
T Please select the place of recruitment?
T
T
T (The interviewer must fill this by himself)
A Leeds train station
A Leeds city bus station
A Crownpoint petrol station
O any other (please specify)
>
*
Q 2 NAME
T Please enter the name of the respondent?
T
T
T (The interviewer must fill this by himself)
>
*
P
S Introduction
Q 0 INTRODUCTION
R 00
T Can I speak to #TITLE# #NAME#.
T
T This is Mr Habib, we met at #PLACE# and  you kindly
agreed to help me in my research by
& _taking part  in a short phone interview about your travel
choices.
T
T I should remind you that the Information you provide will
be used for academic purposes only
& _and will be kept confidential.
T
T Do you have the pack of papers I gave you? Are they
readily to hand?
T
T ^B+(If yes, say "thank you" or ask them to go and get
them).^B-
R 14
>
*
****************************************
*
* Part 1: Questions about your journey
*
****************************************
*
S Part 1: Questions about your journey
Q 1 START
R 00
T OK now we can start..
T
T Firstly some  Questions about the journeys you make ....
T
T How many times a  year do you travel to places in the UK
that are at least 50 miles from Leeds?
T
T ^B+(Do not read the options unless necessary)^B-
T
T 1.   never
T 2.   less than 5 times per year
T 3.   6 to 12 times per year
T 4.   13 to 40 times per year
```

T 5. more than 40 times per year
A never
A less than 5 times per year
A 6 to 12 times per year
A 13 to 40 times per year
A more than 40 times per year
>
*
I #START# EQ 1 REMARKS
*
Q 1 JOURNEY
T When did you last travel from Leeds to somewhere more than 50 miles away (in the UK)?
T
T ^B+(Do not read the options unless necessary)^B-
T
T 1. less than a week ago
T 2. within the last month
T 3. more than a month ago
A less than a week ago
A within the last month
A more than a month ago
>
*
Q 1 SAMEMODE
T Did you come back to Leeds using the same means
& _of transport and without visiting any other destination?
T
T 1. Yes
T 2. No
A Yes
A No
>
*
I #SAMEMODE# EQ 1 DETN
*
Q 1 RECALL
T Can you think of a journey of more than 50 miles (in the UK) when you did
& _come back using the same means of transport and without visiting any other destination?
T
T When did you last make that journey?
T
T ^B+(Do not read the options unless necessary)^B-
T
T 1. less than a week ago
T 2. within the last month
T 3. more than a month ago
T 4. never
A less than a week ago
A within the last month
A more than a month ago
A terminate
>
*
I #RECALL# EQ 4 REMARKS
*
Q 2 DETN
T Where did you go?
>
*
Q 1 DAY
T When did you travel?
T
T ^B+(Read out the options)^B-
T
T 1. weekday peak period ^B+(only explain if required Morning peak = 7:00-10:00 and evening peak =16:00-1900)^B-
T 2. weekday off peak
T 3. weekend
A weekday peak period
A weekday off peak
A weekend
>
*
Q 1 MAINPUR
T What was the ^C09main^C00 purpose of your journey to #DETN#?
T
T ^B+(Do not read the options unless necessary)^B-
T
T ^B+(Business)^B-
T 1. Travelling to/from work
T 2. Travelling in the course of business
T 3. Personal business
T ^B+(Leisure)^B-
T 4. Shopping
T 5. Visiting friends/relatives
T 6. Sports/entertainment
T 7. Other leisure
T ^B+(Education)^B-
T 8. Education
T 9. Other please specify
A Travelling to/from work
A Travelling in the course of business
A Personal business
A Shopping
A Visiting friends/relatives
A Sports/entertainment
A Other leisure
A Education
O Other please specify
>
*
Q 10 SECONDPUR
T Was there any ^C09other^C00 purpose of your journey to #DETN#?
T
T 1. Travelling to/from work
T 2. Travelling in the course of business
T 3. Personal business
T 4. Shopping
T 5. Visiting friends/relatives
T 6. Sports/entertainment
T 7. Other leisure
T 8. Education
T 9. Other please specify
T 10. No other purpose
D #MAINPUR#
A No other purpose
>
*
Q 1 CHOSENMODE
T What mode of transport did you use for your journey to #DETN#?
T
T ^B+(Do not read out unless they don't understand, and if they say they used more than one
& _mode then ask for the one they used for the longest distance).^B-
T
T 1. car
T 2. coach
T 3. train
T 4. airplane
T 5. other please specify
A car
A coach
A train
A airplane
O other please specify
>
*
Q 1 MINDMODE
T When you were ^C09first planning^C00 your journey, did you have any particular mode in mind?
T
T ^B+(Do not read the options unless necessary)^B-
T
T 1. car

T 2. coach
T 3. train
T 4. airplane
T 5. other (please specify)
T 6. no particular mode in mind
A car
A coach
A train
A airplane
O other (please specify)
A no particular mode in mind
>
*
I #MINDMODE# EQ 6 ISCONSULTED
*
Q 12 REASMODE
T Why were you thinking of going by #MINDMODE#?
T
T ^B+(do not read out the options, let him say and then anything else)^B-.
T
T 1. travel cost
T 2. travel time
T 3. bad weather
T 4. congestion/other problem with preferred mode
T 5. safety
T 6. comfort and convenience
T 7. travelling with friends/family
T 8. other (please specify)
A travel cost
A travel time
A bad weather
A congestion/other problem with preferred mode
A safety
A comfort and convenience
A travelling with friends/family
O other (please specify)
>
*
* Q 1 MAINREASM
* T Which one do you think was the most important reason?
* C #REASMODE1# EQ 1 #REASMODE1#
* C #REASMODE2# EQ 1 #REASMODE2#
* C #REASMODE3# EQ 1 #REASMODE3#
* C #REASMODE4# EQ 1 #REASMODE4#
* C #REASMODE5# EQ 1 #REASMODE5#
* C #REASMODE6# EQ 1 #REASMODE6#
* C #REASMODE7# EQ 1 #REASMODE7#
* C #REASMODE8# EQ 1 #REASMODE8#
* >
*
Q 0 DUMMY1
I #CHOSENMODE# NE #MINDMODE# AND #MINDMODE# LT 6 REACMODE
I #CHOSENMODE# EQ #MINDMODE# OR #MINDMODE EQ 6 ISCONSULTED
F 0 #DUMMY1#
*
Q 12 REACMODE
T Why did you change your mind (changing the mode)?
T
T ^B+(he first thought about #MINDMODE# and actually travelled by #CHOSENMODE#)^B-
T
T 1. travel cost
T 2. travel time
T 3. bad weather
T 4. congestion/other problem with preferred mode
T 5. safety
T 6. comfort and convenience
T 7. travelling with friends/family
T 8. other (please specify)
A travel Cost
A travel time
A bad weather
A congestion/other problem with preferred mode
A safety
A comfort and convenience
A travelling with friends/family
O other (please specify)
>
*
* Q 1 MAINREACM
* T From the reasons for changing the mode, which one
* & _do you think was the most important?
* C #MAINREACM1# EQ 1 #MAINREACM1#
* C #MAINREACM2# EQ 1 #MAINREACM2#
* C #MAINREACM3# EQ 1 #MAINREACM3#
* C #MAINREACM4# EQ 1 #MAINREACM4#
* C #MAINREACM5# EQ 1 #MAINREACM5#
* C #MAINREACM6# EQ 1 #MAINREACM6#
* C #MAINREACM7# EQ 1 #MAINREACM7#
* C #MAINREACM8# EQ 1 #MAINREACM8#
* >
*
Q 12 ISCONSULTED
T With reference to ^C09Card 1^C00 in the pack of papers, what information source(s) did you consult before travelling?
T
T
T 1. advice from friend(s)
T 2. map(s)
T 3. a printed timetable
T 4. a telephone enquiry line
T 5. website(s)
T 6. SMS alert(s)
T 7. radio bulletin or TV tele-text
T 8. other (please specify)
T 9. I didn't use any information source
A advice from friend(s)
A map(s)
A a printed timetable
A a telephone enquiry line
A website(s)
A SMS alert(s)
A radio bulletin or TV tele-text
O other (please specify)
A I didn't use any information source
>
*
I #ISCONSULTED9# EQ 1 AND #ISCONSULTED1# EQ 0 AND #ISCONSULTED2# EQ 0 AND #ISCONSULTED3# EQ 0 AND #ISCONSULTED4# EQ 0 AND #ISCONSULTED5# EQ 0 AND #ISCONSULTED6# EQ 0 AND #ISCONSULTED7# EQ 0 AND #ISCONSULTED8# EQ 0 NOTUSINGIS
*
* Q 1 MAINSOURCE
* T Which one do you think was the most useful source?
* C #ISCONSULTED1# EQ 1 #ISCONSULTED1#
* C #ISCONSULTED2# EQ 1 #ISCONSULTED2#
* C #ISCONSULTED3# EQ 1 #ISCONSULTED3#
* C #ISCONSULTED4# EQ 1 #ISCONSULTED4#
* C #ISCONSULTED5# EQ 1 #ISCONSULTED5#
* C #ISCONSULTED6# EQ 1 #ISCONSULTED6#
* C #ISCONSULTED7# EQ 1 #ISCONSULTED7#
* C #ISCONSULTED8# EQ 1 #ISCONSULTED8#
* C #ISCONSULTED9# EQ 1 #ISCONSULTED9#
* C #ISCONSULTED10# EQ 1 #ISCONSULTED10#
* C #ISCONSULTED11# EQ 1 #ISCONSULTED11#
* >
*
I #ISCONSULTED5# EQ 1 WEB
*
Q 0 DUMMY3
I #ISCONSULTED1# EQ 0 OR #ISCONSULTED1# EQ 1 DUMMY8
F 0 #DUMMY3#
*
Q 12 WEB
T By websites, do you mean......

T
T 1. .....website(s) providing information about ^B+only^B- car journeys and road conditions
T 2. or website(s) providing information about ^B+both^B- public transport and car journeys
T 3. or website(s) providing information about ^B+only^B- coach journeys/services
T 4. or website(s) providing information about ^B+only^B- train journeys/services?
A website-about car journeys only
A website-multimodal
A website-coach journeys only
A website-train journeys only
>
*
Q 2 WEBNAME
T Which website(s)/journey planner(s) did you consult?
T
T ^B+(If they are stuck then read out some of the following as examples).^B-
T
T
T Transport Direct,
T Multi Map,
T AA,
T Traffic Master,
T Traffic-I,
T Traveline,
T Nationalrailenquiries
T infotransport
T nationalexpress
T virgintrains
L 1
H 25
>
*
Q 3 WEBCONSUL
T How many days before the actual departure date did you consult the website(s)?
L 0
H 90
>
*
Q 12 REASONIS
T For what reasons did you consult websites for your journey to #DETN#?
T
T ^B+(Ask first about the two groups car or public transport)^B-
T ^C09to get information about car journey^C00
T 1. to plan/select a route for car journey
T 2. to confirm/check travel distances
T 3. to confirm/check travel costs
T 4. to check parking facilities
T 5. to look for live traffic news
T 6. to compare a car with public transport
T ^C09to get information about public transport^C00
T 7. to plan a public transport journey
T 8. to book public transport ticket(s)
T 9. to reserve seats on train or coach
T 10. to consult public transport timetables
T 11. any other (please specify)
A to plan a car journey/selecting a route for car journey
A to confirm/check travel distances
A to confirm/check travel costs
A to check parking facilities
A to look for live traffic news
A to compare a car with public transport
A to plan a public transport journey
A to book public transport ticket(s)
A to reserve seats on train or coach
A to enquire public transport timetables
O any other (please specify)
>
*
* Q 1 MAINREASIS
* T Which one do you think was the most important?
* C #REASONIS1# EQ 1 #REASONIS1#
* C #REASONIS2# EQ 1 #REASONIS2#
* C #REASONIS3# EQ 1 #REASONIS3#
* C #REASONIS4# EQ 1 #REASONIS4#
* C #REASONIS5# EQ 1 #REASONIS5#
* C #REASONIS6# EQ 1 #REASONIS6#
* C #REASONIS7# EQ 1 #REASONIS7#
* C #REASONIS8# EQ 1 #REASONIS8#
* C #REASONIS9# EQ 1 #REASONIS9#
* C #REASONIS10# EQ 1 #REASONIS10#
* C #REASONIS11# EQ 1 #REASONIS11#
* >
*
*
I #REASMODE# GT 0 DUMMY8
*
Q 12 NOTUSINGIS
T Why did you ^C09not^C00 use any information sources?
T
T ^B+(first allow respondent to mention and then direct her/him to the following options).^B-
T
T 1. I didn't know about any of the sources
T 2. I know my journey thoroughly so had no need for extra information
T 3. I used them before but didn't like it (confusing)
T 4. I used them before but didn't like it (time consuming)
T 5. information provided was ambiguous and inconsistent
T 6. my journey was organised for me (e.g. by tour operator/office)
T 7. any other, please mention
A I didn't know about any of the sources
A I know my journey thoroughly so had no need for extra information
A I used them before but didn't like it (confusing)
A I used them before but didn't like it (time consuming)
A information provided was ambiguous and inconsistent
A my journey was organised for me (e.g. by tour operator/office)
O any other, please mention
>
*
* Q 1 MAINREANIS
* T From the reasons for not using the information sources, which one
* & _do you think was the most important?
* C #NOTUSINGIS1# EQ 1 #NOTUSINGIS1#
* C #NOTUSINGIS2# EQ 1 #NOTUSINGIS2#
* C #NOTUSINGIS3# EQ 1 #NOTUSINGIS3#
* C #NOTUSINGIS4# EQ 1 #NOTUSINGIS4#
* C #NOTUSINGIS5# EQ 1 #NOTUSINGIS5#
* C #NOTUSINGIS6# EQ 1 #NOTUSINGIS6#
* C #NOTUSINGIS7# EQ 1 #NOTUSINGIS7#
* >
*
S Part 1b: Questions about your journey by #CHOSENMODE#
Q 0 DUMMY8
I #CHOSENMODE# EQ 1 TTIMECAR
I #CHOSENMODE# EQ 2 OR #CHOSENMODE# EQ 3 REACHOSENMODE
I #CHOSENMODE# EQ 4 REAIRPORT
I #CHOSENMODE# EQ 5 TTIMECAR
F 0 #DUMMY8#
*
Q 6 TTIMECAR
T How long did it take to get from Leeds to your destination in #DETN# by #CHOSENMODE# (door to door)?
L 30
H 900
>
I #TTIMECAR# GT 0 COSTCAR
*
Q 6 REACHOSENMODE

T How long did it take to get from Leeds to your destination by #CHOSENMODE# (door to door)?
T ^B+(Do not read out only unless necessary ..... including getting to the #CHOSENMODE# station, waiting,
& _on the #CHOSENMODE# and then onward to final destination)^B-
H 960
>
*
I #REACHOSENMODE# GT 0 COSTMODE
*
Q 6 REAIRPORT
T How long did it take to get from Leeds to your destination by #CHOSENMODE# (door to door)?
T ^B+(Do not read out unless necessary ..... including getting to the airport, waiting,
& _on the #CHOSENMODE# and then onward to final destination)^B-
H 960
>
*
Q 4 COSTMODE
T What did the round trip (from door to door and back again) cost you?
& _(the fare there and back, plus any other travel costs incurred)?
L 100
H 50000
>
J #COSTMODE# GT 0 1
*
Q 4 COSTCAR
T What did the round trip (from door to door and back again) cost you (please include petrol, parking, any tolls paid etc)?
L 100
H 50000
>
*
I #CHOSENMODE# EQ 1 CREACHCOACH
I #CHOSENMODE# GT 1 CTTIMECAR
*
Q 6 CTTIMECAR
S Part 1c: Questions about your journey by car
T How long do you think it would have taken you to travel to your destination in #DETN# by car (door to door)?
L 30
H 1000
T
T
T
T
T
T
T
T
T
T
T
T
T
T 15=NI, 16=NINC
>
*
I #CTTIMECAR# EQ 960 CREACHCOAH
*
Q 4 CCOSTCAR
T What do you think the round trip (from door to door and back) would have cost you if you had gone by car? (please include petrol, parking, any tolls etc)
T
T
T
T
T
T
T
T
T
T
T
T
T
T 888=NI, 999=NINC
L 100
H 100000
>
*
I #CHOSENMODE# EQ 2 CREACHTRAIN
*
Q 6 CREACHCOACH
S Part 1c: Questions about your journey by coach
T How long do you think it would have taken you to travel to your destination in #DETN# by coach (door to door)?
T ^B+(Do not read out only unless necessary ..... including getting to the coach station, waiting,
& _on the coach and then onward to final destination)^B-
T
T
T
T
T
T
T
T
T
T
T
T
T
T 15=NI, 16=NINC
H 1000
>
*
I #CREACHCOACH# EQ 960 CREACHTRAIN
*
Q 4 CCOSTCOACH
T What do you think the round trip (from door to door and back) would have cost you if you had travelled by coach?
& _(please include fare there and back, plus any other travel costs you would have incurred)
T
T
T
T
T
T
T
T
T
T
T
T
T 888=NI, 999=NINC
L 100
H 100000
>
*
I #CHOSENMODE# EQ 3 CREAIRPORT
*
Q 6 CREACHTRAIN
S Part 1c: Questions about your journey by train
T How long do you think it would have taken to travel from Leeds to your destination in #DETN# by train (door to door)?
T ^B+(Do not read out only unless necessary ..... including getting to the train station, waiting,
& _on the train and then onward to final destination)^B-
T
T
T
T

```
T
T
T
T
T
T
T
T
T
T 15=NI, 16=NINC
H 1000
>
*
I #CREACHTRAIN# EQ 960 CREAIRPORT
*
Q 4 CCOSTRAIN
T What do you think the round trip (from door to door and
back) would have cost you if you had travelled by train?
& _(please include fare there and back, plus any other travel
costs you would have incurred)
T
T
T
T
T
T
T
T
T
T
T
T
T 888=NI, 999=NINC
L 100
H 100000
>
*
I #CHOSENMODE# EQ 4 DUMMY1
*
Q 6 CREAIRPORT
S Part 1c: Questions about your journey by air
T How long do you think it would have taken to travel from
Leeds to your destination in #DETN# by air (door to door)?
T ^B+(Do not read out only unless necessary ..... including
getting to the airport, waiting,
& _on the plane and then onward to final destination)^B-
T
T
T
T
T
T
T
T
T
T
T
T
T
T 15=NI, 16=NINC
H 1000
>
*
I #CREAIRPORT# EQ 960 DUMMY1
*
Q 4 CCOSTAIR
T What do you think the round trip (from door to door and
back) would have cost you if you had gone by plane?
& _(please include fare there and back, plus any other travel
costs you would have incurred)
T
T
T
T
T
T
T
T
T
T
T
T
T
T 888=NI, 999=NINC
L 100
H 100000
>
*
Q 1 WEATHER
S Part 1d: Questions about your journey
T Was bad weather expected on the day of your journey to
#DETN#?
T
T 1. Yes
T 2. No
A Yes
A No
>
*
Q 1 INCIDENT
T Before you started your journey, did you have advance
knowledge about any incident/congestion along the route?
T
T
T 1. Yes
T 2. No
A Yes
A No
>
*
*
*
*****************************************
*
* Part 2: Questions about an imaginary journey to Cardiff
*
*****************************************
*
P
S Part 2: Questions about an imaginary journey to Cardiff
*
Q 0 INTROPART2
T Thank you I have no more questions about your journeys
to #DETN#.
T
T
T I would now like you to imagine that you need to travel
from Leeds to Cardiff.
T
T The journey is described on Card 3 in the papers I gave
you
>
*
Q 1 PASTJOURNEY
T Have you ever travelled to Cardiff from Leeds?
T
T 1. Yes
T 2. No
A Yes
A No
>
*
Q 12 CCMODE
T What modes of transport would you consider for the
journey to Cardiff described on Card 3?
T
T ^B+(read out options)^B-
T
T 1. car
T 2. coach
T 3. train
T 4. airplane
```

A car
A coach
A train
A airplane
>
*
I #CCMODE1# EQ 0 CCMODEA
I #CCMODE2# EQ 0 CCMODEB
I #CCMODE3# EQ 0 CCMODEC
I #CCMODE4# EQ 0 CCMODED
I #CCMODE1# EQ 1 DUMMYS1
*
Q 1 CCMODEA
T Do you mean that there is ^C09no possibility^C00 of your using car for a journey to Cardiff in 3 days time?
T
T ^B+(read out options)^B-
T
T 1. yes there is no possibility
T 2. its unlikely
A yes there is no possibility
A its unlikely
>
*
I #CCMODE2# EQ 1 DUMMYS1
*
Q 1 CCMODEB
T Do you mean that there is ^C09no possibility^C00 of your using coach for a journey to Cardiff in 3 days time?
T
T ^B+(read out options)^B-
T
T 1. yes there is no possibility
T 2. its unlikely
A yes there is no possibility
A its unlikely
>
*
Q 0 DUMMYS1
I #CCMODE3# EQ 1 DUMMYS2
F 0 #DUMMYS1#
*
Q 1 CCMODEC
T Do you mean that there is ^C09no possibility^C00 of your using train for a journey to Cardiff in 3 days time?
T
T ^B+(read out options)^B-
T
T 1. yes there is no possibility
T 2. its unlikely
A yes there is no possibility
A its unlikely
>
*
Q 0 DUMMYS2
I #CCMODE4# EQ 1 MODEUSE
F 0 #DUMMYS1#
*
Q 1 CCMODED
T Do you mean that there is ^C09no possibility^C00 of your using airplane for a journey to Cardiff in 3 days time?
T
T ^B+(read out options)^B-
T
T 1. yes there is no possibility
T 2. its unlikely
A yes there is no possibility
A its unlikely
>
*
*
Q 1 MODEUSE
T Which mode of transport do you think you ^C09would^C00 use?
T
T ^B+(read out options)^B-
T
T 1. certainly car
T 2. certainly coach
T 3. certainly train
T 4. certainly airplane
T 5. or not sure (you would need more information)
A certainly car
A certainly coach
A certainly train
A certainly airplane
A not sure (you would need more information)
>
*
I #MODEUSE# EQ 1 OR #MODEUSE# EQ 5 ROUTEUSE
*
Q 1 TICKETBOOK
T How would you book your ticket?
T
T ^B+(read out options)^B-
T
T 1. certainly online
T 2. certainly from a ticket machine at station
T 3. certainly form a booking office
T 4. certainly from a travel agent
T 5. or not sure (you would need more information)
A certainly online
A certainly from a ticket machine at station
A certainly form a booking office
A certainly from a travel agent
A not sure (you would need more information)
>
*
I #TICKETBOOK# EQ 2 OR #TICKETBOOK# EQ 3 OR #TICKETBOOK# EQ 4 SP2INTRO1
I #TICKETBOOK# EQ 1 OR #TICKETBOOK# EQ 5 CINFOSOURCE
*
Q 1 ROUTEUSE
T If you travel by car, which route do you think you ^C09would^C00 use?
T
T ^B+(read out options)^B-
T
T 1. a Motorway route
T 2. a non Motorway route
T 3. or not sure (you would need more information)
A a Motorway route
A a non Motorway route
A or not sure (you would need more information)
>
*
Q 1 PARK
T Do you know exactly where you would park in Cardiff?
T
T 1. yes you know
T 2. or not sure (you would need more information)
A yes you know
A or not sure (you would need more information)
>
*
I #MODEUSE# EQ 5 OR #ROUTEUSE# EQ 3 OR #PARK# EQ 2 CINFOSOURCE
*
Q 1 PARK1
T where?
T
T 1. certainly Castle Mews
T 2. certainly North Road
T 3. certainly Grey friars Road
T 4. certainly Dumfries Place
T 5. certainly Capitol
T 6. other (please specify)
A certainly Castle Mews
A certainly North Road
A certainly Grey friars Road
A certainly Dumfries Place

A certainly Capitol
O other (please specify)
>
*
Q 0 CDUMMY1
I #MODEUSE# NE 5 AND #ROUTEUSE# NE 3 AND #PARK# NE 2 SP2INTRO1
F 0 #CDUMMY1#
*
Q 12 CINFOSOURCE
T You said you would need more information.
T
T What information source(s) you would ^C09not^C00 use?
T
T Please look again at Card No.2 and tell me which information sources you could not use.
T
T 1. advice from a friend
T 2. map(s)
T 3. a paper timetable
T 4. a telephone enquiry line
T 5. website(s)
T 6. SMS alert(s)
T 7. radio bulletin or TV tele-text
A advice from a friend
A map(s)
A a paper timetable
A a telephone enquiry line
A website(s)
A SMS alert(s)
A radio bulletin or TV tele-text
>
*
* Q 1 CINFOSOURCEE
* T From the information sources that you could use, which one is the most
* & _likely to be used by you for your journey to Cardiff?
* C #CINFOSOURCE1# EQ 0 advice from a friend
* C #CINFOSOURCE2# EQ 0 map(s)
* C #CINFOSOURCE3# EQ 0 a paper timetable
* C #CINFOSOURCE4# EQ 0 a telephone enquiry line
* C #CINFOSOURCE5# EQ 0 website-car journeys and road conditions
* C #CINFOSOURCE6# EQ 0 website-public transport and car journeys
* C #CINFOSOURCE7# EQ 0 website-coach journeys/services only
* C #CINFOSOURCE8# EQ 0 website-train journeys/services only
* C #CINFOSOURCE9# EQ 0 SMS alert(s)
* C #CINFOSOURCE10# EQ 0 radio bulletin or TV tele-text
* >
*
Q 0 DUMMYE
I #CINFOSOURCE5# EQ 1 SP2INTRO1
F 0 #DUMMYE#
*
P
S FIRST SP GAME INTRODUCTION
Q 0 SP1INTRO
T Please look at Card No. 4 in the pack,
T
T Can you just confirm the number on the top right had corner of Card .
T ^B+(Check the correct card number)^B-
T
T Please imagine you are trying to collect information from a website for the journey to Cardiff in three days time.
T
T Card 4 shows 8 choices between pairs of websites, 1-4 on the front and 5-8 on the back.
T
T ^B+(PAUSE)^B-
T
T Starting with pair number 1, which website would you prefer, A or B?
T
>
*
G B 3 SP1
U C
G V 4
G A 16
G T 1 1 1 Government Source
G T 1 1 2 No advertisements on the webpage
G T 1 1 3 No subscription required
G T 1 2 1 Commercial source
G T 1 2 2 Advertisements on the webpage
G T 1 2 3 No subscription required
G T 1 3 1 Commercial source
G T 1 3 2 Advertisements on the webpage
G T 1 3 3 Subscription £5 that you have already paid
G T 1 4 1 Commercial source
G T 1 4 2 No advertisements on the webpage
G T 1 4 3 No subscription required
G D 1 1 1
G D 1 2 1
G D 1 3 1
G D 1 4 1
G D 1 5 2
G D 1 6 2
G D 1 7 2
G D 1 8 2
G D 1 9 3
G D 1 10 3
G D 1 11 3
G D 1 12 3
G D 1 13 4
G D 1 14 4
G D 1 15 4
G D 1 16 4
G L 2 3
G T 2 1 1 General information, relevant to your journey,
G T 2 1 2 about routes/services/prices
G T 2 2 1 Specific information, relevant to your journey,
G T 2 2 2 about routes/services/prices
G T 2 3 1 Specific information, based on your own criteria,
G T 2 3 2 for routes/services/prices
G D 2 1 1
G D 2 2 2
G D 2 3 3
G D 2 4 2
G D 2 5 2
G D 2 6 1
G D 2 7 2
G D 2 8 3
G D 2 9 3
G D 2 10 2
G D 2 11 1
G D 2 12 2
G D 2 13 2
G D 2 14 3
G D 2 15 2
G D 2 16 1
G L 3 3
G T 3 1 1 Search time 5 minutes
G T 3 2 1 Search time 10 minutes
G T 3 3 1 Search time 15 minutes
G D 3 1 1
G D 3 2 3
G D 3 3 2
G D 3 4 2
G D 3 5 2
G D 3 6 2
G D 3 7 3
G D 3 8 1
G D 3 9 3
G D 3 10 1
G D 3 11 2
G D 3 12 2

G D 3 13 2
G D 3 14 2
G D 3 15 1
G D 3 16 3
G L 4 3
G T 4 1 1 Information Changes/updates in real time
G T 4 2 1 Information Changes/updates daily
G T 4 3 1 Information Changes/updates weekly
G D 4 1 1
G D 4 2 2
G D 4 3 2
G D 4 4 3
G D 4 5 2
G D 4 6 3
G D 4 7 1
G D 4 8 2
G D 4 9 3
G D 4 10 2
G D 4 11 2
G D 4 12 1
G D 4 13 2
G D 4 14 1
G D 4 15 3
G D 4 16 2
G C 8
G F 1 8 1 1 8
G F 1 8 2 9 16
G P < Now for pair number _, which website would you prefer, _ or _? >
G X 1 Website (A)
G X 2 Website (B)
G R 2
G Y 1 Prefer (A)
G Y 2 Prefer (B)
G Z 1 1
G Z 2 2
G H 3
G >
U L
**
*
*
P
S SECOND SP GAME INTRODUCTION
Q 0 SP2INTRO1
T Thank you, now I am going to ask you to make some more choices.
T
T Keeping in mind the journey described in Card 3, please select Card 5, and read out the number on top right hand corner.
T
T ^B+(Check the correct card number)^B-
T
T This card has 9 choices 1-4 on the front and 5-9 on the back.
>
*
I #CINFOSOURCE5# EQ 0 OR #CINFOSOURCE6# EQ 0 OR #CINFOSOURCE7# EQ 0 OR #CINFOSOURCE8# EQ 0 SP2INTRO2
*
*
G B 2 SP2
U C
G V 4
G A 18
G L 1 6
G L 2 6
G L 3 6
G L 4 6
G W 2
G T 1 1 1 ^B+A Friend^B- tells you
G T 1 1 2 ^B+ __________Time:^B- 4 hrs 30 min
G T 1 2 1 ^B+A Friend^B- tells you
G T 1 2 2 ^B+ Time:^B- 5 hrs
G T 1 2 3 (including 20 min delay)
G T 1 3 1 ^B+A Friend^B- tells you
G T 1 3 2 ^B+ Time:^B- 6 hrs
G T 1 3 3 (including 30 min delay)
G T 1 4 1 ^B+A telephone enquiry^B- tells you
G T 1 4 2 ^B+ Time:^B- 3 hrs 30 min
G T 1 4 3 (scheduled station to station)
G T 1 5 1 ^B+A telephone enquiry^B- tells you
G T 1 5 2 ^B+ Time:^B- 3 hrs 50 min
G T 1 5 3 (scheduled station to station)
G T 1 6 1 ^B+A telephone enquiry^B- tells you
G T 1 6 2 ^B+ Time:^B- 4 hrs 10 min
G T 1 6 3 (scheduled station to station)
G T 2 1 1 ^B+A friend^B- tells you
G T 2 1 2 ^B+Out of pocket cost:^B- ?60
G T 2 2 1 ^B+A friend^B- tells you
G T 2 2 2 ^B+Out of pocket cost:^B- ?75
G T 2 3 1 ^B+A friend^B- tells you
G T 2 3 2 ^B+Out of pocket cost:^B- ?90
G T 2 4 1 ^B+A friend^B- tells you
G T 2 4 2 ^B+Out of pocket cost:^B- ?110
G T 2 5 1 ^B+A telephone enquiry^B- tells you
G T 2 5 2 ^B+Out of pocket cost:^B- ?90
G T 2 6 1 ^B+A telephone enquiry^B- tells you
G T 2 6 2 ^B+Out of pocket cost:^B- ?130
G T 3 1 1 ^B+A SMS alert^B- tells you
G T 3 1 2 ^B+ Time:^B- 4 hrs 20 min
G T 3 2 1 ^B+A SMS alert^B- tells you
G T 3 2 2 ^B+ Time:^B- 4 hrs 40 min
G T 3 3 1 ^B+A SMS alert^B- tells you
G T 3 3 2 ^B+ Time:^B- 5 hrs 10 min
G T 3 3 3 (including 30 min delay)
G T 3 4 1 ^B+Recent experience^B- tells you
G T 3 4 2 ^B+ Time:^B- 3 hrs 40 min
G T 3 4 3 (scheduled station to station)
G T 3 5 1 ^B+Recent experience^B- tells you
G T 3 5 2 ^B+ Time:^B- 4 hrs
G T 3 5 3 (scheduled station to station)
G T 3 6 1 ^B+Recent experience^B- tells you
G T 3 6 2 ^B+ Time:^B- 4 hrs 5 min
G T 3 6 3 (scheduled station to station)
G T 4 1 1 ^B+A SMS alert^B- tells you
G T 4 1 2 ^B+Out of pocket cost:^B- ?50
G T 4 2 1 ^B+A SMS alert^B- tells you
G T 4 2 2 ^B+Out of pocket cost:^B- ?60
G T 4 3 1 ^B+A SMS alert^B- tells you
G T 4 3 2 ^B+Out of pocket cost:^B- ?70
G T 4 4 1 ^B+Recent experience^B- tells you
G T 4 4 2 ^B+Out of pocket cost:^B- ?115
G T 4 5 1 ^B+Recent experience^B- tells you
G T 4 5 2 ^B+Out of pocket cost:^B- ?95
G T 4 6 1 ^B+Recent experience^B- tells you
G T 4 6 2 ^B+Out of pocket cost:^B- ?125
G C 5
G F 1 5 1 1 5
G F 1 5 2 10 14
G P < Which modes would you prefer? >
G X 1 CAR
G X 2 TRAIN
G R 2
G Y 1 Prefer car
G Y 2 Prefer train
G Z 1 1
G Z 2 2
G H 3
G >
U L
**
*
*
*
*
*
G B 2 SP3
U C
G V 4

```
G A 18
G L 1 6
G L 2 6
G L 3 6
G L 4 6
G W 2
* Levels to define coach
U L
G T 1 1 1 ^B+Time table^B- tells you
U C
G T 1 1 2 ^B+ Time:^B- 6 hrs 30 min
G T 1 1 3 (scheduled station to station)
U L
G T 1 2 1 ^B+Time table^B- tells you
U C
G T 1 2 2 ^B+ Time:^B- 7 hrs
G T 1 2 3 (scheduled station to station)
U L
G T 1 3 1 ^B+Time table^B- tells you
U C
G T 1 3 2 ^B+ Time:^B- 8 hrs 30 min
G T 1 3 3 (scheduled station to station)
U L
G T 2 1 1 ^B+Time table^B- tells you
U C
G T 2 1 2 ^B+Out of pocket cost:^B- ?50
U L
G T 2 2 1 ^B+Time table^B- tells you
U C
G T 2 2 2 ^B+Out of pocket cost:^B- ?60
U L
G T 2 3 1 ^B+Time table^B- tells you
U C
G T 2 3 2 ^B+Out of pocket cost:^B- ?70
U L
G T 3 1 1 ^B+A multimodal website^B- tells you
U C
G T 3 1 2 ^B+ Time:^B- 6 hrs 35 min
G T 3 1 3 (scheduled station to station)
U L
G T 3 2 1 ^B+A multimodal website^B- tells you
U C
G T 3 2 2 ^B+ Time:^B- 7 hrs 30 min
G T 3 2 3 (scheduled station to station)
U L
G T 3 3 1 ^B+A multimodal websiteB- tells you
U C
G T 3 3 2 ^B+ Time:^B- 8 hrs 20 min
G T 3 3 3 (scheduled station to station)
U L
G T 4 1 1 ^B+A multimodal website^B- tells you
U C
G T 4 1 2 ^B+Out of pocket cost:^B- ?55
U L
G T 4 2 1 ^B+A multimodal website^B- tells you
U C
G T 4 2 2 ^B+Out of pocket cost:^B- ?62
U L
G T 4 3 1  ^B+A multimodal website^B- tells you
U C
G T 4 3 2 ^B+Out of pocket cost:^B- ?70
* Levels to define train
U L
G T 1 4 1 ^B+A telephone enquiry^B- tells you
U C
G T 1 4 2 ^B+ Time:^B- 3 hrs 30 min
G T 1 4 3 (scheduled station to station)
U L
G T 1 5 1 ^B+A telephone enquiry^B- tells you
U C
G T 1 5 2 ^B+ Time:^B- 3 hrs 50 min
G T 1 5 3 (scheduled station to station)
U L
G T 1 6 1 ^B+A telephone enquiry^B- tells you
U C
G T 1 6 2 ^B+ Time:^B- 4 hrs 10 min
G T 1 6 3 (scheduled station to station)
U L
G T 2 4 1 ^B+A telephone enquiry^B- tells you
U C
G T 2 4 2 ^B+Out of pocket cost:^B- ?110
U L
G T 2 5 1 ^B+A telephone enquiry^B- tells you
U C
G T 2 5 2 ^B+Out of pocket cost:^B- ?90
U L
G T 2 6 1 ^B+A telephone enquiry^B- tells you
U C
G T 2 6 2 ^B+Out of pocket cost:^B- ?130
U L
G T 3 4 1 ^B+Recent experience^B- tells you
U C
G T 3 4 2 ^B+ Time:^B- 3 hrs 40 min
G T 3 4 3 (scheduled station to station)
U L
G T 3 5 1 ^B+Recent experience^B- tells you
U C
G T 3 5 2 ^B+ Time:^B- 4 hrs
G T 3 5 3 (scheduled station to station)
U L
G T 3 6 1 ^B+Recent experience^B- tells you
U C
G T 3 6 2 ^B+ Time:^B- 4 hrs 5 min
G T 3 6 3 (scheduled station to station)
U L
G T 4 4 1 ^B+Recent experience^B- tells you
U C
G T 4 4 2 ^B+Out of pocket cost:^B- ?115
U L
G T 4 5 1 ^B+Recent experience^B- tells you
U C
G T 4 5 2 ^B+Out of pocket cost:^B- ?95
U L
G T 4 6 1  ^B+Recent experience^B- tells you
U C
G T 4 6 2 ^B+Out of pocket cost:^B- ?125
G C 4
G F 1 4 1 6 9
G F 1 4 2 15 18
G P < Which modes would you prefer? >
G X 1 COAH
G X 2 TRAIN
G R 2
G Y 1 Prefer coach
G Y 2 Prefer train
G Z 1 1
G Z 2 2
G H 3
G >
U L
*
Q 0 SP2INTRO11
I #CINFOSOURCE5# EQ 1 AND #CINFOSOURCE6# EQ
1 AND #CINFOSOURCE7# EQ 1 AND
#CINFOSOURCE8# EQ 1 INTROPART3
F 0 #SP2INTRO11#
*
**
Q 0 SP2INTRO2
I #CCMODEA# EQ 1 SP2INTRO3
I #CCMODEB# EQ 1 SP2INTRO4
I #CCMODEC# EQ 1 SP2INTRO5
F 0  #SP2INTRO2#
Q 0 SP2INTRO2A
T First 5 set of choices ask whether you would choose to go
to Cardiff by Car or by Train and shows information about
each of them .
T ^B+(PAUSE)^B-
T At choice 1 it shows  information from two sources about
going by car and from two sources about going by train.
T ^B+(PAUSE)^B-
```

```
T For the car journey,  a friend says that it will take about .... and will cost about ..... and a multimodal website says it would take about ... and costs about ......
T ^B+(PAUSE)^B-
T For the train journey, again you have 2 information sources, a train website says that it will take about .... and will cost about ..... but your recent experience is that it had taken ....... and had cost you ...........
T
T Which mode would you choose car or train? A is for car and B is for train?
>
*
G B 2 SP2
U C
G V 4
G A 18
G L 1 6
G L 2 6
G L 3 6
G L 4 6
G W 2
G T 1 1 1 ^C02^B+CAR^B-^C00
U L
G T 1 1 2 ^B+A multimodal website^B- tells you
U C
G T 1 1 3 ^B+Time:^B- 4 hrs 30 min
G T 1 2 1 ^C02^B+CAR^B-^C00
U L
G T 1 2 2 ^B+A multimodal website^B- tells you
U C
G T 1 2 3 ^B+ Time:^B- 5 hrs
G T 1 2 4 (including 20 min delay)
G T 1 3 1 ^C02^B+CAR^B-^C00
U L
G T 1 3 2 ^B+A multimodal website^B- tells you
U C
G T 1 3 3 ^B+ Time:^B- 6 hrs
G T 1 3 4 (including 30 min delay)
G T 1 4 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 4 2 ^B+A train website^B- tells you
U C
G T 1 4 3 ^B+ Time:^B- 3 hrs 30 min
G T 1 4 4 (scheduled station to station)
G T 1 5 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 5 2 ^B+A train website^B- tells you
U C
G T 1 5 3 ^B+ Time:^B- 3 hrs 50 min
G T 1 5 4 (scheduled station to station)
G T 1 6 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 6 2 ^B+A train website^B- tells you
U C
G T 1 6 3 ^B+ Time:^B- 4 hrs 10 min
G T 1 6 4 (scheduled station to station)
G T 2 1 1 ^B+Out of pocket cost:^B- ?60
G T 2 2 1 ^B+Out of pocket cost:^B- ?75
G T 2 3 1 ^B+Out of pocket cost:^B- ?90
G T 2 4 1 ^B+Out of pocket cost:^B- ?110
G T 2 5 1 ^B+Out of pocket cost:^B- ?90
G T 2 6 1 ^B+Out of pocket cost:^B- ?130
U L
G T 3 1 1 ^B+A friend^B- tells you
U C
G T 3 1 2 ^B+ Time:^B- 4 hrs 20 min
U L
G T 3 2 1 ^B+A friend^B- tells you
U C
G T 3 2 2 ^B+ Time:^B- 4 hrs 40 min
U L
G T 3 3 1 ^B+A friend^B- tells you
U C
G T 3 3 2 ^B+ Time:^B- 5 hrs 10 min
G T 3 3 3 (including 30 min delay)
U L
G T 3 4 1 ^B+Recent experience^B- tells you
U C
G T 3 4 2 ^B+ Time:^B- 3 hrs 40 min
G T 3 4 3 (scheduled station to station)
U L
G T 3 5 1 ^B+Recent experience^B- tells you
U C
G T 3 5 2 ^B+ Time:^B- 4 hrs
G T 3 5 3 (scheduled station to station)
U L
G T 3 6 1 ^B+Recent experience^B- tells you
U C
G T 3 6 2 ^B+ Time:^B- 4 hrs 5 min
G T 3 6 3 (scheduled station to station)
G T 4 1 1 ^B+Out of pocket cost:^B- ?50
G T 4 2 1 ^B+Out of pocket cost:^B- ?60
G T 4 3 1 ^B+Out of pocket cost:^B- ?70
G T 4 4 1 ^B+Out of pocket cost:^B- ?115
G T 4 5 1 ^B+Out of pocket cost:^B- ?95
G T 4 6 1 ^B+Out of pocket cost:^B- ?125
G C 5
G F 1 5 1 1 5
G F 1 5 2 10 14
G P < Now for choice ___, (again based on all the information provided), would you choose Car or Train? >
G X 1 CAR
G X 2 TRAIN
G R 2
G Y 1 Prefer car
G Y 2 Prefer train
G Z 1 1
G Z 2 2
G H 3
G >
U L
**
*
*
*
Q 0 SP2INTRO2B
T Now the last 4 set of choices ask whether you would choose to go to Cardiff by Coach or by Train and shows information about each of them .
T ^B+(PAUSE)^B-
T At choice 6 it shows  information from two sources about going by coach and from two sources about going by train.
T ^B+(PAUSE)^B-
T For the coach journey,  a coach website says that it will take about .... and will cost about ..... and a multimodal website says it would take about ... and costs about ......
T ^B+(PAUSE)^B-
T For the train journey, again you have 2 information sources, a train website says that it will take about .... and will cost about ..... but your recent experience is that it had taken ....... and had cost you ...........
T
T Which mode would you choose?
>
*
*
G B 2 SP3
U C
G V 4
G A 18
G L 1 6
G L 2 6
G L 3 6
G L 4 6
G W 2
* Levels to define coach
G T 1 1 1 ^C02^B+COACH^B-^C00
U L
G T 1 1 2 ^B+Coach website^B- tells you
U C
G T 1 1 3 ^B+ Time:^B- 6 hrs 30 min
```

```
G T 1 1 4 (scheduled station to station)
G T 1 2 1 ^C02^B+COACH^B-^C00
U L
G T 1 2 2 ^B+Coach website^B- tells you
U C
G T 1 2 3 ^B+ Time:^B- 7 hrs
G T 1 2 4 (scheduled station to station)
G T 1 3 1 ^C02^B+COACH^B-^C00
U L
G T 1 3 2 ^B+Coach website^B- tells you
U C
G T 1 3 3 ^B+ Time:^B- 8 hrs 30 min
G T 1 3 4 (scheduled station to station)
G T 2 1 1 ^B+Out of pocket cost:^B- ?50
G T 2 2 1 ^B+Out of pocket cost:^B- ?60
G T 2 3 1 ^B+Out of pocket cost:^B- ?70
U L
G T 3 1 1 ^B+A multimodal website^B- tells you
U C
G T 3 1 2 ^B+ Time:^B- 6 hrs 35 min
G T 3 1 3 (scheduled station to station)
U L
G T 3 2 1 ^B+A multimodal website^B- tells you
U C
G T 3 2 2 ^B+ Time:^B- 7 hrs 30 min
G T 3 2 3 (scheduled station to station)
U L
G T 3 3 1 ^B+A multimodal websiteB- tells you
U C
G T 3 3 2 ^B+ Time:^B- 8 hrs 20 min
G T 3 3 3 (scheduled station to station)
G T 4 1 1 ^B+Out of pocket cost:^B- ?55
G T 4 2 1 ^B+Out of pocket cost:^B- ?62
G T 4 3 1 ^B+Out of pocket cost:^B- ?70
* Levels to define train
G T 1 4 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 4 2 ^B+Train website^B- tells you
U C
G T 1 4 3 ^B+ Time:^B- 3 hrs 30 min
G T 1 4 4 (scheduled station to station)
G T 1 5 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 5 2 ^B+Train website^B- tells you
U C
G T 1 5 3 ^B+ Time:^B- 3 hrs 50 min
G T 1 5 4 (scheduled station to station)
G T 1 6 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 6 2 ^B+Train website^B- tells you
U C
G T 1 6 3 ^B+ Time:^B- 4 hrs 10 min
G T 1 6 4 (scheduled station to station)
G T 2 4 1 ^B+Out of pocket cost:^B- ?110
G T 2 5 1 ^B+Out of pocket cost:^B- ?90
G T 2 6 1 ^B+Out of pocket cost:^B- ?130
U L
G T 3 4 1 ^B+Recent experience^B- tells you
U C
G T 3 4 2 ^B+ Time:^B- 3 hrs 40 min
G T 3 4 3 (scheduled station to station)
U L
G T 3 5 1 ^B+Recent experience^B- tells you
U C
G T 3 5 2 ^B+ Time:^B- 4 hrs
G T 3 5 3 (scheduled station to station)
U L
G T 3 6 1 ^B+Recent experience^B- tells you
U C
G T 3 6 2 ^B+ Time:^B- 4 hrs 5 min
G T 3 6 3 (scheduled station to station)
G T 4 4 1 ^B+Out of pocket cost:^B- ?115
G T 4 5 1 ^B+Out of pocket cost:^B- ?95
G T 4 6 1 ^B+Out of pocket cost:^B- ?125
G C 4
G F 1 4 1 6 9
G F 1 4 2 15 18
G P < Now for choice ___, (again based on all the
information provided), would you choose Coach or Train? >
G X 1 COAH
G X 2 TRAIN
G R 2
G Y 1 Prefer coach
G Y 2 Prefer train
G Z 1 1
G Z 2 2
G H 3
G >
U L
*
I #START# GT 0 INTROPART3
*
*
Q 0 SP2INTRO3
T Each choice asks whether you would choose to go to
Cardiff by Coach or by Train and shows information about
each of them .
T ^B+(PAUSE)^B-
T At choice 1 it shows information from two sources about
going by coach and from two sources about going by train.
T ^B+(PAUSE)^B-
T For the coach journey, a coach website says that it will
take about .... and will cost about ..... and a multimodal
website says it would take about ... and costs about ......
T ^B+(PAUSE)^B-
T For the train journey, again you have 2 information
sources, a train website says that it will take about .... and
will cost about ..... but your recent experience is that it had
taken ....... and had cost you ..........
T
T Which mode would you choose?
>
*
G B 2 SP4
U C
G V 4
G A 18
G L 1 6
G L 2 6
G L 3 6
G L 4 6
G W 2
* Levels to define coach
G T 1 1 1 ^C02^B+COACH^B-^C00
U L
G T 1 1 2 ^B+Coach website^B- tells you
U C
G T 1 1 3 ^B+ Time:^B- 6 hrs 30 min
G T 1 1 4 (scheduled station to station)
G T 1 2 1 ^C02^B+COACH^B-^C00
U L
G T 1 2 2 ^B+Coach website^B- tells you
U C
G T 1 2 3 ^B+ Time:^B- 7 hrs
G T 1 2 4 (scheduled station to station)
G T 1 3 1 ^C02^B+COACH^B-^C00
U L
G T 1 3 2 ^B+Coach website^B- tells you
U C
G T 1 3 3 ^B+ Time:^B- 8 hrs 30 min
G T 1 3 4 (scheduled station to station)
G T 2 1 1 ^B+Out of pocket cost:^B- ?50
G T 2 2 1 ^B+Out of pocket cost:^B- ?60
G T 2 3 1 ^B+Out of pocket cost:^B- ?70
U L
G T 3 1 1 ^B+A multimodal website^B- tells you
U C
G T 3 1 2 ^B+ Time:^B- 6 hrs 35 min
G T 3 1 3 (scheduled station to station)
U L
G T 3 2 1 ^B+A multimodal website^B- tells you
U C
```

G T 3 2 2 ^B+ Time:^B- 7 hrs 30 min
G T 3 2 3 (scheduled station to station)
U L
G T 3 3 1 ^B+A multimodal websiteB- tells you
U C
G T 3 3 2 ^B+ Time:^B- 8 hrs 20 min
G T 3 3 3 (scheduled station to station)
G T 4 1 1 ^B+Out of pocket cost:^B- ?55
G T 4 2 1 ^B+Out of pocket cost:^B- ?62
G T 4 3 1 ^B+Out of pocket cost:^B- ?70
* Levels to define train
G T 1 4 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 4 2 ^B+Train website^B- tells you
U C
G T 1 4 3 ^B+ Time:^B- 3 hrs 30 min
G T 1 4 4 (scheduled station to station)
G T 1 5 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 5 2 ^B+Train website^B- tells you
U C
G T 1 5 3 ^B+ Time:^B- 3 hrs 50 min
G T 1 5 4 (scheduled station to station)
G T 1 6 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 6 2 ^B+Train website^B- tells you
U C
G T 1 6 3 ^B+ Time:^B- 4 hrs 10 min
G T 1 6 4 (scheduled station to station)
G T 2 4 1 ^B+Out of pocket cost:^B- ?110
G T 2 5 1 ^B+Out of pocket cost:^B- ?90
G T 2 6 1 ^B+Out of pocket cost:^B- ?130
U L
G T 3 4 1 ^B+Recent experience^B- tells you
U C
G T 3 4 2 ^B+ Time:^B- 3 hrs 40 min
G T 3 4 3 (scheduled station to station)
U L
G T 3 5 1 ^B+Recent experience^B- tells you
U C
G T 3 5 2 ^B+ Time:^B- 4 hrs
G T 3 5 3 (scheduled station to station)
U L
G T 3 6 1 ^B+Recent experience^B- tells you
U C
G T 3 6 2 ^B+ Time:^B- 4 hrs 5 min
G T 3 6 3 (scheduled station to station)
G T 4 4 1 ^B+Out of pocket cost:^B- ?115
G T 4 5 1 ^B+Out of pocket cost:^B- ?95
G T 4 6 1 ^B+Out of pocket cost:^B- ?125
G C 9
G F 1 9 1 1 9
G F 1 9 2 10 18
G P < Now for choice ___, (again based on all the information provided), would you choose Coach or Train?>
G X 1 COAH
G X 2 TRAIN
G R 2
G Y 1 Prefer coach
G Y 2 Prefer train
G Z 1 1
G Z 2 2
G H 3
G >
U L
*
I #START# GT 0 INTROPART3
*
Q 0 SP2INTRO4
T Each choice asks whether you would choose to go to Cardiff by Car or by Train and shows information about each of them .
T ^B+(PAUSE)^B-
T At choice 1 it shows information from two sources about going by car and from two sources about going by train.
T ^B+(PAUSE)^B-
T For the car journey, a friend says that it will take about .... and will cost about ..... and a multimodal website says it would take about ... and costs about ......
T ^B+(PAUSE)^B-
T For the train journey, again you have 2 information sources, a train website says that it will take about .... and will cost about ..... but your recent experience is that it had taken .... and had cost you ....
T
T Which mode would you choose?
>
*
*
G B 2 SP5
U C
G V 4
G A 18
G L 1 6
G L 2 6
G L 3 6
G L 4 6
G W 2
G T 1 1 1 ^C02^B+CAR^B-^C00
U L
G T 1 1 2 ^B+A multimodal website^B- tells you
U C
G T 1 1 3 ^B+Time:^B- 4 hrs 30 min
G T 1 2 1 ^C02^B+CAR^B-^C00
U L
G T 1 2 2 ^B+A multimodal website^B- tells you
U C
G T 1 2 3 ^B+ Time:^B- 5 hrs
G T 1 2 4 (including 20 min delay)
G T 1 3 1 ^C02^B+CAR^B-^C00
U L
G T 1 3 2 ^B+A multimodal website^B- tells you
U C
G T 1 3 3 ^B+ Time:^B- 6 hrs
G T 1 3 4 (including 30 min delay)
G T 1 4 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 4 2 ^B+A train website^B- tells you
U C
G T 1 4 3 ^B+ Time:^B- 3 hrs 30 min
G T 1 4 4 (scheduled station to station)
G T 1 5 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 5 2 ^B+A train website^B- tells you
U C
G T 1 5 3 ^B+ Time:^B- 3 hrs 50 min
G T 1 5 4 (scheduled station to station)
G T 1 6 1 ^C02^B+TRAIN^B-^C00
U L
G T 1 6 2 ^B+A train website^B- tells you
U C
G T 1 6 3 ^B+ Time:^B- 4 hrs 10 min
G T 1 6 4 (scheduled station to station)
G T 2 1 1 ^B+Out of pocket cost:^B- ?60
G T 2 2 1 ^B+Out of pocket cost:^B- ?75
G T 2 3 1 ^B+Out of pocket cost:^B- ?90
G T 2 4 1 ^B+Out of pocket cost:^B- ?110
G T 2 5 1 ^B+Out of pocket cost:^B- ?90
G T 2 6 1 ^B+Out of pocket cost:^B- ?130
U L
G T 3 1 1 ^B+A friend^B- tells you
U C
G T 3 1 2 ^B+ Time:^B- 4 hrs 20 min
U L
G T 3 2 1 ^B+A friend^B- tells you
U C
G T 3 2 2 ^B+ Time:^B- 4 hrs 40 min
U L
G T 3 3 1 ^B+A friend^B- tells you
U C
G T 3 3 2 ^B+ Time:^B- 5 hrs 10 min
G T 3 3 3 (including 30 min delay)

U L
G T 3 4 1 ^B+Recent experience^B- tells you
U C
G T 3 4 2 ^B+ Time:^B- 3 hrs 40 min
G T 3 4 3 (scheduled station to station)
U L
G T 3 5 1 ^B+Recent experience^B- tells you
U C
G T 3 5 2 ^B+ Time:^B- 4 hrs
G T 3 5 3 (scheduled station to station)
U L
G T 3 6 1 ^B+Recent experience^B- tells you
U C
G T 3 6 2 ^B+ Time:^B- 4 hrs 5 min
G T 3 6 3 (scheduled station to station)
G T 4 1 1 ^B+Out of pocket cost:^B- ?50
G T 4 2 1 ^B+Out of pocket cost:^B- ?60
G T 4 3 1 ^B+Out of pocket cost:^B- ?70
G T 4 4 1 ^B+Out of pocket cost:^B- ?115
G T 4 5 1 ^B+Out of pocket cost:^B- ?95
G T 4 6 1 ^B+Out of pocket cost:^B- ?125
G C 9
G F 1 9 1 1 9
G F 1 9 2 10 18
G P < Now for choice ___, (again based on all the information provided), would you choose Car or Train? >
G X 1 CAR
G X 2 TRAIN
G R 2
G Y 1 Prefer car
G Y 2 Prefer train
G Z 1 1
G Z 2 2
G H 3
G >
U L
*
I #START# GT 0 INTROPART3
*
Q 0 SP2INTRO5
T Each choice asks whether you would choose to go to Cardiff by Car or by Coach and shows information about each of them .
T ^B+(PAUSE)^B-
T At choice 1 it shows information from two sources about going by car and from two sources about going by coach.
T ^B+(PAUSE)^B-
T For the car journey, a friend says that it will take about .... and will cost about ..... and a multimodal website says it would take about ... and costs about ......
T ^B+(PAUSE)^B-
T For the coach journey, a coach website says that it will take about .... and will cost about ..... and a multimodal website says it would take about ... and costs about ......
T
T Which mode would you choose?
>
*
*
G B 2 SP6
U C
G V 4
G A 18
G L 1 6
G L 2 6
G L 3 6
G L 4 6
G W 2
G T 1 1 1 ^C02^B+CAR^B-^C00
U L
G T 1 1 2 ^B+A multimodal website^B- tells you
U C
G T 1 1 3 ^B+Time:^B- 4 hrs 30 min
G T 1 2 1 ^C02^B+CAR^B-^C00
U L
G T 1 2 2 ^B+A multimodal website^B- tells you
U C
G T 1 2 3 ^B+ Time:^B- 5 hrs
G T 1 2 4 (including 20 min delay)
G T 1 3 1 ^C02^B+CAR^B-^C00
U L
G T 1 3 2 ^B+A multimodal website^B- tells you
U C
G T 1 3 3 ^B+ Time:^B- 6 hrs
G T 1 3 4 (including 30 min delay)
G T 1 4 1 ^C02^B+COACH^B-^C00
U L
G T 1 4 2 ^B+A coach website^B- tells you
U C
G T 1 4 3 ^B+ Time:^B- 6 hrs 30 min
G T 1 4 4 (scheduled station to station)
G T 1 5 1 ^C02^B+COACH^B-^C00
U L
G T 1 5 2 ^B+A coach website^B- tells you
U C
G T 1 5 3 ^B+ Time:^B- 7 hrs
G T 1 5 4 (scheduled station to station)
G T 1 6 1 ^C02^B+COACH^B-^C00
U L
G T 1 6 2 ^B+A coach website^B- tells you
U C
G T 1 6 3 ^B+ Time:^B- 8 hrs 30 min
G T 1 6 4 (scheduled station to station)
G T 2 1 1 ^B+Out of pocket cost:^B- ?60
G T 2 2 1 ^B+Out of pocket cost:^B- ?75
G T 2 3 1 ^B+Out of pocket cost:^B- ?90
G T 2 4 1 ^B+Out of pocket cost:^B- ?50
G T 2 5 1 ^B+Out of pocket cost:^B- ?60
G T 2 6 1 ^B+Out of pocket cost:^B- ?70
U L
G T 3 1 1 ^B+A friend^B- tells you
U C
G T 3 1 2 ^B+ Time:^B- 4 hrs 20 min
U L
G T 3 2 1 ^B+A friend^B- tells you
U C
G T 3 2 2 ^B+ Time:^B- 4 hrs 40 min
U L
G T 3 3 1 ^B+A friend^B- tells you
U C
G T 3 3 2 ^B+ Time:^B- 5 hrs 10 min
G T 3 3 3 (including 30 min delay)
U L
G T 3 4 1 ^B+A multimodal website^B- tells you
U C
G T 3 4 2 ^B+ Time:^B- 6 hrs 35 min
G T 3 4 3 (scheduled station to station)
U L
G T 3 5 1 ^B+A multimodal website^B- tells you
U C
G T 3 5 2 ^B+ Time:^B- 7 hrs 30 min
G T 3 5 3 (scheduled station to station)
U L
G T 3 6 1 ^B+A multimodal website^B- tells you
U C
G T 3 6 2 ^B+ Time:^B- 8 hrs 20 min
G T 3 6 3 (scheduled station to station)
G T 4 1 1 ^B+Out of pocket cost:^B- ?50
G T 4 2 1 ^B+Out of pocket cost:^B- ?60
G T 4 3 1 ^B+Out of pocket cost:^B- ?70
G T 4 4 1 ^B+Out of pocket cost:^B- ?55
G T 4 5 1 ^B+Out of pocket cost:^B- ?62
G T 4 6 1 ^B+Out of pocket cost:^B- ?70
G C 9
G F 1 9 1 1 9
G F 1 9 2 10 18
G P < Now for choice ___(again based on all the information provided), would you choose Car or Coach? >
G X 1 CAR
G X 2 COACH
G R 2
G Y 1 Prefer car

G Y 2 Prefer coach
G Z 1 1
G Z 2 2
G H 3
G >
U L
*
****
*
*Part 3: General Background
*
****
*
*
P
S Part 3: General Background
*
Q 0 INTROPART3
T Thank you, that's the worst part over!
T
T I would now like to ask you about your real life use of information sources
>
*
*Q 12 USEDSOURCES
*T Which of the following sources have you ever used?
*T
*T Please select all that apply from the sources shown below.
*T
*T 1. advice from a friend
*T 2. map(s)
*T 3. a paper timetable
*T 4. a telephone enquiry line
*T 5. website(s) providing information about ^B+only^B-car journeys and road conditions
*T 6. website(s) providing information about ^B+both^B-public transport and car journeys
*T 7. website(s) providing information about ^B+only^B-coach journeys/services
*T 8. website(s) providing information about ^B+only^B-train journeys/services
*T 9. SMS alert(s)
*T 10. radio bulletin or TV tele-text
*T 11. other (please specify)
*A advice from a friend
*A map(s)
*A a paper timetable
*A a telephone enquiry line
*A website-car journeys and road conditions
*A website-public transport and car journeys
*A website-coach journeys/services only
*A website-train journeys/services only
*A SMS alert(s)
*A radio bulletin or TV tele-text
*O other (please specify)
*>
*
*Q 12 BADWSOURCES
*T Which sources would you use if bad weather is expected?
*T
*T Please select all that apply from the sources shown below.
*T
*T 1. advice from a friend
*T 2. map(s)
*T 3. a paper timetable
*T 4. a telephone enquiry line
*T 5. website(s) providing information about ^B+only^B-car journeys and road conditions
*T 6. website(s) providing information about ^B+both^B-public transport and car journeys
*T 7. website(s) providing information about ^B+only^B-coach journeys/services
*T 8. website(s) providing information about ^B+only^B-train journeys/services
*T 9. SMS alert(s)
*T 10. radio bulletin or TV tele-text
*T 11. other (please specify)
*A advice from a friend
*A map(s)
*A a paper timetable
*A a telephone enquiry line
*A website-car journeys and road conditions
*A website-public transport and car journeys
*A website-coach journeys/services only
*A website-train journeys/services only
*A SMS alert(s)
*A radio bulletin or TV tele-text
*O other (please specify)
*>
*
*Q 12 INCISOURCES
*T Which source would you use if you learn that there has been an incident on your intended route?
*T
*T Please select all that apply from the sources shown below.
*T
*T 1. advice from a friend
*T 2. map(s)
*T 3. a paper timetable
*T 4. a telephone enquiry line
*T 5. website(s) providing information about ^B+only^B-car journeys and road conditions
*T 6. website(s) providing information about ^B+both^B-public transport and car journeys
*T 7. website(s) providing information about ^B+only^B-coach journeys/services
*T 8. website(s) providing information about ^B+only^B-train journeys/services
*T 9. SMS alert(s)
*T 10. radio bulletin or TV tele-text
*T 11. other (please specify)
*A advice from a friend
*A map(s)
*A a paper timetable
*A a telephone enquiry line
*A website-car journeys and road conditions
*A website-public transport and car journeys
*A website-coach journeys/services only
*A website-train journeys/services only
*A SMS alert(s)
*A radio bulletin or TV tele-text
*O other (please specify)
*>
*
Q 12 SUBSCRIPTION
T Which, if any, of the following sources do you subscribe to?
T
T 1. a telephone enquiry line
T 2. website(s)
T 3. SMS alert(s)
T 4. Any other (please specify)
T 5. none
A a telephone enquiry line
A website(s)
A SMS alert(s)
O other (please specify)
A none
>
*
J #SUBSCRIPTION8# EQ 1 2
I #SUBSCRIPTION2# NE 1 NAMESUBS
*
Q 12 SUBS1
T which website?
T

T 1. website(s) providing information about ^B+only^B- car journeys and road conditions
T 2. website(s) providing information about ^B+both^B- public transport and car journeys
T 3. website(s) providing information about ^B+only^B- coach journeys/services
T 4. website(s) providing information about ^B+only^B- train journeys/services
A website-car journeys and road conditions
A website-public transport and car journeys
A website-coach journeys/services only
A website-train journeys/services only
>
*
Q 2 NAMESUBS
T Can you name the website(s)?
>
*
Q 1 WEBSITEFREQ
T For what proportion of your long journeys do you use websites as a source of travel information?
T
T
T 1. all of them
T 2. some of them
T 3. none of them
A all of them
A some of them
A none of them
>
*
J #WEBSITEFREQ# LE 3 1
*
Q 12 REASNWEB
T Which of the following reasons do you have for not using websites as a source of travel information...
T
T 1. You don't know of any websites ^B+(pause)^B-
T 2. you have sufficient knowledge about your journeys and so you have no need for extra information ^B+(pause)^B-
T 3. you have used them before but found them confusing ^B+(pause)^B-
T 4. you have used them before found them time consuming ^B+(pause)^B-
T 5. you have used them before but thought the information provided was ambiguous and inconsistent ^B+(pause)^B-
T 6. your journeys are organised for you (e.g. by tour operator or by your office) ^B+(pause)^B-
T 7. any other reason? ^B+(ask for details)^B-
A I didn't know about any of the websites
A I know my journey thoroughly so had no need for extra information
A I used them before but didn't like it (confusing)
A I used them before but didn't like it (time consuming)
A information provided was ambiguous and inconsistent
A my journey was organised for me (e.g. by tour operator/office)
O any other, please mention
>
*
I #REASNWEB# GT 0 WILLINGTPAY
*
Q 1 FINDWEB
T If you use website(s) for travel information, have you designated them as "^C09favourites^C00" in your computer?
T
T 1. Yes
T 2. No
A yes
A no
>
*
I #FINDWEB# EQ 1 WILLINGTPAY
*
Q 12 FIND2
T Which method do you use to access the website.... Do you....
T
T type in the website address
T type a source name (such as "national express") into a search engine (like Google)
T or type a generic phrase (e.g. "travel information" or "time table") into a search engine?
A I type in the website address
A I type a source name into a search engine (e.g. Google)
A I type a generic phrase (e.g. travel information, time table, national express etc)
>
*
I #FIND23# EQ 1 GPHRASE
*
Q 0 DUMMY3
I #FIND21# EQ 1 OR #FIND22# EQ 1 WILLINGTPAY
F 0 #DUMMY3#
*
Q 2 GPHRASE
T What generic phrases would you use for the journey to Cardiff mentioned earlier?
T
T ^B+(Use comma between different phrases if more than one)^B-
>
*
Q 4 WILLINGTPAY
T How much a month would you be prepared to spend on subscribing to a website which provides you with accurate & _and up-to-date information about..., costs, travel time, expected delays, current speeds, and recommended & _departure times?
T
T ^B+("PAUSE")^B-
T
T How much would you pay a month for such a website?
L 1
H 20000
>
*
Q 3 EMPBUISNESS
T What proportion of the trips that you make over 50 miles are on your employer's business?
T
T ("PAUSE")
T
T 1. all
T 2. most
T 3. some
T 4. few
T 5. none
A all
A most
A some
A few
A none
>
*
Q 1 ADVERTWEB
T How important to you is it that a website should be free from irrelevant advertisements and pop ups?
T
T ^B+("PAUSE")^B-
T
T 1. essential
T 2. important
T 3. not important
T 4. irrelevant
A essential
A important
A not important
A irrelevant
>
*

Q 10 REALTIME
T How important to you is it that the source of information provides real-time information (with up-to-the-minute
& _information about delays, congestion, incidents etc)?
T
T ^B+("PAUSE")^B-
T
T 1. essential
T 2. important
T 3. not important
T 4. irrelevant
>
*
Q 10 GETSINFO
T How important to you is it to know where the information source gets its information from?
T
T
T ^B+("PAUSE")^B-
T
T 1. essential
T 2. important
T 3. not important
T 4. irrelevant
>
*
Q 10 SAFETY
T When you are choosing a mode, how important to you are safety aspects of the modes?
T
T
T ^B+("PAUSE")^B-
T
T 1. essential
T 2. important
T 3. not important
T 4. irrelevant
>
*
Q 10 COMFORT
T When you are choosing a mode, how important to you is the ^C09comfort^C00 of each mode?
T
T
T ^B+("PAUSE")^B-
T
T 1. essential
T 2. important
T 3. not important
T 4. irrelevant
>
*
Q 10 SEATS
T When you are choosing a public transport mode, how important to you is it the ^C09availability of seats^C00?
T
T
T ^B+("PAUSE")^B-
T
T 1. essential
T 2. important
T 3. not important
T 4. irrelevant
>
*
Q 10 OPTIMIZE
T When you are choosing a mode, how important to you is the ^C09travel time^C00 by each mode?
T
T
T ^B+("PAUSE")^B-
T
T 1. essential
T 2. important
T 3. not important
T 4. irrelevant
>
*
Q 10 MINIMIZEC
T When you are choosing a mode, how important to you is the ^C09cost^C00 of each mode?
T
T
T ^B+("PAUSE")^B-
T
T 1. essential
T 2. important
T 3. not important
T 4. irrelevant
>
*
Q 1 DISCOVERING
T When you are choosing how to travel, would you say that....
T
T
T 1. you like to test out new options
T 2. or that you prefer to stick with what you know?
A you like to test out new options
A you prefer to stick with what you know
>
*
****
*
*Part 4: Personal Information
*
****
*
*
P
S Part 4: Closing Questions
*
Q 0 INTROPART4
T Finally, I would like to ask a few questions about yourself and your household.
>
*
Q 1 AGE
T Please look at card 6, How old are you?
T
T ^B+(do not prompt unless necessary)^B-
T
T 1. less than 20
T 2. 20 to 29
T 3. 30 to 39
T 4. 40 to 49
T 5. 50 to 64
T 6. 65 and over
A less than 20
A 20 to 29
A 30 to 39
A 40 to 49
A 50 to 64
A 65 and over
>
*
Q 3 EDUCATION
T At what age did you leave full time education? (in years)
L 13
H 40
>
*
Q 1 EMPLOYMENT
T Are you employed for 20 hours per week or more?
T
T 1. yes
T 2. no
A yes
A no
>
*
J #EMPLOYMENT# EQ 1 1

```
*
Q 1 OCCUPATION
T Are you then ......
T
T (read out the options)
T
T 1. employed part time
T 2. retired
T 3. a student
T 4. looking after home/housewife
T 5. without work
T 6. disabled
T 7. other (please specify)
A employed part time
A retired
A a student
A looking after home/housewife
A without work
A disabled
O other (please specify)
>
*
J #OCCUPATION# GT 0 1
*
Q 1 EMPL1
T Are you .....
T
T (read out the options)
T
T 1. an employee
T 2. self employment
A employee
A self employment
>
*
Q 1 CARS
T How many vehicles (cars, company vehicles, motorcycles,
vans or trucks) are available
& _to members of your household?
T
T 1. none
T 2. one
T 3. two or more
A none
A one
A two or more
>
*
Q 1 SIZEOFHH1
T How many people, aged 17 or older, live in your
household ..including yourself?
T
T 1.   1
T 2.   2
T 3.   3
T 4.   4
T 5.  more
A 1
A 2
A 3
A 4
A more
>
*
Q 1 SIZEOFHH2
T Are there any people from 6 to 17 in your household?
T
T 1.  Yes
T 2.  No
A Yes
A No
>
*
Q 1 SIZEOFHH3
T Are there any people under 6 in your household?
T
T 1.  Yes
T 2.  No
A Yes
A No
>
*
Q 1 INCOME
T Please look at card 7, what is your household annual
income before tax?
T
T ^B+(do not prompt unless necessary)^B-
T
T h.  £10,000 or less
T i.  £10,001 to £20,000
T j.  £20,001 to £30,000
T k.  £30,001 to £40,000
T l.  £40,001 to £50,000
T m.  £50,001 to £60,000
T n.  over £60,000
A £10,000 or less
A £10,001 to £20,000
A £20,001 to £30,000
A £30,001 to £40,000
A £40,001 to £50,000
A £50,001 to £60,000
A over £60,000
>
*
Q 1 REMARKS
T Thank you for your time. That was the last question of the
survey
T
T Do you have any further remarks about your travel or
about this interview?
O yes
A no
>
*
```

Appendix C
Variable Used in the Models

Variables used in the models

$Time_n$	*(Generic, in Minutes)*
$Cost_n$	*(Generic, in Pence)*
$DCar_n$	*(Dummy, if car is available in household = 1, otherwise = 0)*
$DBPur_n$	*(Dummy, Business Purpose = 1, otherwise = 0)*
$DLFreq_n$	*(Dummy, Trip Frequency less than 13/year =1, otherwise = 0)*
$DGWeather_n$	*(Dummy, Good weather =1, otherwise = 0)*
$DIncident_n$	*(Dummy, Incident occurred =1, otherwise = 0)*
$DPeakP_n$	*(Dummy, Travelled in the Peak Period = 1, otherwise = 0)*
$DReasMode_n$	*(Dummy, Reason for selecting mode is Time/Cost = 1, otherwise = 0)*
$DTravAlone_n$	*(Dummy, Travelling alone = 1, otherwise = 0)*
$DImpSfty_n$	*(Dummy, Safety important = 1, otherwise = 0)*
$DImpCmfrt_n$	*(Dummy, Comfort important = 1, otherwise = 0)*
$DImpSeats_n$	*(Dummy, Seat availability important = 1, otherwise = 0)*
$DMale_n$	*(Dummy, If male =1, otherwise = 0)*
$DEduc_n$	*(Dummy, Left full time education at or after 20 = 1, otherwise = 0)*
$DEduc25_n$	*(Dummy, Left full time education at or after 25 = 1, otherwise = 0)*
$DFEmpl_n$	*(Dummy, Full time employed =1, otherwise = 0)*
$DIncome_n$	*(Dummy, If income over £30,000 =1, otherwise = 0)*
$DAge_n$	*(Dummy, If Age less than 50 =1, otherwise = 0)*
$DSubs_n$	*(Dummy, Subscribed to a website = 1, otherwise = 0)*
$DImpAd_n$	*(Dummy, free from Advertisements important = 1, otherwise = 0)*
$DImpRealt_n$	*(Dummy, Real time information important = 1, otherwise = 0)*
$DImpGInfo_n$	*(Dummy, General information important = 1, otherwise = 0)*
$Carfrit_n$	*(Time by car information by friend, in Minutes)*
$Carmult_n$	*(Time by car information by multimodal website, in Pence)*
$Carfric_n$	*(Cost by car information by friend, in Minutes)*
$Carmultc_n$	*(Cost by car information by multimodal website, in Pence)*
$Trainwt_n$	*(Time by train information by train website, in Minutes)*
$Trainext_n$	*(Time by train information by past experience, in Minutes)*
$Trainwc_n$	*(Cost by train information by train website, in Pence)*
$Trainexc_n$	*(Cost by train information by past experience, in Pences)*
$Coawt_n$	*(Time by coach information by coach website, in Minutes)*
$Coawc_n$	*(Cost by coach information by coach website, in Pences)*
$Coamult_n$	*(Time by coach information by multimodal website, in Minutes)*
$Coamulc_n$	*(Cost by coach information by multimodal website, in Pence)*
$DComAd_n$:	*(dummy that equals 1 if Commercial Ads No Sub; otherwise 0)*
$DComAdSub_n$:	*(dummy that equals 1 if Commercial Ads Subs; otherwise 0)*
$DCom_n$:	*(dummy that equals 1 if Commercial No Ads No Sub; otherwise 0)*
$Seatime_n$	*(5 min, 10 min, 15 min)*
$DUpdaily_n$:	*(dummy that equals 1 if Website updates daily; otherwise 0)*
$DUpweekly_n$:	*(dummy that equals 1 if Website updates weekly; otherwise 0)*
$DInfoSpec_n$:	*(dummy that equals 1 if Specific info available; otherwise 0)*
$DInfoPerc_n$:	*(dummy that equals 1 if Info w.r.t own criteria; otherwise 0)*
$DYoung_n$	*(Dummy, If Age less than 40 =1, otherwise = 0)*