Delivering Sustainable Mobility Targets: Investigating the characteristics of car users most likely to switch to sustainable modes

SCHOOL OF ENGINEERING

NEWCASTLE UNIVERSITY

By

Fazilatulaili Ali

A thesis submitted for the degree of

Doctor of Philosophy in Transport at

Newcastle University

June 2019

Declaration

This thesis is the original work of the author except where acknowledgement has been given. The material presented has never been submitted to the Newcastle University or to any other educational establishment for purposes of obtaining a higher degree.

June 2019 Fazilatulaili Ali

Abstract

Dependence on car use generates adverse effects on the sustainability of our urban environment due to air and noise pollution and global warming. Along with technological and infrastructure improvements, an overall solution to this problem would require dramatic change to the way we chose to travel. Behavioural change of car users to have them move to other sustainable travel options making less impact to our environment is crucial. Behavioural change of car users is not straightforward given the huge range of factors governing their travel decisions. Local authorities with constrained budgets require evidence to help them introduce policies that will have significant impact on car use. Therefore, we need to understand which specific groups of car users are more inclined to change to sustainable modes so that ultimately targets can be met.

The global aim of this research is to characterise target groups of car users (as a driver or passenger) who are more likely to switch from private transport to sustainable modes. A comprehensive analysis of the British Social Attitudes (BSA) survey data collected during 2011 to 2014 was used in this research. The research identified: *important attitudinal factors through dimension reduction analysis; groups of car users' travel mode choice decisions based on socio-demographic variables; relationships between car users' groups and attitudinal factors which relate to modal shift potential and finally deriving a mathematical model to predict the likely uptake of sustainable modes.* The statistical analysis techniques included Descriptive Analysis (DA), Factor Analysis (FA), Cluster Analysis (CA), Multiple Correspondence Analysis (MCA), and Multinomial Logistic Regression (MLR) coupled with a Multivariate Probit Model (MPM). When developing MPM, Bayesian inference was taken into consideration because it allows the uncertainty in the parameters to be incorporated in the model. Finally, the MPM was used to investigate the relationships between the responses of car users to the different attitudinal questions.

The FA is particularly useful to reduce a wide range of variables into a smaller number of factors and three main factors labelled as *Attitudes to transport and the environment; Traffic awareness;* and *Modal shift potential* emerged from 14 attitudinal variables. The CA is a method to group the car users based on their socio-demographics and travel related variables. Five clusters emerged namely *middle–aged (35-44), female, full–time*

employee (Cluster 1)*; middle–aged (35-44), male, full–time employee* (Cluster 2)*; mature adults (45-54), male, full–time employee* (Cluster 3)*; older–aged (65+), male, retired* (Cluster 4)*; and middle–aged (35-44), female, looking after the home* (Cluster 5). Whilst the majority of respondents are strongly car-orientated either as a driver or as a passenger, the car users associated with Cluster 2 were found to be more likely to cycle once a week already (29%) and travel by train less often than once per month (39%) compared to other car users in other clusters.

The MLR investigates the relationships between the factors and the clusters exploring the details of the change of attitudes over the years, for instance from 2011 to 2014. The outcome of the results show that Cluster 2 has considerably higher environmental awareness compared to other groups of respondents. Therefore, this group is likely to have potential to switch travel modes. The MPM was developed in this study specifically for ordinal responses and enabled responses to several questions, which can be correlated to be considered in a single model. This approach is different to MLR, which does not consider these correlations.

The MPM suggested that younger and older cohorts are the least likely to be susceptible to change whilst the middle-aged population is more likely to mode shift to cycling or public transport. However, the reverse is true for respondents with larger household size. The willingness to switch from the car to walking and cycling for short journeys of less than 2 miles appears to increase depending on the increasing number of people living in the household. Females will be less likely to switch mode from cars to cycling for short journeys. However, switching to walking and going by bus were more or less equally acceptable for both males and females. In addition, there was a greater tendency to agree to use cars with lower CO_2 emissions for the sake of the environment among respondents with one car per household compared to respondents with four or more. This research demonstrated that fitting MPM using Bayesian inference is both a practical and effective way to analyse ordinal survey data and is a novel aspect in this study.

Acknowledgement

I express my deepest gratitude to all who have contributed to the completion of my thesis. Among few cited key persons, my supervisors are the most remarkable. I have been very fortunate to have supervisors who consistently support me with their encouragement, helpful comments and feedback and tireless effort during all stages of my study. Truly, I owe my sincere appreciation to my first supervisor, Dr Dilum Dissanayake. My truthful gratitude is also addressed to my second and third supervisors, Professor Margaret Bell and Dr Malcolm Farrow who have supported me tremendously.

On top of that, I was delighted to interact with both of my panel examiners, Dr Neil Thorpe and Dr Anil Namdeo, without their input and suggestions my thesis would not be the same. I extend my appreciation to all staff of School of Engineering: Melissa Ware, Emma Simblet, Susan Earley and others. I also would like to acknowledge of helping hands to the members of Transport Operation Research Group (TORG) especially my PhD colleagues: Jack, Jamie, Nurten, ShuoLi, Saad, Shadman and Raymond for their constructive comments and suggestions concerning my research as well as our memories together sharing our difficulties during this journey.

Special thanks are due to the Ministry of Higher Education Malaysia (MOHE) and my employer National Defence University Malaysia (NDUM) for the financial assistance, without which I will not be able to even start my study.

Above all, I would like to express my warmest gratitude to my parents–Ali bin Ismail and Fatimah Hj Draman, my siblings and in-laws, my husband–Mohd Anif, my lovely children–Syafiqah, Syafirah and Saifullah, and all my friends especially Malaysian community in Newcastle, who stood beside me and supported me with their patience, motivations and prayers.

To enumerate all those who have contributed variously to the completion of this study is an impossible task. Failure to recognise, however, would be an injustice on my side. Indeed, for all that they have done. Thank you! ☺

And all the praises and thanks be to Allah, the Lord of the Universe.
"He Who taught (the use of) the pen".
"Taught man that which he knew not
(The Noble Qur-an, Al Alaq: 4-5)

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BSA	British social attitudes
CC	Climate change
CO_2	Carbon dioxides
EFA	Exploratory factor analysis
GDP	Gross domestic product
GHG	Greenhouse gas
JAGS	Just another Gibbs Sampler
LAs	Local authorities
MCA	Multiple correspondence analysis
MCMC	Markov Chain Monte Carlo
MLR	Multinomial logistic regression
MSP	Modal shift potential
MPM	Multivariate probit models
PCA	Principle component analysis
PAF	Principal axis factoring
TDM	Transport demand management
VKT	Vehicle kilometres travelled

Publications and conferences

Publications

1. F. Ali, D. Dissanayake, M. Bell, and M. Farrow (2018) 'Investigating Car Users' Attitudes to Climate Change Using Multiple Correspondence Analysis', *Journal of Transport Geography*, 72, pp. 237-247.

2. J. Douglass, D. Dissanayake, B. Coifman, W. Chen, and F. Ali (2018) 'Measuring the effectiveness of a transit agency's social-media engagement with travellers', Transportation Research Record: Journal of the Transportation Research Board, 2672(50), 46–55.

3. F. Ali, D. Dissanayake, M. Bell, and M. Farrow 'Characterising public attitudes that demonstrate potential to switch transport modes with environmental benefit: a Bayesian approach' To be submitted for possible publication to a journal, *In preparation.*

4. Delivering sustainable mobility target: Which individuals are more likely to switch from private to greener transportation options? To be submitted for possible publication to a journal, *In preparation*.

Conferences

1. School of Civil Engineering and Geosciences Conference, Newcastle University, 6th–7th July 2016 - Poster Presenter

2. World Conference of Transport Research (WCTR) Conference, Shanghai, China, 10th–15th July 2016 - Poster Presenter.

3. UTSG Conference, Dublin, Ireland, 4th – 6th Jan 2017 - Oral Presenter.

4. TRB Meeting, Washington DC, USA, 8th – 12th Jan 2017 - Poster Presenter.

5. School of Civil Engineering and Geosciences Conference, Newcastle University, 5th-6th July 2017 - Oral Presenter.

6. TRB Meeting, Washington DC, USA, 7th – 11th Jan 2018 - Oral Presenter.

Chapter 1 Introduction

1.1 Background

Climate change is a major environmental challenge that currently confronts civilisation (Stern, 2006; Hensher, 2008; Koetse and Rietveld, 2009; Liu *et al.*, 2016). The effect of greenhouse gas (GHG) emissions on climate has given rise climate change issues including extreme temperatures, flooding and air pollution (IPCC, 2015). From 1950 to 2011, the USA, European Union, China, Russian Federation and Japan were considered to be the five largest emitters of carbon dioxide (CO_2), which together contributed to 67% of the world's CO_2 emissions (Muradov, 2013). According to a more recent report by Le Quéré *et al.* (2016), the world pumped an estimated 39.8 billion tons (36.1 billion metric tons) of CO_2 into the air in 2016 by burning coal, oil and carbon based gases. This amounted to 778 million tons (706 metric tons) or 2.3% more than in the previous year, despite increasingly crucial warnings over the need to reduce GHG emissions.

When it comes tackling climate change, the transport sector in the UK is still failing to play its part (Anable *et al.*, 2006; Chapman, 2007; Koetse and Rietveld, 2009). The increasing needs of people have led to a higher demand for transportation services. Many countries, such as the USA, UK, China, Russia and India have been struggling to cope with this challenge of an increasing demand for transport over several decades (Tight *et al.*, 2005; Stern, 2006; Hensher, 2007; Koetse and Rietveld, 2009; Van *et al.*, 2014). As a result, a range of transport modes have emerged in the transport system (public and private) making roads more congested. Many countries, and especially those with well-developed or expanding economies, experience chronic traffic congestion due to increasing travel demand and vehicle ownership (O'Flaherty, 1997; Hensher, 2008; Dissanayake and Morikawa, 2010). Transport was considered as the second largest contributor after those from energy supplies (35%) in the UK (DfT, 2014) where 21% of GHG emissions were attributable to road transport. Despite the efforts of governments to mitigate these problems, the car still dominates the modal choice for many people, especially in developed countries.

Cars represent a fast, door to door, convenient and comfortable mode of transport. Hence, travellers prefer to use their cars over other public transport modes. Jakobsson (2004) stated that car use for work, shopping and other purposes has increased over the years due to its flexibility in performing activities in different places. The car also acts as a status symbol for many (Dissanayake and Morikawa, 2001), where its ownership is a reflection of status beyond any other benefit (Bergstad *et al.*, 2011). Even though the car is a popular mode choice among travellers, there are adverse impacts due to car use, for example travel delays, longer journey times due to traffic congestion, consequential tail pipe emissions causing deterioration to the urban environment and associated detrimental health impacts. In addition, car use contributes to urban sprawl, excessive noise and accidents. Based on statistics published by the UK Department for Transport (DfT, 2014), cars and taxis comprise a greater proportion at 40% of transport emissions compared to other transport options such as aeroplanes (21%), heavy goods vehicles (HGVs)(15%), vans (10%) and others (13%). More than half (64%) of people travel by private cars in the UK, whilst in the USA, 87% of daily trips take place in private vehicles (Bureau of Transportation Statistics, 2012).

Current trends in mobility and their adverse impacts on the environment and peoples' health and well-being are of increasing concern. In order to minimize these problems, the transport authorities in the UK have been making an effort to improve public transport facilities and services over the past few decades to maintain and increase patronage. This is also common in other developed countries; for instance, in Australia where public transport modes including light rapid transit (LRT), mass rapid transit (MRT) and monorail have been introduced to cater for people's travel needs and to improve accessibility (Hensher, 2007). This thesis is expected to generate new knowledge to support the review by Anable *et al*. (2006), Schwanen *et al.* (2011), and Skinner *et al*. (2011). Also it will provide scientific evidence to help support local authorities, particularly their target setting for future mobility plans by introducing policies that will have significant impact on car use and cutting down carbon for a greener future (DfT, 2011).

Local transport plans (LTPs) are important part of transport planning in the UK to reduce the problems of traffic and inequality in transport. For example, in Newcastle upon Tyne, four initiatives have been introduced to reduce the city's carbon footprint by encouraging peoples to travel by more sustainable modes. These initiatives are Go Smarter to School, Go Smarter to Work, Go Zero, and Newcastle Park and Ride.

On the other hand, the Climate Change Committee, set a mandatory target for the UK of a 67% reduction in CO_2 emissions by 2050 over 2010 levels (CCC, 2010). In order to meet the carbon target, it is necessary to reduce the numbers of VKT by motorised vehicles by achieving a shift towards sustainable modes. Therefore, the specific groups of car users are more inclined to change to sustainable modes should be addresed in future policies, so that ultimately targets can be met.

Researchers suggest that policies need to be formulated and targeted to address the issues related to traffic congestion and its impact on the environment (Stradling *et al.*, 2000; Hull, 2008; Santos *et al.*, 2010; Geng *et al.*, 2017). It is fair to say that, future policies should be informed by the results of investigation into how car users' attitudes and perceptions can be changed. Because of various concerns associated with car use, more attention has recently been paid to environmental awareness and encouraging travellers to use more sustainable modes including public tranport and non-motorised forms such as walking and cycling. The emphasis in the policies has been changing, with more attention paid to environmental aspects and modal shift has become much higher on the agenda in recent years. However, despite these efforts, car ownership and the use of cars continues to rise (DfT, 2017). In order to address the reasons why people prefer cars over other transport modes or vice versa there is a need to investigate using a carefully designed methodological approach to provide scientific evidence to reverse the trend.

1.2 Motivation of Study

Whereas many studies have attempted to develop theories and to investigate the practicality of a potential shift from private to sustainable modes, little attention has been given to investigating attitudes and behaviour towards environmental issues, especially climate change, as a driver for change. An evidence based review found that attitudes, climate change and travel behaviour have not been comprehensively examined in any consistent, robust and integrated way which warrants a comprehensive analysis of the links between them (Schwanen *et al.*, 2011; Skinner *et al.*, 2011).

An attempt to investigate which cohort are more likely to change travel behaviour was carried out by Curtis and Headicar (1997), who found that those who were susceptible to mode change are more likely to be male, in their 30s. In contrast, Miralles-Guasch *et al.* (2016) found that women's mobility knowledge and practices are typically related to the most sustainable means of transport. Furthermore, gender differences were found by Waygood and Avineri (2016) to be an influence in general environmental concern and knowledge suggesting that women were more willing to pay to reduce their personal impacts and proposed that women are either more willing to change their behaviour or that their response to information on climate change is stronger. These two results suggest that over a period of 18 years, the target group for modal shift has changed from male to female. This raises an interesting research question to be answered in this Ph.D. thesis. Moreover, other characteristics such as age, employment status, household size and car ownership not considered together previously, will also be taken into account to complement the investigation in this study.

In the context of attitudes and travel behaviour studies, the author found a lack of empirical evidence from mathematical modelling to support the idea of incorporating the car users with socio-demographic and attitudinal factors, which requires research with a comprehensive and sound methodological approach. Therefore, this study attempted to fill this research gap by employing carefully planned methodological process that gave due consideration to their attitudes to and perceptions of environmental factors and sustainability. The evidence obtained from this study is expected to inform future policies and inform the decisions of local authorities (LAs), national governments and associated industries to formulate strategies to achieve targets related to sustainability.

Previous research used frequentist probability methods to make inferences, but these do not rely on specific prior data but only conditional distributions of data depending on particular hypotheses. For example, using maximum likelihood estimator, researchers worked with data without considering personal information and therefore had no chance to consider their prior belief before the analysis of actual data was taken into account. In order to come up with effective models, the inference method must be capable of incorporating uncertainties in parameters to identify significant inferences which is essential to quantify the strength of the findings. Such models are addressed by this study.

1.3 Research Questions

Given this gap in knowledge, four overarching research questions to be answered by this research can be posed:

1. What are the attitudinal factors that characterise the uses of sustainable transport modes?
2. What are the key socio-demographic variables that affect travel mode choices and decisions within the cohort of car users?
3. Which car users' groups should be targeted in campaigns that promote the uptake of non-car transport alternatives?
4. Can key factors be used to derive a model to predict the likely uptake of sustainable modes?

These research questions were answered by carrying out detailed state-of-art-review and carefully planned multi-faceted analyses. As part of the review, the relevant datasets were examined to be used in the data analysis.

1.4 Research Aim

This study aims to provide insights into which target groups of individuals within the car user population are more likely to switch from private transport to sustainable modes in order to present scientific evidence to support the decision making of LAs and policymakers who are responsible for the design of marketing strategies and new green transportation policies.

1.5 Research Objectives

The specific objectives of the study were as follows:

1. To investigate important attitudinal factors through dimension reduction analysis;
2. To examine and form groups of car users' travel mode choice decisions based on socio-demographic variables;
3. To investigate the relationships between car users' groups and attitudinal factors which relate to modal shift potential; and
4. To derive a mathematical model to predict the likely uptake of sustainable modes.

The objectives stated above were achieved by conducting an initial search for relevant datasets at the early stage of research. Accordingly, the British Social Attitudes (BSA) dataset was identified as the most appropriate to achieve the aim and objectives in this study. This is a survey representative of adults aged 18 and above. The BSA survey used random probability sampling to select respondents to participate in the survey to ensure that the data are not biased and representative of the British population. The samples were collected from a large population over a given time period and classified as cross-sectional data.

1.7 Research Tasks

In order to achieve the research objectives, the following tasks were proposed:

1. Conduct a critical literature review of previous studies of sustainable transport, which includes mathematical approaches to modelling travel behaviour, modal shift potential and travel demand management policies so as to be able to identify the gap in knowledge and to inform the design of a structured methodological framework to fill the gap.

2. Assemble and explore a range of statistical methods (such as descriptive, factor and cluster analysis) to establish the characteristics of sample populations of respondents, using the existing British Social Attitudes (BSA) survey datasets.

3. Specify the role of the existing dataset in addressing the research gap, identifying any shortfalls, and formulate the approach to mathematical modelling to be used.

4. Investigate attitudinal variables with regard to environmental issues to establish factors that are important in and may govern modal shift from private to sustainable transport.

5. Characterise potential target groups using a clustering method and examine the relationships between factors and clusters to quantify their potential to lead to shifts in modes of travel based on environmental awareness.

6. Validate the model developed using an independent source or data timeframe and constantly review the data's suitability to deliver the objectives.

7. Draw conclusions and recommendations for further research into the delivery of sustainable mobility targets.

Tasks underpin the delivery of all objectives. Objectives 1 and 2 were delivered by Tasks 2 and 3, whilst objectives 3 and 4 were delivered respectively by Tasks 4, 5 and 6. Task 7 brings together the results achieved by objectives 1 to 4.

1.8 Thesis Outline

As an introduction to the key concepts of travel behaviour and its relationship to sustainable transport and modal shift, an extensive literature review is presented in Chapters 2 and 3. Chapter 2 focuses on identifying the gaps in knowledge in the topics considered. The discussion includes overviews of the transport demand management (TDM), transportation and environmental issues, travellers' attitudes to the environment, addressing behavioural changes through transport policies, attitudes towards and behaviour in travelling, and achieving a sustainable transportation system. Chapter 3 focuses on finding a structured methodology for the research to fill the gaps in knowledge identified. Furthermore, a critical review is presented of an extensive collection of methodologies used in previous related studies.

Background information on the gaps in knowledge found and the limitations of methods used in previous studies is provided and forms a sound foundation for the methodological approach used in this research as described in Chapter 4. A flow diagram is used to explain the methodological framework in detail.

The general characteristics of the BSA respondents are described in Chapter 5. In this chapter, a preliminary data analysis looks into the detail in the data in order to gain an understanding of the BSA dataset. The data cleaning, coding and preparation processes used are also explained. An exploration of the structure of data and the factors obtained from the exploratory factor analysis is presented in Chapter 6, including details of respondents' attitudes towards and perceptions of transport and the environment, traffic awareness, current travel behaviour and willingness to switch travel modes.

Chapter 7 provides an investigation of differences among respondents in their perceptions of environmental issues. An investigation into the structure of the data using multiple correspondence analysis (MCA) and hierarchical cluster analysis (HCA) is discussed. Multinomial logistic regression models resulting from the relationships found between factors and clusters are identified and the evaluation of the statistical performance of the models is explained.

The use of the Bayesian inference approach to a log-linear model for categorical data and multivariate probit models is described in Chapter 8. The results estimated from these models are then presented. The car user groups who demonstrate a higher propensity to take action to adjust their mode choice or activity to address environmental issues are discussed.

Finally, the thesis concludes with the main findings of the study presented in Chapter 9. The policy implications of this research, the limitations of the study, and directions for further work are also discussed.

Chapter 2 Transport Attitudes and Travel Behaviour – Critical Review

2.1 Introduction

This chapter provides a critical review of the literature from previous studies to determine the direction of this research. The literature review uses a top to bottom approach where the topic is viewed from a broader perspective and then the discussion is narrowed down as the core issues are revealed. The sequence of the topics reviewed begins by elaborating on Transport Demand Management (TDM) strategies as described in Section 2.2, followed by a discussion of challenges in transport management, including environmental issues, in Section 2.3. An overview of travellers' behavioural change with respect to environmental issues is then presented in Section 2.4, followed by a discussion of policy interventions in Section 2.5. Section 2.6 describes a recent study on travel attitudes and travel behaviour, and Section 2.7 explains ways of achieving a sustainable transportation system. Finally, the research gap in this study is presented in Section 2.8 along with a conclusion of this chapter.

2.2 Transport Demand Management (TDM)

Since the 1970s, Transport Demand Management (TDM) has been a popular tool to promote changes in travel patterns in transport planning contexts. Consequently, many policies and conceptual frameworks have been developed over the years (Litman, 2003; Habibian and Kermanshah, 2011; Habibian and Kermanshah, 2013). More recently, researchers have placed emphasis on Modal Shift Potential (MSP) to understand the influence on travellers' choices when presented with sustainable options so that travel creates only a minimal impact on the environment (Nilsson and Küller, 2000; Beirão and Cabral, 2007; Cairns *et al.*, 2008; Cass and Faulconbridge, 2016).

TDM, also known as mobility management, is a common term used for programmes and strategies that inspire more efficient consumption of transport resources such as road space, parking bays, vehicle capacity, funding and energy. TDM includes measures that improve the transport options available to travellers, offers inducements that encourage travellers to use more efficient and less polluting transport alternatives, more accessible land use patterns, and designs planning transformation policies (Litman, 2003). An example can be seen in Gärling and Schuitema (2007) of how TDM is applied in targeting reduced private car use in terms of three factors: effectiveness, public acceptability and political feasibility.

By managing traffic demand more efficiently by mode shift to public transport, numerous benefits are achieved. These include reduction in congestion, travel time savings, less need for parking facilities, fewer road accidents and less vehicle emissions (Litman, 2003; Litman and Fitzroy, 2009). As highlighted by the DREDF (2014), the advantages of the TDM measures to consumers are improved access to transportation, more reliable public transport services, provision of real-time information, alternative mode choice option and improved transportation performance.

2.3 Transportation and Environmental Issues: The Challenges

Transport is considered to be a driver for economic growth given its role in moving people and goods across regions and cities (DfT, 2014), and yet transport is a major source of greenhouse gas emissions (GHG). The transport sector was responsible for 14% of global GHG in 2010 and accounted for approximately 23% of overall energy-related emissions (SKTP, 2016). Emissions from transport have increased rapidly despite significant carbon dioxide (CO_2) reduction measures implemented to different degrees across the world (Bharadwaj *et al.*, 2017). Low-carbon transportation systems offer significant mitigation potential, enabling growth in travel and the development of urban areas and regions, while reducing the economic, environmental and social costs (Davison and Knowles, 2006; Chapman, 2007; Jochem *et al.*, 2016). The SKTP (2016) emphasised that transportation and mobility are essential components of the economic and social development of any area and recommended that municipal governments, and local and regional authorities, need to be proactive in initiating sustainable transport systems when developing a vision for the reduction of carbon emissions.

With the passing of the Climate Change Act 2008, the UK became the first country in the world to set up a long-term target for reducing GHG emissions, for example, pursuing at least 80% reductions by 2050 compared to 1990 levels (Boardman, 2005). While changes to the climate can occur due to natural processes, the effect of man-made GHG is responsible for recent upward shifts in global temperature, and transport use plays a major role.

In 2017, 26% of the UK domestic GHG emissions were due to road transport (DfT, 2017). This high proportion means that no climate action plan can be successful without the inclusion of transport plans. In order to reduce vehicle kilometres travelled (VKT) by motorised vehicles, delivering a modal shift towards sustainable modes will be necessary to meet the carbon target (DfT, 2009). Undoubtedly the growth in car ownership, demand for personal mobility and traffic congestion all generate negative impacts on the environment (Susilo *et al.*, 2012). In addition to car use, road freight and aviation are perceived to be major contributors to GHG (Chapman, 2007).

Increasing awareness and nurturing an understanding of climate change and other environmental issues, and their effect on each individual in our society, is a great challenge. On the other hand, it is questionable whether raising awareness alone would be sufficient to change travel behaviour. Previous research gives mixed answers to this question, for example Anable *et al.* (2006), Susilo *et al.* (2012), and Ho *et al.* (2017), which will be discussed in Section 2.4.

The 'Stern Review' of the economics of climate change emphasised the need for immediate action to alleviate GHG emissions as the benefits of bold, early action considerably outweigh the costs if delayed (Stern, 2006). Accordingly, the total cost of climate change was estimated to be equal to a fall of at least 5% in global Gross Domestic Product (GDP) each year if no further action was taken. Climate change is one of the biggest challenges that we face today with the UK taking the lead in setting mandatory limits.

Waterson *et al.* (2003) identified a significant modal shift to public transport as an option to reduce CO_2 emissions from road transport, and 'zero-carbon' alternatives such as walking and cycling are worthy alternatives. However, Waterson *et al.* (2003) acknowledged that changing attitudes concerning the dependency on private transport will be a great challenge.

In the UK, a quarter of all car trips were short distance journeys under two miles (Mackett, 2000). Such short trips can easily be made by a high proportion of the population in a more sustainable way by walking and cycling. Ryley (2001) stated that 89% of drivers surveyed agreed with the statement, ''I would find it very difficult to adjust my lifestyle to being without a car". This statement suggests that behavioural change is a great challenge as society has largely become over-dependent on car use.

Integrating transport policy and social psychology are recognised as being important to help drivers move away from their cars onto more sustainable transport modes (Stradling *et al.*, 2000). While technological solutions offer ways to achieve climate change targets over the longer term, there is a pressing need to achieve change in behaviour and travel habits in the short-term through current policies (Anable, 2005; Boardman, 2005; Chapman, 2007).

2.4 Travellers' Attitudes to Environmental Issues: Do They Influence Behavioural Change?

In order to achieve the GHG emissions targets from the standpoint of transport, behavioural change brought about by suitable policies will be vital (Chapman, 2007). People may fail to achieve anticipated behavioural changes if they have no relevant beliefs or perceptions. Perceptions lead to attitudes which ultimately influence behaviour. Schade and Schlag (2003) found that there were positive associations among those that view urban transport pricing approaches acceptable according to social norms, personal effect potentials and perceived effectiveness. Besides that, Beirão and Cabral (2007) and Shiftan *et al.* (2008) examined the relationship between mode choices and travellers' attitudes, and found that potential users were attracted to the services that could accommodate the travellers' needs, such as punctuality and reliability.

Researchers have paid particular attention to the habit of using a car, as this is fundamental in understanding mode choice behaviour (Anable, 2005; Susilo *et al.*, 2012; Susilo and Cats, 2014; Ho *et al.*, 2017). Anable (2005) demonstrated that psychological variables should be considered when studying the propensity to switch travel mode because awareness of environment issues and concerns and car dependency are increasing. In addition, in order to promote sustainable transport, such as encouraging walking and cycling, it is imperative to raise awareness of available sustainable mobility services, including public transport, with emphasis on key aspects such as accessibility constraints, basic safety and security, convenience and cost, and enjoyment, all of which were identified as important by Schneider (2013). Meanwhile, Nkurunziza *et al.* (2012) and Li *et al.* (2013) demonstrated that bicycle commuting, in particular, is strongly related to personal motivation and attitudes to environment.

Some people think that they must change their behaviour in order to use more sustainable alternatives, whereas others do not (Steg and Vlek, 2009). However, in a study by Susilo *et al.* (2012) almost all respondents to a questionnaire agreed that all people are required to change their actions to guarantee a sustainable future rather than only themselves taking action. Travellers are aware of environmental problems, but their interpretations do not necessarily match their travel behaviour (Tertoolen *et al.*, 1998; Anable *et al.*, 2006; Susilo *et al.*, 2012). This is consistent with studies which recommend that even though facts about the negative environmental effects of car use raises some consciousness, it is usually insufficient to change behaviour (Susilo and Cats, 2014).

According to Anable *et al.* (2006), there are certain car-owner groups of travellers whose environmental concerns, as well as sense of responsibility, is greater than others. This suggests that there is potential for mode choice behaviour change. A recent study by Ho *et al.* (2017) clearly pointed out the need to provide information to travellers such as the benefits of sustainable travel planning initiatives to raise awareness and to generate positive attitudes towards green travel initiatives. They also went on to recommend that targeting relevant segments in society would be more effective than adopting a generic approach aimed at the whole population (Ho *et al.*, 2017). This was a key motivator to this research aiming to identifying cohorts of the population most likely to change behaviour.

2.5 Addressing Behavioural Changes through Transport Policies: What is Lacking in the Current Process?

The transport sector creates much environmental pressure. Many policies targeted at reducing the pressure on the environment have not been fully effective because the behavioural characteristics of travellers are inadequately acknowledged (Urwin and Jordan, 2008). In this sense, different strategies targeting different segments in society are needed as different populations are inspired by different aspects and are affected in different ways by policy (Anable, 2005; Banister *et al.*, 2012).

Previous psychological studies conducted on this subject further confirm that policies to date have generally had only a limited effect as they target the entire population in a uniform way (Hunecke *et al.*, 2010; Prillwitz and Barr, 2011). More recently, Semanjski *et al.* (2016) generated population segments with similar attitudes and monitored them separately for the behavioural changes of 3400 respondents using smartphones as mobility. According to their results, segments of the population behave differently in terms of modal shift and route choice decisions and those changes are often based on the purpose of their journeys.

Advances in Information and Communication Technologies (ICT) may help enhance transport choices, as they respond to specific needs, improve the mobility, and create a safer environment for vulnerable groups such as women, people with limited mobility, and people with disabilities to engage in economic activities (Giannopoulos, 2004; Banister, 2008; Cohen-Blankshtain and Rotem-Mindali, 2016). Using increasingly widespread and affordable ICT, transport users can now check and report any delays or disruption before or during the journey (Harris *et al.,* 2015). ICT can help the planning and accessibility of transport. Transport users can get information to assist them for better planning when making a journey.

As reported by Chapman (2007), long-term technological solutions are given considerable attention when addressing transport and climate change concerns. However, short-term behaviour change is seen to be important if the advantages of new technology are to be fully realised (Chapman, 2007). Previously, Anable (2005) and Boardman (2005) confirmed that policies to change behaviour and travel habits are crucial and potentially more important than technological approaches in the short-term. This indicates that a shift to sustainable mobility is unlikely if technological enhancements and modifications in the built environment are not combined with behavioural change.

2.5.1 Current policies on behaviour change – do they work well?

Several policies and interventions such as Travel Demand Management (TDM) have been introduced in the UK and overseas over recent years to minimise travel-related impacts on society and the environment. The emphasis in these policies has changed significantly in recent years with particular attention now being paid to environmental aspects (Tight *et al.*, 2005; Chapman, 2007; Marsden and Rye, 2010).

As the result, the modal shift towards sustainable alternatives, including cycling and walking, as well as public transportation, has gained more attention in policy proposals in the UK (DCLG, 2012). Under the Local Transport Plan (LTP) initiatives, the local authorities (LAs) in the UK work towards achieving emission reduction targets by promoting public transport alternatives, and supporting the market for low-carbon transport and encouraging travellers to move from fuel-hungry vehicles to low-carbon vehicles (DCLG, 2012). For example, Department of Environment, Transport, and the Regions (DETR) published information to Local Authorities London (LAL) about encouraging walking, walking and cycling action plan for sustainable travel guide.

Previous research has placed much emphasis on exploring the undesirable side effects of car use, for instance traffic congestion, carbon dioxides (CO_2) emissions and air quality issues, and health related impacts (Steg and Gifford, 2005; Graham-Rowe *et al.*, 2011) leading to suitable policies being recommended to minimise these effects by encouraging people to change their travel modes from private cars to public or non-motorised transport (Davison and Knowles, 2006; Banister, 2008; Schneider, 2013). However, on many occasions the suggested policies have failed to meet the expectations of the policymakers in integrating climate change concerns to reduce CO_2 pollution and traffic congestion (Rayner *et al.*, 2008; Urwin and Jordan, 2008). One possible reason for this is that the policies were developed without paying adequate attention to travellers' attitudes and perceptions. Policy execution is believed to be successful if implemented for the appropriate targeted groups (Curtis and Headicar, 1997; Campbell *et al.*, 2012).

This research suggests the promotion of modal shift needs a well-defined audience. The identification of a precise target individual or group implies the likely effectiveness of new policy implementations. Marketing strategies for travel behaviour have changed to include segmentation techniques enriched with qualitative attributes to explore behavioural aspects. Needs, beliefs and expectations vary significantly between market segments and evidence suggests that well-designed policies should target specific groups (Jensen, 1999; Anable, 2005; Steg, 2005). Recent advances in travel market segmentation indicate the need for new user segments incorporating perceptions and attitudes (Jensen, 1999; Anable, 2005).

These earlier studies suggest that in order to develop sustainable travel policies, it is useful to identify travellers who chose to use different modes to meet their travel needs. Of particular interest, in the research reported in this thesis is the identification of car users who also use public transport, bicycles, or walking, as this group is more likely to respond to policies that promote the use of those modes.

2.6 Attitudes Towards and Behaviour in Travelling

It has been argued that travel behaviour is not only influenced by peoples' preferences but also results from compromises of many other factors, including individual, household characteristics, and socio-demographic factors (Curtis and Headicar, 1997; Susilo *et al.*, 2012; Susilo and Cats, 2014). The influence of different factors on user preferences, satisfaction and decision-making processes has been given careful attention to investigate the key determinants of travel choices for multi-modal trips by different traveller groups. The results reported by Susilo and Cats (2014) indicate that, for certain groups of travellers such as women, young and low-income or unemployed, there are distinctive determinants of satisfaction with trip stages for various travel modes.

Personalised Transport Planning (PTP) is a well-established method to encourage people to make more sustainable travel choices (Haq et al., 2008). Some cities supported PTP in the reduction of congestion, cleaner air, healthier nations, and reduced CO2 levels. In the Europe, the Changing Habits for Urban Mobility Solutions (CHUMS) project conducting research with strategic aim of car-pooler. The CHUMS expected the changes in travel behaviour mind-sets for commuting leading to more energy efficient transport, shift towards sharing the journey for the working population who currently drive alone to work and to attract more employees to use carpooling for their commute to work trips. However, Waterson et al. (2003) acknowledged that greater rates of walking and cycling will only be achieved when car use becomes significantly more costly and less convenient. It can be argued that policies to increase walking and cycling do not require transport solutions but, rather, need more fundamental changes in society and urban structures that allow more flexibility in how and when people travel, so that walking and cycling can be more easily fitted into household routines.

Researchers also have paid considerable attention to how people's attitudes can be influenced through awareness-raising, social marketing and other interventions in order to change their travel behaviour (Nordlund and Garvill, 2003; Scheiner and Holz-Rau, 2007; Ory and Mokhtarian, 2009). A major issue identified in relation to this is the extent to which people's behaviour is actually influenced by their attitudes and whether changing people's attitudes necessarily will lead to an associated change in their travel choices (Anable *et al.*, 2006).

According to Line *et al.* (2010) and Line *et al.* (2012), attitudes and transport behaviour intentions, from the perspective of climate change, ultimately develop from the knowledge and values held by young people. They prefer cars due to the freedom, and privacy they give, and as a symbol of status, even though they generally recognize the detrimental impact that this mode has on climate change in comparison to other modes (Line *et al.*, 2012). On the other hand, Anable (2005) claims that many studies have used established psychological theories of attitude-behaviour relations, such as the theory of planned behaviour, to predict mode choice and concluded that the choice of travel mode is largely a reasoned decision related particularly to attitudes and perceived barriers to behaviour change.

In terms of perceptions and attitudinal influences on travel behaviour, factors which have been considered in previous research include travel time and cost (Noland and Polak, 2002), beliefs about safety and the health benefits of cycling (Heinen and Handy, 2012) and psychological factors which have a relatively strong impact on mode choice (Heinen *et al.*, 2011). The role of socio-demographic variables as factors influencing individual travel patterns was further examined and a summary is presented in Table 2.1. Some variables have been shown to have a significant relationship with travel patterns, for instance, higher income level, increases in vehicle ownership, higher rate of possession of driver's licences, and more people working in a household, all of which were shown to lead to higher frequency of trips as well as an increase in miles driven.

From previous investigations, certain factors, such as age and car ownership, have been found to have a significant relationship with travel patterns. Among studies of travel behaviour, the most frequently adopted research method was to analyse household travel survey data giving information about individuals' characteristics alongside travel patterns (Domencich and McFadden, 1975). In the next section, the challenges faced in achieving sustainable transportation from the perspective of climate change and environmental issues are discussed.

Socio-demographic factors	Travel trend	Example of studies
Age ↑	Trip frequency →	Hanson (1982)
	Proportion of car journey →	Flannelly and McLeod Jr (1989)
	Transport energy consumption ↑	Naess and Sandberg (1996); Burton *et al.* (2013)
Gender	Trip frequency →	Hanson (1982); Veterník and Gogola (2017)
Employee per household ↑	Trip frequency (per household) ↑	Ewing, De Anna et al. (1996); Boarnet and Crane (2001a)
	Travel time ↑	Ewing (1995); Burton *et al.* (2013)
Level of education ↑	Proportion of car journey↑	Flannelly and McLeod Jr (1989)
	Proportion of public transport use ↑	Kockelman (1997); Boarnet and Sarmiento (1998); Stead (2001)
Household size ↑	Trip frequency↑	Hanson 1982, Kockelman (1997); Boarnet and Crane (2001b)
	Travel time ↑	Ewing (1995); Noland (2001)
	Transport energy consumption ↑	Banister (1997); Watson et al. (1997); Musti *et al.* (2011)
Household income ↑	Trip frequency ↑	Hanson (1982); Victor and Ponnuswamy (2012); Burton *et al.* (2013)
	Travel distance ↑	Cervero (1996); Naess and Sandberg (1996); Farber *et al.* (2014); Ruhe *et al.* (2016)
	Proportion of car journey ↑	Flannelly and McLeod Jr (1989); Burton *et al.* (2013)
	Transport energy consumption ↑	Naess (1993); Hickman *et al.* (2015)
Car ownership ↑	Trip frequency ↑	Hanson (1982); Sillaparcharn (2007)
	Trip frequency →	Prevedouros and Schofer (1991)
	Travel distance ↑	Naess and Sandberg (1996); Kockelman (1997); Wang (2016)
	Proportion of car journey ↑	Naess (1993); Burton *et al.* (2013)
	Travel time ↑	Ewing (1995); Burton *et al.* (2013); Stapleton *et al.* (2017)
Possession of driver's licence per household ↑	Using car ↑	Flannelly and McLeod Jr (1989); Van Acker and Witlox (2010)

Table 2.1: Example of how socio-demographic factors affect travel trend where "↑" means increasing and "→" no change.

Source: Modified from Wang (2015)

2.7 Achieving a Sustainable Transportation System

Travellers' attitudes were found by Sunkanapalli *et al.* (2000) and Parkany *et al.* (2004) to have a close relationship with travel mode choice. Based on theory, attitude is a proxy indicator to assess behaviour, which in turn organises a person to act in a specific way (Ajzen, 1987). Attitudes have been described as part of the decision-making process and, therefore, have significant impact on travel choices (Outwater *et al.*, 2003) and, predominantly, the choice of travel mode(s) (Domarchi *et al.*, 2008). For example, when selecting public transportation services, demographics and travel needs were found to be less important compared to attitudes (Garling *et al.*, 1998; Fujii and Garling, 2003).

Ideas concerning sustainable transportation emerged from discussions of sustainable development, and relate to the need to focus on the modes of transport used and the transport planning system (Litman, 2009). The qualities defining a sustainable transport system would include social, economic and environmental characteristics (Zhou, 2012). In the literature, the research generally focuses on achieving broadly four qualities for a sustainable transport system; namely, accessibility, efficiency, equitability and environmental friendliness. These are as listed in Table 2.2.

No.	Qualities	Focus	Literature
1	Accessibility	Meets the basic access and needs of individuals	Goldman and Gorham (2006) Hull (2008) Habibian and Kermanshah (2013) Eriksson *et al.* (2008)
2	Efficiency	Operating efficiently, including support for a competitive economy	Rietveld and Stough (2005) Litman and Fitzroy (2009) Hensher (2008) Goh (2002)
3	Equitability	Affordable and operating fairly for everyone	Goldman and Gorham (2006) Stern (2006)
4	Environmental friendliness	Uses less energy or renewable resources and limits emissions to minimize the impact on the environment	Bertolini and Martin (2003) Rietveld and Stough (2005) Steg (2005) Anable and Gatersleben (2005)

Table 2.2: Assessing a transport system for sustainability

It is clear that previous research has led to evidence that one or two of these qualities have emerged as important but none have identified all four. Previous research suggests that sustainable transport needs to meet user access (physical and economic) to an efficiently operated system which offers affordable travel in vehicles using renewable energy with limited toxic emissions.

How to achieve sustainability in the way we travel has been widely discussed recently, especially among policy makers, due to the benefits offered to our society, economy and the environment. Various strategies have been designed to promote a shift from car use to more sustainable modes of transport (Stradling *et al.*, 2000; Gärling and Schuitema, 2007; Saleh, 2007). Stradling *et al.* (2000), in a study of behavioural change of English motorists in the UK, carried out "pull" and "push" measures to help motorists out of their cars. They advocate that if people are helped to change, rather than being forced, the resulting changes will be more sustainable. On the other hand, Gärling and Schuitema (2007) suggested in their review that incentives could be offered for the use of non-car transport in order to decrease car use. These include road pricing, which offers monetary incentives for reduction in car use (Saleh, 2007).

Given the potential benefits that are offered to the environment, economy and public health, cities and regions around the world have set ambitious goals for increasing the use of such policies, including non-motorised and public transport focusing on commuter trips. In a study of Ho *et al.* (2017) to assess community awareness, interest and involvement with a number of green initiatives and to understand how sustainable travel planning has been absorbed, 378 residents aged 14 years or over in 2011 were used. The research, using zero-inflated ordered probit (ZOIP) models, emphasised that travel planning by employers that promotes more sustainable travel has delivered less car-dependent behaviour among commuters in New South Wales (NSW), Australia.

Promoting sustainability has been seen as a core aspect in recent research in the transport domain (Prillwitz and Barr, 2011; Susilo *et al.*, 2012; Schneider, 2013; Xenias and Whitmarsh, 2013). Prillwitz and Barr (2011), in their study, classified travellers into two different groups to explore attitudes and behavioural change towards sustainability, and found that 'consistent green travellers' who walk and cycle are mainly young professionals and more likely to use motorised sustainable modes than other participants. The other group is named ‘persistent car users’, often belonging to middle-aged cohorts who support and use cars most frequently. On the other hand, Susilo *et al.* (2012), in a study of individuals’ environmental attitudes using 659 completed questionnaires from residents of 13 developments with some sustainable features in the UK, emphasized that age, health, security, availability and household size are important factors for mode choice options including walking, cycling and using public transport.

Shepherd *et al.* (2006) studied three types of policy instruments, including changes in fares and frequency of public transport, in the city of Edinburgh in order to identify optimal transport strategies. Eriksson *et al.* (2008) studied reasons to reduce car use for the work commute of 1218 employees in a Swedish city. Redman *et al.* (2013), in their review, contributed a better understanding of sustainable development aspects to attract car users to switch to public transport. These researches highlighted those aspects which are important in encouraging car drivers to use public transport. Among them, the introduction of discounted fares (Shepherd *et al.*, 2006; Eriksson *et al.*, 2008; Redman *et al.*, 2013) and increased frequency of services (Shepherd *et al.*, 2006; Eriksson *et al.*, 2008) are highlighted. Moreover, Miralles-Guasch *et al.* (2016) examined the differences between gender mobility through age, modal split and trip purposes using mobility data from a large travel survey taken in 2006 in Spain and suggested that any strategy promoting sustainable growth and attempting to reduce the impact of air pollution should focus on gender. They found that females were using sustainable transport modes more often than males.

Other researchers discuss different important factors that encourage people to participate in the support of sustainable transportation (Dickinson *et al.*, 2003; Anable and Gatersleben, 2005; Beirão and Cabral, 2007; Schneider, 2013). Dickinson *et al.* (2003) investigated travel plan measures to improve cycling activities to work in terms of gender through 2065 completed questionnaires in the UK. Anable and Gatersleben (2005) examined factors influencing work and leisure journeys by different travel modes using 235 participants in an on-line questionnaire. Meanwhile, Beirão and Cabral (2007) conducted a qualitative study of public transport and car users in order to gain an in-depth understanding of travellers’ attitudes and perceptions of public transport service quality. Also, Schneider (2013) used in-depth interview responses from the San Francisco Bay Area to improve the sustainability of their transportation systems by shifting routine automobile travel to walking and bicycling. The most crucial attributes to be taken into consideration when choosing travel modes emerged as flexibility, convenience, cost, environment, security, enjoyment, habits and health.

From the transport provision perspective, Xenias and Whitmarsh (2013), in their study of sustainable transport policies and technologies using open-ended questionnaires of experts (N = 53) and British public (N = 40), identified three main approaches to fostering sustainable transport. These were improving efficiency and reducing the impact of vehicles by promoting more sustainable modes and reducing the need to travel. Additionally, a study of Banister (2008) reported that reducing the need to travel, land-use policy measures, and technological innovations were the main drivers of change towards sustainability.

Kingham *et al.* (2001) suggested that one of the ways to reduce everyday commuting journeys to work is to increase fuel prices and introduce home-based or tele-working jobs. However, in the 4M project modelling, measurement, managing and mapping, funded by the EPSRC in 2010, 575 head of households in Leicester were interviewed and measured energy (gas and electricity) used and VKT by some 1400 individuals living in the homes. Estimates of CO_2 emissions from the reported energy use demonstrated that an average of 25% more CO_2 is emitted by working at home and using electricity and gas for heating, cooking, and appliances during the day, than was saved by not travelling to work (Lomas *et al.*, 2010).

In a study of the effect of transportation policies on modal shift from private car to public transport in Malaysia, Nurdden *et al.* (2007) developed a binary logit model for the three modes (car, bus, and train) using a survey of n=1200 respondents and proposed that appropriate incentives need to be provided for a successful implementation of a policy to switch from private to public transport. This is because the distance from home to workplace has been one of the biggest stumbling blocks in shifting the mode of travel for commuter journeys. Travelling over long distances, especially by public transport, can be difficult and time-consuming. As claimed by Chevalier and Lantz (2013), the longer the distance to travel, the more likely motorised modes are used with lower frequency of ride sharing.

It has been noted that walking and cycling improve general health and save money for those who substitute trips made by other transport modes. Non-motorised modes produce no pollution or CO_2 emissions and ironically can be quicker than motorised forms of transport in congested conditions. Banister (2008) suggests that walking and cycling will become an integral part of a modernised transport environment. In addition, environmental concerns are a significant factor in encouraging individuals to walk within and between neighbourhoods (Susilo *et al.*, 2012).

2.7.1 Factors affecting mode choice decisions

Researchers discuss different factors that affect mode choice decisions in relation to travel alternatives, including car and sustainable modes such as walking, cycling and public transport. Table 2.3 lists chronologically researches that study the factors affecting car use; for example Curtis and Headicar (1997), Steg (2005) and Beirão and Cabral (2007). Specifically, they focus on topics related to private cars including travel time, attachment, dependence, convenience and flexibility, status symbol and environmental concerns. On the other hand, factors that affect sustainable modes such as awareness and availability, safety and security, facility, convenience and cost, enjoyment and habit were discussed by other researchers (Anable and Gatersleben, 2005; Shannon *et al.*, 2006; Schneider, 2013).

Example of studies	Car	Sustainable modes		
		Walking	Cycling	Public transport
Curtis and Headicar (1997)	√			
Steg (2005)	√			√
Anable and Gatersleben (2005)	√	√	√	√
Shannon *et al.* (2006)			√	√
Beirão and Cabral (2007)	√			√
Schneider (2013)		√	√	√
Fuller *et al.* (2013)		√	√	

Table 2.3: Factors affecting car and sustainable modes users

It is interesting to note that only Anable and Gatersleben (2005) considered car along with all other sustainable modes. Previous research on travel behaviour, including Hanson and Schwab (1986), has revealed that when conducting an in-depth analysis of the relevant factors for travel behaviour decisions there is a need to have several contrasting themes such as attitudes, the transport system and the characteristics of travellers. These factors can be considered as external and internal factors. External factors relate to the physical environment experienced by people whilst travelling and comprise built environment, infrastructure, transit service quality, transport policy and the economic situation.

Whilst internal factors include the characteristics of travellers such as income, car ownership, employment status, gender, age group, level of education and household size, are discussed in earlier studies by researchers, for example in (Kingham *et al.*, 2001; Anable, 2005; Beirão and Cabral, 2007; Nurdden *et al.*, 2007; Habibian and Kermanshah, 2013; Susilo and Cats, 2014). Table 2.4 provides a summary of which internal factors have been included in the earlier studies and worthy of note is the variability in the factors considered and that none have dealt with all six factors.

Example of studies	Income	Age	Gender	Employment status	Car ownership	Household size
Kingham *et al.* (2001)		√	√			
Anable (2005)	√		√		√	
Beirão and Cabral (2007)	√	√		√	√	√
Nurdden *et al.* (2007)	√	√	√		√	√
Habibian and Kermanshah (2013)		√		√		√
Susilo and Cats (2014)	√			√		

Table 2.4: Internal factors affecting car users

Beirão and Cabral (2007), Nurdden *et al.* (2007), Eriksson *et al.* (2008) and Bamberg *et al.* (2011), in studies of factors affecting sustainable modes users and more specifically public transport, considered travel time, cost, not having to drive and the opportunity to relax and socialise during bus travel and comfort, as shown in Table 2.5.

Example of studies	Travel time	Travel cost	Travel distance	Trip purpose	Environment
Dickinson *et al.* (2003)			√		√
Shannon *et al.* (2006)	√				√
Nurdden *et al.* (2007)	√	√			
Beirão and Cabral (2007)	√	√			√
Banister (2008)		√	√	√	√
Eriksson *et al.* (2008)	√	√			
Bamberg *et al.* (2011)	√	√		√	√
Schneider (2013)	√				√

Table 2.5: External factors affecting users of sustainable mode

Travel time, travel cost, travel distance, trip purpose and the environment were found to be significant barriers for sustainable modes users. Reducing travel time, travel cost and subsidising fares would have the highest influence on travelling trends (Shannon *et al.*, 2006; Beirão and Cabral, 2007; Nurdden *et al.*, 2007; Eriksson *et al.*, 2008; Bamberg *et al.*, 2011). On the other hand, reducing the distance from home to public transportation stations was discussed as an important variable by Nurdden *et al.* (2007) and short journeys need to be improved so that cyclists are safe from traffic without incurring risk to their personal security Dickinson *et al.* (2003), as well as contributing to trip reduction to modal shift (Banister, 2008).

Trip purposes and the practicalities of the journey were found to be other factors that influenced travellers to use sustainable modes (Banister, 2008; Bamberg *et al.*, 2011). The key findings were that public transport not only serves environmentally friendly objectives, but also benefits travellers in productive use of time and emerged as a safer alternative. Economically, it is also a less expensive option in terms of fuel and parking costs which can potentially lower congestion and transport costs (Redman *et al.*, 2013). It also helps overcome social isolation (Prillwitz and Barr, 2011) and generally enables community interaction, which will benefit social sustainability (Currie and Stanley, 2008).

Measures to promote or encourage the use of cycling as a mode for the commute are suggested by the UK Government in its travel plan resource pack to increase cycling (Steer Davies Gleave and the Association for Commuter Transport, 2000). These include the provision of safe, secure and covered cycle parking; establishing bicycle user groups; providing pool bikes; providing lockers, changing/drying facilities and showers; offering financial incentives such as interest free bicycle loans, discounts for bicycle purchase and preferential cycle insurance rates; providing a cycle mileage allowance to enable financial reimbursement for staff cycling on company business; promoting and publicising cycling; liaising with the local authority to identify the potential for improving cycle links; and considering other initiatives such as a puncture repair service and provision of a spares box for cyclists (Dickinson *et al.*, 2003).

Ho *et al.* (2017) found that women and public transport commuters are more likely to give positive responses to sustainable options. In addition, people aged 30 or below, younger families and those attracted to active travel modes also tend to support cycling. On the other hand, households with greater earnings have a lower propensity to contribute to green initiatives (Ho *et al.*, 2017). Susilo *et al.* (2012) found that a modal shift is dependent on the age of the population towards which the initiative is focused. This research is consistent with the observation that females and older groups contribute greatly to increasing annual car mileage even though average car use has dropped (DfT, 2015).

Local authority policy dictates investment in transport and this in turn influences travel choices. Therefore, one of the key objectives of investigating travel behaviour is to help policy makers and other stakeholders to develop policies that make travel behaviour more sustainable (Banister, 2008). Among other things, this involves reducing car travel and reinforcing travel by public transport, bicycle and walking. In order to take relevant actions on behavioural change, it is important to understand travellers' attitudes, the barriers, and preferences in mode choice decisions. Therefore, the next section presents a review of studies of the potential to switch transport modes from private to sustainable alternatives.

2.7.2 Recent studies of modal shift potential

The increase in numbers of private vehicles has resulted in undesirable side effects, including congestion and air pollution (Hensher, 2008; Stanley *et al.*, 2011). Based on transport statistics in Great Britain in 2014, 68% of people travelled to work by car, 240 billion vehicle miles travelled by car were recorded and only one in ten people walk to work (DfT, 2014).

In order to minimise traffic-related issues, governments in the UK and overseas have developed policies to encourage travellers to use their cars less and public transport more. Continuous on-going discussions as well as reports, guidelines and recommendations demonstrate the ongoing attempts of the UK government to encourage a modal shift from private cars to walking and cycling for short journeys (DfT, 2004; DETR, 2005; DfT, 2010; DoH, 2010). Based on a report published by the UK Department for Transport (DfT, 2011), transport policies are progressively being directed towards shifting travel from car use to walking and cycling.

It is unlikely that significant modal shifts will be achieved without more vigorous engagement to render walking and cycling easier and more attractive, as well as making the option of using cars less attractive (Semanjski *et al.*, 2016). On the other hand, a study by Verplanken and Orbell (2003) suggested that reflections on past behaviour have proved that mode shift is initiated by changing habits. Meanwhile, Ho *et al.* (2017) points out that attitudes to sustainable transportation whilst also extremely important working to change them should also affect the design of greener travel initiatives.

Effective modal shift should aim to contribute either directly or indirectly to the development of a sustainable urban environment. Modal Shift Potential (MSP) comprises either a variety of modes or a specific mode depending on the travellers considered, with a target of optimising temporal, environmental, social and economic benefits (Redman *et al.*, 2013). The indicators that can be used to define sustainable transport goals and to monitor whether or not the current transport system is moving towards sustainability are strongly related to the specific policies of a country. The influences toward healthier and more sustainable patterns of behaviour in travel mode choice research were found as the result why people travel as they do (Thomas and Walker, 2015).

Previous studies appear not to have considered environmental concerns regarding private transport use perhaps because it was not deemed an important factor affecting the travel mode choices of respondents. However, some evidence shows that the combination of different measures of environmental awareness measures leads to views that can be targeted in order to change behaviour (Anable, 2005). Also, Thomas and Walker (2015) suggest that environmental global views are proportionate to the strength of an individual's attitudes concerning environmental problems. Research into the problem of susceptibility to switching modes has begun to raise issues of the environment and the source of car-dependent attitudes, and thus a number of developments have applied psychology to the study of mode choice (Anable, 2005). However, Okushima (2015) and Juho *et al.* (2014) failed to investigate environmental awareness aspects in their study, even though it is one of the important elements that should be considered. It is clear from the literature reviewed above that different views concerning the environment found among travel mode user groups was neglected and the author believes this must be taken into account to help influence behaviour change.

2.8 Research Gap

This chapter has reviewed a wide range of research topics related to the main subject of this study, starting from transport demand management followed by the challenges in transport and environmental issues, travellers' attitudes to environmental issues, behavioural changes through transport policies, and achieving a sustainable transportation system. These studies were considered in four areas: a) attitudes and travel behaviour; b) sustainable transportation; c) environmental issues and d) modal shift potential, acknowledging these areas are interconnected with each other as shown in Figure 2.1. However, in previous studies, researchers were found to focus their investigations in one or more, but not all of these areas as shown in Table 2.6.

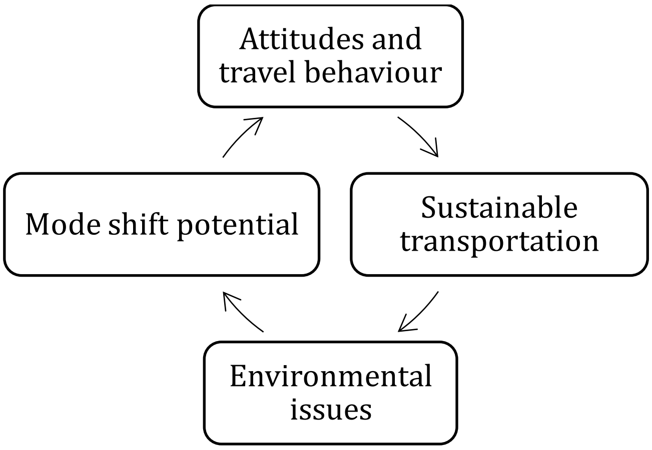


Figure 2.1: Interconnected research areas in this study

Previous research	Topics			
	Attitudes and behaviour	Sustainable transport	Environment / climate change issues	Modal shift potential
Curtis and Headicar (1997)	√			√
Urwin and Jordan (2008)			√	
Thomas and Walker (2015)	√		√	
Anable *et al.* (2006)	√	√	√	
Schneider (2013)	√	√	√	
Kingham *et al.* (2001)	√			√
Clifton and Handy (2003)	√			
Van Wee and Holwerd (2002)	√			
Mikiki and Papaioannou (2012)	√	√	√	
Hunecke *et al.* (2010)	√		√	
Barr *et al.* (2011)	√	√	√	
Prillwitz and Barr (2011)	√	√	√	
Susilo *et al.* (2012)	√	√	√	

Table 2.6: Comparison of the object of investigation of several studies on sustainable mobility targets.

The study by Curtis and Headicar (1997) focused on attitudes and travel behaviour as well as modal shift potential, but did not consider environmental and climate change aspects in their research. Whilst, Urwin and Jordan (2008) investigated climate change and explored policy across different scales of government; however, the study did not explore the potential groups of travellers on modal shift. In addition, Thomas and

Walker (2015) conducted a large-scale survey of drivers, walkers, bicyclists and bus users commuting to a UK university to determine levels of trip satisfaction and the strength of habits for different travel modes, but did not consider environmental issues and potential groups of people to switch modes.

Nevertheless, a study of policy intervention was conducted to generate further cross-mode assessments which consequently will be beneficial for enlightening travel mode involvements (Anable *et al.*, 2006). A more comprehensive study was conducted by Schneider (2013) using the theory of routine mode choice decisions in aiming to increase sustainable transportation. This research suggested steps to increase pedestrian and bicycle users through the environmental benefits of walking or bicycling. However, the study did not investigate policy recommendations in the research on modal shift potential. Studies by Kingham *et al.* (2001) focused on modal shift, attitudes to travel and the segmentation of groups. However, they did not attempt to investigate environmental and climate change issues in their research.

Analyses consistently show that some attitudinal factors are often more significant predictors of travel behaviour than traditional variables (Clifton and Handy, 2003). Combining psychological, socio-demographic and infrastructural variables can enrich predictions (Van Wee and Holwerd, 2002) and including environmental and attitudes may permit the identification of relevant user profiles (Mikiki and Papaioannou, 2012). Such attitude-based approaches can provide important information on environmental measures and aspects of different mobility behaviour (Hunecke *et al.*, 2010). They can thus contribute to the more effective promotion of sustainable behaviour (Barr *et al.*, 2011).

Prillwitz and Barr, (2011) studied the combination of factor and cluster analyses on attitudes towards certain modes of transport, and attitudes towards the environment and sustainability, to enrich explanatory models for individual travel behaviour and deliver helpful additional information for potential policy and planning measures. In terms of attitudes toward and behaviours in relation to the environment, almost all respondents were aware of environmental issues, but their views did not necessarily 'match' their travel behaviour (Susilo et al., 2012).

Although the evidence provided by those studies seems to support the background of this research, some issues still remain as the investigations were conducted separately and there remains a need for further research to be conducted.

The LAs especially Passenger Transport Executives (PTEs) might gain valuable information from this study concerning whether to invest more money to provide more facilities to encourage the positive use of sustainable transport or to invest in public transport services. Driver and Vehicle Licensing Agency (DVLA) might use the result of this study to seriously consider alternatives to higher taxes (road users charging) for private transport and instead lower ticket prices for public transport and other sustainable incentive schemes. Eriksson *et al*. (2008) and Redman *et al*. (2013) present some evidence to suggest that travellers expect to pay less for tickets to encourage them to use train services more. Besides that, the targeted group could be a potential market for car companies to produce lower CO_2 emissions cars. As well as it could be benefited for Department for Communities and Local Government (DCLG) together with DfT for national planning policy framework of future sustainability.

2.9 Conclusions

Numerous studies, as described in this section have attempted to explain the potential of modal shift to sustainable transport. Curtis and Headicar (1997) suggested that travel awareness campaigns should be taken up by LAs as a strategy to encourage people to switch their mode of travelling. Nurdden *et al.* (2007) discussed the efficiency of transport policies in promoting modal shift from private cars to public transport in Malaysia and found that age, gender, car ownership, travel time, travel costs, household size and income are significant factors in influencing the individual's choice of transportation. However, whilst it is acknowledged that targeting initiatives at specific groups is important (Susilo and Cats, 2014), studies on the best groups or individuals to target for new sustainable policy implementations are scarce. Thus, this study aims to develop a methodological approach and investigate which individuals or groups are more likely to move away from private transport and are therefore the best targets for marketing strategies and awareness campaigns to encourage modal shift, considering the UK as a case study.

The BSA dataset has been used to analyse public attitudes to environmental concerns since the early 1990s. Taylor and Brook (1998) showed as time has passed over a period of almost ten years, an improved level of acceptance of sustainable policies has begun to be witnessed. However, despite this observation, they reported that peoples' attitudes to the environment tended not to be used to support policies designed to lessen environmental impacts.

After careful investigation it has been identified that the BSA dataset is a suitable archive for analysis to enable policymakers to understand not only public attitudes to transport policy and how they vary between subgroups, but also why different people have specific attitudes, as this will be fundamental in influencing behavioural change.

The above review has motivated the research proposed in this thesis. It is important to obtain a deeper understanding of the relationship between travel behaviour and the encouragement of sustainable travel. This research expects to obtain new evidence and develops a novel mathematical model as a complement to the evidence already available to move discussion about the topic forwards. In order to develop an effective method to achieve the research objectives, analytical methods applied to categorical research are reviewed and discussed in the next chapter.

Chapter 3 Analysing Attitudinal Data – Critical Review of Methodologies

3.1 Introduction

The previous chapter critically reviewed previous studies relating to attitudes and travel behaviour, transportation and environmental issues, sustainable transportation, and the potential of modal shift. This chapter will carry out a critical review of the methodologies that can be applied to analyse attitudinal, travel and demographic characteristics data.

Attitudinal data generally is categorical. Therefore, a key challenge here is to identify a form of quantitative analysis which can be used to evaluate qualitative measures associated with categorical data. More specifically, any method devised needs to provide an assessment of car users' perceptions and attitudes towards travel aspects, including modes, distances, traffic congestion and environmental issues.

In this chapter, Section 3.2 provides an overview of research approaches. Section 3.3 discusses methodological approaches for categorical data used in previous research. Mathematical methods of modelling travel behaviour are then discussed in Section 3.4, followed in Section 3.5 by a comparison of inferences applied in travel behaviour research. Finally, a conclusion of this chapter is drawn in Section 3.6.

3.2 Quantitative and Qualitative Research Approaches

In general, quantitative approaches place emphasis on developing an approximation of the situation based on a sample of subjects using survey methods and applying statistical techniques to distinguish overall patterns. Instead of drawing conclusions subjectively through perception, consideration or intuition, it is fairly easy to survey people using more scientific and objective methods. This leads to more quantitative approaches which have the potential to provide accurate measurements of people's behaviour. Also, as people's responses are in numerical form, various statistical techniques and operations can be applied. Accordingly, clearer results can be obtained and generalizations can be made.

Qualitative approaches on the other hand focus on understanding a particular situation and examining or interpreting the statistically significant factors that are difficult to quantify using traditional quantitative approaches. In most problems related to people's perceptions, attitudes and behaviour, qualitative approaches are used as a supplementary method to a quantitative approach in order to understand more clearly the subject matter (Clifton and Handy, 2003; Ortúzar and Willumsen, 2011). Qualitative approaches include in-depth interviews, brainstorming, paired interviews, telephone interviews, participant observation, open-ended questioning and focus groups. Methods such as focus groups, interviews, and participant-observer techniques (Clifton and Handy, 2003) and 'action research' transport survey methods (Lucas, 2013) can be used in conjunction with quantitative approaches or on their own to fill the gaps left by quantitative techniques.

Prior to the work of Grosvenor (2000), Pendyala and Bricka (2006), and Clifton (2013), studies focused on the development and application of social survey methods for understanding travel behaviour by using quantitative and qualitative survey instruments. Qualitative designs are more flexible and aim to explore what people think and how they behave and this usually involves knowledge-gathering and observation (Kumar, 2011). For instance, Beirão and Cabral (2007) and Line *et al.* (2010) used qualitative approaches to understand attitudes towards public bus transport and private car use. Meanwhile, Carrasco and Lucas (2015) suggested that when measuring attitudes concerning, and perceptions of, people's travel choices, both quantitative and qualitative approaches are useful methods.

It has been noted that this type of research can be designed using both the quantitative and qualitative types of methods, referred to as mixed methods. In many cases, quantitative and qualitative approaches are ultimately complementary techniques, rather than alternatives. However, considering the objective of this research, which seeks to investigate the characteristics of group(s) of people, based on their socio-demographic characteristics, who may have more potential to shift towards sustainable transportation, both methods are equally important. The qualitative approach offers an effective way to generate attitudinal perceptual data that are predominantly categorical in nature, whilst the quantitative method mostly deals with numerical data. Sections 3.2.1 and 3.2.2 are used to explain the details of attitudinal variables and categorical data respectively.

3.2.1 Attitudinal variables

Attitudinal variables are collected by using questionnaire surveys or structured/semi-structured interviews designed to measure respondents' opinions on a particular subject, either products or services, or to identify their feelings about something (Morey, 2006). The attitudinal variables are generally combined with other types of data, such as the socio-demographic characteristics of the respondents, in order to obtain a more in-depth understanding of a subject.

Qualitative research has long been criticised for its lack of scientific rigour and subjective interpretation (Sandelowski, 1986). In order to avoid these weaknesses, there is a growing trend of using attitudinal questionnaire surveys combined with numerical data to provide a richer understanding of "attitude-caused travel behaviour" (Clifton and Handy 2001). As suggested in psychological research, attitudes are powerful elements of people's actions (Kollmuss and Agyeman, 2002; Howarth, 2006) and taking these into account is crucial for the success of new strategies designed to reduce private car driving and promote pro-environmental travel behaviour (Nilsson and Küller, 2000).

3.2.2 Categorical data

In many fields such as psychology, science and transportation, categorical variables are commonly used when designing surveys. A set of non-overlapping variables are called categorical variables (Salkind, 2010). In transportation research, multinomial or binary logistic regression seem to be popular, especially when analysing categorical data (Al-Ghamdi, 2002; Li et al., 2016) and in log-linear modelling (Jang, 2006; Olmuş and Erbaş, 2012; Samimi, 2012). User preferences for mode choice, journey related variables, socio-demographic characteristics, and attitudinal factors have been used when analysing categorical data (Anable, 2005; Anable and Gatersleben, 2005; Steg, 2005). Categorical data can be divided into two types, ordinal and nominal.

A measurement scale is ordinal if the categories can be ranked, such as perception variables with options ranging from "strongly agree" to "strongly disagree", or "not important" to "very important". However, a measurement scale is nominal if the categories have no ordering, such as colour (for example: red, blue, and green) and gender (male and female). Rating scales are normally used to define categorical variables over a range from lower to upper values. Several kinds of evaluation measures have been developed to rate attitudes, and the most commonly used is the Likert scale. Accordingly, responses using Likert scales are usually treated as ordinal data (Bertram, 2006), where respondents can give a numerical response within a range of incremental scores. These are often on a scale with a range of for example 1 to 5 or 1 to 7 where 1 is labelled "strongly agree" and the upper value "strongly disagree". An odd number of intervals is given so that a respondent can be neutral in their responses. In general, this technique is easy to manage and adopt and is a suitable method for gathering numerical data for non-physical latent variables such as of respondents' awareness, perceptions, opinions, attitudes, intentions and preferences.

Latent variables are referred to as variables that cannot be directly observed. The variables of this type therefore are used in the questionnaires as the indicators to measure the perceptions and attitudes of respondents with respect to their preferences and intentions.

3.3 Methods Used in Analysing Categorical Data

Various methods have been employed to explore categorical data in travel behaviour research. A number of techniques have been developed to analyse categorical data such as factor analysis for dimension reduction and cluster analysis and multiple correspondence analysis to allocate respondents to groups. Multinomial logistic regression and multivariate probit model are generally used to explore any relationships that exist between dependent and independent variables. Each method is discussed in detail in the following sections.

3.3.1 Dimension reduction using exploratory factor analysis (EFA)

Factor analysis is used to simplify large sets of data in order to reduce the number of variables and to explore in further detail any structures in the relationships between the variables, establishing those that are independent and those that are not independent. Variables highly correlated are collected together into a new variable called a factor (Costello and Osborne, 2005). So factor analysis is more "model based". PCA can be seen as a first step in factor analysis. There is no rotation in PCA but rotation is used in factor analysis (Fabrigar *et al.*, 1999). There are four requirements or assumptions for a dataset to be suitable for factor analysis which are normality, linear relations, factorability, and sample size.

Two factor analysis methods were considered for use in this research, namely Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). EFA and CFA techniques are similar, but their purposes are different. EFA is used to provide a number of factors with specified variables that are assimilated into factors, whilst CFA is used to examine whether expectations concerning the factor structure are indeed present. EFA is briefly described in this section as it is more applicable to the objectives of this study. This is because EFA is data focussed, whereas CFA is based on theory or empirical research. In the initial steps of scale development, EFA is suitable since items loading on the non-hypothesised factors did not show in CFA. The origins of factor analysis can be traced back to Pearson (1901) and Spearman (1904), and the term was first introduced by Thurstone (1931). A review of the subject can be found in Gorsuch (1983).

The seven steps involved in EFA, suggested by Yong and Pearce (2013) have been used to create a conceptual diagram for the approach adopted in this research as shown in Figure 3.1.

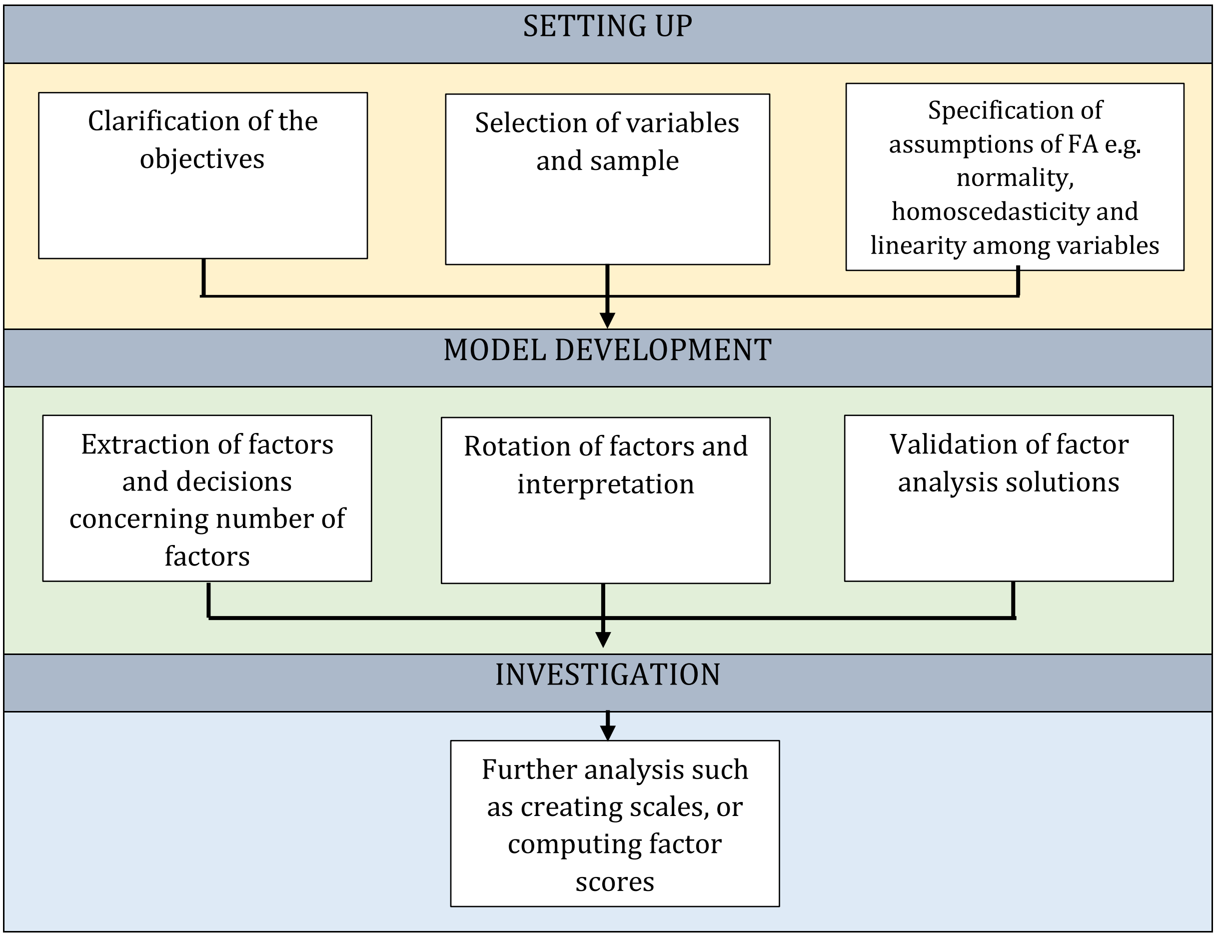


Figure 3.1: Steps involved in factor analysis

EFA is conducted using the correlation coefficients between variables and factors referred to as factor loading (Field, 2009). The squared factor loading represents the percentage of variance explained by a factor. If the observed variables are $X_1, X_2, \dots, X_n$, the common factors are $F_1, F_2, \dots, F_m$ and the unique factors are $U_1, U_2, \dots, U_n$, where the variables may be expressed as linear functions of the factors:

$$X_1 = a_{11}F_1 + a_{12}F_2 + a_{13}F_3 + \cdots + a_{1m}F_m + a_1U_1$$

$$X_2 = a_{21}F_1 + a_{22}F_2 + a_{23}F_3 + \cdots + a_{2m}F_m + a_2U_2$$

$$\vdots$$

$$X_n = a_{n1}F_1 + a_{n2}F_2 + a_{n3}F_3 + \cdots + a_{nm}F_m + a_nU_n$$

Each of these equations is a regression equation; EFA is used to find the coefficients $a_{11}, a_{12}, \dots, a_{nm}$ which best reproduce the observed variables from the factors. The coefficients $a_{11}, a_{12}, \dots, a_{nm}$ are weights in the same way as regression coefficients.

In EFA, the coefficients are called factor loadings and when the factors are not correlated, they also show the correlation between each variable and a given factor. In the model above, a_{11} is the loading for variable X_1 on F_1, a_{23} is the loading for variable X_2 on F_3, and so on.

The sum of the squares of the loadings for variable X_1, labelled as ${a_{11}}^2 + {a_{12}}^2 + \cdots + {a_{nm}}^2$, shows the proportion of the variance of variable X_1 which is accounted for by the common factors. This is called communality. EFA solutions are more successful if larger values of communality are obtained for each variable.

There are various measures used to identify the inter-correlation among variables. The Kaiser Meyer Olkin (KMO) measure is used to assess the degree of correlation among variables (Field, 2009). Hair *et al.* (2006) recommended this measure as being appropriate to deliver a specific level of confidence of the prediction value ranging from 0.9 – 1.00 to be perfectly predicted down to a value less than 0.50 for unacceptable results. The minimum sample required for PAF is 50 (Hair *et al.*, 2006).

Costello and Osborne (2005) argued that principal component analysis is one of the factor analysis techniques used for data reduction which produces “components” while principal axis factoring produces “factors”. However, both are acceptable and used depending on the research questions and ease of interpretation of the results (Yong and Pearce, 2013). Also, they pointed out that maximum likelihood extraction method is more suitable for CFA and is used to estimate the factor loading for a population.

There are two rotation methods in factor analysis, orthogonal rotation and oblique rotation. In orthogonal rotation, factors are assumed not to be correlated and are rotated 90^0 from each other. Quartimax and Varimax are commonly selected in orthogonal method. In general, Quartimax rotation concentrates on variables, whereas Varimax focuses on factors. More specifically, Quartimax and Varimax rotations maximise the variance of the squared factor loadings in variable and factor, respectively.

In oblique rotation technique, factors are assumed to be correlated and are not rotated 90^0 from each other. Oblimin and Promax are two common oblique rotation methods. Oblimin attempts to simplify the structure of the output. However, Promax is advantageous because of its speediness in raising the loading in a larger dataset to reach a simple structure (Gorsuch, 1983).

In transportation studies, factor analysis has been widely applied to analyse categorical data, for example Anable (2005); Steg (2005); Van *et al.* (2014); Kamruzzaman *et al.* (2016); Molin *et al.* (2016); and Batool and Carsten (2017). Table 3.1 presents the differences in choice of extraction and rotation methods that have been applied in previous studies to conduct EFA.

Example of study	Topics	Details of the research	Extraction method		Rotation method		
			PCA	PAF	Varimax	Promax	Oblimin
Steg (2005)	Car use; Motives; Symbolic function.	Examined important factors that influence the level of car use and found three main attributes: symbolic and affective, instrumental, and independence.	√		√		
Anable (2005)	Attitudes; Cluster analysis; Market segmentation; Mode choice; Theory of planned behavior.	Generated 19 factors out of 105 attitudinal statements and they included: moral norms, general attitudes towards the car, environmental beliefs, social and behavioural norms, and perceived behavioural control. The results demonstrate that different reasons could influence the same behaviour, however different behaviours could lead to the same attitudes.	√		√		
Cao *et al.* (2009)	Land use; Seemingly unrelated regression; Self-selection; Smart growth; Travel behavior; Urban design.	Collected 32 attitudinal variables concerning travel in Northern California using Likert scale options from “strongly agree” to “strongly disagree” and performed factor analysis. Generated six attitudinal factors: pro-bike or walk, pro-transit, pro-travel, travel minimizing, car dependent and safety of car. They discovered that the car dependent factor is positively associated with higher frequency of car trips and lower frequency of transit trips.	√		√		
Molin *et al.* (2016)	Attitude; Factor analysis; Latent class cluster analysis; Mode choice; Mode frequency; Multimodality.	Identified travellers’ attitudes using 19 statements mostly about public transport (PT) and acknowledged seven perception factors, namely; PT transfer acceptability, PT waiting acceptability, car inexpensive, PT timeliness, PT seat availability, PT planning ease and PT inexpensive. They concluded that changes to the built environment influence walking using the cross-sectional analysis, where there were significant relationships between environment factors with travel attitudes and perceptions and socio-demographics.		√			√

Continued on the next page

Table 3.1 (continued)

Example of study	Topics	Details of the research	Extraction method		Rotation method		
			PCA	PAF	Varimax	Promax	Oblimin
Kamruzzaman *et al.* (2016)	Residential self-selection; Travel attitudes; Urban form; Walking as a mode of transport.	Conducted factor analysis by using 16 questions of a 5-point Likert scale survey data regarding travel attitudes and perceptions. Four factors were identified and they were: respondents' perceptions about PT, sensitivity to environmental externalities, car dependency and safety of car travel. They concluded that changes to the built environment influence walking using the cross-sectional analysis, where there was a significant relationship between environment factors with travel attitudes and perceptions and socio-demographics.		√			√
Batool and Carsten (2017)	Driving behaviour; Extended violations scale; Developing countries; Road safety.	Investigated drivers' unusual actions in Pakistan and found that personal characteristics resulted in four behavioural factors which are aggressive driving, unlawful driving, risky driving and egoistic driving.	√			√	
Fatmi and Habib (2017)	Attitude; Changes in household state; Mode switch; Past travel behaviour; Random parameters logit model; Residential relocation.	Explored the long-term choices of commuters in terms of travel modes shift when relocating to another residential area. Factor analysis was used to produce 6 attitudinal factors based on 295 samples from the Household Mobility and Travel Survey (HMTS). The model results suggest that mode shift decisions are significantly influenced by previous travel experiences. Moreover, larger household size, closer location to transit stops and driver's licence ownership also influence travel mode switch decisions.	√			√	

Notes: *PCA: Principal component analysis **PAF: Principal axis factoring

Table 3.1: Previous studies using different extraction and rotation methods in dimension reduction technique.

Along the lines of the discussions from previous studies, it is clear from Table 3.1 that EFA is used in a wide range of areas that embrace, for example, attitudes, travel choices, driving behaviour and land use. All seek to reduce the number of variables measured in qualitative surveys such as questionnaires and interviews, to remove commonality and any inherent correlation. Interestingly the sample size and number of variables vary widely from 105 to 19 and 295 to 6 respectively.

Based on previous research, PCA and PAF are among the popular dimension reduction techniques chosen by researchers. Steg (2005), Anable (2005), and Van *et al.* (2014) used PCA with Varimax rotation method. Whilst, Molin *et al.* (2016) and Kamruzzaman *et al.* (2016) selected PAF extraction method with Oblimin rotation method to conduct EFA in their studies. Whereas, recently, Batool and Carsten (2017) and Fatmi and Habib (2017) applied PCA with Promax to find factors involving driving behaviour and attitudes towards switching travel modes, respectively.

Several attempts have been made, using PCA and PAF, to investigate attitudinal variables to obtain important factors regarding car use (Steg, 2005; Batool and Carsten, 2017), public transport (Molin *et al.*, 2016), travel attitudes and behaviour (Anable, 2005; Kamruzzaman *et al.*, 2016) and travel mode shift (Fatmi and Habib, 2017). Based on previous research, this method demonstrated its suitability for the analysis of Likert scale survey data regarding travel attitudes and perceptions in order to investigate behavioural and attitudinal related factors (Kamruzzaman *et al.*, 2016) and, more specifically, to investigate travel mode choice decisions (Fatmi and Habib, 2017).

3.3.2 Exploring the data structures with socio-demographic and travel behaviour - data clustering

Cluster analysis is an exploratory statistical tool to identify the existence of similarities or patterns in responses (Tabachnick and Fidell, 2013). It is used to create and identify groups that are homogeneous within the total dataset and to segregate respondents based on similar characteristics. In other words, cluster analysis groups respondents with particular characteristic(s) based on their socio–demographic and travel profiles. By identifying the sub-groups, it becomes simpler to analyse and reveal any statistically significant relationships which may or may not emerge among the sample but within and between the groups. Homogeneity or similarity is measured using the strength of a relationship by using the distance between the pair of objects. The smaller the distance, the more similar the objects are within the cluster.

In travel behaviour research, there are many studies conducted using cluster analysis to explore different groups of respondents by using socio-demographics and travel related variables. Table 3.2 shows examples of previous studies that have employed three different cluster analysis techniques: K-means, hierarchical, and two-step clustering predominantly using categorical data.

Examples of study	Topics	Details of the research	Cluster analysis	
			K-means	Hierarchical
Kandt *et al.* (2015)	Cluster analysis; Comparative study; Transport policy; Travel attitudes; Travel behavior.	Grouped travellers into 6 categories based on travel attitudes in Berlin and London which vary in sustainable mode shift potential. The variables considered were socio-demographic, travel behaviour, residential location, future intentions and behavioural change when investigating actual travel behaviour. The 6 clusters were traditional car-oriented, pragmatic transit-sceptics, green travel-oriented, pragmatic transit-oriented, technology-focused individualists and innovative access-oriented. The results revealed a strong relationship between attitudes and behaviour.	√	√
Bösehans and Walker (2016)	Active travel; Commuting; Public transport; Segmentation; Sustainable transport; Travel behavior.	Investigated bus users' experiences to identify the characteristics of groups within the total sample which have potential to switch to sustainable modes. An on-line survey of students and staffs who use the bus into a UK university was carried out. 6 different groups of bus users were classified, namely mode mixers, wannabe walkers, all fine on the Weston Front, first fans, car curtailed and daily drags. Among those 6 groups, they concluded that the "wannabe walkers" bus user group would be the best target for the future interventions and sustainable campaigns.	√	√
Papadimitriou *et al.* (2017)	Cluster analysis; Human factors; Pedestrian behaviour; principal component analysis.	Investigated human factors of pedestrian walking and crossing behaviour in urban areas. They conducted a survey using 54 questions with 5-point Likert scale options. Around 75 respondents (young and middle-aged) completed the questionnaire. 2 clusters of pedestrians were identified. These were: "positive and motivated – those who have positive attitudes and strong motivations to walking" and "negative and unmotivated – those who have negative attitudes and weak motivations to walking".		√
Rodriguez Cote and Diana (2017)	Cluster analysis; Market segmentation; Multimodality; Traveller profiles.	Collected 164 responses through an online survey to systematically measure attitudes and behaviours towards different modes of transportation. Four different clusters were obtained: low multimodal, open to change; car monomodalists, eager to change; multimodal, carpool interested and low multimodal creatures of habit. Each cluster was treated differently when promoting more sustainable travel behaviours.	√	

Continued on the next page

Table 3.2 (continued)

Examples of study	Topics	Details of the research	Cluster analysis	
			K-means	Hierarchical
Briand *et al.* (2017)	Longitudinal analysis; Mixture model; Passenger clustering; Public transit; Smart card.	Presented a two-level generative model that applies the Gaussian mixture model to regroup passengers based on their temporal habits in their public transportation usage. Five years of data collected by the Société de transport de l'Outaouais were used. The results show that the efficiency of the proposed approach in identifying a reduced set of passenger clusters linked to their fare types.		√
Mondschein and Parkany (2017)	Demographics; Emissions; Sustainability; Travel behaviour; Travel choices; Voting precincts.	In an investigation carried out over the past half-decade using a survey data of 2373 respondents. The differences in sustainability and socio-demographics within the clusters displayed an increasing trend within the state over time. Distance and CO_2 emissions were combined to achieve a proxy for green travel and identified 4 precinct-based sustainability clusters: sweet spots, emerging sweet spots, neutral and non-sustaining. The results showed that a significant relationship was found between different demographics and politics when analysing the demographic dissimilarities among the clusters.	√	
Jin *et al.* (2018)	Attitudinal Aspects; Cluster Analysis; Factor Analysis; Managed Lanes; Market Segmentation; Value of Reliability (VOR); Value of Time (VOT).	Observed how drivers' selections in utilising managed lane facilities were affected by behavioural attitudes. Based on the data from the South Florida Expressway Stated Preference Survey, 4 clusters of users regarding their attitudes toward tolling, schedule shifting, time savings, and congestion emerged from the k-mean cluster analysis. These were derived from four latent attitudinal factors: willingness to pay, willingness to shift travel schedule, utility (cost/time) sensitivity, and congestion tolerance resulting from ten attitudinal statements reduced by factor analysis.	√	

Table 3.2: Examples of previous studies applying the use of cluster analysis techniques

According to Mooi and Sarstedt (2011), the steps in conducting cluster analysis are as shown in Figure 3.2, these being the steps which are adopted in the research.

Identifying variables
Personal characteristics such as socio-demographic data (age, employment status, household size, car ownership, and gender) and travel behaviour data (frequency of using car as a driver or passenger, frequencies of using bus, train and bicycle).

Select clustering procedure
Variables used are ordinal and nominal, therefore hierarchical cluster analysis is the appropriate method.

Deciding the number of clusters
The number of clusters is decided differently depending on the clustering methods chosen. In the hierarchical method, the number of clusters is obtained by cutting a dendrogram at a certain level. K-Means method was not used because there was no reason/evidence to constrain the number of clusters.

Validating and interpreting cluster solutions
The stability and validity of the cluster analysis solution obtained needs to be examined. When the applications of different clustering techniques on the same data produce statistically significantly similar results, the cluster analysis solution is considered to be stable. Validation was achieved by running the K-means cluster analysis using the number of cluster ±1 defined by HCA.

Figure 3.2: The steps in cluster analysis applied to the research

Since attitudinal variables can be considered as categorical data, both K-means and hierarchical clustering methods have been applied in previous research instead of a two-step clustering method. This is because two-step clustering methods are applicable for analyses with a combination of both categorical and continuous data. Given the application criteria of each clustering approach and mindful that this research uses categorical data, the decision was made that hierarchical cluster analysis is the most suitable approach to be used in this study when compared to K–means. This is because in K-means, the number of clusters have to be defined in advance and this is unknown. Hierarchical clustering method is a method which identified the number of statistically significant groups and is discussed in more detail in next section.

Hierarchical cluster analysis

In the hierarchical clustering procedure, the objects consecutively form clusters. Hierarchical clustering has two categories which are agglomerative and divisive clustering. In agglomerative clustering a bottom-up approach, each object represents a cluster and then clusters are combined according to their similarity. In contrast, in divisive clustering, a cluster is produced top-down from a single cluster which subsequently is divided increasingly to create further clusters at lower levels. The results of hierarchical clustering are usually presented in a dendrogram. Agglomerative clustering is more popular and is often used in market research (Everitt *et al.*, 2011) whilst divisive clustering is used in scientific studies.

Euclidean distance is used as a measurement of similarity in hierarchical cluster analysis. A straight line between two objects is used to measure their proximity in terms of a Euclidean distance, and G and H are two individual measures of a variable availability. The closeness, referred to as $d(G,H)$, of the variables x and y is calculated using equation 3.1 (Liberti *et al.*, 2014):

$$d_{Euclidean}(G,H) = \sqrt{(X_G - X_H)^2 + (Y_G - Y_H)^2} \quad (3.1)$$

where:

$d_{Euclidean}(G,H)$ = Euclidean distance between individuals G and H

X_G and X_H = values of variable X for individuals G and H respectively

Y_G and Y_H = values of variable Y for individuals G and H respectively

The same computation method can be used to calculate the distance between all pairs of data and written by means of a distance matrix. City-block-distance and the Chebyshev distance are other methods used in assessing distance. The city-block-distance uses the sum of the variable's absolute differences, as shown in the equation 3.2 (Souza and Carvalho, 2004):

$$d_{ij} = \sum_{k}^{n} |X_{ik} - X_{jk}| = |X_G - X_H| + |Y_G - Y_H| \quad (3.2)$$

where:

d_{ij} = City block distance between *i* and *j*

X_{ik} = coordinate attribute *k* of point *i*

X_{jk} = coordinate attribute *k* of point *j*

X_G and Y_G = coordinate G

X_H and Y_H= coordinate H

The Chebyshev distance $d_{Chebyshev}\ i,j$ uses the maximum difference of the absolute difference in the clustering variables, as described in the equation 3.3 (Cantrell, 2000):

$$d_{Chebychev}\ i,j\ = Max_i\ X_k - Y_k \tag{3.3}$$

Where X_i and Y_j are the values of the *i*th variable at points X and Y respectively.

The CA methodology adopted in this research is based on a statistical process that uses the parameters of distance and similarity, where ‘distance’ is a measure of how far apart two objects are, whilst ‘similarity’ considers how similar the two objects are.

Multiple Correspondence Analysis (MCA)

Multiple correspondence analysis (MCA) is particularly suitable to explore unpredicted dimensions and relationships in the tradition of exploratory data analysis. MCA is a data analysis technique for categorical data, used to detect and represent underlying structures in a data set by representing data as points in a low-dimensional Euclidean space (Costa *et al.*, 2013; Greenacre and Blasius, 2006). MCA aid the visualision where the data represented individually, allowing improved understanding of each group of people surveyed. The MCA is based on positive integer data. All variables have the multiple nominal scaling levels and the data must contain at least three valid cases. The results of the MCA can be presented analytically (results tables) and visually (graphs and charts). In addition, MCA positions each variable as a point in a low-dimensional space and helps to describe patterns of relationships, particularly using geometrical methods.

In this context, the respondent's answer to an attitudinal question is considered to be a categorical variable. Hence, a respondent is characterised into a cluster by the answers given. Two respondents are closer if they have given common answers in a higher number of categories. In MCA, the distance between two respondents i and l is given by equation 3.4 (Greenacre and Blasius, 2006):

$$d^2(i,l) = \sum_k \frac{NJ}{I_k}\left(\frac{x_{ik}}{J}-\frac{x_{lk}}{J}\right)^2 = \frac{N}{J}\sum_k \frac{N}{I_k}\left(x_{ik}-x_{lk}\right)^2 \tag{3.4}$$

where:

x_{ik} = is equal to 1 if the respondent i has taken the category k and 0 otherwise,
x_{lk} = is equal to 1 if the respondent l has taken the category k and 0 otherwise,
N = is the total number of respondents,
I_k = is the number of respondents who are assigned to category k, and
J = is the number of variables.

The expression $(x_{ik}-x_{jk})^2$ is either equal to 0 or 1. The distance $d^2(i,j)$ increases with the number of different categories for both respondents. A category k takes part in this distance formula with a weight equal to $\frac{N}{I_k}$, which corresponds to the inverse of the category's frequency. This means that respondents having an uncommon category are separated from all other respondents.

More recently, MCA has appeared in transport research covering the topics of vehicle accidents (Das and Sun, 2015; Xu et al., 2015; Billot-Grasset et al., 2016; Jalayer and Zhou, 2017), airline services (Wen and Chen, 2011) and public transportation (Grison et al., 2015; Truong and Somenahalli, 2015) which mostly deal with latent variables. Table 3.3 shows example of studies related to transport that have used MCA and presents brief details of each study.

Study	Author	Topic	Details of the research
Factor Association with Multiple Correspondence Analysis in Vehicle–Pedestrian Crashes	Das and Sun (2015)	Crash injuries; Demographics; Fatalities; Highway design; Injury severity; Pedestrian-vehicle crashes.	MCA was used to analyse the data collected during 2004-2011 of vehicle–pedestrian crashes in Louisiana, USA to identify the relative closeness of the key association of factors. The findings indicated the nontrivial focus groups such as drivers with high-occupancy vehicles, female drivers in bad weather conditions, and drivers distracted by mobile phone use. They concluded that male drivers were seen to be relatively susceptible to severe and moderate injury crashes.
Quasi-Induced Exposure Method for Pedestrian Safety at Signalized Intersections	Xu *et al.* (2016)	Pedestrian crash; Signalised intersection; Quasi-induced exposure method; Multiple correspondence Analysis.	MCA was used to analyse pedestrian-vehicle crashes data in Las Vegas from 2004 to 2008 to investigate how pedestrian crash severity is influenced by the interactions among a range of variables. The results demonstrate that overall crash severity, light conditions, and weather conditions are potentially dangerous factors in pedestrian crashes, and the time of day or day of the week are less risky.
How cyclist behaviour affects bicycle accident configurations?	Billot-Grasset *et al.* (2016)	Cycling accident typology; MCA; Medical register; Survey.	MCA was applied to check the consistency of identified topology and provided additional insight concerning road user behaviour by projection of supplementary variables. Surveyed data of injured cyclists in the period 2009–2011, as identified in a medical database of road trauma victims in a French territorial department were used to give a full picture of cycling accidents, as well as to understand how cyclist behaviour interacts with other factors in causing accidents. They concluded that the risk of accidents and fear of injuries leads a number of cyclists to give up cycling.
Continued on the next page			

Table 3.3 (continued)

Study	Author	Topic	Details of the research
A multiple correspondence analysis of at-fault motorcycle-involved crashes in Alabama	Jalayer and Zhou (2017)	Critical Analysis Reporting Environment (CARE) database; Exploratory data analysis; Motorcycle crash.	MCA was used to identify key factors contributing to at-fault motorcycle-involved crashes to develop effective feedback on safety. This study used the data from 2009-2013 collected in the state of Alabama. The most significant contributors to the frequency and severity of at-fault motorcycle-involved crashes were found to be light conditions, time of day, driver condition, and weather conditions.
Using Multiple Correspondence Cluster Analysis to Map the Competitive Position of Airlines	Wen and Chen (2011)	Airline market positioning; competition; Market dominance.	MCA used the data collected from air passengers of the international airlines operating on the Taipei-Tokyo and Taipei-Osaka routes to recognize the main strengths and weaknesses of their services. The findings suggest that each airline can simultaneously adopt strategies to maintain and enhance its current strengths and to strengthen attributes in which it is lacking.
Exploring Factors Related to Users' Experience of Public Transport Route Choice: Influence of Context and User Profiles.	Grison *et al.* (2015)	Urban mobility; Public transport; Users' experience; Multimodality.	MCA was conducted to investigate an in-depth understanding of the users-related factors that affect the choice of routes in PT. The experience of 19 users of PTs was used to provide an insight into the importance of the contextual factors and the users' profile in route choice.
Exploring frequency of public transport use among older adults: a study in Adelaide, Australia	Truong and Somenahalli (2015)	Public transport.	MCA was used to capture non-linear relationships between the frequency of PT use and influencing factors among older adults in Adelaide, Australia in 2010. Results show that frequent use of PT is closely related to high perceived importance of PT, residential locations, easy access to PT in neighbourhoods, and mobile phone possession.

* PT: public transport

Table 3.3: Example of studies related to transport that have used MCA

The MCA has been widely applied in a wide range of transport studies mostly to analyse latent variables from questionnaires and survey data. All seek to obtain a representation and the typology of the individuals based on the components. The MCA is important in those studies because it can deal with interactions among all the categorical variables straightforwardly and reduce the complicated interactions into two dimensions, which is helpful in explaining the relationships among the variables and guiding decision-making directions.

Abdi (2007) identified MCA as an extension of univariate correspondence analysis methods of managing and demonstrating the patterns of relationships among several categorical dependent variables from rows and columns of the data. It has been established that dimension reduction techniques applied to continuous and categorical data and latent variable models are strongly linked in the sense that they group variables according to patterns or relationships within them (Bartholomew et al., 2011).

One important difference between dimension reduction technique (factor analysis) and the MCA is that factor analysis groups data into factors according to the variable and cannot allocate a specific individual data item to the factor. On the other hand, the MCA allocates specific individuals to a category/group identified by the characteristic of and relationship between the variables. In the research reported in this thesis, both of these approaches are applied given that all data is categorical. This will be discussed further in Chapter 4 methodology and in the results Chapter 7.

3.4 Logit and Probit Models

Logit and probit models are suitable to model either dichotomous dependent variables such as “yes” or “no”, “like” or “dislike” (also called binary logit/probit model) or more than two independent variables (called multinomial logit/probit model). However, logit and probit differ in how the function is defined to explain the range of values the variable can take to yield a predicted probability. Their outcomes are expressed in terms of odds ratio (for logit) or marginal effect (for probit).The logit and probit models use the cumulative function respectively of the logistic distribution and the standard normal distribution (Klieštik *et al.*, 2015).

Table 3.4 is a useful overview showing the advantages and drawbacks of these models. In essence, the choice of logit and probit is a compromise between simplicity and failing to accommodate all variables and complexity leading to an outcome which is difficult to interpret.

Logit		Probit	
Advantages	Disadvantages	Advantages	Disadvantages
• Closed-form solution • Provides one set of globally optimal parameter estimates • Simple calculation • Widely understood and used in practice • Easy to interpret parameter estimates • Easy to calculate probability outcome • Less demanding data quality requirements	• Highly restrictive error assumptions • Ignores firm/specific observed and unobserved heterogeneity which can lead to inferior model specification and spurious interpretation of model outputs • Parameters are point estimates with little behavioural definition • Often provide good aggregate fits, but can be misleading given the simple form of the model • Tends to be less behaviourally responsive to changes in attribute levels	• The link function for both the probit and the multivariate is a normal distribution • Can jointly estimate several response variables at a time and apply adjustments to the covariance matrix	• Normal cumulative distribution function contains unquantified integral – complex model • Inverse transformation of probit has no direct interpretation

Table 3.4: Advantages and disadvantages of logit and probit models

The previous research uses both binary/multinomial logit/probit models as well as bivariate/multivariate logit/probit models (Linardakis and Dellaportas, 2003; Nurdden *et al.*, 2007; Yamamoto, 2009; Habibian and Kermanshah, 2013). Therefore, it is worth clarifying the differences between them. Binary logit models are used to estimate the probability of a binary response (such as "yes" or "no" / "good" or "bad") when those answers are governed by one or more independent variables. Whereas multinomial logit models are applied when the dependent variable is nominal with more than two levels, for example types of cars, colours and ethnic groups (Greene, 2003). Bivariate analysis uses two paired variables to investigate the association that occurs between them and multivariate analysis uses more than two variables to analyse whether they are predictive of a certain outcome (Greene, 2003).

The following sections 3.4.1 and 3.4.2 discuss further the details of two different models: multinomial logistic regression and multivariate probit model which uses frequentist (likelihood) inference and Bayesian inference respectively.

3.4.1 Multinomial logit (multinomial logistic regression)

In previous research, multinomial logistic regression (MLR) was used as the next step of analysis after factors and clusters are obtained in the initial analysis (Zandvliet *et al.*, 2006). As discussed in Kwak and Clayton-Matthews (2002), MLR is suitable for a nominal dependent variable where the number of levels is more than two. Multinomial regression is a predictive analysis, similar to all linear regressions (Greene, 2003). However, MLR is used to describe data and to clarify the relationships between one dependent nominal variable and one or more continuous independent variables, which may be either interval or ratio.

MLR models explain the association between a set of predictors and unordered multi-category nominal outcomes. Consistent with common practice, a conditional probability of the logistic model, which is a generalisation of the multinomial outcome of standard logistic regression, is chosen as a reference against which others are compared.

According to Ben-Akiva and Lerman (1985), the procedure of this model is as follows. Assume that an individual n is related to every travel mode of the choice set. Level

$j\ (j \in \{1,2,\dots,n\})$ and level $k\ (k \in \{1,2,\dots i-1, i+1,\dots,n\})$ are two levels, such that level j is the reference level, and k is the selected level. $\pi_k = P\left(y = \frac{k}{x}\right)$ is the conditional probability that an individual chooses alternative k and $\pi_j = P\left(y = \frac{j}{x}\right)$ is the reference conditional probability that an individual chooses the alternative *j.* The multinomial logistic regression model is defined by:

$$ln\left(\frac{\pi_k}{\pi_j}\right) = ln\left(\frac{P\left(y=\frac{k}{x}\right)}{P\left(y=\frac{j}{x}\right)}\right) = \alpha_k + \beta_{k1}x_1 + \beta_{k2}x_2 + \cdots + \beta_{km}x_l = \alpha_k + \sum_{k=1}^{m} \beta_{kl}x_l \quad (3.5)$$

$$k \in \{1,2,\dots i-1. i+1,\dots,n\}$$

where:

x_l = is the independent variable;
l = is the number of independent variables;
α_k = is the estimated intercept;
β_{kl} = is the estimated coefficient.

Table 3.5 provides brief details of examples of the application of MLR in travel behaviour research. These include Zandvliet *et al.* (2006), Geraghty and O'Mahony (2015), Wuerzer and Mason (2015), Geng *et al.* (2017), Gim (2017), Hamersma *et al.* (2017), Zhao *et al.* (2018) and Nutsugbodo *et al.* (2018).

Examples of study	Topic	Details of the research	Dependent variable	Independent variable
Zandvliet *et al.* (2006)	Destination choice; Space-time ecology; Visitor populations.	Applied MLR for car users to analyse the joint effect of personal and household characteristics using 1998 Netherlands National Travel Survey (NTS) dataset. They concluded that different types of visitor population environments attract different kinds of visitors at different times of the day.	Visitor population environment: 1. Central-place 2. Contemporary-node 3. Self-contained 4. Mobile-children type 5. Local-children	1. Personal variables (gender, age, and educational level) 2. Household variables (car ownership, household income and household type) 3. Workday 4. Day-of-the-week 5. Location
Geraghty and O'Mahony (2015)	Urban noise; Transport; Household characteristics.	Conducted MLR to investigate noise levels regarding location, month of the year, weekday, and hour of the day using a database of urban noise measurements collected from April 2013 to March 2014 at 10 sites in Dublin, Ireland. Location emerged as the most significant variable in forecasting urban noise levels, whilst transport and household characteristics were not significant.	Noise levels - Leq(A)dBA	1. Location 2. Month 3. Weekday 4. Hour of the day
Wuerzer and Mason (2015)	GIS; Cycling; College students; Distance.	Conducted a health survey among college students (n = 949) to explore distance and socio-demographic factors towards susceptibility to cycling. MLR was used to examine their demographic and personal characteristics and distance variable in cycling. The results showed that age and susceptibility to cycling for transportation mediate the adverse influence of distance on the probability of cycling.	Cycling distance	1. Age 2. Gender 3. Healthy weight 4. Cycling for transportation 5. Cycling for recreation

Continued on the next page

Table 3.5 (continued)

Examples of study	Topic	Details of the research	Dependent variable	Independent variable
Geng *et al.* (2017)	Motivation; Sustainable urban transportation; Travel mode choice; Urban residents.	Developed MLR based on survey data from 1244 urban residents in the Jiangsu, China to observe the effects of motivations, government measures and demographic characteristics on residents' travel mode choice behaviours. The result shows that pro-environment has a significant and positive role in promoting sustainable transportation (walking, bicycling, and PT) compared to car use. Furthermore, the effects of gender, age, income, vehicle ownership, travel distance, and government instruments show significant differences among travel mode choices. The results proposed that in order to warrant sustainable urban transportation, pro-environmental motivation has to be emphasized. Policies targeted to only increase public awareness of the environment are not sufficient. Instead they should be targeted to specific groups with diverse key inspirations.	1. Walking 2. Bicycle 3. Public transport	1. Pro-environmental motivation 2. Self-interested motivation 3. Government instruments 4. Demographic variables, vehicle ownership, and travel distance
Gim (2017)	Travel utility; Linear regression; Mode choice; Travel time.	Investigated three travel methods – the automobile, public transit, and non-motorised based on the most frequently chosen mode in the year by using survey data collected in Seoul, South Korea in June 2013. MLR results indicated that in comparison to life situation and land-use variables, utility elements are among the strongest travel influencing factors. Non-motorised travel and modal shift (walking/biking) are strongly influenced by services, facilities, and trip timeliness. The results were significant to public health and transportation planners by providing evidence to inform policies that promote more walking or use bicycles in order to achieve modal shift target.	1. Automobile 2. Public transport 3. Non-motorised	1. Life situation 2. Travel-related utility

Continued on the next page

Table 3.5 (continued)

Examples of study	Topic	Details of the research	Dependent variable	Independent variable
Hamersma *et al.* (2017)	Highway development; Perception; Residential satisfaction; Residential self-selection; Residents.	Studied the perceptions of residents regarding the impact of new highway development. MLR was used to explore the comparison between *before* and *after* a highway development among the original population with those who have moved into the area. The results indicated that residents who had moved into the area after the highway improvement had a slightly more 'highway-oriented' profile than the original residents.	1. Environmental quality 2. Car accessibility 3. Highway proximity	1. Age 2. Number of children 3. Income 4. House ownership
Zhao *et al.* (2018)	Attitude; Future cycling and car purchasing; Mobility policy; Perceived cycling environment.	Analysed the perception of cycling environment, current travel behaviour, urban form and socio-demographic variables to forecast attitudes towards future cycling and car purchasing by using MLR. The results of surveyed data (n=1427) collected in 8 Beijing neighborhoods suggested that respondents have a greater tendency towards cycling rather than buying a car. However, travel distance was more likely to affect respondents' attitudes towards future cycling. Also, the results showed that respondents with low levels of education and income were less likely to buy a car in the future.	1. Cyclists 2. Non-cyclists 3. Non-car owners	1. Perceived cycling environment 2. Current travel behaviour 3. Urban form 4. Socio-demographics
Nutsugbodo *et al.* (2018)	Mode; Preference; Public transport; Tourists.	Investigated public transport mode preferences of international tourists in Ghana using data collected from 479 out-bound international tourists at the departure hall of the Kotoka International Airport in Ghana. MLR was used to examine the relationships between tourists' socio-demographic characteristics and their mode preference. The results showed that they have strong relationships.	Travel mode preference	Socio-demographics

Table 3.5: Examples of study using MLR

It can be seen from the table that MLR has been applied across the world in countries as diverse as Ghana, China, South Korea, Netherlands, Ireland and USA. They conducted investigations involving a wide range of modes for instance car, cycling, public transport and air travel. MLR was applied for impact assessments, highways, noise and other areas. Data for this research were collected by means of attitudinal surveys, questionnaires and direct measurement.

MLR has been widely used in transport research: to analyse the combined effect of personal and household characteristics of car users (Zandvliet *et al.*, 2006); to investigate impacts of traffic noise levels (Geraghty and O'Mahony, 2015); to examine demographic and personal characteristics together with distance travelled by cycle (Wuerzer and Mason, 2015); to study mode choice and more specifically the frequency of use of the automobile, public transit, and non-motorised on the year to year basis (Gim, 2017); to explore the comparison between before and after highway development (Hamersma *et al.*, 2017); to observe the effects of different incentives, government measures, and demographic characteristics on residents' travel mode choice behaviours (Geng *et al.*, 2017); to analyse the perception of cycling environment, current travel behaviour, urban form and socio-demographic variables to forecast attitudes towards future cycling and car purchasing (Zhao *et al.*, 2018) and to examine the relationships between tourists' socio-demographic characteristics and their mode preference (Nutsugbodo *et al.*, 2018).

Previous research has demonstrated the value and appropriateness of this method to investigate the relationships between socio-demographic and travel related variables (Geng *et al.*, 2017; Nutsugbodo *et al.*, 2018; Zhao *et al.*, 2018) and also to understand the association of socio-demographic with the environmental variables (Zandvliet *et al.*, 2006; Hamersma *et al.*, 2017). In addition, among all the studies mentioned above, none have used clustered groups as dependent variables and factors as independent variables to investigate relationships between them. Given the above evidence and underlying benefits, MLR was chosen to study the relationships between factors obtained in EFA and the clusters obtained in HCA, as illustrated in Figure 3.3. Hence, this is a novel area pursued in the research reported in this thesis.

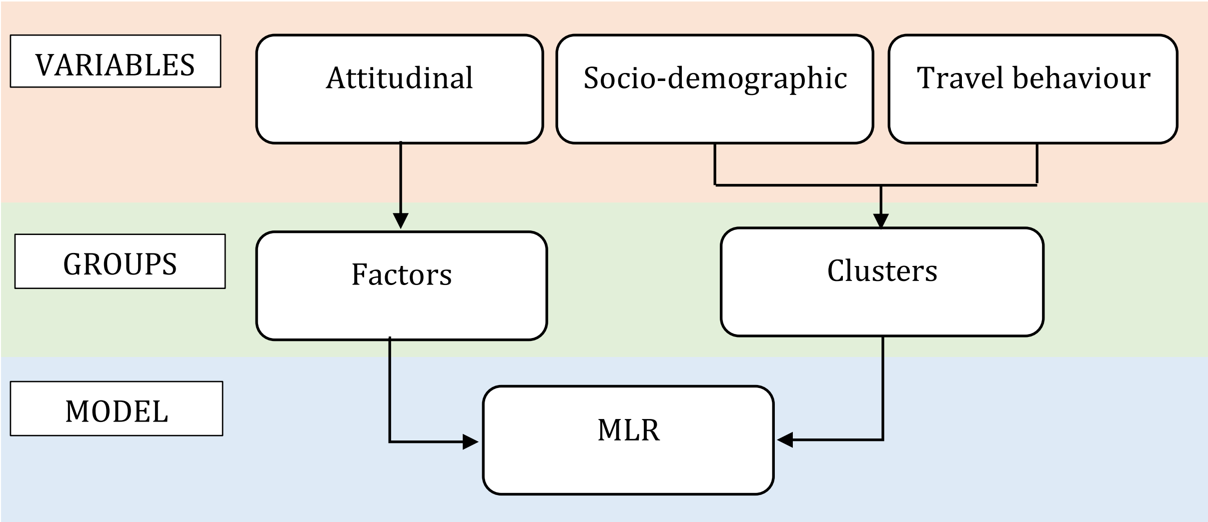


Figure 3.3: MLR approach in this study

The practical advantages of using multinomial logistic regression, as claimed by Tabachnick and Fidell (2013), are as follows:

i. more robust to violations of assumptions of multivariate normality and equal variance-covariance matrices across groups,
ii. similar to linear regression, but gives more easily interpretable diagnostic statistics.

Furthermore, advantages of this analysis that have increased its reputation come from the following assumptions:

i. a linear relationship between the dependent and independent variables is not assumed;
ii. independent variables need not be interval;
iii. no requirement that the independent variables to unbounded, and finally;
iv. normally distributed error terms are not assumed.

3.4.2 Multivariate probit model (MPM)

An ordered probit model can be estimated by the maximum likelihood technique from the frequentist perspective. In contrast, a Bayesian analysis by using a Markov Chain

Monte Carlo (MCMC) approach which develops latent variable representation is another option.

A number of studies have made use of the MPM approach in transport research by using frequentist inference. Frequentist inference is a form of statistical inference that draws conclusions from sample data by emphasising the frequency or proportion of the data. Under the frequentist approach, parameters and hypotheses are viewed as unknown but fixed (non-random) quantities, and consequently there is no possibility of making probability statements about these unknowns (Wakefield, 2013). Maximum likelihood estimator is an example of frequentist inference (Choo and Mokhtarian, 2008; Ferdous *et al.*, 2010; Milioti *et al.*, 2015; Becker *et al.*, 2017). The application of MPM in transportation research has been adopted due to its capability of predicting choices on an individual-specific basis. However, these choices are assumed to be certain. In reality, there can be uncertainty in these choices. Bayesian inference approach has become popular in recent research due to the advantages attached to it.

Choo and Mokhtarian (2008) developed a multivariate probit model with the aim of exploring the relationships between the consideration and adoption of three travel-strategy bundles, namely travel maintaining/increasing, travel reducing and major location/lifestyle change, and linking them to a variety of explanatory variables using 1,300 commuting workers living in three distinct San Francisco Bay area neighbourhoods in May 1998. Whilst Ferdous *et al.* (2010) proposed a multivariate ordered-response structure to deliver significant perceptions into the factors of adults' weekday activity behaviour. Whereas Milioti *et al.* (2015) applied the equation proposed by Greene (2003) in their study regarding the passengers' influencing factor of airline choices using a sample of 853 respondents and developed a multivariate probit model. The common description for the MPM with dependent variables is given in equation 3.6.

$$Y_i^* = \beta_i X_i + \varepsilon_i, \quad i = 1, \dots, n \tag{3.6}$$

where:

Y_i^* = the latent utility or propensity for considering alternative i,

X_i = vector of observed characteristics determining choice alternative i,

β_i = vector of unknown coefficients to be estimated,

ε_i = vector of error terms that are normally distributed with zero mean and constant variance.

Milioti *et al.* (2015) found that fare and safety are the two most important factors in choosing an airline. Others included reliability and, friendly helpful staff during flight. These factors were found to be influenced by socio-demographic (income, age, nationality, education level and gender) and trip characteristics (booking method, final destination, purpose of the trip and cost of the ticket).

Becker *et al.* (2017) studied a combined model of the ownership of car-sharing and other different mobility tools, such as public transport by using data from Swiss national travel surveys of 2005 and 2010. They proved that the multivariate approach is significantly more efficient than univariate approaches due to its ability to investigate the effects of the explanatory variables.

As stated earlier, Bayesian inference is becoming popular due to its capability of handling uncertainties in the models. It is well known and there are good explanations of theory and applications published (Congdon, 2005; Congdon, 2006; Lee, 2012; Gelman *et al.*, 2013; Kruschke, 2015; Bolstad and Curran, 2016).

Linardakis and Dellaportas (2003) used a stated preference data to explain and predict passengers' behaviour towards three main types of transportation (metro, car, bus) in the city of Athens. In order to determine whether a policy has positive or negative net benefits, there are several key evaluation factors such as walking, waiting, and journey time, and 95% credible intervals of the probability of selecting a specific mode of transportation. The MPM was estimated using Bayesian inference in their study as specified in equation 3.7.

$$W_i = X_i\beta + \varepsilon_i \tag{3.7}$$

where:
$W_i =$ is a latent vector of dimension $J - 1 \times 1$,
$X_i =$ is the matrix of dimension $J - 1 \times K$,
$\beta =$ is a vector of dimension $K \times 1$,
$\varepsilon_i =$ is the error term of dimension $J - 1$.

The traditional approach uses the maximum likelihood technique and focuses on the likelihood $p \; y \; \theta$ where θ is conditional probability without introducing a prior distribution, whereas Bayesian analysis uses the prior information about θ coupled with the information contained in the data in an iterative way. Wang and Kockleman (2009) also obtained a broad understanding of the practical benefits of a Bayesian context over a frequentist method. The Bayesian method proved to be more straightforward, advanced and much easier to apply compared to maximum likelihood estimation specifically for complicated models. By using "conditional" distributions, the Bayesian approach decomposes the joint estimation of many variables into much simpler and sequential simulations. Furthermore, an efficient Bayesian inference approach to multivariate probit models with sparse inverse correlation matrices, taking into account the associational structure between binary observations, was proposed by Talhouk *et al.* (2010).

In addition, MPM estimated with a Bayesian approach allows for the remaining uncertainty in the model parameter. However, the use of a Bayesian multivariate probit modelling approach to study attitudes and travel behaviour, which consists of ordinal categorical data, is very scarce in research, in sustainability research in particular.

In MPM, *Just Another Gibbs Sampler* (JAGS) uses Markov chain Monte Carlo (MCMC) to generate a sequence of dependent samples from the posterior distribution of the model parameters (Plummer, 2016). Specifically, the "rjags" package is used which provides an interface from R to the JAGS library for Bayesian data analysis (R Core Team, 2017).

3.5 Bayesian Vs Frequentist (likelihood) Inferences Applied to Travel Behaviour Research

Bayesian and frequentist (likelihood) inferences are two different approaches used to evaluate evidence about competing hypotheses or models. Bayes formula can be used to compute the posterior if the prior and likelihood are known for each hypothesis. However, on many occasions, the prior probabilities on hypotheses are not known. Therefore, the options here would be to use Bayesian with prior belief or to use only the frequentist (likelihood) to make an inference.

The capability to make inferences based on Bayesian subjected to the significance level in the preferred prior and the alternative prior distributions may be significant and essential to the strength of the findings. On the other hand, the frequentist inference does not rely on a specific prior and only uses conditional distributions of data given particular hypotheses. In-depth discussions of the pros and cons of Bayesian analysis have been published by other researchers (Berger, 1985; Bernardo and Smith, 1994; Carlin and Louis, 2000; Robert, 2001).

In recent years, there has been an increasing amount of literature on the use of Bayesian approaches in travel behaviour research because of the benefits that they offer. Bayesian statistics help people to make decisions under uncertainty, allow for the remaining uncertainty in the model parameter, and to quantify the strength of our beliefs when actual data have been taken into account. Table 3.6 shows the wide range of applications of the use of Bayesian in travel behaviour research such as mode choice decisions, travel behaviour, modal shift, and electric vehicles. All sought to analyse variables measured in qualitative surveys such as questionnaires and interviews to aid decision process. However, none of them applied MPM for ordinal data in their study. Given that advantage Bayesian inference was chosen for this study.

No.	Paper	Topics	Data	Method	Application
1.	Stark *et al.* (2018)	Mode choice; Children; School trips; Non-school trips.	- 186 pupils in the 7th grade of eight classes from four secondary schools in Austria and Germany. - Categorical scale of response variables.	Bayesian approach for nonlinear Structural Equation Modelling (SEM)	Used binary response variables to assess the effects of external factors on the choice of travel modes. School trips are quite affine to transit even in rural areas, given a sufficient service quality. Long school trips increase the frequency of transit use. Non-school trips, however, are much more affine to car ridership, if trip length exceeds the range for walking and cycling.
2.	Xiong and Zhang (2017)	Mode choice; Hidden Markov; Modal preference; Decision process.	- Memory-recall survey. - 146 respondents from the University of Maryland. - 3 year period from 2009-Fall to 2012-Summer.	Bayesian estimation with Markov chain Monte Carlo (MCMC)	Estimated the Hidden Markov Model (HMM) and described the heterogeneity of the influence of the local mode share. Driving licence possession makes carpool/transit-loving individuals more sensitive to travel time and as such, it significantly encourages travellers to switch to a car-loving state. It is also found that lower-income travellers are more sensitive to travel cost while female travellers are more sensitive to travel time.
3.	Yang *et al.* (2017)	Residential relocation; Travel behaviour; Modal shift.	258 relocated residents in Nanjing, China in 2014.	Bayesian network	Explored the relationships between residential relocation and shifts in transport mode choice for the work commute respondents who were walking or biking to work before their relocation. The results show that better transit accessibility and services, walking distances to subway stations and the convenience of subway commuting have a significant effect on the modal shift. Purchase of a second (new) car, current car ownership, and personal income strongly influence a modal shift to the car after the relocation.

Continued on the next page

Table 3.6 (continued)

No.	Paper	Topics	Data	Method	Application
4.	Wu and Hong (2017)	Travel mode choice; Subway expansion; Multilevel model.	9462 valid respondents' surveys in Beijing in 2005 and 2009. Cross-sectional data.	Bayesian multilevel binary logistic models with spatial random effects	Examined the potential effects of subway system expansion on commuting behaviour. The results suggest that there is a significant rise in subway commuting trips while non-motorised and bus commuting trips are reduced with the new subway expansion. The results show that young adults (i.e., under 40) are more likely to take the subway for commuting while driving less than old adults. Household size is significantly associated with automobile and other modes used, indicating that people from a large household are less likely to drive and more likely to use other modes such as walking, cycling and bus. Middle-income people tend to use the subway more for commuting and rich people tend to use private cars more for commuting than the poor.
5.	Brady and O'Mahony (2016)	Electric vehicles; Charging patterns.	Real world electric vehicle driving database	Stochastic modelling, Bayesian inference	Synthesises precise journey schedules and models charging decision making behaviour. Contrary to single realisations, charging profiles would be useful to electric vehicle grid integration studies such as aggregated power demand, power systems services and charging optimisation analyses.

Table 3.6: Recent Bayesian applications in travel behaviour

3.6 Conclusions

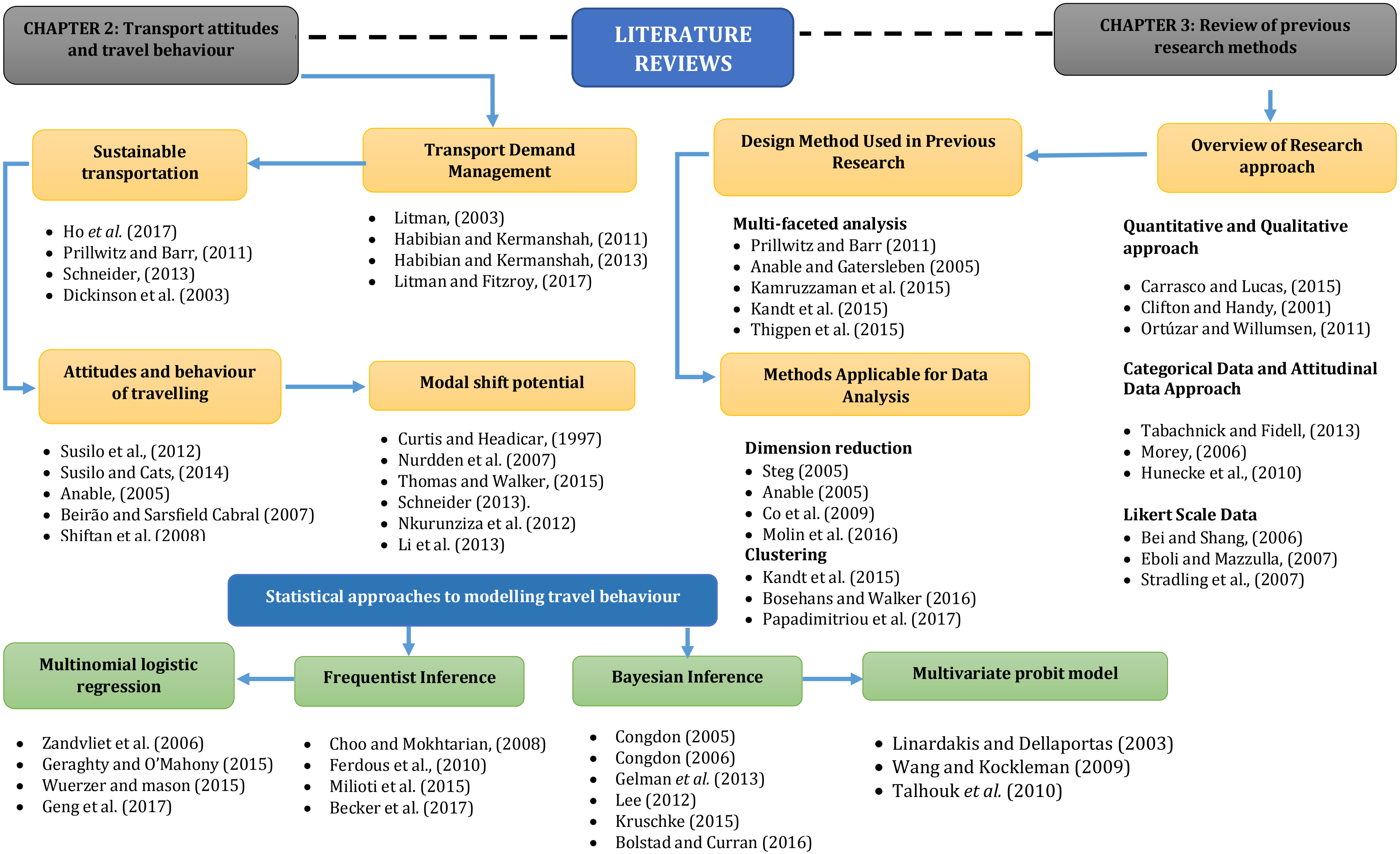
The analytic methods that have been used in previous research have been reviewed in this chapter, including mathematical approaches to modelling travel behaviour. Cluster analysis and factor analysis were found to be the most common approaches applied when analysing attitudinal data. The choice of a specific technique such as cluster and factor analysis will be dependent on the data types such as continuous or categorical and other aspects including the purpose of the study.

Descriptive, factor, and cluster analyses were popular methods to investigate relationships and patterns in socio-demographic and travel behaviour data (Anable, 2005; Prillwitz and Barr, 2011; Kamruzzaman *et al.*, 2015; Kandt *et al.*, 2015; Thigpen *et al.*, 2015). In this study, a multi-faceted approach was used including factor, cluster, MCA, MLR and MPM analyses to investigate the link within attitudes, travel behaviour and environmental aspects.

Previous research demonstrated that PCA and PAF are the most popular methods when it comes to factor extraction. A rotated factor loading should not be less than 0.32 and the sample size should be greater than 300 to be considered statistically meaningful (Costello and Osborne, 2005). Most studies suggest that all factors with eigenvalue greater than 1 should be retained (Kaiser, 1960; Costello and Osborne, 2005; Yong and Pearce, 2013). According to Costello and Osborne (2005) and Yong and Pearce (2013), the number of factors to be retained should be above the "break" in the scree plot.

It is clear that previous studies have not dealt with associations of the five main methods as mentioned above. Therefore, in this study, MPM with Bayesian inference has grown in importance in the light of recent investigation as it considered to be a potentially advantageous approach. All the studies reviewed in this chapter support the appropriateness of MLR and MPM analysis as the most suitable method in developing a model with categorical data to achieve the objectives of the study.

Based on the review of methodology in this chapter, a methodological framework which adopts several methods has been developed to achieve the study's objectives. This is presented in the next chapter.



CHAPTER 2: Transport attitudes and travel behaviour
LITERATURE REVIEWS
CHAPTER 3: Review of previous research methods
Sustainable transportation
Ho et al. (2017)
Prillwitz and Barr, (2011)
Schneider, (2013)
Dickinson et al. (2003)
Transport Demand Management
Litman, (2003)
Habibian and Kermanshah, (2011)
Habibian and Kermanshah, (2013)
Litman and Fitzroy, (2017)
Attitudes and behaviour of travelling
Susilo et al., (2012)
Susilo and Cats, (2014)
Anable, (2005)
Beirão and Sarsfield Cabral (2007)
Shiftan et al. (2008)
Modal shift potential
Curtis and Headicar, (1997)
Nurdden et al. (2007)
Thomas and Walker, (2015)
Schneider (2013).
Nkurunziza et al. (2012)
Li et al. (2013)
Design Method Used in Previous Research
Multi-faceted analysis
Prillwitz and Barr (2011)
Anable and Gatersleben (2005)
Kamruzzaman et al. (2015)
Kandt et al. (2015)
Thigpen et al. (2015)
Methods Applicable for Data Analysis
Dimension reduction
Steg (2005)
Anable (2005)
Co et al. (2009)
Molin et al. (2016)
Clustering
Kandt et al. (2015)
Bosehans and Walker (2016)
Papadimitriou et al. (2017)
Overview of Research approach
Quantitative and Qualitative approach
Carrasco and Lucas, (2015)
Clifton and Handy, (2001)
Ortúzar and Willumsen, (2011)
Categorical Data and Attitudinal Data Approach
Tabachnick and Fidell, (2013)
Morey, (2006)
Hunecke et al., (2010)
Likert Scale Data
Bei and Shang, (2006)
Eboli and Mazzulla, (2007)
Stradling et al., (2007)
Statistical approaches to modelling travel behaviour
Multinomial logistic regression
Zandvliet et al. (2006)
Geraghty and O'Mahony (2015)
Wuerzer and mason (2015)
Geng et al. (2017)
Frequentist Inference
Choo and Mokhtarian, (2008)
Ferdous et al., (2010)
Milioti et al. (2015)
Becker et al. (2017)
Bayesian Inference
Congdon (2005)
Congdon (2006)
Gelman et al. (2013)
Lee (2012)
Kruschke (2015)
Bolstad and Curran (2016)
Multivariate probit model
Linardakis and Dellaportas (2003)
Wang and Kockleman (2009)
Talhouk et al. (2010)

Chapter 4 Methodological Framework

4.1 Introduction

A critical review of methods for analysing categorical data was covered in Chapter 3. Factor analysis, cluster analysis, multiple correspondence analysis, logistic regression and multivariate probit model have been applied in previous research into attitudes and travel behaviour. However, previous research has not combined these techniques. Therefore, in this chapter, the discussion is focused on the methodological framework for the proposed multi-faceted approach designed for this study. The most appropriate techniques for categorical data analysis were identified and explored to achieve the aim and objectives of this study.

The methodological framework that explains all stages of the research is presented in Section 4.2. This is followed by a description of data collection in Section 4.3. Section 4.4 focusing on the preliminary data analysis. Section 4.5 describes the steps in the main data analysis that consists of exploratory data analysis (EFA), data clustering using multiple correspondence analysis (MCA) and hierarchical clustering analysis (HCA), multiple logistic regression (MLR), and multivariate probit model (MPM) with a Bayesian inference approach. Finally, conclusions of this chapter are presented in Section 4.6.

4.2 Methodological Framework

A methodological framework developed to achieve the objectives of this study is illustrated in Figure 4.1. Each section in the diagram was mapped according to the chapters of the thesis. As part of the data search at the beginning of the research, three datasets were identified as being relevant; these included the Tyne and Wear Household Survey (TWHS), the National Travel Survey (NTS) and the British Social Attitude (BSA) survey.

The NTS is the primary source of data on personal travel patterns in Great Britain. It is an established household survey which has been running continuously since 1988 and is reported annually. It is designed to monitor long-term trends in personal travel. The survey collects information on how, why, when and where people travel as well as information of day, household, individual, long distance journey, stage, ticket, trip, and vehicle. In 2013, the survey coverage changed from sampling residents across all residents of Great Britain covering England only.

The TWHS survey complements other transport monitoring collected across Tyne and Wear in order to provide essential knowledge regarding travel patterns and thus better inform the planning of long-term transport strategies. First introduced in February 2003, the household survey is a rolling programme of person-to-person interviews giving details of travel movements. In the TWHS and NTS dataset, there were no attitudinal questions towards travelling and environment that can be used to achieve the objectives of the research. Therefore, the BSA datasets from 2011 – 2014 were identified as being the most suitable to conduct the analysis in this study.

Four years is a short period of time to investigate attitudinal change. However, the relevant questions regarding climate change and the environment, and peoples' attitudes toward them, were only introduced for the first time in 2011 and therefore, four years of data covering 2011-2014 were considered in the study.

The BSA dataset covers several decades of data since 1983. The BSA survey is designed to yield a representative sample of adults aged 18 or over. Interviewers called at each address selected from the list of addresses compiled by the post office and listed all those qualified for inclusion in the BSA sample – that is, all resident at the selected address. The interviewer then selected one respondent using a computer-generated random selection procedure. Where there were two or more dwelling units at the selected address, interviewers first had to select one dwelling unit using the same random procedure. They then followed the same procedure to select a person for interview within the selected house unit.

The methodological framework is divided into four sections: data collection and preparation, preliminary data analysis, main data analysis, and conclusions. Each step of this process is explained in detail in the following sections.

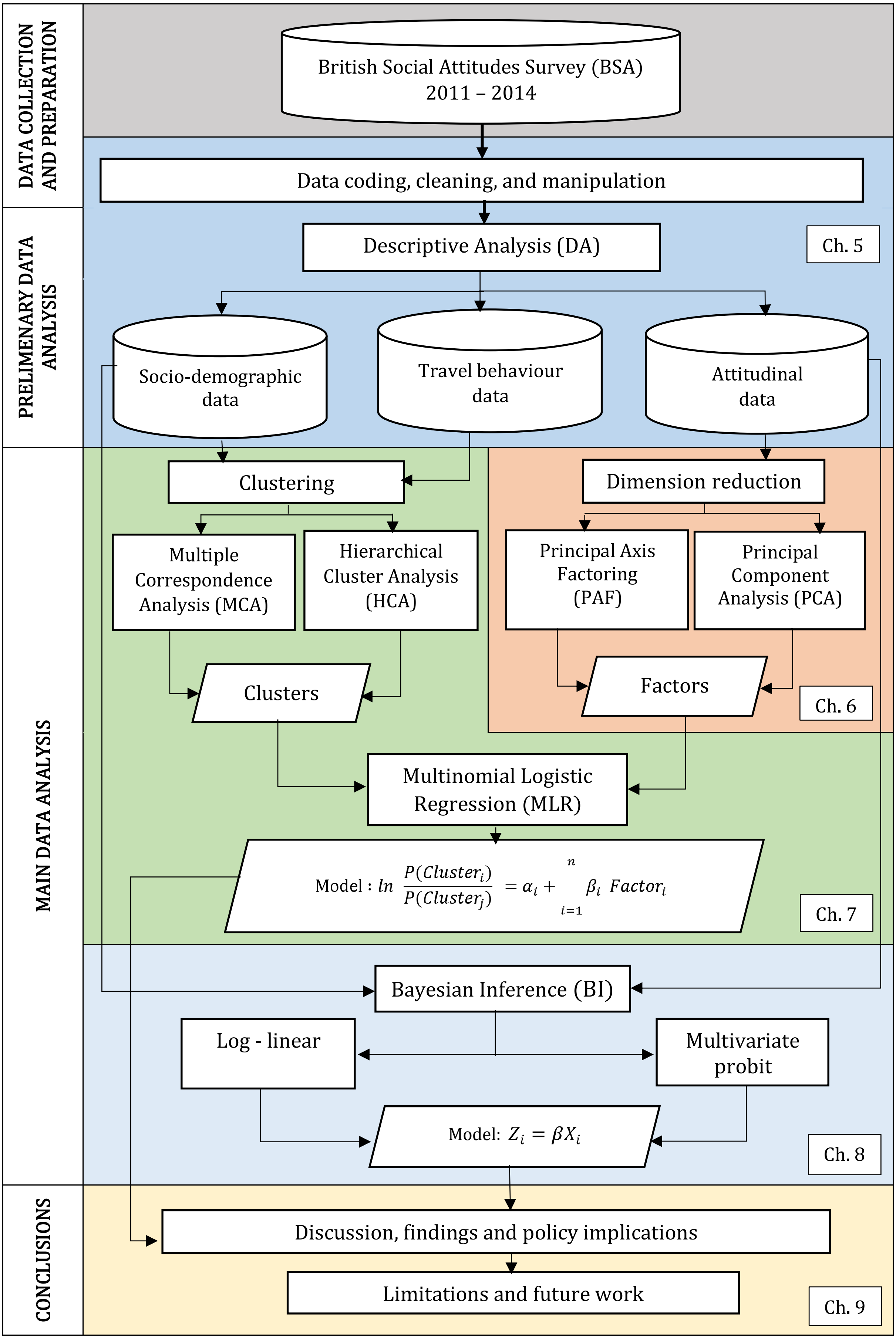


Figure 4.1: Methodology framework of this study

4.3 Data Collection and Preparation

As mentioned earlier, data from the British Social Attitudes (BSA) survey was used in this study. The datasets that were collected during 2011 to 2014 were chosen for the analysis. The BSA survey has been carried out annually since 1983 in Great Britain to collect a wide range of information and facts in different areas such as politics, health, education, and social factors (Social and Community Planning Research, 1983). Attitudes towards and behaviour with regards to travelling are among the information collected through the BSA survey in order to examine the changes of attitudinal trends over time.

In 2010, the transport section in the BSA questionnaire was changed so as to collect data regarding people's attitudes towards, and opinions on, climate change and the environment. Therefore, the BSA data collected after 2010 were considered in this study and all data records appertaining to car users, whether as drivers or passengers, were selected from the BSA using a filter.

Before preliminary analysis was conducted, the data was cleaned, coded and prepared in the appropriate formats needed for the analysis. These formats include IBM-SPSS version 23 (IBM, 2013) and the Rstudio package (R Core Team, 2017). At the initial stage of this analysis, incomplete samples were removed during the data cleaning process, bringing the final dataset to a total of 1509 respondents of adults aged 18+ who were car users whether as drivers or passengers or both.

According to Alison Park *et al.* (2013), in assembling the BSA data the collection methods adopted were a combination of face-to-face interviews and self-completion questionnaires. A random sampling technique was used to select people to participate in the survey to ensure that everyone had an equal chance of being selected to take part and therefore, the results are representative of the British population (Alison Park *et al.*, 2013). The transportation part of the BSA survey was funded by the UK Department for Transport (DfT).

4.4 Preliminary Data Analysis

The preliminary data analysis was conducted to discover patterns in the data using basic statistics to provide information that describes the characteristics of the sample taken from the BSA survey. After careful screening for errors or missing values, all categorical data from the 1509 car users were divided into three parts, namely socio-demographics, attitudinal, and travel behaviour variables as listed below:

A. Socio–demographic variables:

1. Age
2. Gender
3. Number of people living in the household, including respondent
4. Number of cars owned and regularly used in the household
5. Employment status

B. Travel behaviour variables:

1. How often nowadays do you usually travel by car as a driver?
2. How often nowadays do you usually travel by car as a passenger?
3. How often nowadays do you usually travel by local bus?
4. How often nowadays do you usually travel by train?
5. How often nowadays do you usually travel by bicycle?

C. Attitudinal variables:

1. How serious a problem for you is congestion on motorways?
2. How serious a problem for you is traffic congestion in towns and cities?
3. How serious a problem for you are exhaust fumes from traffic in towns and cities?
4. Next time I buy a car, I would be willing to buy a car with lower CO_2 emissions.
5. I am willing to reduce the amount I travel by car (to help reduce the impact of climate change).

6. View on climate change and causes.
7. Many of the short journeys that I now make by car I could just as easily walk.
8. Many of the short journeys that I now make by car I could just as easily go by bus.
9. Many of the short journeys that I now make by car I could just as easily cycle.
10. For the sake of the environment, car users should pay higher taxes.
11. People should be allowed to use their cars as much as they like, even if it causes damage to the environment.
12. For the sake of the environment, everyone should reduce how much they use cars.
13. There is no point in reducing my car use to help the environment unless others do the same.
14. People who drive cars that are better for the environment should pay less to use roads.

Socio-demographic variables were coded as nominal data, whereas travel behaviour and attitudinal variables were ordinal data. More detail is provided in Chapter 5. The next step was to bring these data together and to understand whether or not any correlations or links between attributes existed in the dataset by carrying out correlation analysis. Firstly, the normality test was conducted on each variable to establish whether or not parametric or non-parametric tests should be used in statistical testing. Furthermore, the descriptive analysis gives a basic understanding of the population based on the data gathered from the sample such as percentages, mean, median and mode. Statistical tables, graphs and charts consisting of information regarding car users are presented and discussed in detail in Chapter 5.

4.5 Main Data Analysis

Based on the framework shown in Figure 4.1, there are four main steps in analysing categorical data that are mapped directly onto each thesis chapter as follows:

1. Dimension reduction using principal axis factoring (PAF) and principal component analysis (PCA);
2. Data clustering using multiple correspondence analysis (MCA) and hierarchical clustering analysis (HCA);
3. Multinomial logistic regression (MLR) to explore the relationship between clusters and factors;
4. Log – linear and multivariate probit models (MPM) with a Bayesian inference (BI) approach.

Each of these analyses is discussed in detail in the following sections.

4.5.1 Dimension reduction by using exploratory factor analysis (EFA)

EFA was used to reduce the number of variables. In this study, data from 14 attitudinal variables were used to examine the factors. These were based on the questions and statements provided in the BSA dataset and listed in Section 4.4. In coding the data, they labelled Q1 through Q14.

As a result of the EFA the number of variables (in this study that relate to attitudes) is reduced to fewer factors. A name or label was identified to distinguish factors made up of variables that were grouped to go together (Yong and Pearce, 2013). There are four main components or steps in conducting EFA as shown in Figure 4.2.

Factor extraction

- Most common factor extraction methods are unweighted least squares, generalized least squares, maximum likelihood, principal axis factoring (PAF), principal component analysis (PCA), alpha factoring, and image factoring. Previous research demonstrated that PCA and PAF are most popular methods when it comes to factor extraction.

- ***PAF and PCA were selected to check the most suitable for this study and decide the right approach.***

Rotation method

- The aim of the rotations is to aspire to achieve a correlation close to 1 to reveal the variables which indicate significant contribution to the reduced factors. There are two types of rotation, namely orthogonal and oblique. Orthogonal rotation is when the factors are rotated through 90^0 and it is assumed that the factors are uncorrelated. Conversely, oblique rotation is when the factors are not rotated 90^0 from each other, and the factors are considered to be correlated (Costello and Osborne, 2005).

- ***The most appropriate rotation method will be decided when analysing the data in Chapter 6.***

Interpretations of factor loadings

- The strength of the relationship of each factors can be decided by observing the loadings when interpreting the factors obtained in EFA. A rotated factor loading should not be less than 0.32 and the sample size should be greater than 300 to be considered statistically meaningful (Costello and Osborne, 2005).

- ***In this study, 1509 samples were used in the analysis. The rotation factor loading will be carefully checked to make a meaningful set of factors.***

Number of factors to retain

- After extraction, decision have to be made for number of factors to retain. The number of factors are extracted based on the scores from the factor analysis (Costello and Osborne, 2005). The eigenvalue and scree plot are used to determine how many factors to retained. Most studies suggest that all factors with eigenvalue greater than 1 should be retain (Kaiser, 1960; Costello and Osborne, 2005; Yong and Pearce, 2013). According to Costello and Osborne (2005) and Yong and Pearce (2013) the number of factors to be retained should be above the "break" in the scree plot.

- ***The scree plot will be carefully examined after analysing the data and the number of factors to be retained will be decided considering recommendation of previous research.***

Figure 4.2 : Components or steps in conducting EFA

4.5.2 *Data clustering using multiple correspondence analysis (MCA) and hierarchical clustering analysis (HCA)*

MCA and HCA were chosen to examine the structure of the dataset. MCA is particularly suitable to explore unpredicted dimensions and relationships. MCA positions each variable as a point in a low-dimensional space and helps to describe patterns of relationships, particularly using geometrical methods (LeRoux and Rouanetm, 2010). Whilst cluster analysis is an exploratory statistical tool to identify the existence of the similarity patterns in responses (Tabachnick and Fidell, 2013), it is used to create and identify groups that are homogeneous within the data. In the research reported in this thesis, cluster analysis is used to segregate car users based on similar characteristics. By identifying the sub-groups, it becomes simpler to analyse and reveal any relationships among the sample within and between groups.

<u>Multiple correspondence analysis (MCA)</u>

The next stage of analysis is Multiple Correspondence Analysis (MCA). MCA is one of the statistical techniques used to discover and represent basic structures in a dataset, especially for categorical data (Murtagh, 2007). Data is represented as points in a low-dimensional Euclidean space.

The main application of MCA in this study is to visualise the interrelationships between response categories for each question. For this purpose, socio-demographic (age, gender, household size, employment status and car ownership) and travel behaviour (frequency of travel by car as a driver, frequency of travel by car as a passenger, frequency of travel by local bus, frequency of travel by train and frequency of travel by bicycle) variables have been selected to classify the car users with similar characteristics.

As discussed previously in Chapter 3, equation 3.4, the distance between two individuals i and l can be calculated by the following equation according to Greenacre and Blasius (2006):

$$d^2(i,l) = \sum_k \frac{IJ}{I_k}\left(\frac{x_{ik}}{J} - \frac{x_{lk}}{J}\right)^2 = \frac{I}{J}\sum_k \frac{I}{I_k}\left(x_{ik} - x_{lk}\right)^2$$

where:

x_{ik} = is equal to 1 if the individual i has taken the category k and 0 otherwise,

x_{lk} = is equal to 1 if the individual l has taken the category k and 0 otherwise,

I = is the total number of individuals,

I_k = is the number of individuals who have taken the category k, and

J = is the number of variables.

In this study, correlation with categorical variables to reveal behaviour trend was investigated and the categories were illustrated in a 2-dimensional discrimination plot and MCA factor maps. Three main steps are involved in MCA as shown in Figure 4.3.

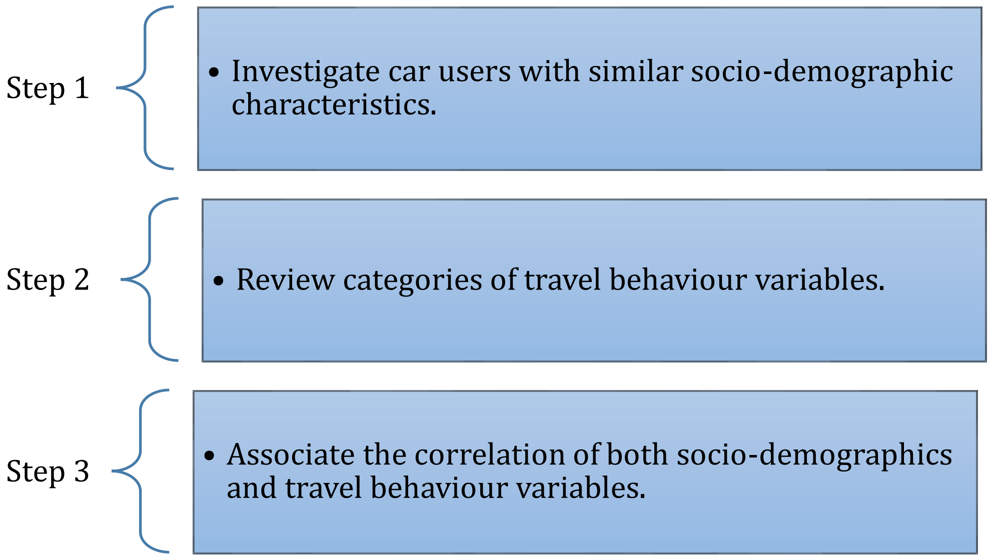


Figure 4.3: Steps in conducting MCA

The data obtained from the BSA survey is in .sav format, which is a file extension type used to store data for SPSS analysis. Hence, the data was reconstructed, as this study used both SPSS and R studio to visualize the results of MCA. The input for .sav files required the survey data to be reformatted as comma separated values (.csv), given that the original datasets were represented in numeric codes.

The discrimination measure and joint plot were obtained by using SPSS. Subsequently the R studio package was used to enrich the cluster classification. MCA follows Ward's criterion in conducting hierarchical clustering for categorical data. In this study by using MCA, clearer cluster visualisation was achieved when car users were segregated by groups in different colours. The detailed results of MCA are presented in Chapter 7.

Hierarchical cluster analysis (HCA)

Following MCA, HCA was carried out to investigate the data organisation and to group the car users according to their similarity in socio-demographics and travel behaviour variables. There are five steps in conducting HCA as suggested by Yim and Ramdeen (2015) and shown in Figure 4.4.

Step 1: Choosing Cluster Variables

- HCA was conducted using socio–demographic and travel behaviour data with 10 variables chosen to be relavant to this study as follows:
  - Age
  - Gender
  - Household size
  - Employment status
  - Car ownership
  - Frequency of travel by car as a driver
  - Frequency of travel by car as a passenger
  - Frequency of travel by local bus
  - Frequency of travel by train
  - Frequency of travel by bicycle

Step 2: Selecting Cluster Method

- All categorical variables were either ordinal (frequencies of each modes choice) or nominal data (socio-demographics), therefore, as mentioned above, HCA was used. This was achieved using the SPSS software tool and "EnquireR" package using R programming.

Step 3: Specifying Parameters

- Following to the MCA method in previous analysis, Ward's criterion was selected to conduct HCA. This criterion allows to decompose the total inertia or total variance in between and within-group variance (Husson et al., 2010).

Step 5: Organising Data into Subgroups and Interpreting the Output

- The number of clusters was decided by cutting the dendogram plot at the point of inflection. The characteristics of each cluster/ group was examined. The agglomeration display an output from HCA aids the interpretation.
- In order to classify the data into the cluster groups, cluster memberships single solution was saved in SPSS analysis steps. This step will produce a new variable in data view window which assigns each case into one of the cluster groups. Next, each different cluster groups was analysed and organised by examine the proportion of each variables. The highest shares is linked in to the cluster classification.

Figure 4.4: Steps in conducting HCA

4.5.3 Multinomial logistic regression (MLR) to explore the relationship between clusters and factors

After the attitudinal factors were obtained from the EFA and the car users were clustered into groups depending on a combination of their socio-demographic and travel-related characteristics from the HCA, the next step was to take a closer look at how the clusters were linked to the factors in a model. Therefore, the next step in the data analysis was to conduct MLR to investigate the significant relationships between clusters and for each of the four successive years in the period from 2011 to 2014. This enabled any changes in the relationships between clusters and factors over time to be identified. The MLR analysis was carried out using IBM SPSS Statistics V23 (IBM, 2013).

According to Long (2012), the equation for the model is written in terms of the logit of the outcome, which is a comparison of a particular category to the reference category, denoted by π_i and π_j as follow:

$$ln \ \frac{\pi_i}{\pi_j} = \alpha_j + \beta_j X$$

Therefore, the model for this study would be:

$$ln \ \frac{P(Cluster_i)}{P(Cluster_j)} = \alpha_i + \sum_{i=1}^{n} \beta_i \ Factor_i$$

where:

$Factor_i$ = is the independent variable i;

$Cluster_i$ = is the dependent variable i;

$Cluster_j$ = is the baseline variable j;

α_i = is the estimated intercept;

β_i = is the estimated coefficient.

As discussed previously in Chapter 3, the MLR model in equation 3.5, as defined by Ben-Akiva and Lerman (1985), was used and reproduced here for completeness.

$$ln\left(\frac{\pi_k}{\pi_j}\right) = ln\left(\frac{P\left(y = \frac{k}{x}\right)}{P\left(y = \frac{j}{x}\right)}\right) = \alpha_k + \beta_{k1}x_1 + \beta_{k2}x_2 + \cdots + \beta_{km}x_l = \alpha_k + \sum_{k=1}^{m} \beta_{kl}x_l$$

$$k \in \{1,2, \ldots\ i-1. i+1, \ldots, n\}$$

where:

π_k = is the conditional probability that an individual chooses alternative k;

π_j = is the reference conditional probability that an individual has chooses the alternative *j;*

x_l = is the independent variable;

l = is the number of independent variables;

α_k = is the estimated intercept;

β_{kl} = is the estimated coefficient.

In SPSS, cluster variables were used as the categorical dependent and attitudinal factor scores were used as the explanatory variables. One of the cluster variables was selected as the baseline or reference category. This exposed significant changes in peoples' attitudes from 2011 to 2014 and the results obtained by the MLR analysis is based on the baseline category chosen. The detail and results of MLR analysis are discussed in Chapter 7.

MLR uses a maximum likelihood estimator, where researchers have no chance to consider their prior belief before the actual data is taken into account in the analysis. Therefore, Bayesian inference approach was chosen as the next step of analysis because Bayesian combines two separate sources of information (prior and actual data) to obtain the posterior distribution.

4.5.4 Log – linear and multivariate probit models (MPM) with Bayesian inference approach

The final step of the analysis in this study was to develop a log–linear model using only one variable and a multivariate probit model using 14 attitudinal variables with Bayesian inference approach. Bayesian analysis lets us compute exactly how much our beliefs have changed after the actual data is taken into account. There are four main reasons why Bayesian inference approach was selected in the research: to produce a statistical model to link data to parameters; to formulate prior information about parameters; to combine the two sources of information using Bayes' theorem; and to use the resulting posterior distribution to derive inference about the parameters.

In this procedure, two models are constructed: the log-linear model and the multivariate probit model. Firstly, one attitudinal variable (q12 = "for the sake of the environment, everyone should reduce how much they use cars") is chosen to construct a log-linear model to observe groups of people of different age and gender. Age and gender were found as significant factors in influencing an individual's choice of transportation (Nurdden *et al.*, 2007).

This model consists of two parts:

1) Separate models for age and gender, and
2) Model with interaction between age and gender effect.

Secondly, the probit model for ordinal responses is constructed with a mean structure, allowing for covariates. The model was extended to consider more than one question by introducing a multivariate normal vector latent variable, which then generates a multivariate probit model (MPM).

Five socio-demographic variables were chosen to construct this model: age, gender, household size, car ownership and employment status; whilst 14 variables were selected from the attitudinal variables. Chapter 8 gives further explanation of, and the results obtained from the multivariate probit analysis.

4.6 Conclusions

This chapter has presented the methodological framework used and outlined the data collection, preparation and processing employed in order to achieve the research objectives. The appropriate applicable methods of data analysis used in this research have been presented. These include factor analysis, multiple correspondence analysis, cluster analysis, multinomial logistic regression, and Bayesian inference approach methods.

Factor analysis reduces the attitudinal variables into a smaller number of variables and groups them as factors. Whereas, cluster and multiple correspondence analyses are used to segregate the car users into groups, and then explore differences between groups of car users to investigate the structure of the data.

After factors and clusters were obtained from EFA and HCA, the next step was to investigate the significant relationships between them using MLR. From the model, perceptions and attitudes of travelling can be investigated for possible change over a 4 year period.

Log-linear and multivariate probit models, using Bayesian inference approach for categorical data, were developed for exploring relationships between and within the data, consisting of 14 attitudinal and socio-demographic variables. This model has the potential to provide insights into transport mode choice decisions according to socio-demographic characteristics and is also useful to demonstrate the reliability and aptness of the BSA information.

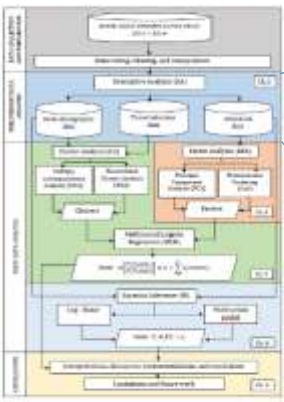
Finally, after all the analyses mentioned above are successfully conducted, the interpretation and discussions are presented in Chapter 9. The output of the MLR models and MPM models are brought together in the interpretations, discussion, and conclusions. The discussion is expected to answer the research questions posed in Chapter 1. In addition, limitations of the study are discussed and future research is addressed to complement the findings. Furthermore, the implication of the results to inform future policy implementations for LAs are highlighted.

Chapter 5 Preliminary Data Analysis

5.1 Introduction

A step by step description of the multi-faceted analysis presented in the context of the methodological framework proposed for this research was presented in Chapter 4. This study uses the BSA dataset collected during 2011-2014. The first step was to collect, clean and code the data before carrying out simple descriptive statistics. 14 attitudinal variables were reduced to factors in parallel with cluster analysis which allowed for grouping of data considering 5 socio-demographic and 5 travel related variables. Multinomial logistic regression and multivariate probit models with Bayesian inference were used to explore and investigate the relationships. In this chapter, the preliminary data analysis using descriptive statistics is designed to gain a fundamental understanding of and develop knowledge from the BSA dataset.

In Section 5.2, the data cleaning, coding and manipulation of the data in preparation for analysis is explained, followed by descriptions and comparison of the characteristics of car users for each year from 2011 through to 2014 inclusive, in terms of their socio-demographic variables and travel behaviour in Section 5.3. The responses to the attitudinal questions involved in this study are detailed in Section 5.4. This is followed by the correlation analysis using Spearman's rank-order correlation in Section 5.5. Finally, a summary of this chapter is presented in Section 5.6.



Step 1: Data cleaning, coding and manipulation
Step 2: Descriptive statistics
- Proportion of car users by years
- Proportion of car users by socio-demographic background
- Proportion of car users by travel behaviour

Step 3: Exploring attitudinal variables
- 3, 4 and 5 Likert scales
- Mean, median and mode

Step 4: Correlation analysis
- Spearman correlation coefficient

Figure 5.1: Steps involved in Chapter 5

5.2 Data Cleaning, Coding and Manipulation

There was a sizeable proportion of respondents in the BSA datasets from 2011 to 2014 who declared as non-car users at the time of completing the survey, which created some redundancy in the datasets. Therefore, those samples were removed as this research focussed on the attitudes of car users, whether as a driver or passenger. The data available for the car users were grouped into three types, namely socio-demographic characteristics, travel behaviour, and travel attitudes, based on the questionnaires completed during the annual surveys. At the initial stage of this analysis, responses such as “not answered”, “skip this question” and “can’t choose” were removed during the data cleaning process, bringing the final dataset to a total of 1509 car users for the period of four years commencing in 2011. Removal of data was considered not to cause bias, based on the assumption that they occurred at random. Table 5.1 shows the coding system that was used for each variable in the dataset and, once each questionnaire was checked, they were assigned a serial number. These records were retained in a database which formed the basis for all statistical analyses.

Coding	Variables /Questions	Categories /Answer coding
Serial	Serial Number	-
HH#	Number living in household, including respondent	1 = One, 2 = Two, 3 = Three, 4 = Four or more
Gender	Gender	1 = Male, 2 = Female
Age	Age	1 = 18-24, 2 = 25-34, 3 = 35-44, 4 = 45-54, 5 = 55-64, 6 = 65+
EmpStatus	Current economic position of respondents	1 = Employee (FT), 2 = Employee (PT), 3 = Self-employed (FT), 4 = Self-employed (PT), 5 = In work (status not known), 6 = Waiting to take up work, 7 = Unemployed, 8 = Looking after the home, 9 = Retired, 10 = In FT education, 11 = Other.
Car#	How many, if any, cars or vans does your household own or have the regular use of?	1 = One, 2 = Two, 3 = Three, 4 = Four or more.
Income	Household pre-tax income quartiles	1 = less than £1,200 p.m, 2 = £1,201 - 2,200 p.m, 3 = £2,201 - 3,700 p.m, 4 = £3,701 or more p.m.
Drive	May I just check, do you yourself drive a car at all these days?	1 = Yes, 2 = No.
Cong_MWs	How serious a problem for you is congestion on motorways?	1 = A very serious problem, 2 = A serious problem, 3 = Not a very serious problem, 4 = Not a problem at all.
Cong_cities	How serious a problem for you is traffic congestion in towns and cities?	1 = A very serious problem, 2 = A serious problem, 3 = Not a very serious problem, 4 = Not a problem at all.
Exhaustfumes	How serious a problem for you are exhaust fumes from traffic in towns and cities?	1 = A very serious problem, 2 = A serious problem, 3 = Not a very serious problem, 4 = Not a problem at all.
Car_driver	How often nowadays do you usually travel by car as a driver?	1 = Every day or nearly every day, 2 = 2-5 days a week, 3 = Once a week, 4 = Less often but at least once a month, 5 = Less often than that, 6 = Never nowadays.
Car_passenger	How often nowadays do you usually travel by car as a passenger?	1 = Every day or nearly every day, 2 = 2-5 days a week, 3 = Once a week, 4 = Less often but at least once a month, 5 = Less often than that, 6 = Never nowadays.
Bus_usage	How often nowadays do you usually travel by local bus?	1 = Every day or nearly every day, 2 = 2-5 days a week, 3 = Once a week, 4 = Less often but at least once a month, 5 = Less often than that, 6 = Never nowadays.
Continued on the next page		

Train_usage	How often nowadays do you usually travel by train?	1 = Every day or nearly every day, 2 = 2-5 days a week, 3 = Once a week, 4 = Less often but at least once a month, 5 = Less often than that, 6 = Never nowadays.
Bike_usage	How often nowadays do you usually travel by bicycle?	1 = Every day or nearly every day, 2 = 2-5 days a week, 3 = Once a week, 4 = Less often but at least once a month, 5 = Less often than that, 6 = Never nowadays.
BuyLowEmi	Next time I buy a car, I would be willing to buy a car with lower CO_2 emissions.	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly.
ReducTravCar	I am willing to reduce the amount I travel by car (To help reduce the impact of CC).	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly.
CCView	View on climate change and causes.	1 = I don t believe that CC is taking place, 2 = I believe that CC is taking place but not as a result of human actions, 3 = I believe that CC is taking place and is, at least partly, a result of human actions.
CartoWalk	Many of the short journeys that I now make by car I could just as easily walk.	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly.
CartoBus	Many of the short journeys that I now make by car I could just as easily go by bus.	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly.
CartoBike	Many of the short journeys that I now make by car I could just as easily cycle.	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly.
HiTaxforCarUse	For the sake of the environment, car users should pay higher taxes.	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly.
AllowCarUse	People should be allowed to use their cars as much as they like, even it is cause damage to the environment.	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly
ReducCarUse	For the sake of the environment, everyone should reduce how much they use cars.	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly.
ReducCarUse_NP	There is no point in reducing my car use to help the environment unless others do the same.	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly.
CarBetterPayLess	People who drive cars that are better for the environment should pay less to use roads.	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly
CycDang	It is too dangerous for me to cycle on the roads.	1 = Agree strongly, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Disagree strongly.
Bike_own	Bike ownership.	1 = Own bicycle yourself, 2 = Have regular use of a bicycle owned by someone else, 3 = Have no regular use of a bicycle.
Bike_ride	Have you ridden a bicycle during the last 12 months?	1 = Yes, 2 = No.

*CC = Climate change, FT = Full time, PT = Part time

Table 5.1: Data coding

Being mindful that it is important not to over-parameterise the model when setting up the multivariate probit model (MPM) in Chapter 8. Therefore, to overcome this problem, the employment status variables were grouped and allocated new codes into 4 categories. Table 5.2 presents the new codes for employment status and the description of each category are as follows:

1. In work: this should take the value 1 if the person is in work (current categories 1 – 5), and -1 otherwise (current categories 6 – 11)

2. Employee: takes the value 1 for current categories 1 and 2 and takes the value of -1 for current categories 3 and 4. Those in current category 5 – 11 should get a value of 0 for this.

3. Full time: takes the value 1 for current categories 1 and 3 and takes the value -1 for current categories 2 and 4. Those in current category 5 – 11 should get a value of 0 for this.

4. Non-employed status: treat this as a 3-category variable. Category 1 would be current categories 6 and 7 (unemployed and waiting to take up work). Category 2 would be current categories 8 and 9 (looking after the home and retired). Category 3 would be current categories 10 and 11 (in full-time education and other). Respondents in current categories 1 – 5 get the value 4.

No.	Variable	Coding	Description
1	In work	1	• FT employee • PT employee • FT self-employed • PT self-employed • In work (status not known)
		-1	• Waiting to take up work • Unemployed • Looking after the home • Retired • In FT education • Other
2	Employee	1	• FT employee • PT employee
		-1	• FT self-employed • PT self-employed
		0	• In work (status not known) • Waiting to take up work • Unemployed • Looking after the home • Retired • In FT education • Other
3	Full time	1	• FT employee • FT self-employed
		-1	• PT employee • PT self-employed
		0	• In work (status not known) • Waiting to take up work • Unemployed • Looking after the home • Retired • In FT education • Other
4	Non-employed	1	• Waiting to take up work • Unemployed
		2	• Looking after the home • Retired
		3	• In FT education • Other
		4	• FT employee • PT employee • FT self-employed • PT self-employed • In work (status not known)

*FT = Full time, PT = Part time

Table 5.2: Manipulation of employment status variable

5.3 Descriptive Statistics Analysis

The purpose of descriptive analysis was to present the frequency distribution and characteristics of car users and to gain a fundamental understanding of the responses in the BSA dataset. Tables and graphs were used to present the results, including statistical parameters of the data such as mean, median, mode and percentages.

5.3.1 Proportion of car users by years

The data from 2011 to 2014 were specifically selected to allow sufficient elapsed time so that attitudinal changes may have occurred and could be investigated with a measure of statistical confidence. However, the descriptive statistical analysis first studied the characteristics aggregated over four years. All of the data were collated together in a file and reconstructed into two types of format: SPSS (.sav) and R studio (.csv) in preparation for all steps of the analysis. Table 5.3 shows the proportion of car users in the four successive years.

Year	Frequency	Percentage
2011	365	24.19
2012	400	26.51
2013	341	22.60
2014	403	26.71
Total	1509	100.0

Table 5.3 Car users' distribution according to year

The lowest percentage was recorded in 2013 at 22.60%, whilst the highest percentage was reported in 2014 with 26.71%. The BSA dataset does not have variable such as geographic locations of the respondents, for example city, region, and postcodes. It would have been better if the study would be able to relate such variables when interpreting the research results but this is one of the limitations of the BSA dataset. However, to be able to make sure the car user data sample drawn from the BSA dataset, Chi-square test was conducted. Chi-square χ^2 test with contingency table was conducted to demonstrate whether the data selected (sample) was representative of each year in the BSA.

Year	Sample		BSA population		Total
	O	E	O	E	
2011	365	386.00	2279	2257.99	2644
2012	400	385.28	2239	2253.72	2639
2013	341	383.82	2288	2245.18	2629
2014	403	353.89	2021	2070.11	2424
Total	1509		8827		10336

*O: Observed value, E: Expected value

Table 5.4: Contingency table

The result shows that χ^2 =15.57 > 7.81 (critical value) at 95% significant level. Therefore, this indicates that the distribution of car users in the BSA dataset who participated in the questionnaires survey and interviews is similar to the distribution for all car users in the sample (2011–2014). This outcome indicates that car users in the sample are representative of the all car users in the BSA dataset with respect to the year in which data was collected.

5.3.2 Proportion of car users by socio-demographic background

It can be seen from the data distributions in Table 5.5 that 17.5% of the car users were in the younger age group in the range of 18 – 34 years old, 41.88% in the middle-aged group (35–54 years old), and 40.63% in the older group (55 years old and above). From the perspective of gender, 50.17% of car users were male, whilst 49.83% were female. The employment status of the car users was divided into 4 categories, as mentioned above: in work (63.75%), employee (54.41%), full time (48.51%), and non-employment (36.25%). It is shown that half of the car users (50.10%) owned one car per household, 40.03% owned 2 cars, 7.42% owned 3 cars, and only 2.45% owned four or more cars in a household. However, 90.13% of the male and female respondents owned either one or two cars per household. Detailed breakdowns of each characteristic, including household incomes and household size, are presented in Table 5.5 and illustrated in Figure 5.2.

Characteristics	Interval	Male		Female		Total	
		Count	%	Count	%	Count	%
Household size	One	152	20.08	180	23.94	332	22.00
	Two	**332**	**43.86**	**254**	**33.78**	**586**	**38.83**
	Three	104	13.74	141	18.75	245	16.24
	Four or more	169	22.32	177	23.54	346	22.93
Age	18-24	19	2.51	30	3.99	49	3.25
	25-34	89	11.76	126	16.76	215	14.25
	35-44	130	17.17	**185**	**24.60**	315	20.87
	45-54	159	21.00	158	21.01	317	21.01
	55-64	153	20.21	129	17.15	282	18.69
	65+	**207**	**27.34**	124	16.49	**331**	**21.94**
Income	Less than £1,200 p.m.	115	15.19	152	20.21	267	17.69
	£1,201 – 2,200 p.m.	157	20.74	194	25.80	351	23.26
	£2,201 – 3,700 p.m.	239	31.57	**204**	**27.13**	443	29.36
	£3,701 or more p.m.	**246**	**32.50**	202	26.86	**448**	**29.69**
Car ownership	One	**344**	**45.44**	**412**	**54.79**	**756**	**50.10**
	Two	322	42.54	282	37.50	604	40.03
	Three	70	9.25	42	5.59	112	7.42
	Four or more	21	2.77	16	2.13	37	2.45
In work	In work	**492**	**64.99**	**470**	**62.50**	**962**	**63.75**
	Not in work	265	35.01	282	37.50	547	36.25
Employee	Employee	**398**	**52.58**	**423**	**56.25**	**821**	**54.41**
	Self-employed	94	12.42	46	6.12	140	9.28
	Others	265	35.01	283	37.63	548	36.32
Full time	Fulltime employee	**453**	**59.84**	279	37.10	**732**	**48.51**
	Part time employee	39	5.15	190	25.27	229	15.18
	Others	265	35.01	**283**	**37.63**	548	36.32
No employment	Unemployed / waiting to take up work	32	4.23	26	3.46	58	3.84
	Looking after the home / retired	204	26.95	207	27.53	411	27.24
	In FT education / Other	29	3.83	49	6.52	78	5.17
	In work	**492**	**64.99**	**470**	**62.50**	**962**	**63.75**
Total		757	100.0	752	100.0	1509	100.0

Notes: Bold figures represent the highest proportion

Table 5.5: Characteristics of car users belonging to the dataset selected for the study

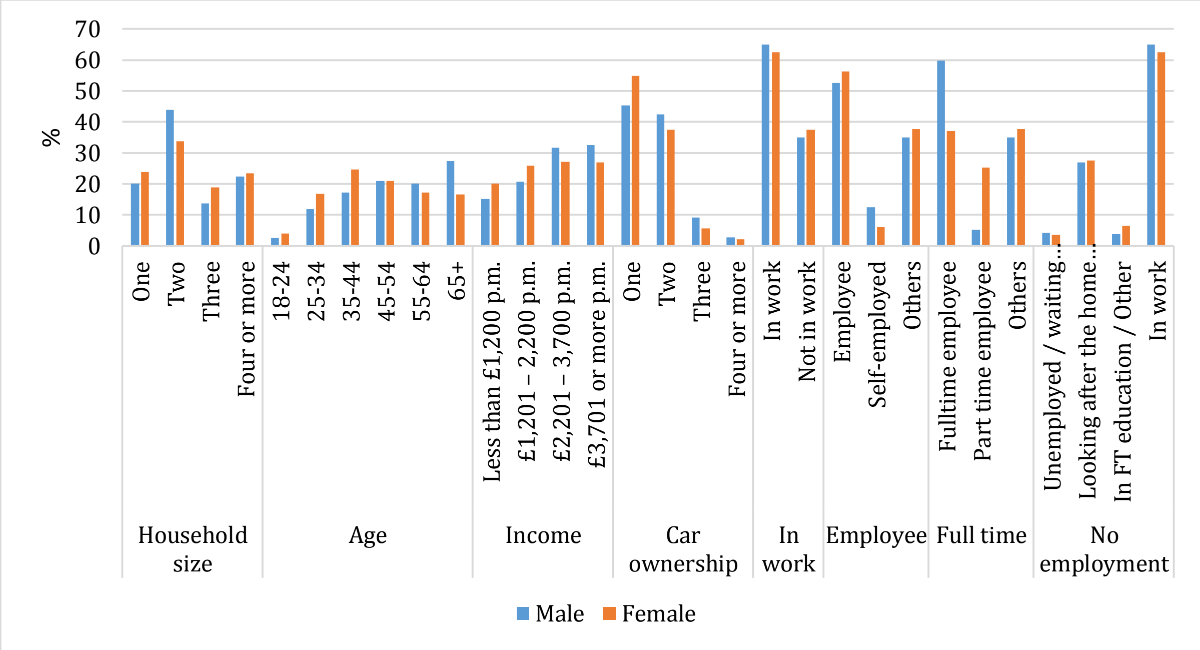


Figure 5.2: Socio-demographic characteristics of car users belonging to the dataset selected for the study

It can be seen clearly from Figure 5.2 that the highest proportion of number of respondents belong to a two members household. Respondents aged more than 18 years old were divided into 6 groups and the highest proportion were from male 65+ years old, whilst the lowest proportion was the younger group (18 – 24 years old). This is probably because they didn't have driving licences or didn't own or use the cars. It is reported that the majority of respondents (90%) were active car users. They owned and used one or two cars per household either as a driver or as a passenger. In addition, 59.05% of the respondents earned ≥ £2201 per month and it is noted that almost half (48.51%) of respondents were in full-time employment.

5.3.3 Proportion of car users by travel behaviour

This study is concerned with identifying cohorts of car users most likely to be persuaded to use their cars less and to use other sustainable travel alternatives. Therefore, those who have no car had been excluded from the analysis. Table 5.6 shows that half of the car users (50.1%) owned only one car or van in their household, 40.03% indicated that they owned two, and 7.42% three. Only 2.45% of the car users had four or more cars or vans in their household.

No. of car	Count	%
One	756	50.10
Two	604	40.03
Three	112	7.42
Four or more	37	2.45
Total	1509	100

Table 5.6: Car or van availability in the households with respect to the car user respondents belonging to the sample selected for the study

With regards to travel behaviour patterns, the BSA questionnaire has 5 questions to collect information on how car users currently travel. It can be seen in Table 5.7 that for each question, there were six options from which car users could choose: "every day or nearly every day", "2-5 days a week", "once a week", "less often but at least once a month", "less often than that" and "never nowadays". The essential step in investigating peoples' perceptions of, and attitudes towards travelling, as well as to identify who is likely to switch their choice of travel modes, is to first understand their current travel behaviour.

It is noted that daily travel by car as the driver was the most frequent option selected by respondents (68.99%), followed by never use bike to travel (64.68%), and never use local bus to travel (58.65%). The results demonstrate that very few (0.53%) never travelled by car as the driver, suggesting that most adults in respect of age own and use a driving licence. In contrast, only 0.8% of the car users travelled by train every day.

Conversely, other modes of transport (bicycle, trains and local buses) were used less frequently. Only 5.7% of car users said that they travelled by local bus at least once a week, 1.99% by train at least once a week, and 6.76% by bicycle at least once a week. Moreover, public transport and bicycles were often not an option for travel as some respondents reported that they never used local buses (58.65%), trains (36.65%), or bicycles (64.68%) for their journey. Bicycles showed the highest percentage of non-use compared to other modes. The frequencies and percentages with respect to travel behaviour characteristics are shown in Table 5.7.

Responses	Frequency	%
How often nowadays do you usually travel by car as a driver? (Car_driver)		
Every day or nearly every day	1041	68.99
2-5 days a week	374	24.78
Once a week	61	4.04
Less often, but at least once a month	11	0.73
Less often than that	14	0.93
Never nowadays	8	0.53
How often nowadays do you usually travel by car as a passenger? (Car_passenger)		
Every day or nearly every day	64	4.24
2-5 days a week	323	21.40
Once a week	412	27.30
Less often, but at least once a month	272	18.03
Less often than that	215	14.25
Never nowadays	223	14.78
How often nowadays do you usually travel by local bus? (Bus_usage)		
Every day or nearly every day	22	1.46
2-5 days a week	55	3.64
Once a week	86	5.70
Less often, but at least once a month	166	11.00
Less often than that	295	19.55
Never nowadays	885	58.65
How often nowadays do you usually travel by train? (Train_usage)		
Every day or nearly every day	12	0.80
2-5 days a week	30	1.99
Once a week	30	1.99
Less often, but at least once a month	245	16.24
Less often than that	639	42.35
Never nowadays	553	36.65
How often nowadays do you usually travel by bicycle? (Bike_usage)		
Every day or nearly every day	27	1.79
2-5 days a week	63	4.17
Once a week	102	6.76
Less often, but at least once a month	127	8.42
Less often than that	214	14.18
Never nowadays	976	64.68
Total	1509	100.0

Table 5.7: Car users' travel behaviour related characteristics

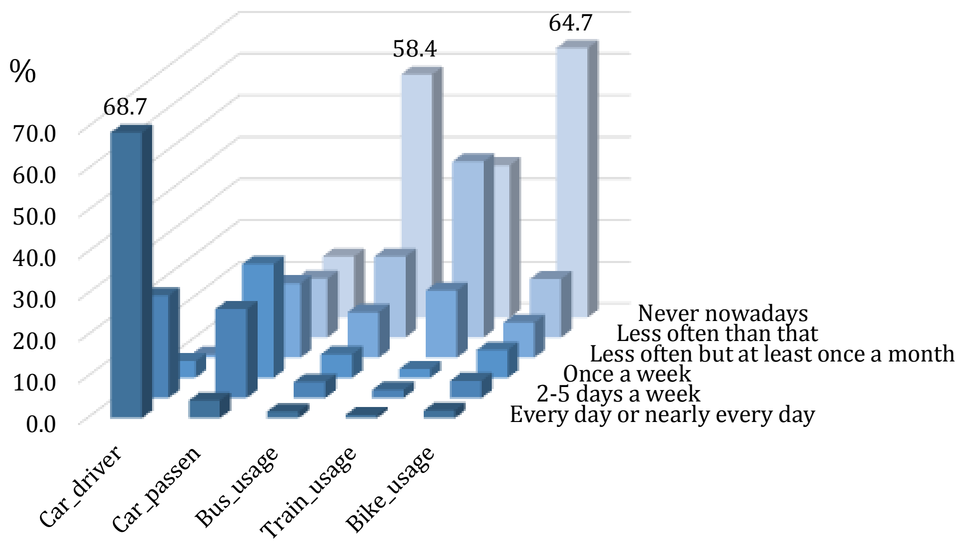


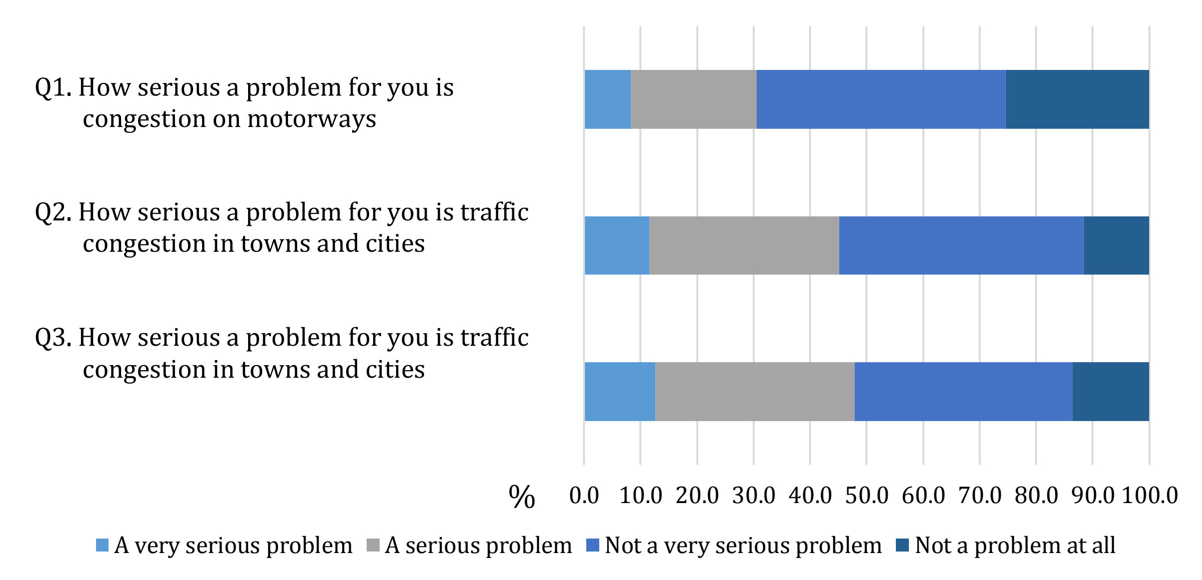
Figure 5.3: Car users' travel behaviour related characteristics

Figure 5.3 is a plot of the data in Table 5.7 which show clearly a trend of decreasing car use as a driver, from every day to never use nowadays. In contrast, the use of public transport, such as buses and trains, shows a reverse trend with a significant rise above daily use to never use for cycling. This data illustrated the car users' reluctance towards and unwillingness to use sustainable modes exposing the challenges of encouraging mode shift. This is discussed in Chapter 6.

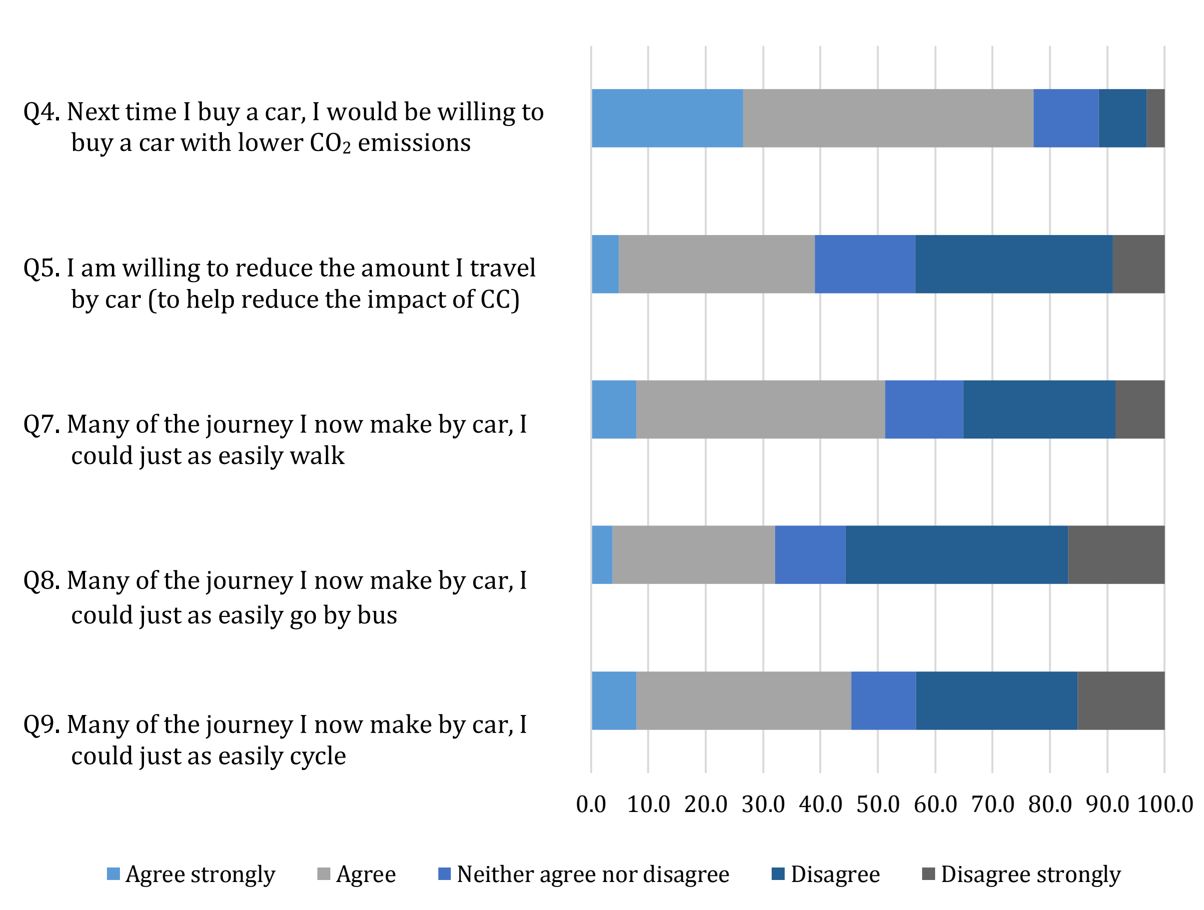
5.4 Exploring Attitudinal Variables

This section sets out to gain an insight into the attitudes of car users towards environmental and climate change issues by analysing responses to the attitudinal variables. In the large scale of the BSA dataset, only specific questions from the transport section were selected as relevant to conduct this study.

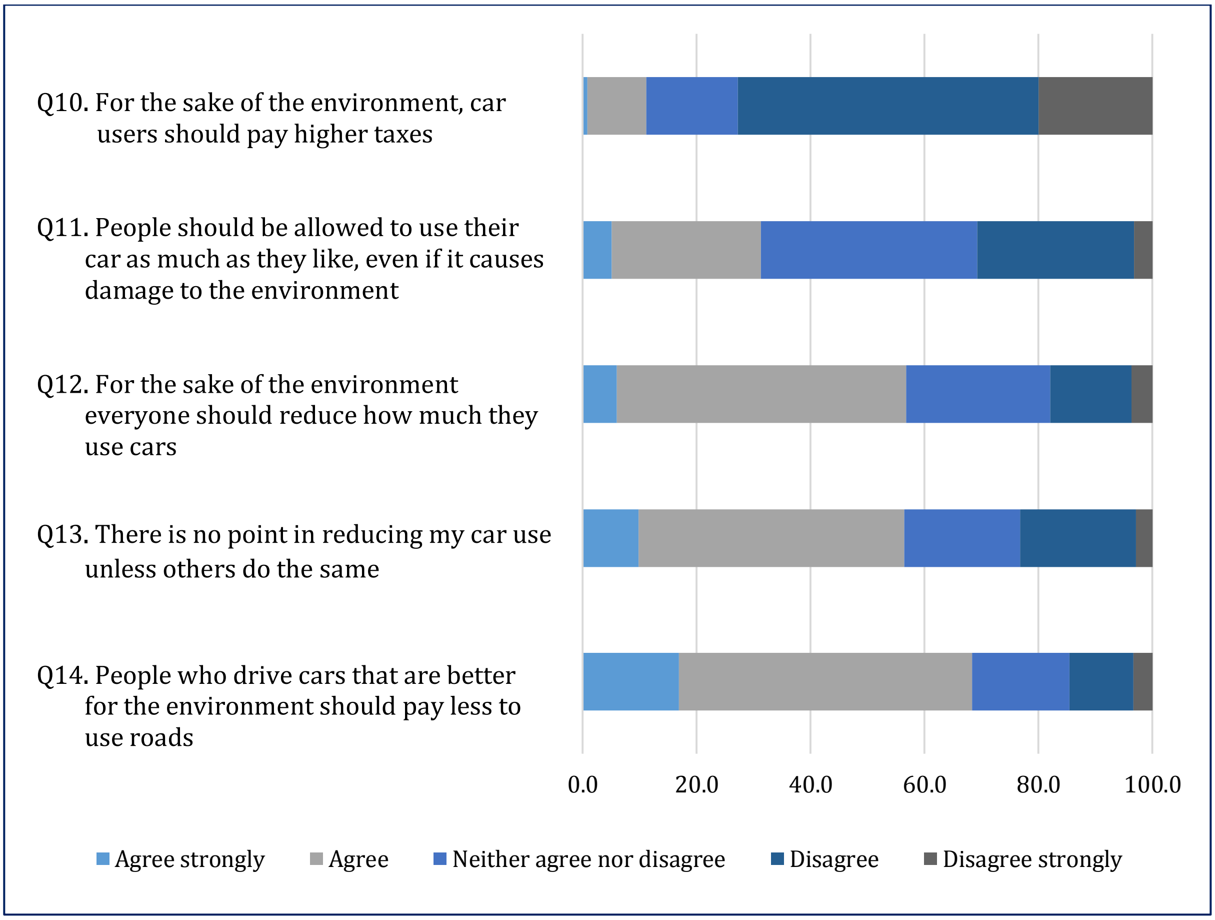
Figure 5.4 shows the attitudinal questions extracted from the BSA datasets along with the proportion of respondents who returned response options one of 4 (questions 1-3) or 5 (questions 4–10) except question 6. These questions used a specified Likert scale to measure the respondent's perception and attitude to a particular statement or question. For 4 options these were scale 1 "A very serious problem", scale 2 "A serious problem", scale 3 "Not a very serious problem" and scale 4 "Not a problem at all"; and for the 5 options, scale 1 means 'strongly agree', scale 2 implies 'agree', scale 3 represents 'neither agree nor disagree', scale 4 illustrates 'disagree', and scale 5 represents 'strongly disagree'. Individual responses using Likert scales are usually treated as ordinal data (Likert, 1932).



(a)



(b)



(c)

Figure 5.4: Attitudinal variables (a) Questions 1 – 3 with 4 Likert-scale options; (b) Questions 4 – 9 except 6 with 5 Likert-scale options; (c) Questions 10 – 14 with 5 Likert-scale options.

Figure 5.4 (a) demonstrates that car users acknowledged that traffic congestion on motorways, in towns and cities is a serious problem. Furthermore, Figure 5.4 (b) indicates that car users were aware of the importance of reducing carbon dioxide emissions and very keen to buy a lower emissions vehicle in the future. However, there were some who were reluctant to give up cars and change behaviour from car to walking, cycling or use local buses. It can be seen clearly in Figure 5.4 (c) that charging taxes is definitely not something that car users were in favour of. They were very hostile towards paying taxes, even for the sake of environment.

Question 6 was approached quite differently in the sense that there were three unique options for assessing an individual's view on the causes of climate change. These were: option 1, "I believe that climate change is taking place and is, at least partly, a result of human actions (81.7% of respondents); option 2, "I believe that climate change is taking place but not as a result of human actions" (13.5%); and option 3, "I don't believe that climate change is taking place" (4.8%). It is clear that the majority of the car user population acknowledge that climate change results from human actions. However, little more can be added by analysing this question in isolation of users of other modes. This is outside the scope of this thesis.

For ordinal data with descriptive codes without numerical value, both the median and the mode can be used as measures of the average. The median is the value which occupies the middle position when all the observations are arranged in increasing or decreasing order, whereas the mode is the most frequently occurring value in a set of nominal data (Sundar Rao and Richard, 2012; Onunwor *et al.*, 2014). Whilst some authors (Aristodemou, 2014) consider the median to be the most suitable measure for ordinal data, others believe that the mode is the only appropriate measure for nominal data.

Using the Shapiro-Wilk (n<2000 samples) statistic to test for normality, it was found that in all cases the distributions were not normally distributed at the 95% statistical significance level. Therefore, non-parametric tests were used throughout the analysis in this thesis. The medians and mean Likert scores over all car users for the attitudinal variables which were designed to reflect perceptions are given in Table 5.8.

The mean for each variable, presented for completeness, assumes that the Likert Scale scores are interval data (Knapp, 1990) although known to be not normally distributed. The reason for providing the mean as well as the median is to allow a more complete discussion regarding the relative differences between the attitudinal variables, because the mean better reflects the distribution. Also, as can be seen from Table 5.8, the medians are mostly the same showing little granularity. It should be noted that standard deviations of data are not presented in the table because the distributions are not normal and therefore, are not used in any of the formal statistical tests which were carried out.

No.	Attitudinal variables	Mean	Median	Mode
Q1	How serious a problem for you is congestion on motorways?	2.87	3.00	3.00
Q2	How serious a problem for you is traffic congestion in towns and cities?	2.55	3.00	3.00
Q3	How serious a problem for you are exhaust fumes from traffic in towns and cities?	2.53	3.00	3.00
Q4	Next time I buy a car, I would be willing to buy a car with lower CO_2 emissions.	2.11	2.00	2.00
Q5	I am willing to reduce the amount I travel by car (To help reduce the impact of climate change).	3.09	3.00	4.00
Q6	View on climate change and causes.	2.77	3.00	3.00
Q7	Many of the short journeys that I now make by car I could just as easily walk.	2.85	2.00	2.00
Q8	Many of the short journeys that I now make by car I could just as easily go by bus.	3.37	4.00	4.00
Q9	Many of the short journeys that I now make by car I could just as easily cycle.	3.05	3.00	2.00
Q10	For the sake of the environment, car users should pay higher taxes.	3.81	4.00	4.00
Q11	People should be allowed to use their cars as much as they like, even if it is a cause of damage to the environment.	2.98	3.00	3.00
Q12	For the sake of the environment, everyone should reduce how much they use cars.	2.59	2.00	2.00
Q13	There is no point in reducing my car use to help the environment unless others do the same.	2.60	2.00	2.00
Q14	People who drive cars that are better for the environment should pay less to use roads.	2.33	2.00	2.00

Table 5.8: Descriptive analysis of attitudinal responses from car users selected for the study

Based on the results presented in Table 5.8, the following observations can be made:

i. Car users generally were not willing (mean=3.09, median=3 and mode=4) to reduce the amount they travel by car in order to help reduce the impact of climate change. In addition, they were reluctant (mean=3.37, median=4 and mode=4) to switch from car to bus for a short journey of less than 2 miles. Furthermore, car users showed hostility (mean=3.81, median=4 and mode=4) towards paying higher taxes, even for the sake of the environment.

ii. Interestingly, car users considered that congestion and exhaust fumes on motorways as well as in towns and cities was not a very serious problem (median=3 and mode=3). Moreover, they were neutral with the statement that people should be allowed to use their cars as much as they like (mean=2.98, median=3 and mode=3), even when acknowledging that it damages the environment. Also, car users are of the belief that climate change is taking place and is, at least partly, a result of human actions.

iii. Car users were willing (mean=2.11, median=2 and mode=2) to buy a car with lower CO_2 emissions in the future and showed an inclination to switch mode from car to either walking (mean=2.85, median=2 and mode=2) or cycling (mean=3.05, median=3 and mode=2) for a short journey of less than 2 miles. For the sake of the environment, they showed willingness (mean=2.59, median=2 and mode=2) to reduce how much they use cars, whilst at the same time indicating that there is no point in reducing car use unless others do the same (mean=2.60, median=2 and mode=2) and users of lower emission vehicles should pay less to use roads (mean=2.33, median=2 and mode=2).

5.5 Correlation Analysis

A correlation between variables is a measure of how closely the variables are related. The most common measure of correlation in statistics is the Pearson Correlation for normally distributed data, which shows a linear relationship between two variables (Kinnear and Gray, 2000; Rumsey, 2007). Since the dataset is not normally distributed, based on the Shapiro-Wilk normality test conducted, the Spearman's rank-order correlation (ρ, also denoted by r_s) for the non-parametric analysis was chosen to measure the strength and direction of association between two variables. Spearman's coefficient is appropriate for both continuous and discrete ordinal variables (Myers and Well, 2003; Lehman *et al.*, 2005).

Spearman's correlation coefficient is a statistical measure of the strength of a monotonic relationship between paired data. The formula for the Spearman's correlation coefficient is as follow:

$$\rho = 1 - \frac{6 \quad d_i^2}{n \;\; n^2 - 1}$$

where:

ρ = is the coefficient of Spearman correlation (rho),

d_i = is the difference in rank between paired values of X and Y,

n = is the sample size or the number of paired values X and Y in the selected sample.

ρ lies between -1 and 1 $\;-1 \leq \rho \leq +1\;$, where -1 means that there is a perfect negative correlation between the two variables, and a result of +1 means that there is a perfect positive correlation between them. The closer the value of ρ gets to zero, the greater the variation of the data points around the line of best fit. While a positive correlation means that the other variable has a tendency to increase with the first variable, a negative correlation means that the other variable has a tendency to decrease.

5.5.1 *Correlation analysis of socio-demographic variables*

The correlation analysis of socio-demographic variables, which are number of people living in the household (HH#), gender, age, number of cars in the household (Car#) and employment status (EmpStatus), is shown in Table 5.9.

Classification		HH#	Gender	Age	Car#	EmpStatus
HH#	Correlation	1	0.01	**-0.40****	**0.42****	-0.22**
	p-value		0.58	0.00	0.00	0.00
	N		1509	1509	1509	1509
Gender	Correlation		1	-0.16**	-0.10**	0.07**
	p-value			0.00	0.00	0.01
	N			1509	1509	1509
Age	Correlation			1	-0.09**	**0.49****
	p-value				0.00	0.00
	N				1509	1509
Car#	Correlation				1	-0.16**
	p-value					0.00
	N					1509
EmpStatus	Correlation					1
	p-value					
	N					

**Correlation is significant at the 0.01 level (2-tailed)

Table 5.9: Correlation analysis of socio-demographic variables

As can be seen from Table 5.9, 'Age' is significantly correlated at the 99% significance value 0.000 (<0.01) with 'Employment status' (ρ=0.49), and "Household size" (ρ =-0.40), whilst "Household size" also correlated with "Car ownership" (ρ=0.42). There is a negative correlation between age and household size, suggesting that older people tend to live in smaller households, perhaps because their children have left home or, in some cases, spouses have died. There is a positive correlation between household size and number of cars. This is perhaps not surprising now that households often have more than one car. There is also, apparently, a positive correlation between age and employment status. This result indicates that different age distributions are associated with different employment statuses. These relationship can be considered fairly strong (Posavac, 2015). Therefore, since the values of correlations for the rest of the variables are less than 0.4, it can be concluded that they exhibit weak correlations.

5.5.2 Correlation analysis of attitudinal questions

In Table 5.10 the Spearman's correlation values for the attitudinal variables are presented. The correlation signs between Q11 (People should be allowed to use their cars as much as they like, even if it causes damage to the environment) and the other variables were negative except for Q1 (How serious a problem for you is congestion on motorways?) and Q6 (View on climate change and causes). This means that there were inverse relationships.

On the other hand, there were significant positive correlations (0.54) between Q1 and Q2 (How serious a problem for you is traffic congestion in towns and cities?), as well as Q2 and Q3 (How serious a problem for you are exhaust fumes from traffic in towns and cities?) which is 0.50. This indicates the existence of positive relationships between aspects of traffic congestion and exhaust fumes.

Meanwhile, the correlations between Q7 (Many of the short journeys that I now make by car I could just as easily walk) with Q8 (Many of the short journeys that I now make by car I could just as easily go by bus), and Q9 (Many of the short journeys that I now make by car I could just as easily cycle) are positive 0.40 and 0.58 respectively. This positive relationship suggests that there is potential to use incentives or provide facilities to overcome barriers for not using sustainable transport options.

An interesting indication is that the sign of the correlation between Q11 and Q13 (There is no point in reducing my car use to help the environment unless others do the same) is weakly positive 0.17, whereas the sign of the correlation between Q11 and Q12 (For the sake of the environment everyone should reduce how much they use cars) is negative (-0.35). It can be understood that car users who have negative perceptions on environmental issues tend to acknowledge similar statements, but deny the importance of reducing car use and are reluctant to do so for the sake of environment. This correlation is understandable. The majority of sample Spearman's Correlation coefficient (ρ) between the variables are less than 0.4 which suggests that correlations between these variables are generally weak.

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Q1 Cong_MWs	ρ	1	**0.54****	0.37**	0.01	-0.01	0.00	0.05*	0.03	0.09**	-0.00	0.02	-0.01	0.01	-0.01
	p-value		0.00	0.00	0.57	0.81	0.97	0.04	0.28	0.00	0.87	0.38	0.61	0.64	0.63
Q2 Cong_Cities	ρ		1	**0.50****	0.02	0.00	-0.03	0.02	0.07*	0.05	-0.04	-0.02	0.02	0.03	0.02
	p-value			0.00	0.39	0.96	0.23	0.43	0.01	0.07	0.18	0.55	0.45	0.21	0.41
Q3 Exhaustfumes	ρ			1	0.15**	0.14**	-0.13**	0.08**	0.12**	0.08**	0.11**	-0.15**	0.17**	-0.06*	0.10**
	p-value				0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00
Q4 BuyLowEmi	ρ				1	0.25**	-0.17**	0.06*	0.07**	0.08**	0.13**	-0.18**	0.24**	-0.12**	0.32**
	p-value					0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Q5 ReducTravCar	ρ					1	-0.18**	0.21**	0.24**	0.21**	0.24**	-0.24**	0.38**	-0.17**	0.10**
	p-value						0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q6 CCView	ρ						1	-0.07**	-0.03	-0.06*	-0.17**	0.18**	-0.24**	0.03	-0.17**
	p-value							0.01	0.18	0.02	0.00	0.00	0.00	0.34	0.00
Q7 CartoWalk	ρ							1	**0.40****	**0.58****	0.15**	-0.12**	0.21**	-0.01	0.08**
	p-value								0.00	0.00	0.00	0.00	0.00	0.79	0.00
Q8 CartoBus	ρ								1	0.32**	0.15**	-0.15**	0.17**	-0.02	0.05
	p-value									0.00	0.00	0.00	0.00	0.43	0.07
Q9 CartoBike	ρ									1	0.14**	-0.12**	0.22**	-0.01	0.07**
	p-value										0.00	0.00	0.00	0.62	0.00
Q10 HiTaxCarUse	ρ										1	-0.26**	0.27**	-0.16**	0.16**
	p-value											0.00	0.00	0.00	0.00
Q11 AllowCarUse	ρ											1	-0.35**	0.17**	-0.15**
	p-value												0.00	0.00	0.00
Q12 ReducCarUse	ρ												1	-0.12**	0.29**
	p-value													0.00	0.00
Q13 ReducCarUse_NP	ρ													1	0.08**
	p-value														0.00
Q14 CarBetterPayLess	ρ														1

**Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

Table 5.10: Correlation analysis of attitudinal variables

5.6 Conclusions

This chapter focuses on carrying out descriptive analysis of the BSA data from 2011 to 2014 to obtain a better understanding of the characteristics of car users, the proportion of responses, and preparation of the data for further analysis. In terms of the yearly distributions of car users, the sample represents a similar proportion each year from 2011 to 2014 and is representative of the BSA population.

From the socio-demographics variables, the distributions of age, gender, car ownership, and household size were presented. It is noted that car users in the younger-aged (18-24 years old) group were the lowest proportion of participants in this study (3.25%). It can be assumed that they didn't have a driving licence or didn't own any car at this age. This is because at this age they are normally in full-time or part-time education. On the other hand, the assessments of travel patterns show that car was by far the most commonly used mode compared to other transportation modes. Hence, the opinions and responses of the subjects will give insight into how to persuade people to use their cars less and switch to alternatives modes to reduce the road burden and the impact of environmental problems.

The correlation analyses of socio-demographic and attitudinal variables were conducted separately. These analyses demonstrated the relationships between and the levels of dependency of each variable. Essentially, an overview of general feelings of car users can be obtained by descriptive statistics of attitudinal variables. However, that information alone cannot be used to identify characteristics of specific cohorts that have potential for behavioural change. Therefore, the next step is to conduct factor and cluster analyses. The structure of the attitudinal variables is investigated in further detail using factor analysis in the next chapter.

Chapter 6 Exploration of the Structure of Data by Factors

6.1 Introduction

Preliminary analysis and general characteristics of respondents were discussed in Chapter 5 showing an insight into the general characteristics of car users using descriptive statistics. This analysis revealed a fundamental understanding of socio-demographics, travel behaviour and attitudinal variables of car users. However, in order to explore patterns of attitudes of car users, the next step is to conduct a dimension reduction technique. Therefore, this chapter investigates the structure of data by factors, using principal axis factoring (PAF) and principal component analysis (PCA). The following Section 6.2 investigates the normality of 14 attitudinal variables. Section 6.3 discusses the method used for dimension reduction. Section 6.4 determines changes evident between 2011 and 2014. Section 6.5 explores the details of attitudes and perceptions of transport and the environment. Section 6.6 presents the perceptions and attitude to traffic awareness. Next, in Section 6.7, respondents' willingness to switch travel behaviour is presented followed by a summary to conclude the chapter in Section 6.8.

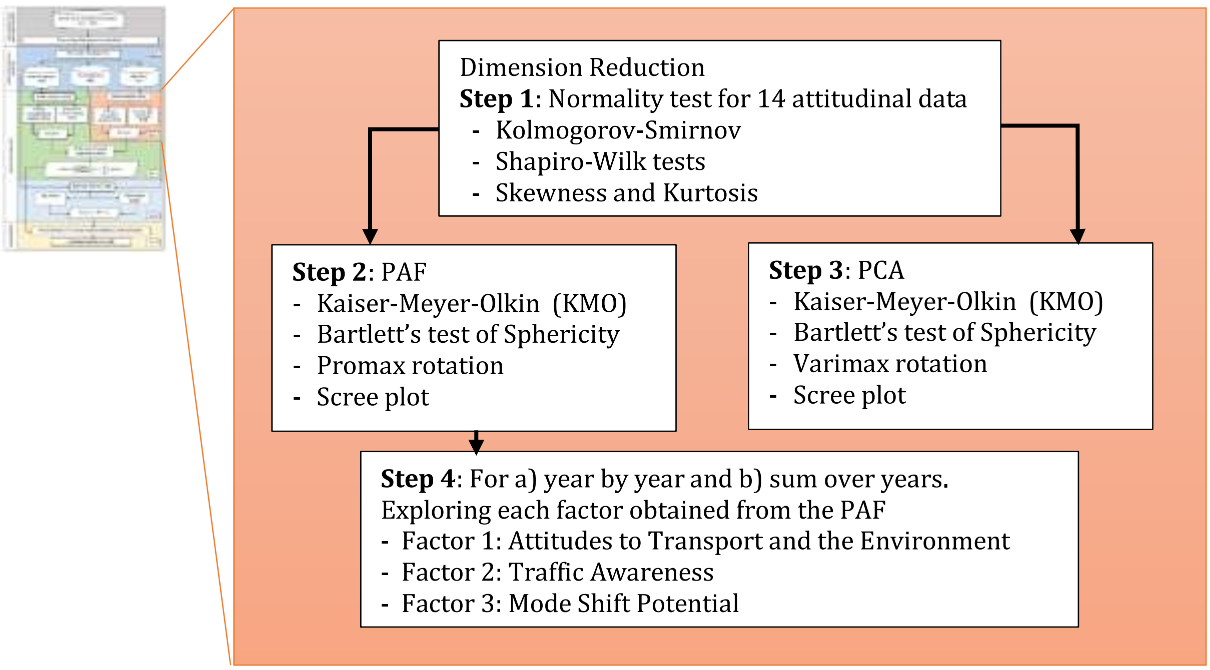


Figure 6.1: Steps involved in Chapter 6

6.2 Normality Test

Firstly, the attitudinal variables were analysed to investigate whether the data are normally distributed. This was achieved in four ways: using the Kolmogorov-Smirnov test ($p<0.05$) (Razali and Wah, 2011); the Shapiro-Wilk test ($p<0.05$) (Shapiro and Wilk., 1965; Razali and Wah, 2011) along with a visual inspection of their histogram; a normal Q-Q plots and box plots.

6.2.1 Kolmogorov-Simornov

The Kolmogorov-Smirnov test is a non-parametric test to determine whether the dataset differs significantly. The Lilliefors significance correction was used to improve or for correcting the Kolmogorov-Smirnov test for small values at the tails of probability distributions. With reference to Table 6.1, p-values are <0.05. This indicated that the variables were not normally distributed.

6.2.2 Shapiro-Wilk

In terms of the Shapiro-Wilk test, the null hypothesis for this test of normality is that the data are normally distributed. Based on Table 6.1, p-values are <0.05, therefore the null hypothesis is rejected and it can be assumed that attitudinal variables are not normally distributed. The results of the Kolmogorov-Smirnov and Shapiro-Wilk tests both show clearly that the distribution of responses for the attitudinal variables were statistically significant by difference at the 99.99% significant level.

No.	Variables	Kolmogorov-Smirnov[a] df = 1509		Shapiro-Wilk df = 1509	
		Statistic	Sig.	Statistic	Sig.
1	Cong_MWs	0.26	0.00	0.86	0.00
2	Cong_Cities	0.25	0.00	0.87	0.00
3	Exhaustfumes	0.22	0.00	0.88	0.00
4	BuyLowEmi.	0.32	0.00	0.81	0.00
5	ReducTravCar	0.23	0.00	0.88	0.00
6	CCViews	0.49	0.00	0.49	0.00
7	CartoWalk	0.28	0.00	0.86	0.00
8	CartoBus	0.26	0.00	0.87	0.00
9	CartoBike	0.25	0.00	0.87	0.00
10	HiTaxForCarUse	0.31	0.00	0.84	0.00
11	AllowCarUse	0.20	0.00	0.90	0.00
12	ReducCarUse	0.30	0.00	0.84	0.00
13	ReducCarUse_NP	0.29	0.00	0.87	0.00
14	CarBetterPayLess	0.31	0.00	0.84	0.00

df: degree of freedom
a: Lilliefors Significance Correction

Table 6.1: Tests of normality for attitudinal data

6.2.3 Skewness and kurtosis

Regarding skewness and kurtosis, Table 6.2 shows the attitudinal variables are skewed $-37.00 < z < 0.17$ and kurtotic $-41.23 < z < 0.03$ for all 14 variables differ significantly from normality. Therefore, these independent tests confirm that the distribution of responses to all 14 attitude questions are not normally distributed. As can be seen in Table 6.2, the z-score values of the skewness and kurtosis were outside the range of $-1.96 < z < 1.96$ (Cramer, 1998; Cramer, 2004; Doane and Seward, 2011) for most of the variables except for "ReducTravCar" (skewness 0.17), HiTaxForCarUse (kurtosis 1.54) and ReducCarUse (kurtosis 0.03) variables.

No.	Variables	Statistics		z-score	
		Skewness SE = 0.06	Kurtosis SE = 0.13	Skewness	Kurtosis
1	Cong_MWs	-0.45	-0.51	-7.50	-3.92
2	Cong_Cities	-0.15	-0.57	-2.50	-4.38
3	Exhaustfumes	-0.05	-0.70	-0.83	-5.38
4	BuyLowEmi.	1.06	0.86	17.67	6.62
5	ReducTravCar	0.01	-1.13	**0.17**	-8.69
6	CCViews	-2.22	3.95	-37.00	30.38
7	CartoWalk	0.33	-1.06	5.50	-8.15
8	CartoBus	-0.25	-1.15	-4.17	-8.85
9	CartoBike	0.12	-1.28	2.00	-9.85
10	HiTaxForCarUse	-0.74	0.20	-12.33	**1.54**
11	AllowCarUse	-0.88	-5.36	-14.67	-41.23
12	ReducCarUse	0.73	0.004	12.17	**0.03**
13	ReducCarUse_NP	0.46	-0.63	7.67	-4.85
14	CarBetterPayLess	0.83	0.28	13.83	2.15

*SE = standard error

Table 6.2: Standard error, skewness and kurtosis for attitudinal variables

6.3 Method used for Dimension Reduction

As discussed earlier in Chapter 3, dimension reduction is a technique to reduce the 14 attitudinal variables to a smaller number. First, the results of the exploratory factor analysis (EFA) by using principal axis factoring (PAF) are presented, followed by principal component analysis (PCA).

6.3.1 Principal axis factoring (PAF)

Figure 6.2 shows the framework of PAF where F is the factor; $Y_1, Y_2, Y_3,$ and Y_4 are observed variables; $u_1, u_2, u_3,$ and u_4 are the random errors; and $b_1, b_2, b_3,$ and b_4 represent the factor loadings of $Y_1, Y_2, Y_3,$ and Y_4.

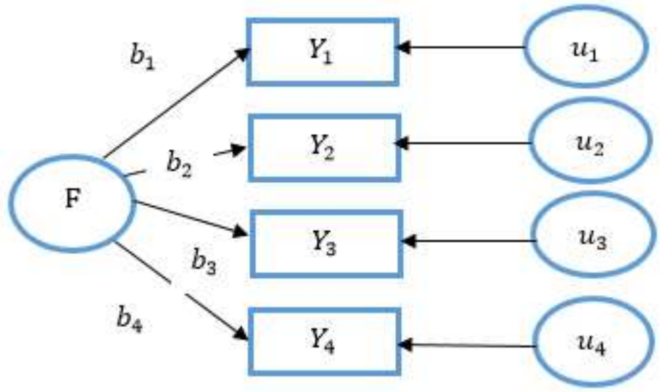


Figure 6.2: Principal axis factoring

PAF evaluates a factor or latent variable which is found to have influence on observed variables. PAF, in identifying a factor, considers common variance of each variable and separating the common variance from the unique variance accounts for co-variation.

PAF was used to capture any relationships that might exist between attitudinal variables. By using PAF, an independent analysis of the data is conducted in an attempt to confirm similarity in the patterns in the data to those revealed by the other statistical approaches. Oblique rotation was used to reduce the 14 attitudinal variables of perceptions and attitudes concerned with travelling listed in the BSA dataset by exposing commonalities within variables. PAF is widely used to simplify large sets of data into reduced numbers of factors by grouping data in a statistical way.

As discussed earlier in Chapter 3, the minimum sample required for PAF is 50 (Hair *et al.*, 2006). This research has a total sample of 1509 respondents which was found to be sufficient to give statistical significance in the results. PAF is a technique that can accommodate commonality in attitudinal variables and reveals multicollinearity between variables. Therefore, by identifying the correlations between factors, variables can be combined into fewer factors and, from the analysis, the 14 attitudinal variables can be reduced in number.

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy using Bartlett's test of sphericity was used to test whether the variables are suitable for structure detection (Hotelling, 1933; Bartlett, 1950; Field, 2009; Hair *et al.*, 2010). Sarstedt and Mooi (2014) highlighted that a KMO value of less than 0.5 is considered unacceptable and over 0.80 is preferable in determining the sampling adequacy of the correlations. Given a KMO of 0.74 (see Table 6.3) in this research, the adequacy of the correlation is considered meritorious, being so close to a value of 0.80.

The Bartlett's test of sphericity was used to find the statistical significance of all the correlations within the correlation matrix as an indicator of the strength of the relationships among variables. This test was used to test the null hypothesis that the variables in the population correlation matrix were uncorrelated. Bartlett's test of sphericity was applied with the Chi-square (χ^2) critical value of 3742.51 for a statistical significant level of 99%, p<0.001, as shown in Table 6.3. This result indicates that there was a redundancy between variables that can be summarised with some factors.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.74
Bartlett's Test of Sphericity	Approximate χ^2	3742.51
	Degree of freedom	91
	Statistical significance	0.00

Table 6.3: KMO and Bartlett's test

In this study, because the variables were not-normally distributed, EFA follows a principal axis factoring (PAF) approach to obtain the best results (Osborne and Castello, 2005) with some specific types of rotation methods (either Promax or Varimax), with the constructs based on those that exceed an eigenvalue of one. Tabachnick and Fidell (2013) suggested that, if any of the absolute values of the factor correlation matrix are greater than 0.32, Promax rotation should be selected for oblique rotation. In contrast, if the absolute value is smaller than 0.32, Varimax rotation should be selected. When applied to these data, the absolute value of the correlation between two factors considered in the analysis was 0.37, as shown in Table 6.4. Since this is greater than 0.32, Promax rotation is recommended and initially was selected for the PAF to identify the absolute value of a factor correlation matrix.

Factor	1	2	3
1	1	**0.37**	0.10
2	**0.37**	1	0.10
3	0.10	0.10	1

Extraction Method: Principal Axis Factoring.
Rotation Method: Promax

Table 6.4: Factor correlation matrix

The scree plot is a sliding curve displaying an output of the eigenvalues on the *y*-axis and the number of factors on the *x*-axis. The number of factors that should be produced by the analysis depends on the point where the slope of the curve is levelling off (Yong and Pearce, 2013). The scree plot in Figure 6.3 shows clearly how the eigenvalues drop sharply after the first four factors, suggesting the extraction of four factors.

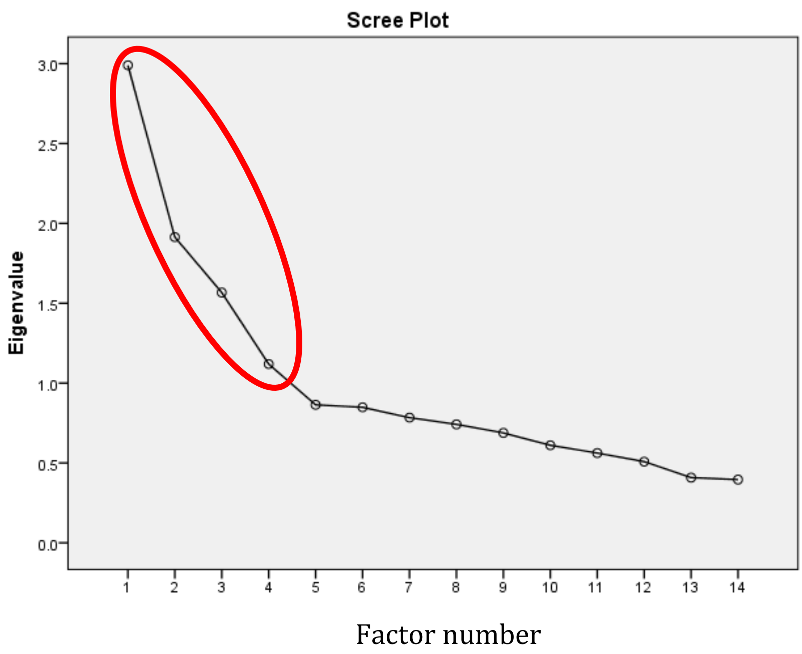


Figure 6.3: Scree plot of PAF with the initially selected factors (Eigenvalue >1) in the oval shape

Based on the Kaiser Criterion (Kaiser, 1960) and the scree plot (Cattell, 1996), the number of factors extracted was decided. All components with eigenvalues below 1.0 were dropped because the first four factors adequately represent the variance in the dataset (Salkind, 2010). However, factor 4, contains only one variable (People who drive cars that are better for the environment should pay less to use roads - CarBetterPayLess) which is not acceptable because there is no covariance to consider except the item's own variance (Raubenheimer, 2004; Osborne and Castello, 2005). They also recommended to use at least 3 items per factor. Based on this advice, the PAF was run again constraining the number of factors to 3; with three factors 34% of variance in the data being accounted for, as shown in Table 6.5.

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings[a]
	Total	% of Variance	Cum %	Total	% of variance	Cum %	Total
1	2.98	21.30	21.30	2.35	16.75	16.75	2.03
2	1.91	13.65	34.95	1.45	10.31	27.06	1.76
3	1.57	11.19	46.14	1.02	7.24	34.30	1.54
4	1.12	8.00	54.15				
5	0.86	6.16	60.30				
6	0.85	6.05	66.35				
7	0.78	5.58	71.93				
8	0.75	5.33	77.26				
9	0.69	4.91	82.16				
10	0.62	4.41	86.58				
11	0.56	4.03	90.60				
12	0.51	3.62	94.22				
13	0.41	2.93	97.16				
14	0.40	2.85	100.00				

Extraction Method: Principal Axis Factoring.
a. When factors are correlated, sums of squared loadings cannot be added to obtain a total variance.

Table 6.5: Percentage of variance explained

Table 6.6 presents the output from the PAF, identifying the variables along with factor loadings and Cronbach alpha values. These factors will be analysed in greater depth in the next section. Based on the dataset of the Likert scale responses of all respondents, the oblique rotation converged in four iterations and three factors emerged from the PAF. These factors were labelled as "Attitudes to transport and the environment", "Traffic awareness" and "Mode shift potential".

The first factor consisted of a willingness to change travel behaviour, to buy cars with low emissions in the future and to pay higher taxes for the sake of the environment; the second factor embraced awareness of traffic congestion and exhaust fumes in towns and cities; and the third consisted of the potential for switching travel modes for short journeys of less than 2 miles.

New factor	Variables	Factor		
		1	2	3
Attitudes to transport and the environment α=0.16	ReducCarUse	0.635		
	AllowCarUse	-0.512		
	BuyLowEmi	0.505		
	ReducTravCar	0.468		
	HiTaxforCarUse	0.442		
	CarBetterPayLess	0.433		
	CCView	-0.384		
	ReducCarUse_NP	-0.246		
Traffic awareness (α=0.73)	Cong_MWs		0.849	
	Cong_cities		0.658	
	Exhaustfumes		0.588	
Mode shift potential (α=0.70)	CartoWalk			0.862
	CartoBike			0.703
	CartoBus			0.459

α: Cronbach's alpha
Extraction method: principal axis factoring.
Rotation Method: Promax

Table 6.6: Pattern matrix (convergence in 4 iterations)

In order to evaluate the reliability of the factors identified in the PAF, Cronbach's alpha is calculated to consider the internal consistency of the grouped statements (Cronbach, 1951). Internal consistency is a method to compute the correlation of each test item with the total score test; items with low correlations (approaching zero) are deleted. If alpha is too high it may suggest that some items are redundant as they are testing the same question but in a different guise. By using Cronbach's alpha, the internal consistency could be measured, that is, how closely related a set of items are as a group.

It is considered to be a measure of scale reliability. Cronbach's alpha is not a statistical test but it is a coefficient of reliability (or consistency). Peterson (1994) indicated that acceptable alpha scores range from 0.5 for a preliminary analysis to 0.9 for applied research. The higher the inter-correlation among the scale items, the greater the reliability of the scale and this can be supported by a high value of Cronbach's alpha. In this study, Cronbach's alpha has been calculated for each new factor identified in the analysis. The Cronbach's alpha values for factors 1. Attitudes to transport and the environment, 2. Traffic awareness and 3. Mode shift potential were 0.16, 0.73, and 0.70 for each factor respectively.

The reason for the low Cronbach's alpha value for the first factor was because it consists of three attitudinal variables with negative factor loading (i. People should be allowed to use their cars as much as they like, even if it causes damage to the environment; ii. There is no point in reducing my car use to help the environment unless others do the same; and iii. View on climate change and causes). After removing these three variables, Cronbach's alpha value was improved to 0.63. In addition, Cronbach alpha is not concerned with factors that this research created, but more with the way that a questionnaire is designed. Since data from a secondary source is used in this study, there was no control over the formulation of the questions.

6.3.2 Principal component analysis (PCA)

Figure 6.4 shows the rather different framework of PCA, where the observed variables Y_1, Y_2, Y_3, and Y_4 use the weights, w_1, w_2, w_3, and w_4 to provide a principal component score C which is a combination of linear variables. The components are selected based on the highest variance scores. PCA is in fact a dimensional reduction technique. A key difference between PAF and PCA is that the variable's variance in PCA is a measure of total variance without separation into a common and unique variance, as is the case in PAF. PCA also accounts for the highest proportion of the variance observed in the variables and breaks down the correlation matrix to discover the principal components.

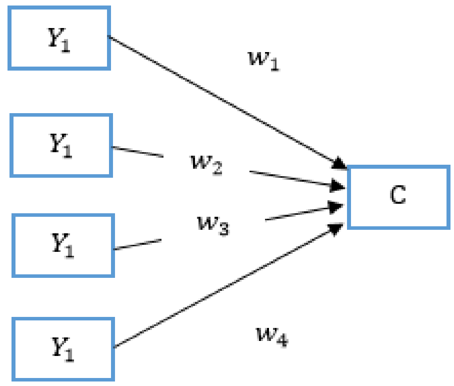


Figure 6.4: Principal component analysis

In order to add credibility to the three factors emerging from the PAF, an independent analysis of the data using PCA was conducted to simplify the 14 attitudinal variables of perceptions and attitudes concerning travelling into a reduced number of factors. PCA is a technique to convert a set of observations of feasibly correlated variables into a set of values of linearly uncorrelated variables (Hotelling, 1933). In essence PCA identifies the correlated variables which can then be combined into fewer factors. Varimax orthogonal rotation is by far the most common choice to simplify and clarify the data structure and produces the results which make it as easy as possible to identify each variable with a single factor (Osborne and Castello, 2005).

The rules in PCA were tested using the Bartlett test of sphericity (Field, 2009; Hair *et al.*, 2010) in the same way as PAF above, employing the χ^2 value of 3742.51 at a statistical confidence level of 99% (p<0.001) as shown in Table 6.7. These results were the same by using PAF and PCA.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.74
Bartlett's Test of Sphericity	Approximate χ^2	3742.51
	Degree of freedom	91
	Statistical significance	0.00

Table 6.7: KMO and Bartlett's test

All components with eigenvalues below 1.0 were dropped. Thus, four factors were recognised which accounted for 54% of variance in the data, as shown in Table 6.8. The scree plot in Figure 6.5 demonstrates that the eigenvalues drop sharply after the first four factors, suggesting the extraction of four factors as before.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cum %	Total	% of Variance	Cum %	Total	% of Variance	Cum %
1	2.99	21.36	21.36	2.99	21.356	21.36	2.05	14.61	14.61
2	1.91	13.67	35.03	1.91	13.670	35.03	2.03	14.47	29.08
3	1.57	11.19	46.22	1.57	11.191	46.21	1.96	13.98	43.06
4	1.12	7.99	54.21	1.12	7.994	54.21	1.56	11.16	54.21
5	0.86	6.17	60.38						
6	0.85	6.06	66.44						
7	0.78	5.60	72.04						
8	0.74	5.30	77.34						
9	0.69	4.91	82.25						
10	0.61	4.36	86.61						
11	0.56	4.02	90.63						
12	0.51	3.63	94.26						
13	0.41	2.92	97.17						
14	0.40	2.83	100.0						

Extraction Method: Principal Component Analysis.

Table 6.8: Total variance explained

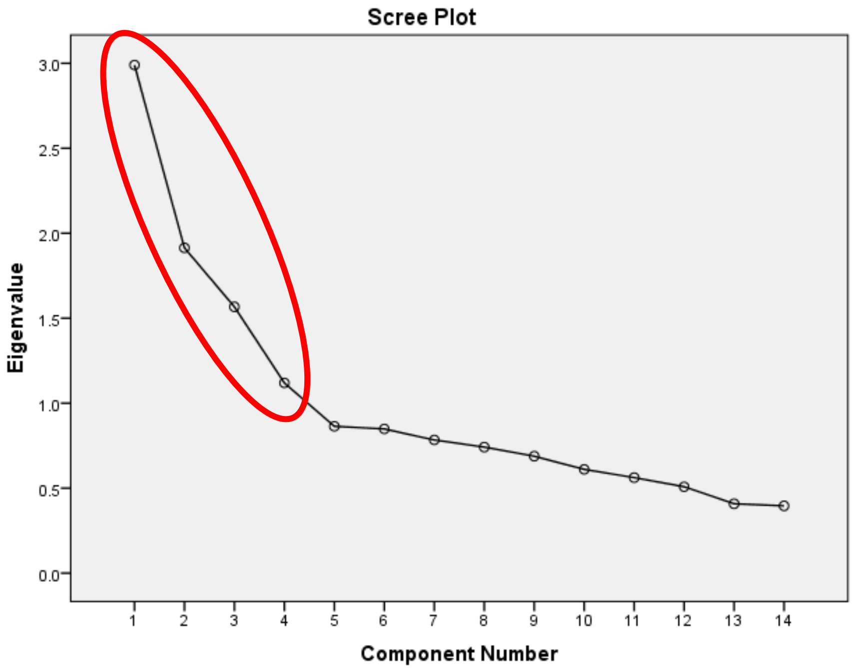


Figure 6.5: Scree plot of PCA with the selected factors in the oval shape

However, unlike PAF, in PCA the orthogonal rotation converged in six iterations and four components instead of the three factors that emerged. Table 6.9 presents the output from the PCA of those variables with component loadings. The first three components, as before, were labelled as "Environmentally sensitive", "Traffic awareness" and "Mode shift potential" because they consisted of the same variables. However, in PCA a fourth component emerged and was labelled as "Attitudes to car use", consisting of a willingness to change travel behaviour and to pay higher taxes for the sake of the environment.

New component	Variables	Component			
		1	2	3	4
Environmentally sensitive $\alpha = 0.334$	CarBetterPayLess	0.765			
	BuyLowEmi	0.617			
	ReducCarUse	0.572			
	CCView	-0.529			
Model shift potential $\alpha = 0.703$	CartoWalk		0.850		
	CartoBike		0.810		
	CartoBus		0.672		
Traffic awareness	Cong_cities			0.863	

$\alpha = 0.730$	Cong_MWs			0.804	
	Exhaustfumes			0.742	
Attitudes to car use $\alpha = -0.336$	ReducCarUse_NP				-0.819
	ReducTravCar				0.494
	AllowCarUse				-0.489
	HiTaxforCarUse				0.412

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax

Table 6.9: Rotated component matrix (convergence in 6 iterations)

A Cronbach's alpha test was carried out on the new factors obtained in the PCA. The higher the inter-correlation among the scale items, the greater the reliability of the scale and this can be supported by a high value of Cronbach's alpha. In order to evaluate the reliability of the factors identified in the PCA, Cronbach's alpha is calculated to consider the internal consistency of the grouped statements (Cronbach, 1951). The Cronbach's alpha values for the four factors were 0.334, 0.730, 0.703 and -0.336 for each factor respectively. Even though the negative components were removed, the Cronbach's alpha values in the first and fourth factors did not improve.

After considering both the PAF and PCA outputs, a detailed scrutiny of results was performed. Table 6.10 lists the comparison of results obtained in both the PAF and PCA dimension reduction methods. Based on the comparison presented in Table 6.10, it was concluded that PAF evidenced better outcomes compared to PCA. Therefore, in the following sections, based on result obtained by using the PAF method, these factors will be analysed in greater depth.

No.	Principal Axis Factoring	Principal Component Analysis
1	The method produced “factor”	The method produced “component”
2	The method took the measurement errors into account	The method did not take into account the measurement errors.
3	The variance was not linked specifically to the factor of an observed variable	Each variable was assumed to be perfectly reliable
4	The method does not produce initial communalities as 1	The method produced initial communalities as 1
5	The method was able to recover weaker factors	The method was not able to discover weaker factors
6	The method aimed to gain an understanding of the structure	The method aimed to determine the structure
7	The results were much more reliable because PAF retained both the unique and error variance	The method removed the unique variance, but did not consider the error variance

Table 6.10: Comparison of PAF and PCA

6.4 Changes Evident Between 2011 and 2014

The analysis carried out so far in this thesis has treated the data collected during the period 2011 to 2014 inclusive as one dataset. Having identified three factors, evidences of changes in attitudes and perceptions from year to year are explored.

Attitudes to the environment: From 2011 to 2014, 81.7% of respondents believed that climate change is taking place and at least partly results from human activities. Questions related to climate change were asked for the first time in 2011 and, with reference to Table 6.11, the awareness was highest in 2011 (84.66%) with small but, based on the χ^2 test, statistically significant differences at the 95% level of confidence being found from year to year.

Year	Count	Total	%
2011	309	365	84.66
2012	317	400	79.25
2013	279	341	81.82
2014	328	403	81.39
Total	1233	1509	81.71

Table 6.11: Proportion of respondents who believe that climate change is taking place and is, at least partly, a result of human actions

However, no statistically significant difference within the 4 years' timescale from 2011 to 2014 was evident for those who did not believe climate change is taking place, albeit a much smaller proportion of the sample (3.84%, 4.75%, 3.81% and 6.70%) for 2011, 2012, 2013 and 2014 respectively.

Attitudes to congestion: A clear trend in opinions regarding serious congestion on the roads can be seen in Figure 6.6, with typically 13-15% more car users finding serious urban road congestion compared to motorways with a rise for both urban and motorways of about 6-8% between 2012 and 2014. Interestingly, car users' concern regarding congestion on motorways and in urban areas reached its lowest point in 2012. These results were not statistically significantly different within the 4 years. Almost one in every two car users in 2014 considered traffic congestion in towns and cities was a very serious or serious problem.

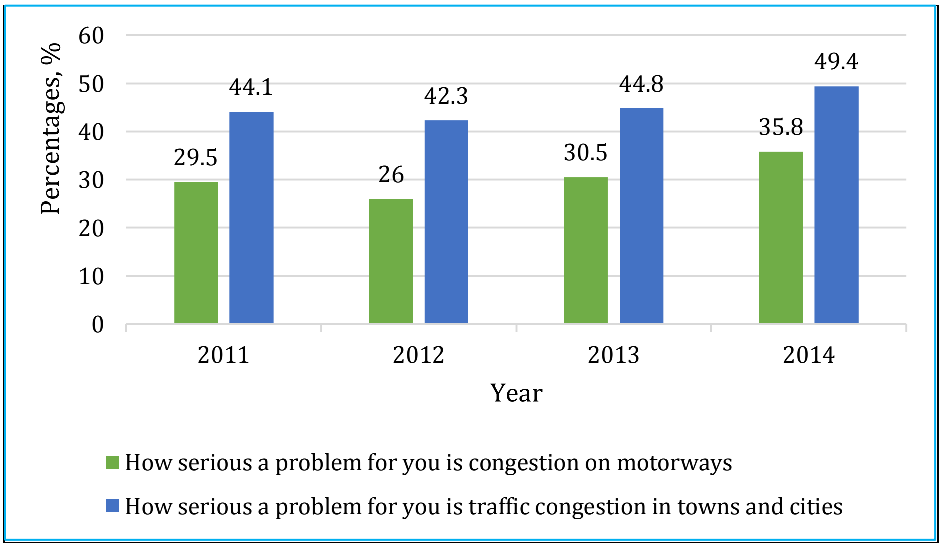


Figure 6.6: Concerns about congestion as a very serious or serious problem

Attitudes to Exhaust Fumes: Traffic is a major source of pollution and is responsible for the majority of air quality management areas declared across the UK road networks. Therefore, respondents' attitudes towards exhaust fumes in towns and cities are an indicator of drivers' awareness of the environmental impact of traffic. Table 6.12 shows how attitudes have fluctuated over time from 2011 to 2013, with increasing concern from 2013 to 2014. The lowest level of concern within the four years was seen in 2012 (45%), and by 2014, half of the respondents (50.12%) considered exhaust fumes from traffic in towns and cities to be a very serious or serious problem.

Year	Count	Total	%
2011	175	365	47.95
2012	180	400	45.00
2013	166	341	48.68
2014	202	403	50.12
Total	723	1509	47.91

Table 6.12: Attitudes towards exhaust fumes in towns and cities as a serious or very serious problem, 2011 – 2014

6.5 Factor 1: Attitudes to Transport and the Environment

Transport and Climate Change: What stands out in Table 6.13, when considering age and gender along with their views on climate change and its causes, is that the youngest (18–24 years old) age group has the lowest percentages believing that climate change is taking place and is a result of human actions, while the middle-aged groups (35–54 year olds) are more likely to believe that climate change is taking place and is, at least partly, due to human actions. In contrast, the proportion of the oldest (65+ years old) aged group (1.99%) who don't believe that climate change is taking place is much higher compared to the other age groups. There is also significant difference by gender, where a higher proportion of females believe that climate change is taking place as a result of human actions compared to males (42.74% compared to 38.97%).

Variable		I don t believe that climate change is taking place		I believe that climate change is taking place but not as a result of human actions		I believe that climate change is taking place and is, at least partly, a result of human actions		Total	
		Count	%	Count	%	Count	%	Count	%
Gender	Male	48	3.18	121	8.02	588	38.97	757	50.17
	Female	25	1.66	82	5.43	645	42.74	752	49.83
Total		73	4.84	203	13.45	1233	81.71	1509	100
Age	18-24	3	0.20	8	0.53	38	2.52	49	3.25
	25-34	10	0.66	27	1.79	178	11.80	215	14.25
	35-44	11	0.73	35	2.32	269	17.83	315	20.87
	45-54	9	0.60	31	2.05	277	18.36	317	21.01
	55-64	10	0.66	44	2.92	228	15.11	282	18.69
	65+	30	1.99	58	3.84	243	16.10	331	21.94
Total		73	4.84	203	13.45	1233	81.71	1509	100

Table 6.13: Variation by gender and age in proportions who believe that climate change is taking place and is, at least partly, a result of human action.

From the χ^2 test, the results indicate that there was a statistically significant association between gender and climate change views and causes and that is, males and females have different thoughts regarding this issue. There was also a statistically significant association across different age groups.

Regarding travel modes considered by car users to have the most impact on climate change aggregating responses over the period 2011 to 2014, the most commonly mentioned mode was cars (41.5%), whilst motorbikes are the lowest mode chosen with a proportion of 0.1% only. The second largest assumed contributor to climate change is aeroplanes (24.6%), followed by vans and lorries (19.3%), buses and coaches (9.6%), ships or ferries (2.1%) and trains (0.4%). There was no significant difference from the BSA population evident from the χ^2 test at 95% significant level.

There were a few respondents who did not believe climate change was taking place or believed that climate change would happen anyway. However, the fact is that the transport sector contributes around 26% of all greenhouse gas (GHG) emissions in the

UK (DfT, 2009). The petrol and diesel used in road transport represents the main source of these emissions. In particular, domestic GHG emissions from cars accounted for 40%, and heavy goods vehicles and light vans for 15% and 10% of UK vehicle emissions respectively (DECC, 2016).

Willingness to Change Travel Behaviour for the Environment: In this analysis, "agree strongly" and "agree" only were combined on the premise that those choosing "neither agree nor disagree", "disagree" and "strongly disagree" were the section of the population unlikely to change behaviour. Attitudes towards a modal shift for environment reasons varied depending on the transport mode used, aggregating across all years 2011 to 2014. Figure 6.7 shows that over two-thirds of respondents (77.2%) reported a willingness to buy a car with lower CO_2 emissions for their next purchase, whereas 56.8% thought that they should reduce the amount of car use for the sake of the environment. However, less than half of those surveyed (39%) reported that they were willing to reduce the amount they travelled by car to help reduce the impact of climate change. Changes from year to year from 2011 to 2014 exhibited a decreasing trend in these attitudes, although this was not found to be statistically significantly different using the χ^2 test within the 4 years.

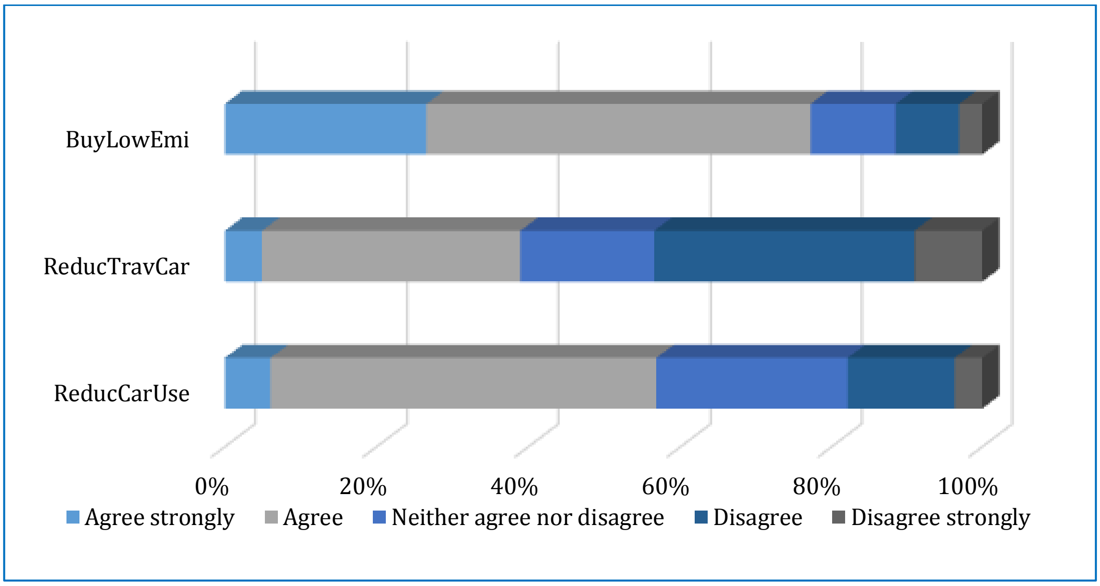


Figure 6.7: Willingness to change travel behaviour for the environment

Considering aggregated data across all years from 2011 to 2014, there are in fact two aspects that are somewhat counterintuitive. Even though the respondents reported that they were willing to take action towards climate change and acknowledge environmental problems, there was a considerable inconsistency with their actual behaviour. Certainly, whilst only 39% of respondents reported that they were willing to reduce the amount they travelled by car to alleviate the impact on climate change, over half of those respondents (57%) agreed that, for the sake of the environment, everyone should reduce how much they use their cars. This suggests that collective action prevails over individual action.

Table 6.14 describes the proportions of respondents who were willing to reduce the amount of travel by car in order to help reduce the impact of climate change, according to age and gender. Middle-aged adults (35–44 years old) and oldest (65+ years old) seem to be the most willing to take action, with the youngest age group (18–24 years old) having the lowest concern. Interestingly, females (21.21%) were more likely to reduce car usage compared to males (17.76%).

Variable		"Strongly agree" or "Agree"		Other responses		Total	
		Count	%	Count	%	Count	%
Gender	Male	268	17.76	489	32.41	757	50.17
	Female	320	21.21	432	28.63	752	49.83
Total		588	38.97	921	61.03	1509	100
Age	18-24	17	1.13	32	2.12	49	3.25
	25-34	94	6.23	121	8.02	215	14.25
	35-44	121	8.02	194	12.86	315	20.87
	45-54	115	7.62	202	13.39	317	21.01
	55-64	106	7.02	176	11.66	282	18.69
	65+	135	8.95	196	12.99	331	21.94
Total		588	38.97	921	61.03	1509	100

Table 6.14: Willingness to reduce the amount of travel by car with respect to age and gender of the respondents

The percentage of respondents who were willing to buy a lower CO_2 emissions car as their next purchase was almost the same proportion as those willing to reduce the amount of travel by car in order to help reduce the impact of travel on climate change, although there is no way in this analysis to suggest these are the same individuals. A gender difference was apparent, with 41.22% of females compared to 35.98% of male, whilst the age factor showed a similar pattern except for the oldest age group (65+ years old) recording the highest percentage (17.76%) and the youngest respondents (18 – 24 year-old) the lowest percentage (1.86%) as shown in Table 6.15.

Variable		"Agree strongly" and "agree"		Other responses		Total	
		Count	%	Count	%	Count	%
Gender	Male	543	35.98	214	14.18	757	50.17
	Female	622	41.22	130	8.61	752	49.83
Total		1165	77.20	344	22.80	1509	100
Age	18-24	28	1.86	21	1.39	49	3.25
	25-34	163	10.80	52	3.45	215	14.25
	35-44	235	15.57	80	5.30	315	20.87
	45-54	249	16.50	68	4.51	317	21.01
	55-64	222	14.71	60	3.98	282	18.69
	65+	268	17.76	63	4.17	331	21.94
Total		1165	77.20	344	22.80	1509	100

Table 6.15: Willingness to buy a car with lower CO_2 emissions in future with respect to age and gender of the respondents

Opinions on the Environment and Car Travel: Aggregating car users' attitudes over the period 2011 to 2014, there was a similar proportion 31% agreeing and 31% disagreeing, with 38% giving neutral answers, when asked whether people should be allowed to use their cars as much as they like, even if it caused a damage to the environment. Surprisingly, there were small differences between attitudes to car travel and damage to the environment, even though people were willing to give up car use if it caused environmental damage. Nonetheless, this needs to be reconciled against 39% of respondents prepared to reduce their own travel by car and 57% who believe that everyone should act (see Figure 6.7).

Figure 6.8 presents respondents' attitudes towards car travel and the environment: just over two thirds, 68% of respondents, agreed that low emission vehicles should pay less to use the roads. Whilst in contrast, 73% disagreed that, for the sake of the environment, car users should pay higher taxes. On the other hand, more than half (56%) felt there was no point in reducing their car use to help the environment unless others do the same. Taken together, these statements suggest that there is an association between attitudes towards environmental issues and transport policies.

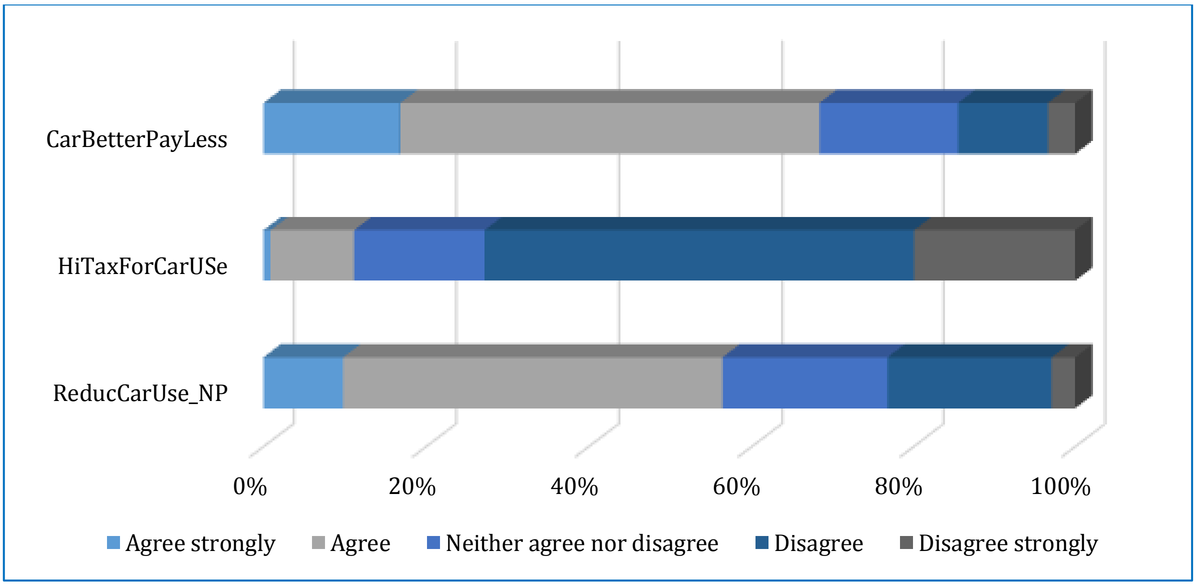


Figure 6.8: Attitudes towards car travel and the environment

Collating together those responses relating to car users changing their behaviour to save the environment, Figure 6.9 illustrates the selfishness of drivers with reluctance to pay taxes to use private cars, tempted by financial incentives to drive low emissions vehicles with no point in reducing car use to help the environment unless others do the same, and with a belief that people should be allowed to use their cars as much as they like, even if it causes damage to the environment.

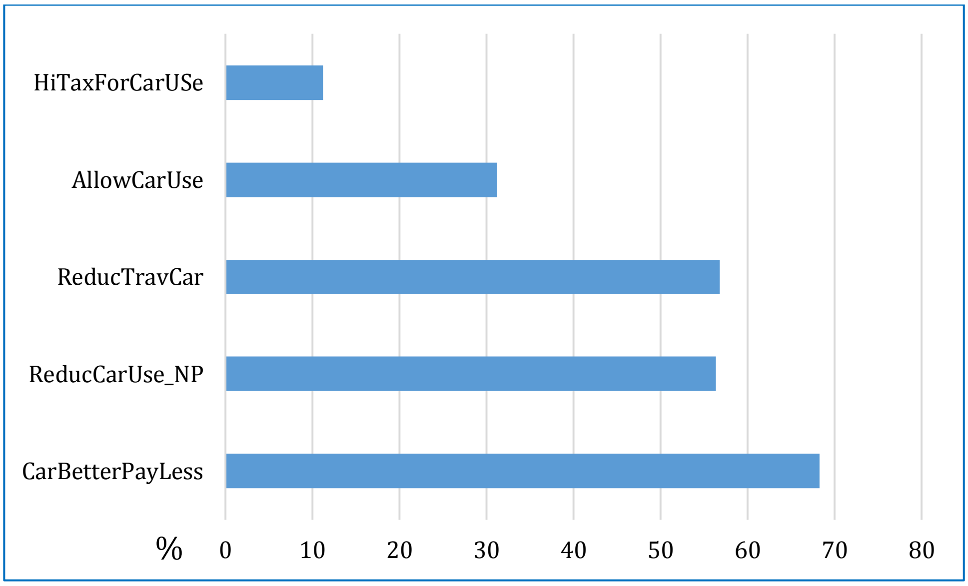


Figure 6.9: The proportion who strongly agreed or agreed among car-users

6.6 Factor 2: Traffic Awareness

Attitudes to congestion: Considering responses across all years 2011, 2012, 2013 and 2014, different respondents gave different opinions on congestion and the response varied across demographics. Whilst 45.2% of respondents mentioned that traffic congestion in towns and cities is a serious or very serious problem compared to the lower proportion (30.5%) on motorways, see Figure 6.10, it is worth noting that 11.5% and 25.4% of car users strongly disagreed there was no problem of congestion in urban areas and on motorways respectively. This suggests that some roads at certain times of the day do have spare capacity. As mentioned before, only car-users were considered in this study. Therefore, attitude differences between drivers and non-drivers were beyond the scope of the study, even though they may have very different views on congestion problems.

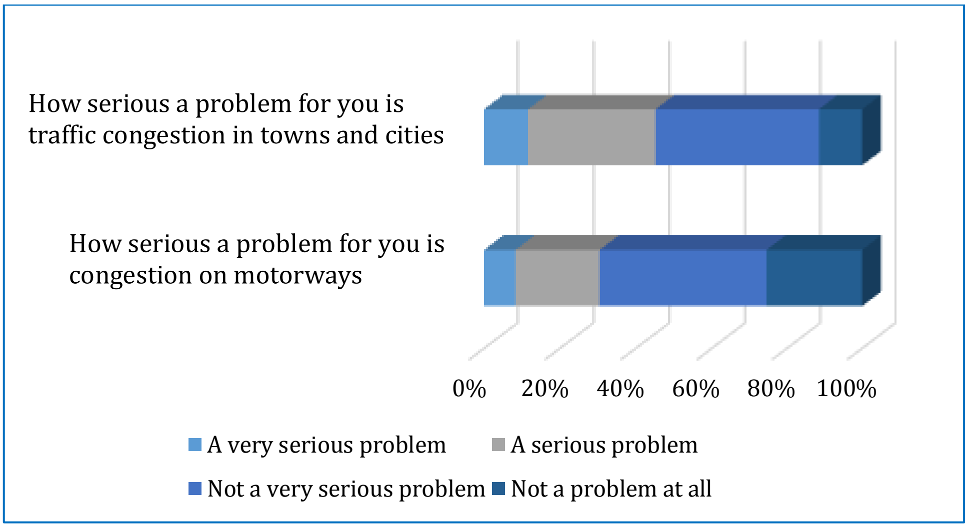


Figure 6.10: Proportions reporting the seriousness of congestion problem in motorways, towns, and cities

Males were more likely than females to consider motorway congestion to be a problem (17.69% compared to 12.79%) and more likely to consider congestion in towns and cities to be a problem (24.72% compared to 20.48%). Table 6.16 shows a clear trend of increasing opinion that the level of traffic congestion is more serious in towns and cities compared to motorways across all age groups from the younger to the older groups.

Variable		How serious a problem for you is congestion on motorways?				How serious a problem for you is traffic congestion in towns and cities?				Total	
		"A very serious problem" or "A serious problem"		Other responses		"A very serious problem" or "A serious problem"		Other responses			
		Count	%	Count	%	Count	%	Count	%	Count	%
Gender	Male	267	17.69	490	32.47	373	24.72	384	25.45	757	50.17
	Female	193	12.79	559	37.04	309	20.48	443	29.36	752	49.83
Total		460	30.48	1049	69.52	682	45.20	827	54.80	1509	100
Age	18-24	13	0.86	36	2.39	17	1.13	32	2.12	49	3.25
	25-34	62	4.11	153	10.14	91	6.03	124	8.22	215	14.25
	35-44	92	6.10	223	14.78	137	9.08	178	11.80	315	20.87
	45-54	98	6.49	219	14.51	143	9.48	174	11.53	317	21.01
	55-64	101	6.69	181	11.99	143	9.48	139	9.21	282	18.69
	65+	94	6.23	237	15.71	151	10.01	180	11.93	331	21.94
Total		460	30.48	1049	69.52	682	45.20	827	54.80	1509	100

Table 6.16: Proportions of opinions on congestions as a serious or a very serious problem by age and gender

Attitudes to exhaust fumes: Aggregating data over all years 2011 to 2014 inclusively, drivers were more likely to say that they considered exhaust fumes from traffic in town and cities to be very serious or a serious problem 12.66 % and 35.25% respectively, with 38.50% and 13.59% believing it was not a very serious problem and not a problem at all respectively. With reference to Table 6.17, 25.78% of female respondents were aware that exhaust fumes from traffic in towns and cities are a problem compared to 22.13% of male. As can be seen in Table 6.17, the opinions on exhaust fumes vary across all age groups from younger to older. These trends reveal the similarities in attitudes obtained from the preliminary analysis of attitudes to congestion where mature groups acknowledged the seriousness of exhaust fumes in urban areas compared to younger groups, who did so to a much less degree. This observation was expected, given that congestion is a major source of air pollution emissions.

Variable		How serious a problem for you are exhaust fumes from traffic in towns and cities?				Total	
		"A very serious problem" or "A serious problem"		"Not a very serious problem" or "Not a problem at all"			
		Count	%	Count	%	Count	%
Gender	Male	334	22.13	423	28.03	757	50.17
	Female	389	25.78	363	24.06	752	49.83
Total		723	47.91	786	52.09	1509	100
Age	18-24	19	1.26	30	1.99	49	3.25
	25-34	115	7.62	100	6.63	215	14.25
	35-44	163	10.80	152	10.07	315	20.87
	45-54	135	8.95	182	12.06	317	21.01
	55-64	148	9.81	134	8.88	282	18.69
	65+	143	9.48	188	12.46	331	21.94
Total		723	47.91	786	52.09	1509	100

Table 6.17: Proportions of opinion on exhaust fumes as a serious or a very serious problem by age and gender

6.7 Factor 3: Mode Shift Potential

Willingness to Switch to Sustainable Modes of Transport: Public transport, walking, and cycling were discussed in detail in Chapter 2 as more sustainable modes of transportation compared to the car. In particular, identifying modal shift potential for short journeys of less than 2 miles (about 3km) made by car is one of the most important points, because of the additional emissions caused by cold starts to be considered in this study. Table 6.18 clearly shows that about 50% of car users made four or more short trips in a week, contributing to cold starts.

Number of short journey	Count	%
0	224	14.84
1-3	546	36.18
4-6	351	23.26
7-9	78	5.17
10 or more	310	20.54
Total	1509	100

Table 6.18: Number of journeys less than 2 miles made by car in a typical week

For the 50% car users with less than four trips per week, there is potential for mode shift to public transport either by cycling or walking to a local bus stop or train station, or using park and ride. Given that 51% reported that they could just as easily walk (see Table 6.19), 45% just as easily cycle, and 32% just as easily use the bus, there is a clear indication that opportunity presents itself within existing services. However, the inconvenience or hostility towards modal shift was evident given that 56% of car users disagreed about using buses, as shown in Table 6.19. Also, there is considerable modal shift potential to enhance the minority of respondents (15%) who did not use cars for short journeys.

Response	Walking		Bus		Cycling	
	Count	%	Count	%	Count	%
Agree strongly	119	7.89	57	3.78	119	7.89
Agree	655	43.41	427	28.30	566	37.51
Neither agree nor disagree	205	13.59	185	12.26	169	11.20
Disagree	400	26.51	587	38.90	426	28.23
Disagree strongly	130	8.61	253	16.77	229	15.18
Total	1509	100	1509	100	1509	100

Table 6.19: Many of the journeys of less than 2 miles that I now make by car, I could just as easily walk, take the bus, or cycle.

Figure 6.11, see also data in Table 6.20, shows differences in the potential for modal shift from private transport to sustainable transport modes according to socio-demographic characteristics such as age, gender, number of people living in the household, and car ownerships per household. In terms of gender (Figure 6.11a), males are more willing to switch from cars to cycling for short journeys compared to females (26.51% compared to 18.89%). This may be due to greater safety concerns amongst females about cycling on the road (Schneider, 2013; Susilo and Cats, 2014).

Another reason may be that purchases of a second car in a household occur with life cycle changes (birth of children or taking up employment) and females are more likely to be involved with the school run and local activities, including shopping, making travel by car more practical (Clark *et al.*, 2016). However, evidence to back this up is outside the scope of the research presented in this thesis. Males and females were more or less equally likely to use buses for short journeys instead of walking and cycling and willingness to switch from the car to green modes (walking and cycling) for short journeys appears to increase depending on the number of people living in the household (Figure 6.13b).

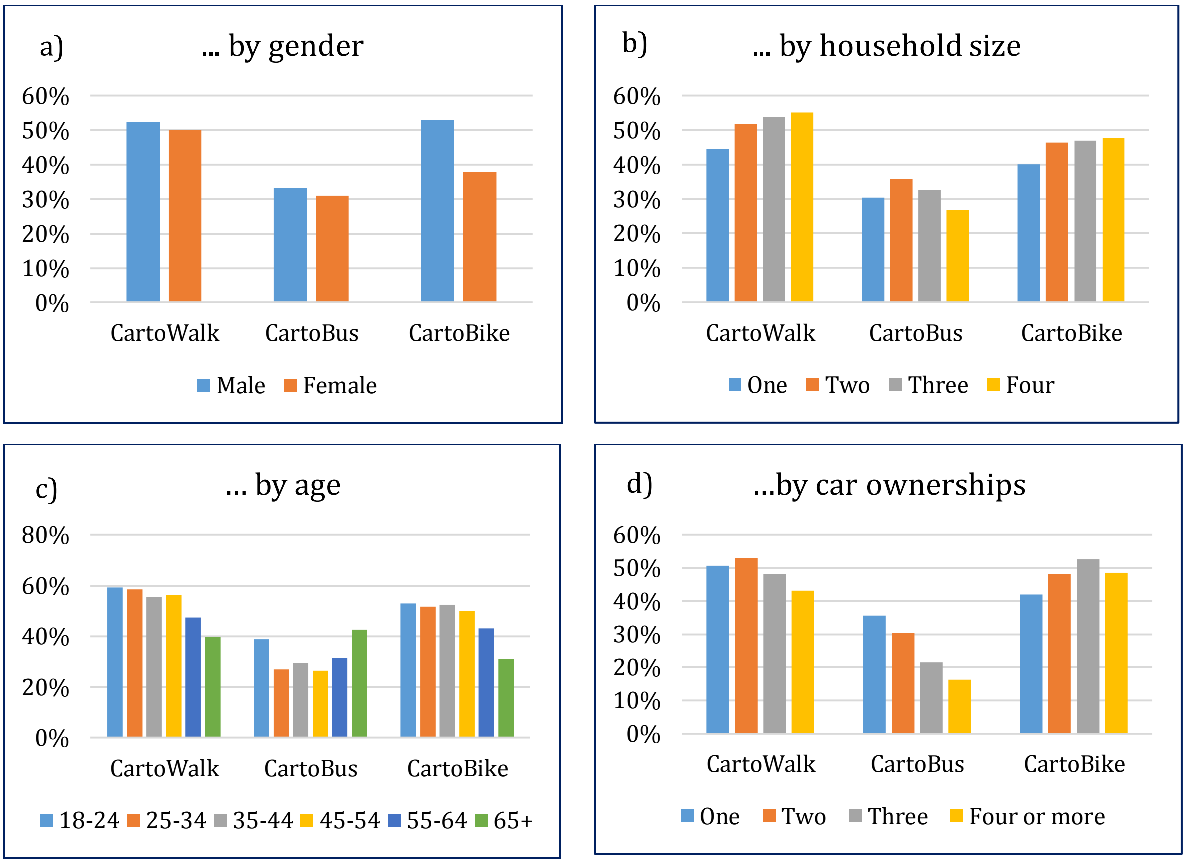


Figure 6.11: The proportion of respondents who agreed or strongly agreed that many of the journeys of less than 2 miles by car, could just as easily be made by walk, take the bus or cycle according a) age b) household size c) gender and d) car ownerships

Figure 6.11c suggests that willingness to switch from the car to non-carbon modes (walking and cycling) for short journeys seems to decrease with age. The older age group is statistically significantly more likely than all others to disagree that they could switch to non-carbon modes. This could be due to mobility difficulties, security and health issues. Interestingly, car users aged 45 and above showed an increasing trend in willingness to switch from cars to buses for short journeys of less than 2 miles and certainly the availability of free bus passes for senior citizens encourages bus travel for the older group (Andrews, 2012).

Respondents who owned one to three cars per household seemed to show an increased willingness to switch to cycling, although these differences are not statistically significant. However, the willingness to switch to local buses seems to decrease with car ownerships, which could be linked to both the affordability and the social acceptability of these modes (Davison and Knowles, 2006). Indeed, those who owned more cars and estimated to have higher incomes are significantly more likely to disagree that they could switch to the bus compared to the group of respondents who owned one car only. This is consistent with the increase in trend of households with young adults purchasing cars when they take up employment and still remain at home in order to save capital for house purchases (Clark *et al.*, 2016). However, evidence that this is the case is beyond the scope of this research.

In terms of car ownership, respondents in households that owned and used one (50%) or two (40%) cars (see Table 6.20) reported a similar potential to switch to walking for short journeys rather than cycling. The most unpopular choice for this group was shifting from driving cars to taking buses, particularly among those who owned four or more cars in a household. It can be seen in Table 6.20 that only 0.4% of this group would switch from car to bus. Finally, there is a general trend for the willingness to switch from car to walk and cycle with increase in household income, with the reverse being true for mode shift to bus.

Variable	Response	Walk				Bus				Cycle				Total	
		"Strongly agree" or "Agree"		Other responses		"Strongly agree" or "Agree"		Other responses		"Strongly agree" or "Agree"		Other responses			
		Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%
Gender	Male	397	26.31	360	23.86	251	16.63	506	33.53	400	26.51	357	23.66	757	50.17
	Female	377	24.98	375	24.85	233	15.44	519	34.39	285	18.89	467	30.95	752	49.83
Total		774	51.29	735	48.71	484	32.07	1025	67.93	685	45.39	824	54.61	1509	100
Age	18-24	29	1.92	20	1.33	19	1.26	30	1.99	26	1.72	23	1.52	49	3.25
	25-34	126	8.35	89	5.90	58	3.84	157	10.40	111	7.36	104	6.89	215	14.25
	35-44	175	11.60	140	9.28	93	6.16	222	14.71	165	10.93	150	9.94	315	20.87
	45-54	178	11.80	139	9.21	84	5.57	233	15.44	158	10.47	159	10.54	317	21.01
	55-64	134	8.88	148	9.81	89	5.90	193	12.79	122	8.08	160	10.60	282	18.69
	65+	132	8.75	199	13.19	141	9.34	190	12.59	103	6.83	228	15.11	331	21.94
Total		774	51.29	735	48.71	484	32.07	1025	67.93	685	45.39	824	54.61	1509	100
Household size	One	148	9.81	184	12.19	101	6.69	231	15.31	133	8.81	199	13.19	332	22.00
	Two	303	20.08	283	18.75	210	13.92	376	24.92	272	18.03	314	20.81	586	38.83
	Three	132	8.75	113	7.49	80	5.30	165	10.93	115	7.62	130	8.61	245	16.24
	Four or more	191	12.66	155	10.27	93	6.16	253	16.77	165	10.93	181	11.99	346	22.93
Total		774	51.29	735	48.71	484	32.07	1025	67.93	685	45.39	824	54.61	1509	100
Income per month	Less than £1,200	113	7.49	154	10.21	103	6.83	164	10.87	90	5.96	177	11.73	267	17.69
	£1,201 - 2,200	179	11.86	172	11.40	135	8.95	216	14.31	135	8.95	216	14.31	351	23.26
	£2,201 - 3,700	246	16.30	197	13.06	142	9.41	301	19.95	218	14.45	225	14.91	443	29.36
	£3,701 or more	236	15.64	212	14.05	104	6.89	344	22.80	242	16.04	206	13.65	448	29.69
Total		774	51.29	735	48.71	484	32.07	1025	67.93	685	45.39	824	54.61	1509	100
Car ownership	One	384	25.45	372	24.65	270	17.89	486	32.21	317	21.01	439	29.09	756	50.10
	Two	320	21.21	284	18.82	184	12.19	420	27.83	291	19.28	313	20.74	604	40.03
	Three	54	3.58	58	3.84	24	1.59	88	5.83	59	3.91	53	3.51	112	7.42
	Four or more	16	1.06	21	1.39	6	0.40	31	2.05	18	1.19	19	1.26	37	2.45
Total		774	51.29	735	48.71	484	32.07	1025	67.93	685	45.39	824	54.61	1509	100

Table 6.20: The proportion of respondents who agreed or strongly agree that many of the journeys of less than 2 miles by car, could just as easily be made by walk, take the bus or cycle according to socio-demographic characteristics.

6.8 Conclusions

In this chapter an independent analysis of 14 attitudinal variables using EFA (and more specifically PAF) and PCA was carried out. A comparison of the results showed that PAF produces a better solution provided that correlation and covariance between factors is considered to represent how strongly two factors are related.

The PAF established multicollinearity among the attitudinal variables and reduced the number that were statistically significant from 14 attitudinal variables grouping into three factors: namely, attitudes to transport and the environment, traffic awareness, and modal shift potential.

These three factors emerged from the PAF on the premise that a single variable factor is not acceptable (Raubenheimer, 2004; Osborne and Castello, 2005). The characteristics of the respondents who strongly agreed or agreed to the sentiments expressed in the questions grouped in each of the factors obtained from PAF were studied in more detail. The analysis revealed a statistically significant difference at a 95% level of confidence, found based on the χ^2 test of the attitudes to the environment, attitudes to congestion, and attitudes to exhaust fumes from year to year analysis.

It was clear from the in-depth analysis of the characteristics of the factors that there were interrelationships between travel choices, awareness of the environment and demographics; therefore, in the next chapter, relationships between these perceptions and attitudes to climate change and environmental issues are explored in more detail. The derived factors from PAF are further investigated using multinomial logistic regression and the results are discussed in Chapter 7.

Chapter 7 Analysing Car Users' Perceptions

7.1 Introduction

Descriptive analysis and an exploratory factor analysis of the socio-demographic and attitudinal variables were presented respectively in Chapters 5 and 6 and used to explore the characteristics of respondents current travel behaviour and attitudes towards changing modes of transport, the environment, congestion and exhaust fumes. The results suggested causal links and interdependency between variables, but did not identify specific relationships within groups of the total population.

The next step was to carry out more detailed investigation to identify patterns in the perceptions towards climate change and environmental issues within the population. The analytical approach adopted used multiple correspondence analysis (MCA), hierarchical cluster analysis (HCA) and multinomial logistic regression (MLR). The following sections begin by analysing categorical variables, followed by presenting the results of each of the analyses and, in the final section, the conclusions drawn from the analysis are summarised.

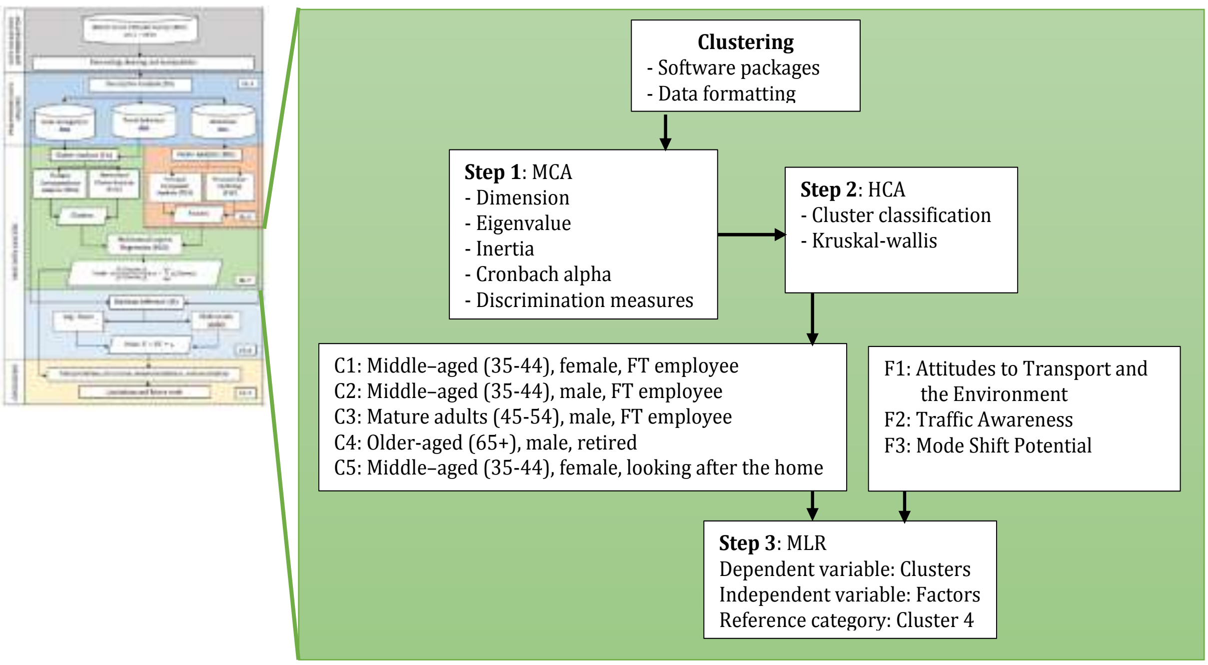


Figure 7.1: Steps involved in Chapter 7

7.2 Clustering Categorical Variables

In travel behaviour research, when conducting surveys, categorical variables are most commonly used. When dealing with categorical variables, the first task is to obtain an overall view of the dataset, including how individuals are distributed among the categories. This is because the aim of the questionnaire is to obtain a typology of surveyed people based on the answers they have provided. This section considers the use of MCA to present the scatter plot of the car users in factor maps and the use of hierarchical clustering to segregate the car users into the same cohorts with similar characteristics.

Specifically, this section focuses on car users with attention to socio-demographics and travel behaviour variables and categories:

1. Study of individual car users: two car users are close to each other if they answered the questions the same way. This study is interested in one car user and also in populations: Are there groups of car users who are similar?

2. Study of variables and categories: Firstly, the relationships between variables and the associations between categories are observed. Two categories are close to each other if they often occur together. Secondly, the study seeks to characterise groups of car users by categories.

<u>Software packages</u>

In order to perform the analysis of categorical variables, the Rstudio software, specifically the *EnQuireR* package (Gwenaelle *et al.*, 2010) and *FactoMineR* package (Sebastien *et al.*, 2008), was used along with the IBM SPSS Statistics 23 software which provides many tools to automate the process of analysing survey data. It includes both univariate (each categorical variable separately) and multivariate (more than one categorical variable) analyses comprising MCA and hierarchical cluster analysis.

<u>Data formatting</u>

The data obtained from the BSA survey is in *.sav* format, which is a file extension type used to store data for SPSS analysis. Hence, the data was imported as a Microsoft Excel spreadsheet, given that this study used both SPSS and Rstudio to visualise the results of MCA. Therefore, the survey data needed to be in comma separated value (*.csv*) format due to the original datasets being represented in numeric codes.

Two subgroups of data were used for the formation of the clusters of car users, one representing variables related to socio-demographic characteristics and the other one representing the car users' travel behaviour. Details of each of the variables, as shown respectively in Tables 5.5 and 5.7 in Chapter 5, consist of car users' age, gender, number of people in the household, car ownership, employment status and daily travel behaviour.

7.2.1 Multiple correspondence analysis (MCA)

The routine used to perform MCA in the *EnQuireR* package (Gwenaelle *et al.*, 2010) is the one implemented in the *FactoMineR* package (Sebastien *et al.*, 2008), which provides a representation of individuals and their answers to questions. This routine has been enhanced in a questionnaire context where missing values are frequently encountered and where large numbers of people may be surveyed. The characteristics of travel behaviour and socio-demographic variables selected from the BSA to perform this analysis is shown in Table 7.1. However, after an investigation using discrimination measures, three variables (Bus_usage, Train_usage, and Bike_usage) were removed from the MCA analysis because these variables were contributing substantially less and redundancy issues were encountered when one car user could also be the same person who used bus, train and bicycle. After removing them, the percentage of variance increased in both dimensions 1 and 2.

Travel behaviour	Socio-demographics
Frequency of travel using car as a driver (Car_driver)	Age
Frequency of travel using car as a passenger (Car_passenger)	Gender
Frequency of travel using local buses (Bus_usage)	Household size
Frequency of travel using trains (Train_usage)	Car ownerships
Frequency of travel using bicycles (Bike_usage)	Employment status

Table 7.1: Variables selected for MCA analysis

Dimensions

A two-dimensional MCA solution was considered in the analysis. The first and second dimensions explained 7.46% and 5.05% variance of the principal inertia respectively. Hence, a two-dimensional plane representing 12.51% of the whole dataset was used to interpret the results from MCA, as shown in Table 7.2. A similar small proportion of total dimensional plane was found in Das and Sun (2015) where they reported 5.42% and 4.68% variance explained in dimensions 1 and 2 respectively. The percentages of inertia in MCA are low and tend to be close to one another and this latter fact might lead to an assumption that individual axes might be unstable (Ayele *et al.*, 2014).

Dimension	Inertia	% of Variance	Cumulative % of Variance
1	0.34	7.46	7.46
2	0.23	5.05	12.51
3	0.21	4.54	17.05
4	0.20	4.11	21.16
5	0.19	4.06	25.22
6	0.17	3.77	28.99
7	0.16	3.57	32.66
8	0.16	3.53	36.09
9	0.16	3.44	39.53
10	0.15	3.37	42.90

Table 7.2: Inertia values for the first 10 dimensions

The inertias and percentages of variance of the first 10 dimensions obtained from the MCA are given in Table 7.2. Based on the results, the first ten dimensions may account for similar amounts of variance and it is expected that 57.1% percent of the inertia will be accounted for by the remaining dimensions. Furthermore, 95% of the association was found to be well represented in 28 dimensions.

Inertia

The percentages of inertia associated with dimensions are often quite low in MCA. Since inertia can be interpreted as the information associated with a dimension, it is important to check whether or not the percentages found really do reveal a meaningful structure of the dataset. In this study, the p-values associated with the test of the significance of the dimensions using "*p_inertia*" function are zero (see Table 7.3). This confirms that even though the percentages of inertia of the dataset are quite low, they are statistically significantly different from what would be obtained as the result of chance.

	Inertia	% of variance	p-value
Dimension 1	0.34	7.46	0.00
Dimension 2	0.23	5.05	0.00
Total	0.57	12.51	0.00

Table 7.3: P-values associated with the test of the significance of the dimensions

Eigenvalue

An eigenvalue is an indication of the magnitude of information explained by each dimension in representing the whole dataset. In this study, the first and second dimensions respectively have the following features: eigenvalues of 2.39 and 1.62, as shown in Table 7.4.

Dimension	Cronbach's Alpha	Variance accounted for		
		Total (Eigenvalue)	Inertia	% of Variance
1	0.68	2.39	0.34	34.09
2	0.45	1.62	0.23	23.08
Total		4.01	0.57	
Mean	0.58[a]	2.01	0.29	28.59

a. Mean Cronbach's Alpha is based on the mean Eigenvalue.

Table 7.4: Model summary

Cronbach's alpha

The MCA output from SPSS presents Cronbach's alpha of 0.68 for dimension 1 and 0.45 for dimension 2, which is equivalent to a mean Cronbach's alpha coefficient of 0.58, as shown in Table 7.4. A satisfactory level for Cronbach's alpha lies between 0.60 – 0.70, and a value smaller than 0.60 is acceptable in exploratory research, as explained by Johnson and Wichern (2007). The possible reasons behind the small value here include a limited number of questions, weak relationships between them or heterogeneous constructs (Loewenthal, 2001; Johnson and Wichern, 2007). In addition, Cronbach alpha is influenced by the design of the questionnaire but as this study used a third party data source, the Cronbach alpha values cannot be controlled.

Discrimination measures

The use of MCA in SPSS produced a table of discrimination measures in conjunction with the joint plot of the variable points. The discrimination measure consists of the square loadings of the variables which are equivalent to the percentages of variance explained by each variable in the dimension (Franco, 2015). The purpose of the discrimination measures shown in Table 7.5 is to indicate, by a comparison of the magnitude of the values, which variables can be discriminated between the first and second dimensions. By way of illustration, the embolden values indicate those variables showing higher discrimination in a given dimension and these indicate the variables with higher correlations with that dimension.

Variables	Dimension	
	1	2
Number living in household, including respondent	**0.50**	0.09
Gender	0.04	**0.50**
Age	**0.72**	0.11
Car ownerships	**0.23**	0.13
Frequency of travel by car as a driver	0.10	**0.19**
Frequency of travel by car as a passenger	0.12	**0.17**
Current economic position of respondent	**0.69**	0.42
Active Total	2.39	1.62
% of Variance	34.09	23.09

Table 7.5: Discrimination measures

The variable categories with larger values contribute the most to the definitions of dimensions. The bold figures indicate the relevance of that variable when explaining dimensions 1 and 2. For example, age (0.72), employment status (0.69), number living in household (0.50) and car ownership (0.23) are the most highly correlated with dimension 1. Likewise, the variables of gender (0.50), frequency of using car as a driver (0.19) and frequency of using car as a passenger (0.17) are the most highly correlated with dimension 2. Accordingly, dimension 1 represents socio-demographic variables and dimension 2 represents level of car use with gender.

MCA minimizes the sum of squared distances between category points and respondents. For each variable, a discrimination measure, which can be regarded as a squared component loading, is computed for each dimension. This measure is also the variance of the quantified variable in that dimension. A maximum value of 1 is achieved if the object scores fall into mutually exclusive groups and all object scores within a category are identical.

The joint plots shown in Figure 7.2 (a) and (b) present alternative ways to visualise the discrimination of the variables on each dimension and to identify the correlation between the categories (points) of the variables. The internal consistency of the categories is identified by the coordination of the variable points, which are plotted on the joint plot in such a way that similar variables associated with each dimension are displayed close together.

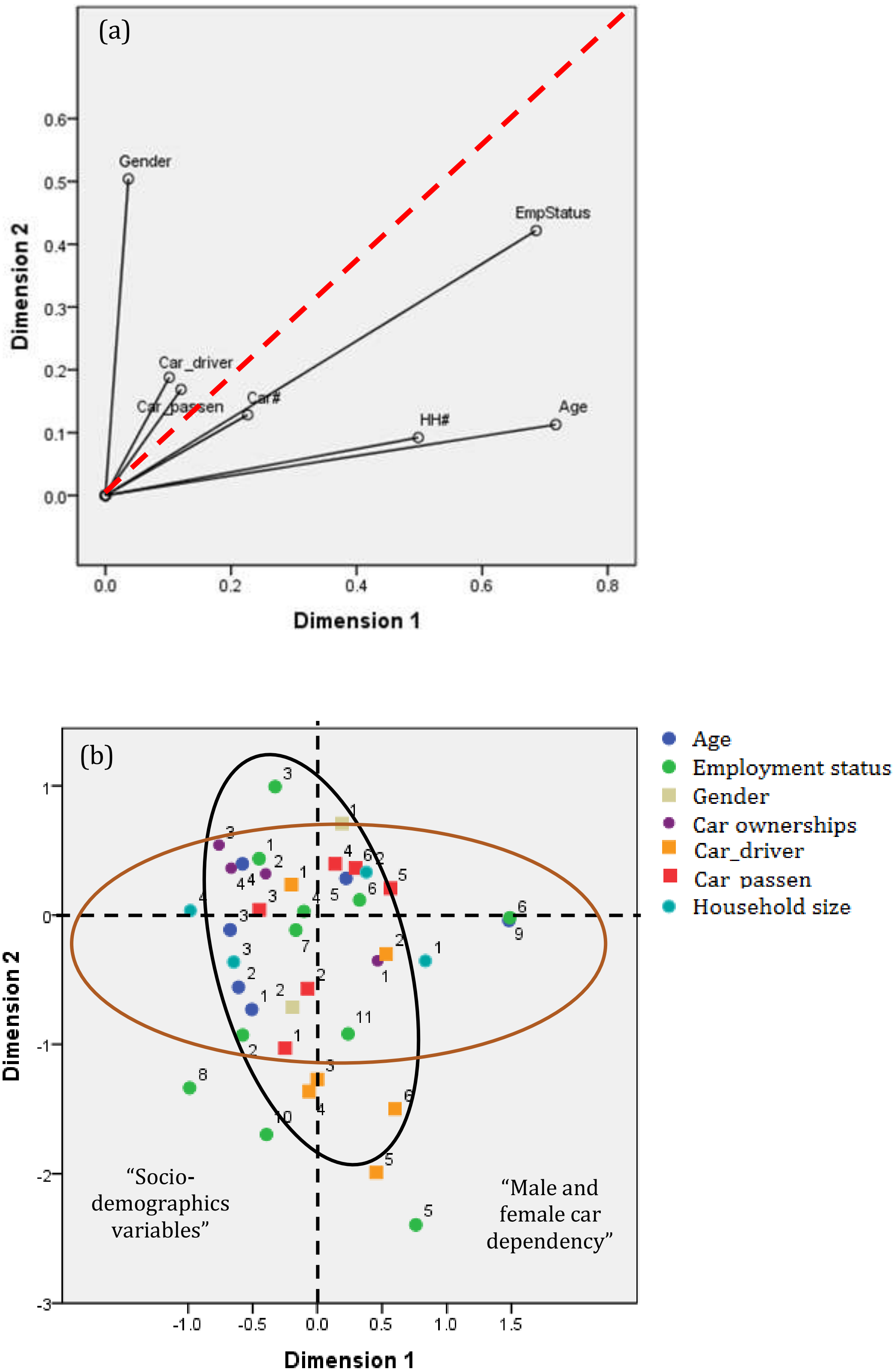


Figure 7.2: Results obtained from MCA: (a) MCA dimensions discrimination measures; (b) joint plot of the categories for each variable illustrating the two dimensions of 'Socio-demographic variables' (Dimension 1 black oval) and 'Male and female car dependency' (Dimension 2 brown oval).

In addition, the length and steepness of lines in the joint plot signify the degree of discrimination measures of each variable for the two dimensions considered (Costa *et al.*, 2013), where variable lines lying above an angle of 45-degrees are more correlated with dimension 2 and vice versa. In terms of the length, a longer line signifies a greater distance between the categories of variables, which is equivalent to a higher percentage of variance explained by a particular dimension (Das and Sun, 2015).

From the results and their graphical visualisation in Figure 7.2 (b), the first dimension was labelled as "Socio-demographics" and the second dimension as "Male and female car dependency". Figure 2 (b) illustrates the categories as different coloured symbols depending on their variable type and it can be seen that the location of the two ovals approximately map on to the two dimensions. The proportion of variation explained by dimensions 1 and 2 respectively were 34% and 23%, yielding a total variance of 57%, as discussed above and presented in Table 7.5. Clearer relationships among positive and negative centroid coordinates for both dimensions are explained in Figure 7.3 and show the characteristics of the two cohorts that emerge from this analysis.

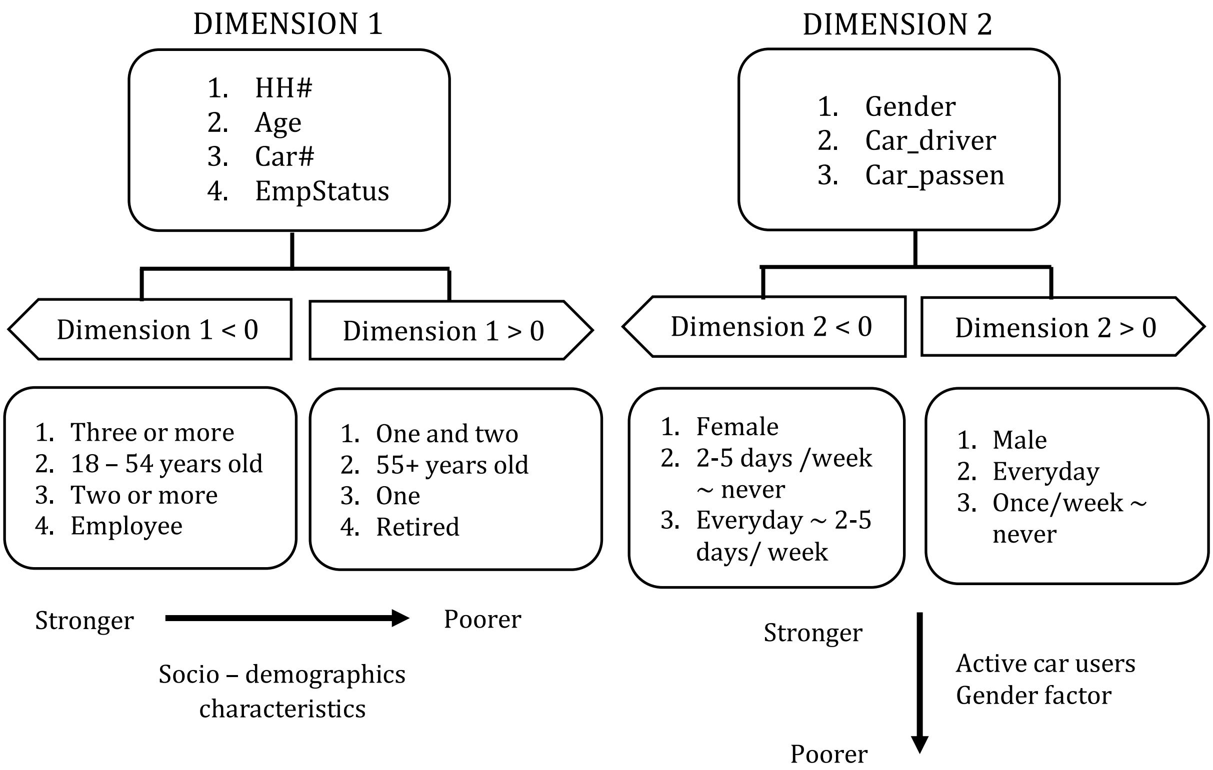


Figure 7.3: Results obtained from MCA: positive and negative centroid coordinates for dimensions 1 and 2

The coordinate for a categorical variable in a dimension is a squared correlation ratio between the dimension and the categorical variable. The R^2 and p-values of each variable along with statistical significance in dimensions 1 and 2 are detailed in Table 7.6. The first and second dimensions are significantly linked to socio-demographic variables. The strongest link in dimension 1 is with age, employment status, household size, and so on, in order of reducing influence. Whilst dimension 2 has the strongest link with gender. The variables are ranked from the most to least related and only the ones with a correlation ratio significantly different to zero are presented.

Variable	R^2	p-value	Variable	R^2	p-value
MCA Dimension 1			*MCA Dimension 2*		
Age	0.72	0.00	Gender	0.50	0.00
EmpStatus	0.69	0.00	EmpStatus	0.42	0.00
HH#	0.50	0.00	Car_driver	0.19	0.00
Car#	0.23	0.00	Car_passen	0.17	0.00
Car_passen	0.12	0.00	Car#	0.12	0.00
Car_driver	0.10	0.00	Age	0.11	0.00
Gender	0.04	0.00	HH#	0.09	0.00

Table 7.6: Significance of key variables

Table 7.7 lists the categories, showing each category and its coordinate values on the dimensions, along with the statistical significance test value. The variables' categories used to describe the axes (x = Dim1, y = Dim2) are also provided. Figure 7.4 plots the output from the MCA for each individual irrespective of category. The points are scattered about (0, 0) according to whether the individual is associated more negatively or positively with one dimension or the other. This presentation of data is useful if the number of individuals or categories are small; however, in this analysis, the number (n=1509) considered is large, therefore some work is needed to aid interpretation, for example by plotting the individuals, variables and categories separately.

Variable	Estimate	p-value	Variable	Estimate	p-value
MCA Dimension 1			*MCA Dimension 2*		
HH_four or more	0.51	0.00	Female	0.34	0.00
Employee (FT)	0.25	0.00	Employee (PT)	0.19	0.00
35-44	0.33	0.00	Car_one	0.27	0.00
Car_two	0.03	0.00	Car_passen_2-5 days a week	0.23	0.00
Car_driver_Every day or nearly every day	0.25	0.00	Looking after home	0.39	0.00
45-54	0.27	0.00	Car_driver_Once a week	0.11	0.00
HH_three	0.32	0.00	25-34	0.20	0.00
Car_passen_Once a week	0.28	0.00	Car_passen_Every day or nearly every day	0.45	0.00
25-34	0.29	0.00	Employee (FT)	0.58	0.00
Employee (PT)	0.33	0.00	Car_driver_Less often than that	0.47	0.00
Looking after home	0.57	0.00	Other	0.18	0.00
Car_three	0.25	0.00	HH_one	0.12	0.00
Female	0.11	0.00	HH_three	0.14	0.00
Car_four or more	0.19	0.00	18-24	0.31	0.00
18-24	0.23	0.00	Car_driver_Less often but at least once a month	0.17	0.00
Self-employee (FT)	0.18	0.00	Car_driver_Never nowadays	0.23	0.00
Car_passen_Every day or nearly every day	0.17	0.04	In work (status not known)	0.88	0.02
Car_passen_Less often but at least once a month	-0.06	0.01	35-44	-0.01	0.03
55-64	-0.19	0.00	Car_four or more	-0.07	0.03
Car_passen_Never nowadays	-0.15	0.00	Car_passen_Less often than that	-0.15	0.00
Male	-0.11	0.00	55-64	-0.20	0.00
Car_passen_Less often than that	-0.31	0.00	Car_three	-0.15	0.00
HH_two	-0.28	0.00	Car_passen_Never nowadays	-0.23	0.00
Car_driver_2-5 days a week	-0.18	0.00	Car_driver_2-5 days a week	-0.36	0.00
HH_one	-0.55	0.00	Car_passen_Less often but at least once a month	-0.24	0.00
Car_one	-0.47	0.00	45-54	-0.26	0.00
65+	-0.93	0.00	Car_two	-0.05	0.00
Retired	-0.88	0.00	HH_two	-0.19	0.00
			Self-employee (FT)	-0.73	0.00
			Car_driver_Every day or nearly every day	-0.62	0.00
			Employee (FT)	-0.47	0.00
			Male	0.34	0.00

Table 7.7: Categories found to be statistically significant in their contribution to Dimensions 1 and 2

One of the challenges in MCA is graphical display of the results to aid interpretation. The "*FactoMineR*" package provides a graphical representation of both individuals and the categories constructed by the MCA function (refer Figure 7.4). With a large number of respondents answering numerous questions, interpretation and decision making is difficult.

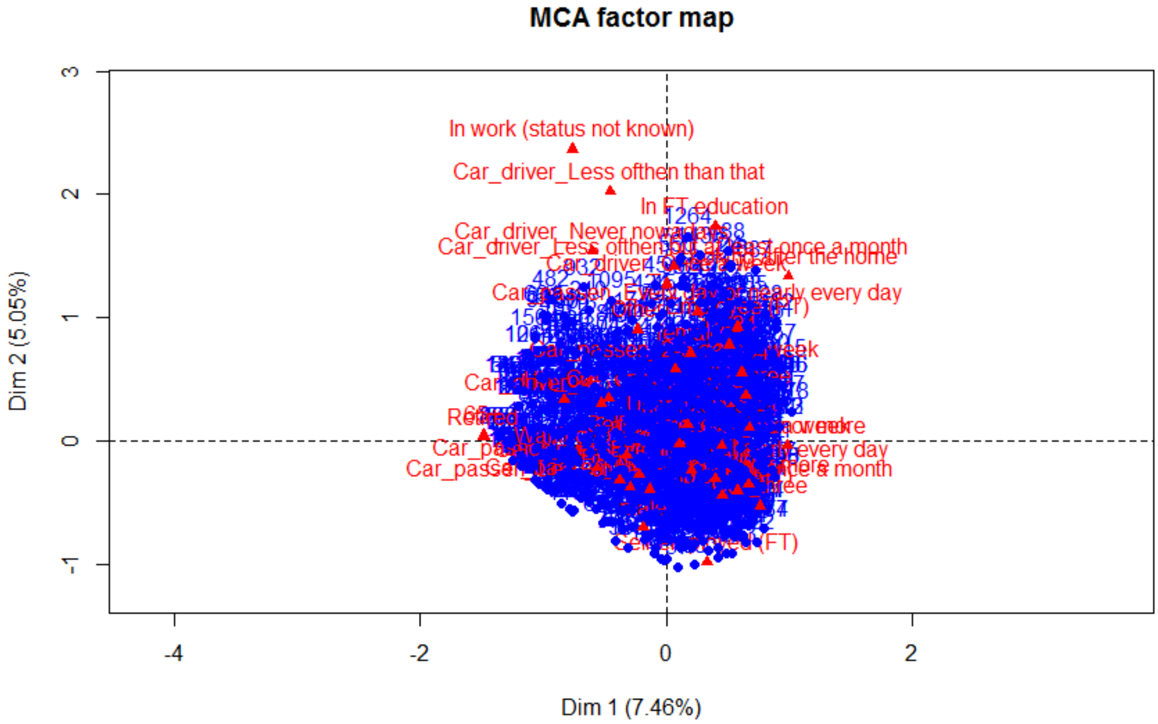
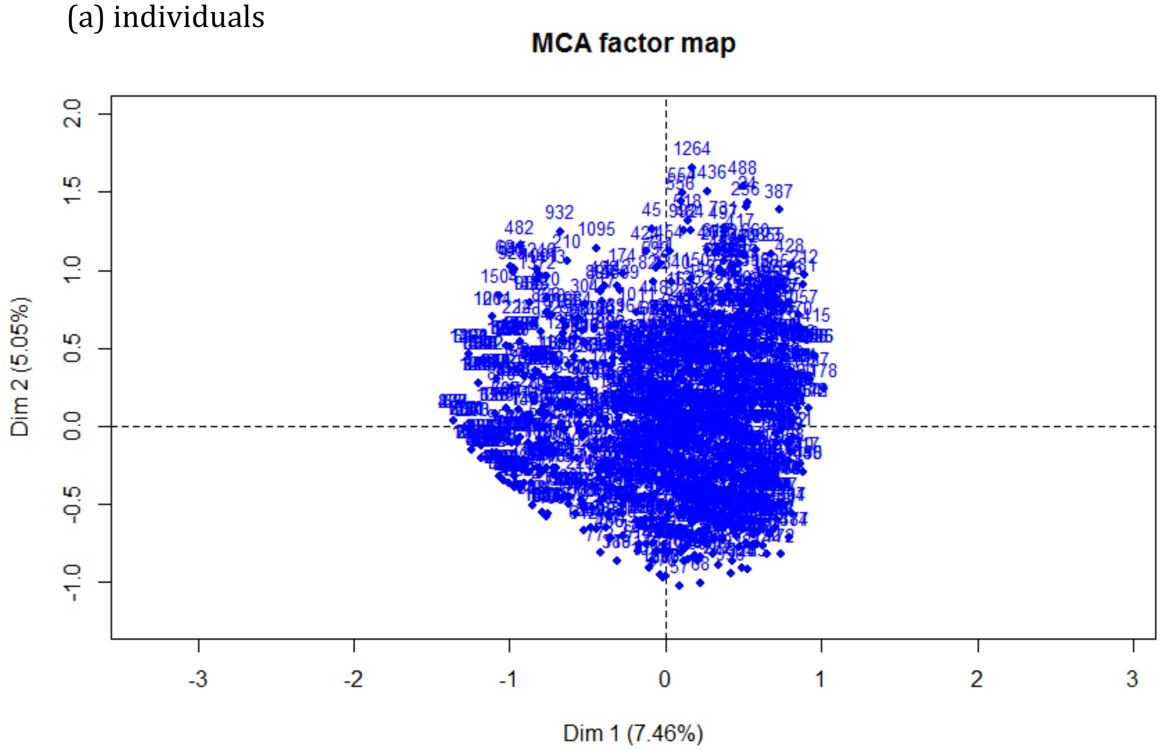


Figure 7.4: Graphical representation of the individuals and the categories

Interpretation can be made easier by selecting individuals with the highest coordinates and for those categories that are significantly linked to the components (dimensions) provided by MCA using "*ENlisib*" function which plots separately (a) individuals and (b) variables as shown in Figure 7.5 and categories, see Figure 7.6. The plot in Figure 7.5 (a) of individuals provides little clarity due to the overlay of 1509 points, whilst the categories in Figure 7.5 (b) is rather oversimplified, but clearly reveals a similar distribution of points as in Figure 7.2, illustrating the consistency between the SPSS and R analysis. However, it could be argued that the discrimination analysis of SPSS allows for a clearer indication of these variables influencing dimensions 1 and 2 by considering their position relative to the 45^0 line.

(a) individuals



(b) variables

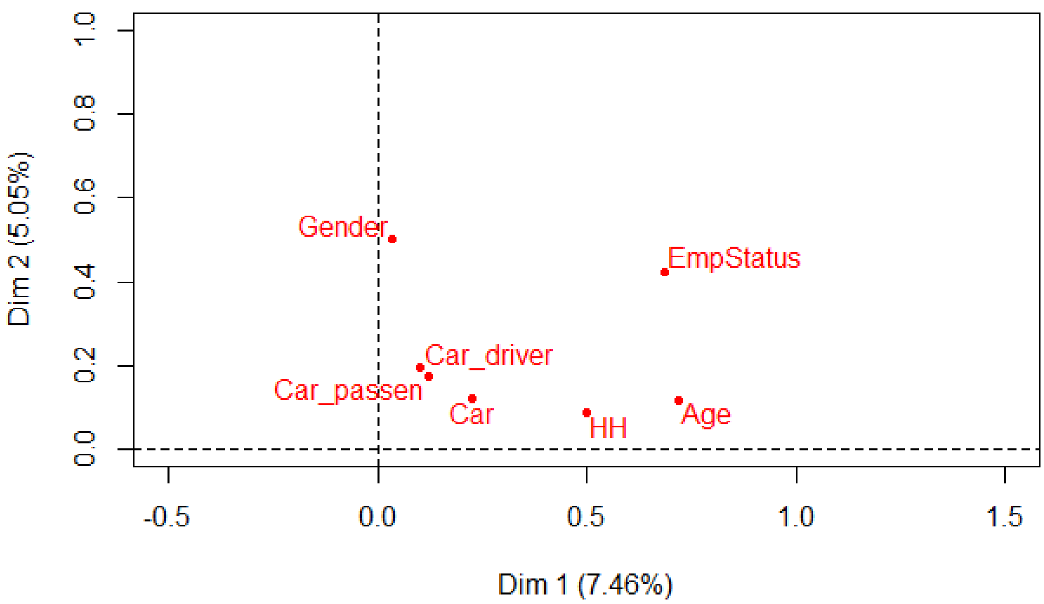


Figure 7.5: MCA “ENlisib” plot for all the study assigned to: (a) individuals and (b) variables

Figure 7.6 on the other hand is a compromise and shows the distribution of all categories in four different quadrants. Quadrant 1 consists of respondents between 18 to 44 years old who were persons in full-time education, mainly at home, part-time employee and unemployed. They were using a car as a passenger every day or nearly every day with 3 or more people living in the household. In quadrant 2, respondents were mainly female, unemployed, single car ownership and living alone. Quadrant 3 consisted mainly of males who rarely used the car as a passenger, were older aged and in retirement. Finally, quadrant 4 located respondents who have three or more cars and active car drivers. They were persons who were in full-time employment or self-employed and mature adults (45–54 years old). Overall, this analysis suggests that the more sustainable groups emerge from quadrant 1 and 3 compared to quadrant 2 and 4.

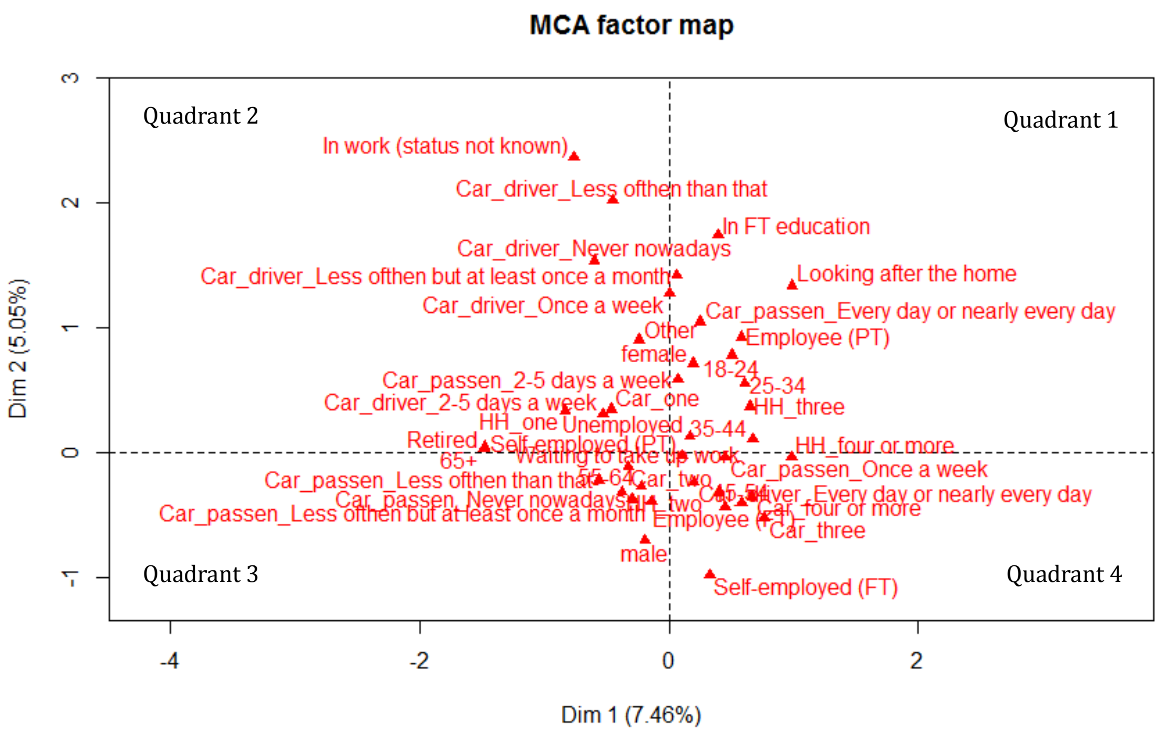


Figure 7.6: MCA "*ENlisib*" plot for all study categories

The quality of representation of the individual can be improved by the $SelectMod =$ "$cos^2 20$" function which is used to draw labels for the 20 categories that are those best represented in the 2 dimensions aiding interpretation. The function $Select =$ "$contrib 10$" is then added to highlight 10 respondents which contribute the most to the dimensions, as can be seen in Figure 7.7. From this figure, the characteristics of those 10 respondents can be traced and identified.

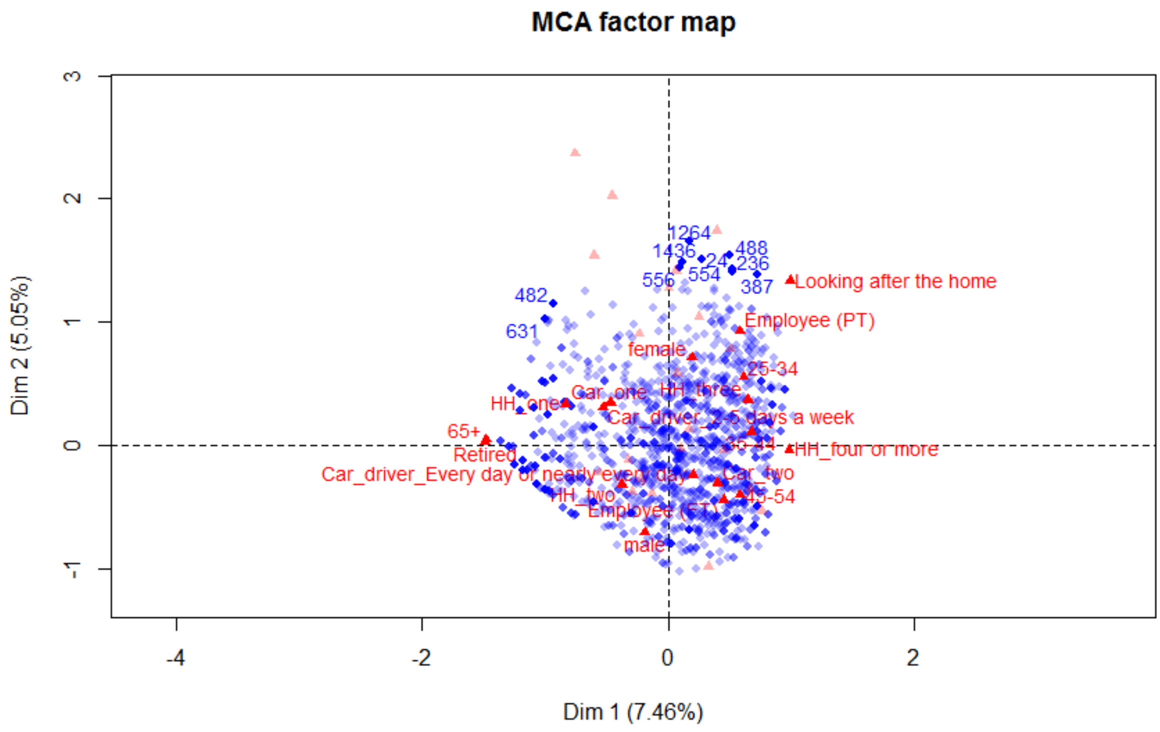


Figure 7.7: The highest contribution of individual and categories

Whilst there are clusters of points in the two dimensions' plot, the overlaying of individual datasets masks potential structure in the data. Another way to avoid the problem of the superposition of the individuals due to their number, is to use density plots as shown in Figure 7.8. The "*ENdensity*" function provides a visualisation of the shape of the scatter plot of the distribution of the coordinates of the individuals within each category.

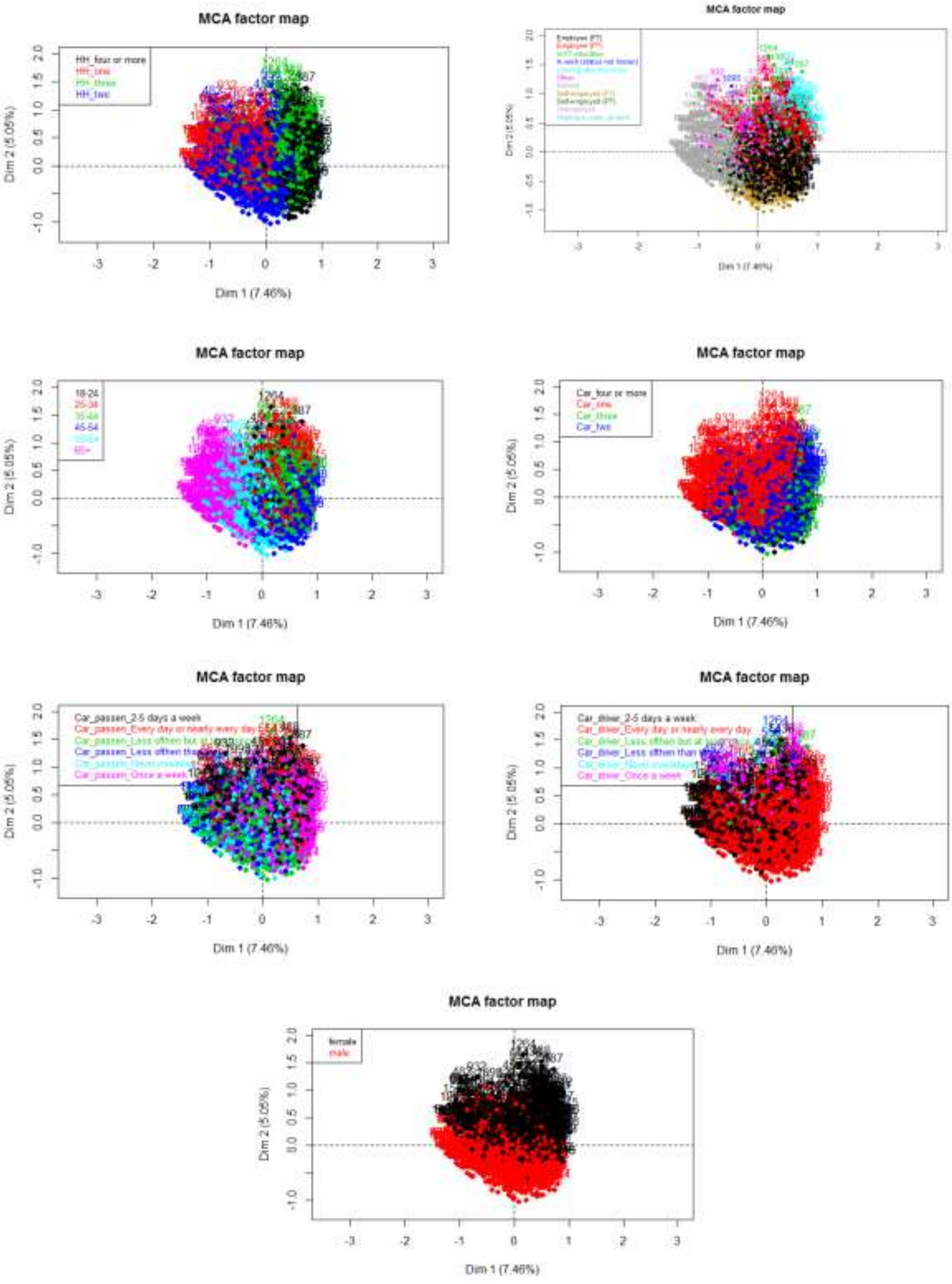
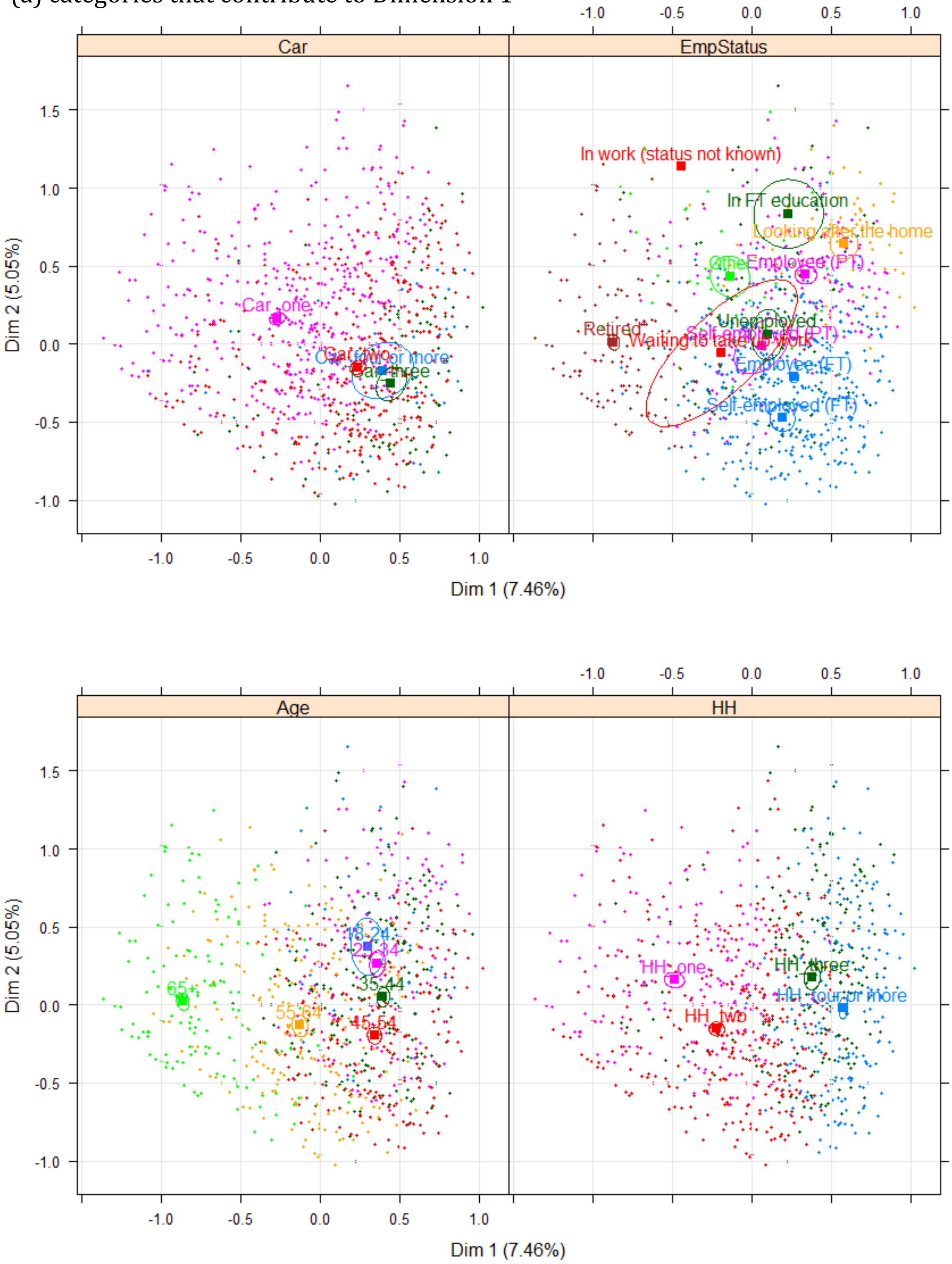


Figure 7.8: Scatter plot for categories of each variable

Output plots with confidence ellipses around the categories of some variables can be plotted. The function used here is called "*plotellipses*". This function gives one plot per variable along with a confidence ellipse around each category of the variable. Figure 7.9 (a) and (b) present the confidence ellipse around the categories that contribute to dimensions 1 and 2 respectively.

(a) categories that contribute to Dimension 1



(b) categories that contribute to Dimension 2

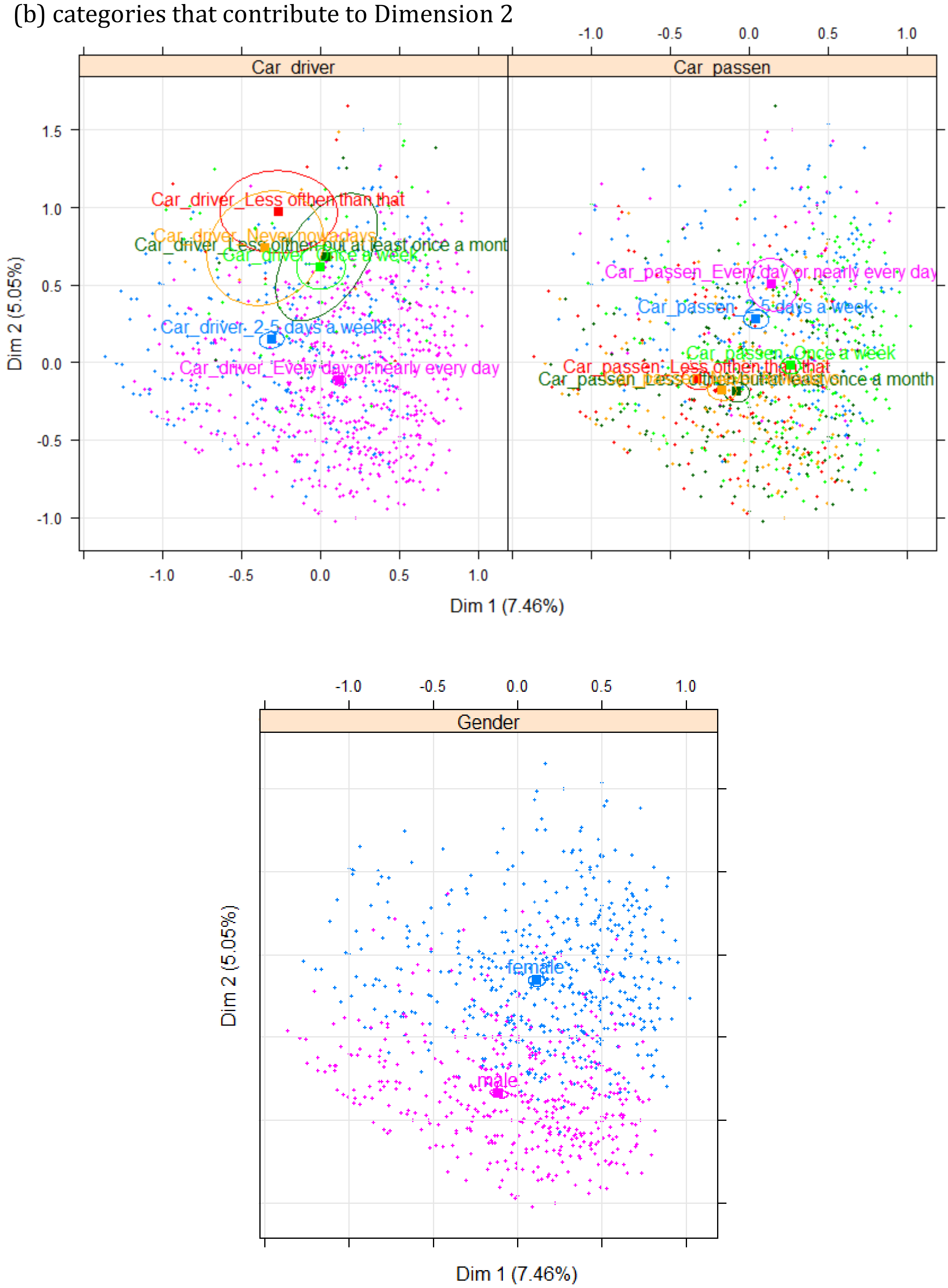


Figure 7.9: Confidence ellipses around the categories: (a) categories that contribute to Dimension 1; (b) categories that contribute to Dimension 2

The confidence ellipses are quite small. In effect, given 1509 respondents, the sub-populations are quite distinct with little overlap, showing that the sub-populations are statistically significantly separated. All the graphs above were plotted using the MCA analysis which was proven to enhance the overall view of the dataset, including how individuals are distributed among the categories, based on the responses given to each question. In the following step, these variables were grouped based on the same characteristics.

7.2.2 Hierarchical clustering analysis (HCA)

The next step of the analysis was to visualise the coordinates of the scatter plots using the *ENMCA* function. This function allows a cluster analysis to be performed following an MCA in R. Many algorithms are available to perform cluster analysis on numeric variables, but it is difficult to find algorithms which perform cluster analysis directly on categorical variables. The basic principle of the *ENMCA* function is that it performs MCA on categorical variables, followed by the classification of the corresponding variables to the coordinates of the individuals.

The *ENMCA* function provides outputs directly related to the clusters, allowing improved understanding of each group of people surveyed. MCA was then used as a pre-processing stage for clustering, where categorical variables were transformed into a set of continuous variables. This allows the coordinates for each respondent, as shown in the scatter plot (Figure 7.8), to be assigned to relevant clusters based on inertia (Ward's criteria). This has been identified as common practice, especially for the analysis of questionnaires with categorical data (Husson et al., 2010).

Ward's criterion to aggregate clusters was chosen to be consistent with MCA itself, whose principle is to maximise the inertia of the cloud of the individuals. Indeed, Ward's criterion aggregates clusters by minimising the inertia within the cluster thus obtained, and so this fits perfectly with the objective of MCA.

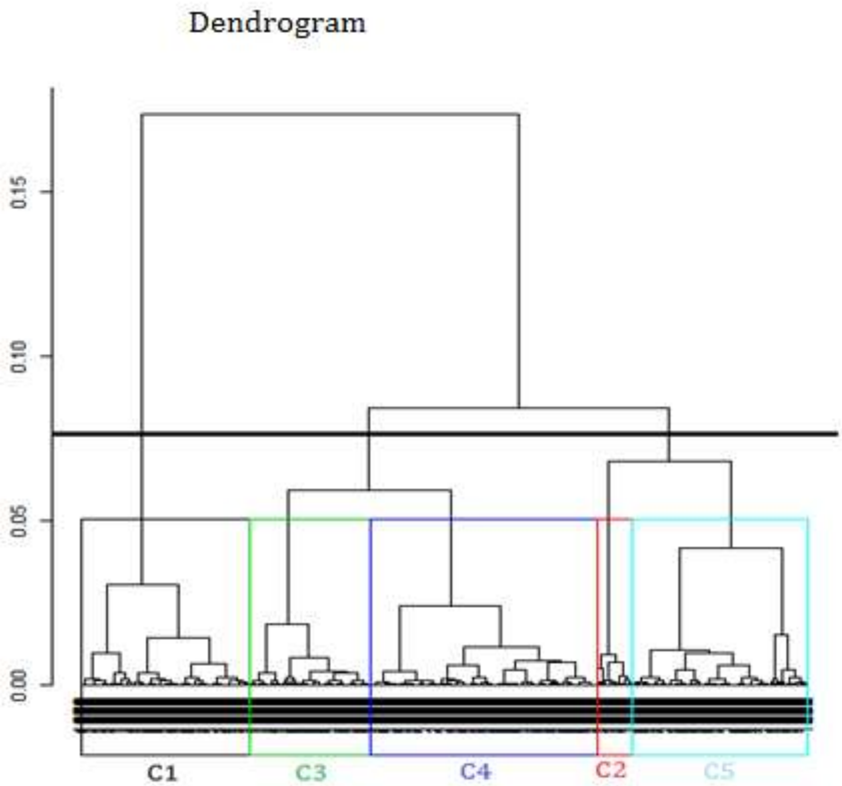


Figure 7.10: Cluster maps from MCA known as dendrogram

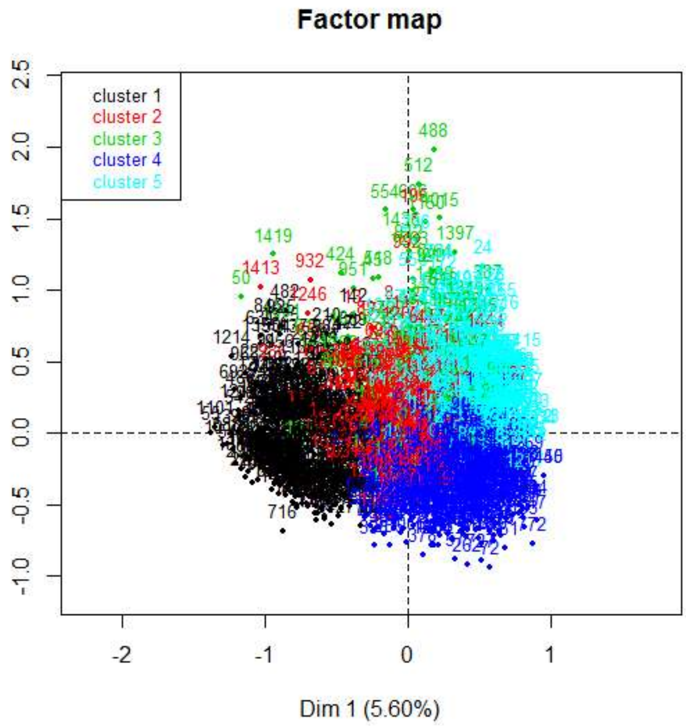


Figure 7.11: Scatter plot with the cluster visualisation

Figure 7.10 shows the results of hierarchical cluster analysis carried out in this study plotted as a dendrogram which illustrates the cut-off point for the retention for the number of clusters emerging from this data. The various levels of cut-off point on the dendrogram sort the data into different numbers of clusters. The MCA then assigns the individuals to different clusters segregated using different colours, as shown in Figure 7.11. The choice of the number of clusters is a compromise given the descriptions of the clusters are more significant and consist of fewer individuals within the clusters when the cut-off point on the dendrogram is at a lower point. However, a lower cut-off point will lead to a higher number of clusters which in this study is not effective because a high number of clusters fails to distinguish groups of individuals.

In depth scrutiny of different cut-off points was used to reveal the most optimum cut-off point, which resulted in five clusters to represent the data. A further check was made on the choice of 5 clusters by using the K-means clustering technique. Indeed, the results produced statistically significantly similar results on the same data. Therefore, validation was achieved and the solution obtained by HCA considered stable.

Cluster classifications

The demographic profiles of the combined datasets of car users from 2011 to 2014 for each of the five clusters identified in both by cluster analysis and the MCA were tabulated in Table 7.8. Since all clusters were dominated by car users, their level of use was used to separate them as follows: Car Engagement (CE–active/ frequent/ inactive use of car) and Sustainable Transport Consumption (STC–never/ infrequent/ occasional/ often types of travellers using public transport/bicycle).

The five clusters are described in Table 7.9. This analysis revealed different groups of car users with different socio-demographic characteristics, car engagement and sustainable transport consumption.

Variables	Categories	Clusters				
		1 28.56%	2 11.07%	3 31.81%	4 20.94%	5 7.62%
Age	18-24	2.78	2.99	1.67	5.70	5.22
	25-34	2.78	19.16	8.75	**33.54**	20.00
	35-44	2.78	**34.73**	18.54	**38.29**	**30.43**
	45-54	5.57	28.74	**31.04**	22.15	22.61
	55-64	24.36	11.38	**30.63**	0.32	8.70
	65+	**61.72**	2.99	9.38	0.00	13.04
Gender	Male	**51.28**	**65.87**	**55.63**	36.71	37.39
	Female	48.72	34.13	44.38	**63.29**	**62.61**
Household size	One	33.64	16.77	23.13	10.13	13.91
	Two	**52.90**	**36.53**	**42.08**	20.89	25.22
	Three	8.58	18.56	15.63	24.05	22.61
	Four or more	4.87	28.14	19.17	**44.94**	**38.26**
Car ownership	One	**65.43**	41.32	**51.04**	32.28	**50.43**
	Two	28.54	**49.70**	38.13	**54.11**	38.26
	Three	4.41	7.78	7.92	9.81	9.57
	Four or more	1.62	1.20	2.92	3.80	1.74
Employment status	Employee (FT)	0.00	**69.46**	**64.38**	**62.97**	0.00
	Employee (PT)	0.00	13.17	19.58	25.63	0.00
	Self-employed (FT)	0.00	12.57	11.88	9.49	0.00
	Self-employed (PT)	0.00	3.59	4.17	1.90	0.00
	In work (status not known)	0.00	0.60	0.00	0.00	0.00
	Waiting to take up work	0.23	0.60	0.00	0.00	2.61
	Unemployed	3.94	0.00	0.00	0.00	31.30
	Looking after the home	3.71	0.00	0.00	0.00	**47.83**
	Retired	**74.48**	0.00	0.00	0.00	16.52
	In FT education	4.41	0.00	0.00	0.00	1.74
	Other	13.23	0.00	0.00	0.00	0.00
Car_driver	Every day or nearly every day	**52.20**	**57.49**	**81.46**	**79.43**	**67.83**
	2-5 days a week	37.59	31.14	16.04	15.82	28.70
	Once a week	6.26	5.99	1.88	3.48	3.48
	Less often but at least once/month	0.70	2.99	0.42	0.32	0.00
	Less often than that	1.86	1.20	0.21	0.95	0.00
	Never nowadays	1.39	1.20	0.00	0.00	0.00
Car_passenger	Every day or nearly every day	4.64	4.19	1.04	7.91	6.09
	2-5 days a week	**25.99**	26.95	9.58	31.01	19.13
	Once a week	20.42	**43.11**	13.75	**47.15**	**32.17**
	Less often but at least once/month	18.10	19.76	21.67	13.29	13.04
	Less often than that	16.94	4.79	24.58	0.63	12.17
	Never nowadays	13.92	1.20	**29.38**	0.00	17.39
Bus_usage	Every day or nearly every day	1.86	6.59	0.63	0.00	0.00
	2-5 days a week	7.19	6.59	2.71	0.00	0.00
	Once a week	9.98	10.18	2.71	2.53	4.35
	Less often but at least once/month	13.46	9.58	9.79	10.76	9.57
	Less often than that	18.56	17.37	19.58	20.25	24.35
	Never nowadays	**48.96**	**49.70**	**64.58**	**66.46**	**61.74**
Train_usage	Every day or nearly every day	0.00	6.59	0.00	0.32	0.00
	2-5 days a week	0.23	14.37	0.63	0.32	0.87
	Once a week	2.55	5.99	1.25	0.95	0.00
	Less often but at least once/month	14.62	20.36	16.04	18.35	11.30
	Less often than that	34.80	**39.52**	**45.63**	**48.10**	**45.22**
	Never nowadays	**47.80**	13.17	36.46	31.96	42.61
Bike_usage	Every day or nearly every day	0.00	9.58	1.04	0.00	5.22
	2-5 days a week	0.46	22.75	1.67	0.00	13.04
	Once a week	2.09	**29.34**	3.96	2.22	15.65
	Less often but at least once/month	3.48	11.38	8.75	14.56	4.35
	Less often than that	6.73	4.79	18.13	23.10	14.78
	Never nowadays	**87.24**	22.16	**66.46**	**60.13**	**46.96**
Total		100.00	100.00	100.00	100.00	100.00

Note: The red figures are related to the highest shares that linked in to the cluster classification

Table 7.8: Cluster characteristics for all respondents aggregated over 2011 to 2014

An in-depth study of the clusters' characteristics suggested that individuals are considered to be fairly homogenous in terms of travel behaviour, mainly because the majority of the individuals either drive a car or are car passengers in the cohort selected for study. Only the second cluster showed a slightly higher usage of bicycle. However, individuals in the different clusters showed differences in socio-demographic characteristics such as age, gender and employment status. These are discussed here in more detail, taking each cluster in turn.

Cluster 1: Older respondents (65+), who were males and in retirement. Respondents in this group used cars every day or nearly every day as drivers, and often as passengers. What stands out in this group is that they never use public transport (buses or trains) or bicycles for travel.

Cluster 2: Middle-aged males (35-44 years old) in full-time employment who mostly use cars as the driver and occasionally as passengers. Respondents in this group live in households of two people, own 2 cars per household and show a similar pattern in sustainable transport consumption as the first cluster. Interestingly, respondents in this cluster reported the frequent use of bicycles.

Cluster 3: Mature males (45–64 years old) in full-time employment recorded as active car drivers, and never car passengers. This cluster includes respondents living in two-person households with single car ownership and use. Their sustainable transport consumption levels are low, given that they never use buses or bicycles.

Cluster 4: Young and middle-aged females (25-44 years old) in full-time employment with larger household sizes of four or more people, who own and use two cars regularly as drivers every day or nearly every day, whilst using cars as passengers once a week only. This group consumes quite low levels of sustainable transport and typically reported that they never used buses and bicycles, and infrequently used trains for travel.

Cluster 5: Middle aged (35-44 years old), female, and unemployed (looking after the home) with larger families (4 or more members living in the household), owning only 1 car per household. The travel behaviour for this group is the same as respondents in Cluster 1.

An overview of the five clusters is presented in Table 7.9. Cluster 4 (and Cluster 2) differ only in their use of bicycle never (often), gender female (male), and number in household ≥4 (2) respectively. The final Cluster 5 is one of the two mainly female groups and is the same in all respects to Cluster 4, except in that they look after the home, owning one car, probably because these households have less disposable income. These clusters are used in the next section to construct a multinomial logistic regression analysis. Car users aged 18-24 years old being 3.25% of the population did not emerge within clusters descriptive as an independent cluster.

Cluster	CE	STC	Descriptions
One	• Active-driver • Often-passenger	• Never use bus • Never use of train • Never use bicycle	• Older–aged (65+) • Male • Retired • 2 people in the HH • Owned 1 car per HH
Two	• Active-driver • Occasionally - passenger	• Never use bus • Infrequent use of train • Often use bicycle	• Middle–aged (35-44) • Male • Full–time employee • 2 people in the HH • Owned 2 cars per HH
Three	• Active-driver • Never-passenger	• Never use bus • Infrequent use of train • Never use bicycle	• Mature adults (45-54) (55-64) • Male • Full–time employee • 2 people in the HH • Owned 1 car per HH
Four	• Active-driver • Occasionally - passenger	• Never use bus • Infrequent use of train • Never use bicycle	• Middle–aged (25-34) (35-44) • Female • Full–time employee • ≥4 people in the HH • Owned 2 cars per HH
Five	• Active-driver • Occasionally-passenger	• Never use bus • Infrequent use of train • Never use bicycle	• Middle–aged (35-44) • Female • Looking after the home • ≥4 people in the HH • Owned 1 car per HH

*HH = Household
CE = Car engagement (active/ frequent/ inactive)
STC = Sustainable transport consumption (never/ infrequent/ occasional/ often)

Table 7.9: Cluster classifications

Kruskal-Wallis

In keeping with this empirical method, a recognised hypothesis testing was carried out using a Kruskal-Wallis non-parametric, which is an assessment to certify that cluster differences are statistically significant at 95% level of confidence. Table 7.10 shows the results of the hypothesis testing for each variable and demonstrates that the five clusters are statistically significantly different.

	Test Statistics[a,b]		
Variable	χ^2	df	Asymp. Sig.
HH#	248.95	4	0.00
Gender	52.77	4	0.00
Age	646.30	4	0.00
Car#	80.59	4	0.00
EmpStatus	1163.74	4	0.00
Car_driver	120.36	4	0.00
Car_passenger	324.30	4	0.00
Bus_usage	63.16	4	0.00
Train_usage	112.24	4	0.00
Bike_usage	342.20	4	0.00

df: degree of freedom
a. Kruskal Wallis Test
b. Grouping Variable: Clusters

Table 7.10: Independent samples Kruskal-Wallis

7.3 Multinomial Logistic Regression Analysis with Clustered Data

The next step of the analysis was to use multinomial logistic regression considering the 5 clusters obtained from MCA and 3 factors from PAF. This analysis explores whether any relationships exist between clusters and factors and if any patterns and trends occur over 4 successive years (2011 to 2014). Given that Cluster 1 was unique in being solely of older car users, for this analysis, therefore, Cluster 1 was chosen as a baseline or reference category to investigate how attitudes towards and perceptions of environmental issues of other car users differed from older car users. Clusters were used as dependent variables, whilst the factors, discussed in Chapter 6, were used as independent variables. Table 7.11 presents the relationships between the clusters and factors relative to the older male retired car users. The following text serves to identify key messages that emerged from this analysis presented in Table 7.11.

The attitudes to transport and environment (factor 1), consisting of car users' perception towards reducing the amount of car use and intention towards buying a car with lower CO_2 emission in the future, of clusters 2, 3, 4 and 5 were statistically significant at 90% - 99% confidence level in 2011 only, but in 2012 to 2014 were seen as less important. In 2013, car users were found to be aware of the sustainability associated with owning or using sustainable modes such as public transport, cycling and walking for a short journey of less than 2 miles for the environmental benefits. This statistically significant relationship was observed for car users in cluster 2 with factor 3 (modal shift potential), confirming middle–aged males in full–time employment have become more aware of the need for mode shift and the car users in this cluster are a favourable target for any campaign to encourage mode switch to sustainable modes.

There was a high level of statistical significance between clusters 2, 3 and 4 with factor 2 (traffic awareness) every year with statistical significance at 95% – 99% across all four years. The reason that all three factors are statistically significant for car users in cluster 3 in 2011 is possibly due to their maturity, being in a higher income bracket, in full–time employment and having a strong affinity towards car use which is exhibited by their attitudes to transport and environment, traffic awareness, and also travel mode shift. Unlike cluster 3, all factors are not significant for respondents of cluster 5 in 2012.

Within cluster 5 (Middle–aged (35-44), female, looking after the home), traffic congestion and exhaust fumes in towns, cities and motorway (factor 2) became more significant from 2013 to 2014. Although the p-value does not change greatly (0.00 to 0.01), it may represent a noteworthy change in attitudes among this particular group. Initially there was a statistically significant positive relationship with attitudes to transport and the environment in 2011. However, due to active use of cars and rare use of sustainable modes, this factor seems to be less important for middle–aged females who were looking after the home in cluster 5. This may be due to more responsibility for children putting them in situation of needing to travel by car at peak times. However, this is a conjecture as the data is not available in the BSA to explore this further.

Cluster	Factors – EFA	2011		2012		2013		2014	
		B	p-value	B	p-value	B	p-value	B	p-value
C2*	Attitudes to transport and environment	1.77	**0.04**[b]	0.85	0.47	1.04	0.89	0.87	0.58
	Traffic awareness	0.37	**0.00**[a]	0.47	**0.00**[a]	0.56	**0.03**[b]	0.51	**0.01**[a]
	Modal shift potential	0.74	0.18	0.85	0.41	0.61	**0.04**[b]	0.98	0.94
	Intercept	-0.79	0.00	-0.65	0.00	-1.11	0.00	-1.31	0.00
C3**	Attitudes to transport and environment	1.54	**0.03**[b]	1.35	**0.08**[c]	1.06	0.78	1.23	0.19
	Traffic awareness	0.51	**0.00**[a]	0.46	**0.00**[a]	0.66	**0.03**[b]	0.64	**0.01**[a]
	Modal shift potential	0.82	**0.03**[b]	0.80	0.16	0.83	0.24	0.82	0.18
	Intercept	0.46	0.00	0.02	0.28	0.11	0.43	-0.07	0.57
C4***	Attitudes to transport and environment	1.70	**0.01**[a]	0.80	0.28	1.40	0.12	1.06	0.77
	Traffic awareness	0.62	**0.01**[b]	0.63	**0.02**[b]	0.72	**0.10**[c]	0.54	**0.00**[a]
	Modal shift potential	0.82	0.24	0.84	0.33	1.00	1.00	1.21	0.28
	Intercept	0.17	0.29	-0.33	0.04	-0.26	0.11	-0.56	0.00
C5****	Attitudes to transport and environment	1.71	**0.07**[c]	0.96	0.87	0.12	0.71	0.54	0.88
	Traffic awareness	0.72	0.24	0.75	0.24	0.39	**0.00**[a]	0.24	**0.01**[a]
	Modal shift potential	0.78	0.32	1.10	0.69	1.24	0.38	0.45	0.26
	Intercept	-1.15	0.00	-1.10	0.00	-1.33	0.00	-1.74	0.00
McFadden Pseudo R^2		0.03		0.03		0.03		0.03	
No of samples		365		400		341		403	

Notes: [a] significant at 99%, [b] significant at 95%, and [c] significant at 90%

* Middle–aged (35-44), male, full–time employee

** Mature adults (45-64), male, full–time employee

*** Young and middle–aged (35-44), female, full–time employee

**** Middle–aged (35-44), female, looking after the home

- Reference category is cluster 1 (older–aged (65+), male, retired)

Table 7.11: MLR analysis examining the effects of respondents' travel behaviour relative to senior citizen through relationships between the clusters and factors

Back in 2011, all clusters were significantly more positive towards attitudes to transport and environment (factor 1) compared to the older male retired group who were really reluctant to give up their cars. They were more likely to be thinking positively about transport and the environment in 2011, and then year on year that slowly waned. However, cluster 3 who were mature male adults in a small household with only one car, were also more positive than the retired group in 2012. Nevertheless, it was no longer statistically significant in the following years.

This brings two messages: either the older generation has become more aware of environment and the need for sustainable transport and environment, given that the majority will possess a free bus pass, or as the population has aged over years, these factors become less important to them. They own cars so they use them.

In 2011, all clusters were found to be more aware of traffic congestion (factor 2) than the older group. This could be because the older group travel less mileages by car compared to the other groups. Respondents in cluster 2 and cluster 3 who were males, 2 person households, likely childless couples owning one or two cars, were statistically significantly consistently aware of traffic congestion throughout the year 2011 – 2014. This probably reflected the amount of travel (probably day-to-day commuting) and dependency on the cars of these two groups. Respondents in cluster 2 could be aware of traffic congestion of cycle users, as they were reported as frequent cycle users.

Interestingly, respondents in Cluster 2 relative to the older group in 2011 were not statistically significant with modal shift potential (factor 3), but by 2013 had begun to accept the need for action to switch travel from private to sustainable modes for a short journey of less than 2 miles, with evidence of occasional use of cycles. Meanwhile, respondents in Cluster 3 show a reverse trend, accepting the need for action to switch transport modes in 2011 only. Therefore, the car users in Cluster 2 are a favourable target for any campaign to encourage mode switch to sustainable modes. They were estimated to be the most prone to take action to help reduce the impact on environmental and climate change problems and were the most likely groups to be willing to change travel modes.

7.4 Conclusions

This chapter has presented an extensive analysis of the categorical variables using MCA and MLR, which aims to investigate the groups within the population with greater levels of concern and awareness of climate change and environmental problems. Changes in attitudes to and perceptions of climate change over time were also sought in order to decide which group of travellers would be willing to take action to help reduce their impact on climate change.

By using MLR, each cluster within the population with the higher levels of mode shift potential over time (significant value) was sought in order to decide which group of car users can be used to produce targeted travel behaviour campaign and acknowledge which cluster is more susceptible for sustainability. The results confirm that the population is not uniform in terms of their attitudes and motivations in relation to reducing CO_2 emissions from personal travel, therefore policies and universal solutions to encourage more sustainable transport behaviours are deemed unlikely to be effective. The results also show that the cluster groups that exist are not defined or differentiated by demographic features alone. Motivations and barriers to change in travel behaviour and to use alternative modes differed widely between the groups. A degree of influence from environmental concerns was found for all groups.

When the datasets from 2011 to 2014 were combined together to gain further insights into groups with similar views, 5 clusters emerged from the analysis. Differences were found in each year, yet similar views were seen spread throughout the five clusters which differed in demographic characteristics and views. The evidence recommends that acknowledgement of the concept of climate change among the car users were high. The results demonstrated that different reasons could influence the same behaviour; however, different behaviours could lead to the same attitudes.

MCA additionally identified various significant correlations of socio-demographic attributes with travel behaviour. The hierarchical cluster analysis distinguished 5 clusters to represent all the individuals in this study. From an interpretation of the clusters, Cluster 2 represent groups of respondents who have moderate consumption of travel and dependency on sustainable transport. Furthermore, MLR analysis identified that the respondents in Clusters 2 and 3 were more sensitive to factor 3 “modal shift potential”. On the other hand, even though almost all of the respondents from 2011 to 2014 strongly agreed that human actions are partly responsible for the impact of climate change, the results also revealed an opposite trend, because almost all groups paid less attention to environmental problems in the later years during this period.

The respondents’ demographic characteristics such as age, gender, employment status, household size and car ownership play an important role and are revealed to have a significant effect on travel behaviour patterns and the willingness to switch to other transportation modes. This is similar to Fatmi and Habib (2017) where they found that bigger household size and driver’s licence also influence travel modes switch decisions.

In the next chapter we will describe the use of log-linear and multivariate probit models, fitted using Bayesian inference.

The model so-far developed using multinomial logistic regression does not consider the correlations between responses to several questions. Therefore, in the next chapter, we will use a multivariate probit model (MPM) which will allow us to include the ordinal responses to several questions, which can be correlated, in a single model. This will allow us to look at the whole collection of responses from an individual and relate this collection to explanatory variables.

Furthermore, in the next chapter, we will adopt Bayesian inference. This will allow us to do such things as computing predictive probabilities of responses, given particular values for explanatory variables, in a way which allows for both the sampling variation between individuals and the remaining uncertainty in the values of model parameters.

The practicality of this approach will be demonstrated and show that the methodology could be applied to different datasets in the future. For example, the methodology could be applied to research conducted using datasets from developing countries to compare with the results from developed countries.

Chapter 8 Bayesian Inference Approach

8.1 Introduction

In Chapter 7, an investigation of the differences in their perceptions to the environment among respondents with dissimilar characteristics and travel behaviours was presented using MLR. This chapter further investigates car users' attitudes in an attempt to demonstrate their potential to switch modes to more sustainable transport and thus reduce their environmental impact. Achieving objective 4 involves the following tasks:

1. To develop log-linear models for categorical data with and without age and gender effects using question 12 ($q12$) only.

2. To construct a multivariate probit model (MPM) using a Bayesian inference approach for ordinal responses, allowing for covariates for each question.

3. To include responses to all 14 attitudinal questions by introducing a multivariate normal vector latent variable ($\underline{Z}_i$).

4. To demonstrate the potential use of the BSA information more effectively for decision making in transport planning activities.

In section 8.2, log-linear models for categorical data constructed in the study are discussed by exploring $q12$ to investigate age and gender effects. Attention is then directed to the development of the MPM with Bayesian inference using attitudinal data in Section 8.3. Next, the results are discussed in detail; and finally, Section 8.4 presents the outcomes of the analysis and conclusions drawn from the study.

8.2 Log-Linear Model for Categorical Data

This section develops a model for categorical data. One specific question, $q12$ namely, “For the sake of the environment, everyone should reduce how much they use their cars”, was selected to carry out the analysis as an example.

8.2.1 Separate model for age and gender groups

This analysis develops separate models for $q12$ ignoring the interaction between age and gender. The respondents were divided into twelve age-gender groups and given six age groups each for males and females. The groups were labelled 1 to 12 as follows:

1	Male	18 – 24 years old
2	Male	25 – 34 years old
3	Male	35 – 44 years old
4	Male	45 – 54 years old
5	Male	55 – 64 years old
6	Male	65+ years old
7	Female	18 – 24 years old
8	Female	25 – 34 years old
9	Female	35 – 44 years old
10	Female	45 – 54 years old
11	Female	55 – 64 years old
12	Female	65+ years old

Responses to the question were categorised into one of the five labelled 1 to 5 as follows:

1 Agree strongly
2 Agree
3 Neither agree nor disagree
4 Disagree
5 Disagree strongly

It was assumed that the respondents in each group were a random sample of people in that group. For the population of people in group i , the proportion who would give response k is given in equation 8.1.

$$\Pr\ \text{response } k|\text{group } i\ \ = \pi_{i,k} \qquad (8.1)$$

So, we write:

$$\pi_{i,k} = \frac{\phi_{i,k}}{{}_{k=1}^{5} \phi_{i,k}} \qquad (8.2)$$

and:

$$\eta_{i,k} = ln\phi_{i,k} \qquad (8.3)$$

where *ln* stands for natural logarithm, so $\phi_{i,k} = \exp\ \eta_{i,k}$, except for the case that $\eta_{i,3} = 0$ which means that the response 3, "neither agree nor disagree", was chosen as the baseline response. It follows that:

$$\phi_{i,k} = \frac{\Pr\ \text{response } k|\text{group } i}{\Pr\ \text{response } 3|\text{group } i} \qquad (8.4)$$

are the *odds* in favour of response k compared to response 3. Then $\eta_{i,k}$ is the corresponding *log odds*. So, if $\eta_{i,1} = ln2 \approx 0.69$, then the response "Agree strongly" would be twice as likely as the response "Neither agree nor disagree" for members of group i. Similarly, if $\eta_{i,5} = -ln2 \approx -0.69$, then the response "Disagree strongly" would be half as likely as the response "Neither agree nor disagree" for members of group i. Furthermore,

$$\eta_{i,k} - \eta_{i,j} = ln\ \ \frac{\pi_{i,k}\ \ \pi_{i,3}}{\pi_{i,j}\ \ \pi_{i,3}}\ \ = ln\ \ \frac{\pi_{i,k}}{\pi_{i,j}} \qquad (8.5)$$

So, if $\eta_{i,k} - \eta_{i,j} = ln2$, then response k is twice as likely as response j for members of group i. Normal prior distributions were given to the parameters $\eta_{1,1}, \dots, \eta_{12,5}$, except for $\eta_{i,3}$, for $i = 1, \dots, 12$, which are fixed as equal to zero. A hierarchical prior specification induces prior correlations between the parameters. For $i = 1, \dots, 12$ and for $k = 1, 2, 4, 5$, given the values of $\beta_1,\ \beta_2, \beta_4,\ \beta_5$, the conditional prior distribution of $\eta_{i,k}$ given β_k is:

$$\eta_{i,k}|\beta_k \sim N\ \ \beta_k, 1.25 \tag{8.6}$$

where $\eta_{i,k}$ is conditionally independent of $\eta_{i',k'}$ given $\beta_k, \beta_{k'}$ unless $i = i'$ and $k = k'$. Then, for $k = 1, 2, 4, 5$:

$$\beta_k \sim N\ \ 0, 5 \tag{8.7}$$

with β_k independent of $\beta_{k'}$ unless $k = k'$. Thus, the marginal prior distribution of $\eta_{i,k}$ for $k = 1, 2, 4, 5$, is:

$$\eta_{i,k} \sim N\ \ 0, 6.25 \tag{8.8}$$

and $\eta_{i,k}$ and $\eta_{i',k}$ have a prior correlation of:

$$\text{Corr}\ \ \eta_{i,k}, \eta_{i',k}\ \ = \frac{5}{6.25} = 0.8 \tag{8.9}$$

For this purpose, JAGS (Plummer, 2003), which is a program used for the analysis of Bayesian graphical models using Gibbs sampling, via the rjags package in R (R Core Team, 2017), has been used to compute the posterior distribution by Markov Chain Monte Carlo (MCMC) sampling (Martin *et al.*, 2011; Plummer, 2016). The model specification, R commands, and content of the data are shown in Appendix A. Table 8.1 presents the sample distribution for $q12$ according to age and gender variables that are used in this model.

Gender	Age	Responses to $q12$					Total
		1	2	3	4	5	
Male	18-24	3	7	5	2	2	19
	25-34	2	42	24	15	6	89
	35-44	12	55	39	20	4	130
	45-54	8	80	31	33	7	159
	55-64	7	74	37	25	10	153
	65+	11	107	48	33	8	207
Female	18-24	1	12	10	6	1	30
	25-34	11	68	31	13	3	126
	35-44	12	99	56	12	6	185
	45-54	8	86	45	17	2	158
	55-64	8	66	34	17	4	129
	65+	8	70	22	22	2	124

Note:
1=agree strongly, 2=agree, 3=neither agree nor disagree, 4=disagree, 5=disagree strongly

Table 8.1: Gender and age groups for $q12$.

Two parallel MCMC chains were used to fit the log-linear model. Convergence was checked with trace plots and found to be satisfactory, as can be seen in Appendix A. A burn-in[1] for the MCMC sampler of 1000 iterations of both chains was used and samples of the values of η were collected from 2000 further iterations in two chains. Summaries of the posterior distributions of the various η parameters were obtained, based on 4000 samples, from the posterior distribution, as shown in Table 8.2. The results are presented in Table 8.2, where $\eta_{i,3}$, which is fixed at zero for all i, has been removed.

[1] A colloquial term that describes the practice of throwing away some iterations at the beginning of an MCMC run.

Parameter	Mean	SD	Parameter	Mean	SD
Group 1 (Male, 18-24)			**Group 2** (Male, 24-34)		
$\eta_{1,1}$	-1.28	0.32	$\eta_{2,1}$	-1.68	0.29
$\eta_{1,2}$	0.51	0.28	$\eta_{2,2}$	0.60	0.19
$\eta_{1,4}$	-0.66	0.31	$\eta_{2,4}$	-0.51	0.23
$\eta_{1,5}$	-1.84	0.34	$\eta_{2,5}$	-1.76	0.30
Group 3 (Male, 35-44)			**Group 4** (Male, 45-54)		
$\eta_{3,1}$	-1.28	0.24	$\eta_{4,1}$	-1.48	0.26
$\eta_{3,2}$	0.44	0.17	$\eta_{4,2}$	0.78	0.16
$\eta_{3,4}$	-0.61	0.21	$\eta_{4,4}$	-0.22	0.19
$\eta_{3,5}$	-2.04	0.29	$\eta_{4,5}$	-1.85	0.28
Group 5 (Male, 55-64)			**Group 6** (Male, 65+)		
$\eta_{5,1}$	-1.56	0.26	$\eta_{6,1}$	-1.48	0.24
$\eta_{5,2}$	0.65	0.16	$\eta_{6,2}$	0.75	0.14
$\eta_{5,4}$	-0.49	0.20	$\eta_{6,4}$	-0.46	0.18
$\eta_{5,5}$	-1.68	0.27	$\eta_{6,5}$	-1.91	0.27
Group 7 (Female, 18-24)			**Group 8** (Female, 25-34)		
$\eta_{7,1}$	-1.51	0.32	$\eta_{8,1}$	-1.24	0.25
$\eta_{7,2}$	0.51	0.25	$\eta_{8,2}$	0.77	0.17
$\eta_{7,4}$	-0.50	0.29	$\eta_{8,4}$	-0.76	0.23
$\eta_{7,5}$	-1.97	0.34	$\eta_{8,5}$	-2.06	0.30
Group 9 (Female, 35-44)			**Group 10** (Female, 45-54)		
$\eta_{9,1}$	-1.44	0.24	$\eta_{10,1}$	-1.53	0.25
$\eta_{9,2}$	0.68	0.14	$\eta_{10,2}$	0.75	0.16
$\eta_{9,4}$	-1.09	0.22	$\eta_{10,4}$	-0.77	0.21
$\eta_{9,5}$	-2.03	0.27	$\eta_{10,5}$	-2.21	0.29
Group 11 (Female, 55-64)			**Group 12** (Female, 65+)		
$\eta_{11,1}$	-1.44	0.26	$\eta_{12,1}$	-1.34	0.27
$\eta_{11,2}$	0.69	0.17	$\eta_{12,2}$	0.93	0.18
$\eta_{11,4}$	-0.64	0.22	$\eta_{12,4}$	-0.32	0.21
$\eta_{11,5}$	-2.01	0.29	$\eta_{12,5}$	-2.10	0.31

Table 8.2: Posterior summaries: means and standard deviations

The table above presents the posterior means and standard deviations of 12 groups for $q12$ – "for the sake of the environment, everyone should reduce how much they use their cars". Next, plots of the posterior probability functions were produced by extracting the sampled values first and then using the standard R plot functions to produce the plots. For illustration, group 3, male, 35-44 years old, "agree strongly" was chosen and a plot for $\eta_{3,1}$ was produced. The resulting graph is shown in Figure 8.1.

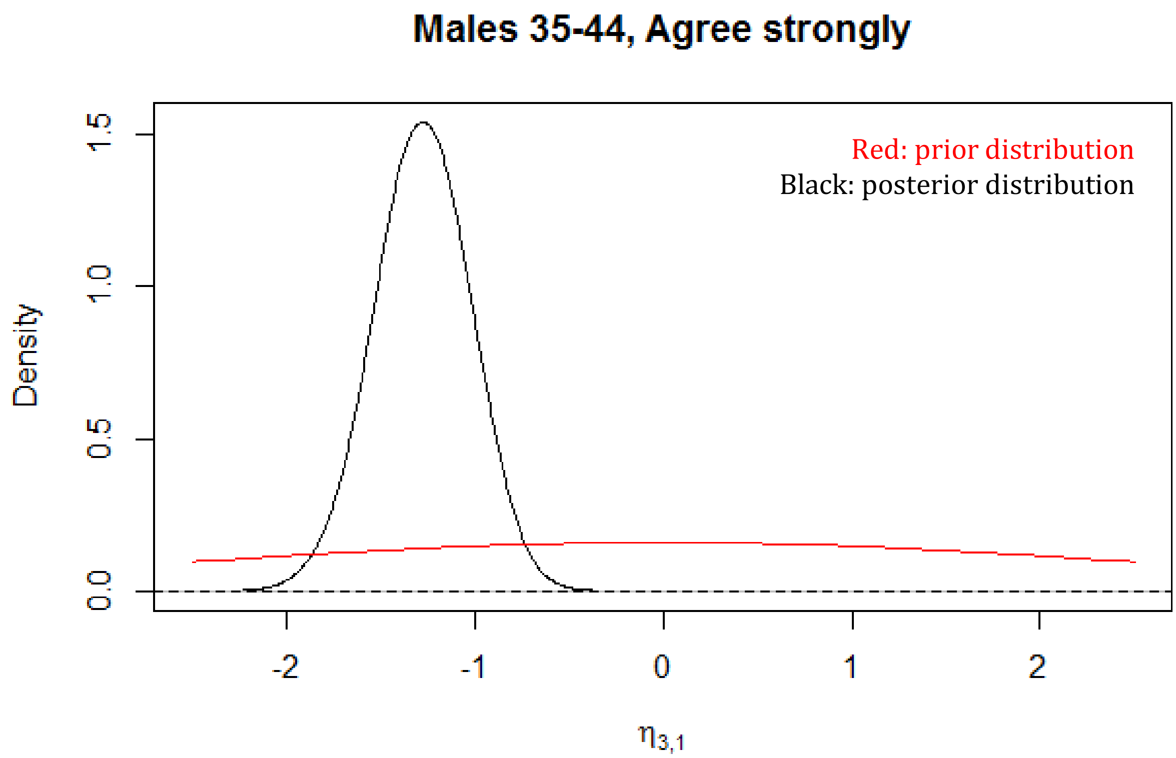


Figure 8.1: Probability density functions for $\eta_{3,1}$, the log odds for "agree strongly" in $q12$ males belonging to age group 35-44.

Looking at Figure 8.1, it is apparent that the value of $\eta_{3,1}$ is almost certainly negative. This finding suggests that members of this group were less likely to make the response "agree strongly" than to respond as "neither agree nor disagree" when they were asked to reduce how much they used their cars for the sake of the environment.

More generally, two responses can be compared by looking at the difference between the log odds. For this case, the same group was chosen to compare their level of perceptions for the same question, $q12$, for the responses $\eta_{3,2}$ ("agree") with $\eta_{3,4}$("disagree").

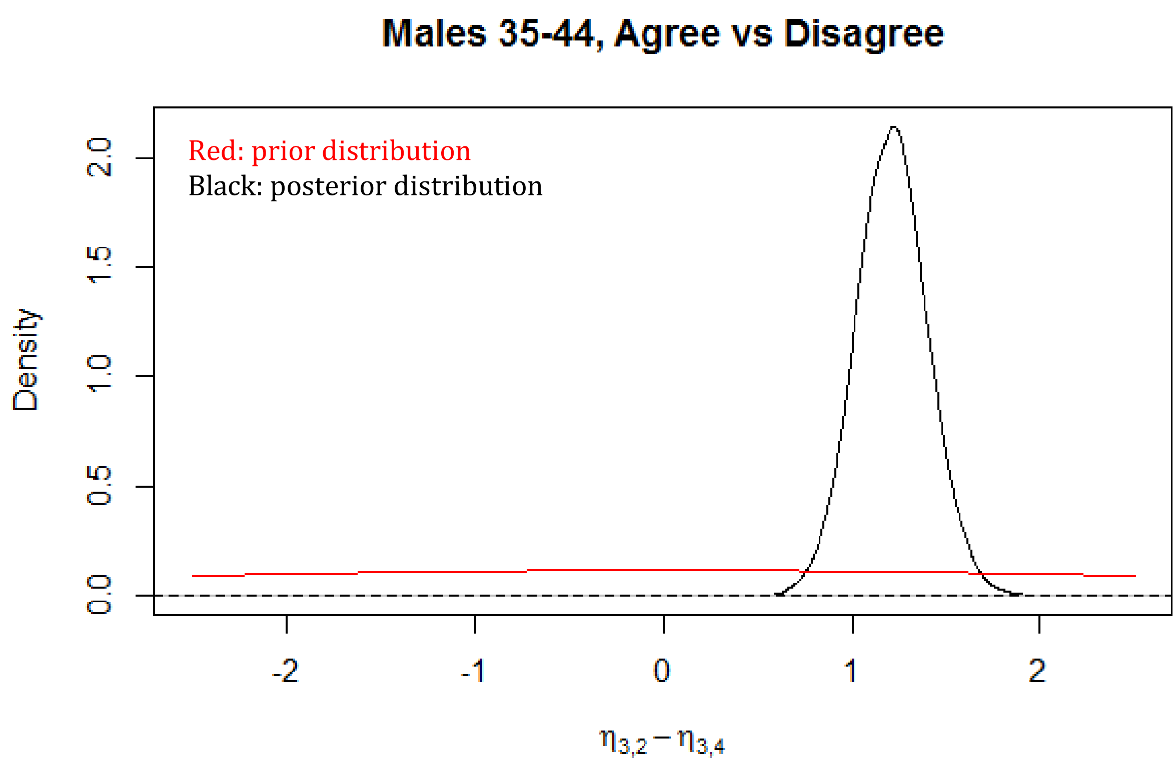


Figure 8.2: Probability density functions for $\eta_{3,2} - \eta_{3,4}$, the log odds for "agree" compared to "disagree" for males aged 35-44.

The prior distribution for $\eta_{3,2} - \eta_{3,4}$ is normally distributed assumed to have a mean 0 and variance 12.5 $\sim N\ 0{,}12.5$, since they are independent in the prior. The standard deviation is thus $\sqrt{12.5} = 3.54$. The resulting graph is shown in
Figure 8.2. It can be clearly seen that a person in this group (male, 35-44 years old) was more likely to respond "agree" than "disagree" when they were asked to reduce the amount of car use for the sake of the environment.

8.2.2 Model with interaction between age and gender effects

In this section, the effect of age and gender on responses to $q12$ "For the sake of the environment, everyone should reduce how much they use their cars" was investigated. The respondents are categorised as either "males" or "females" associated with six age-groups. Interest lies in the effects of the respondents' age and gender on the outcome for the responses. The effects of age and gender are introduced using six orthogonal contrasts, labelled as gender, a_1, a_2, a_3, a_4, and a_5, as shown in Table 8.3.

Gender	Age	Responses to $q12$					Total	Gender	a_1	a_2	a_3	a_4	a_5
		1	2	3	4	5							
Male	18-24	3	7	5	2	2	19	1	1	1	-1	0	0
	25-34	2	42	24	15	6	89	1	1	-1	-1	0	0
	35-44	12	55	39	20	4	130	1	1	0	2	0	0
	45-54	8	80	31	33	7	159	1	-1	0	0	1	-1
	55-64	7	74	37	25	10	153	1	-1	0	0	-1	-1
	65+	11	107	48	33	8	207	1	-1	0	0	0	2
Female	18-24	1	12	10	6	1	30	-1	1	1	-1	0	0
	25-34	11	68	31	13	3	126	-1	1	-1	-1	0	0
	35-44	12	99	56	12	6	185	-1	1	0	2	0	0
	45-54	8	86	45	17	2	158	-1	-1	0	0	1	-1
	55-64	8	66	34	17	4	129	-1	-1	0	0	-1	-1
	65+	8	70	22	22	2	124	-1	-1	0	0	0	2

Note:
Responses to q12; 1= agree strongly, 2= agree, 3= neither agree nor disagree, 4= disagree, 5= disagree strongly

Table 8.3: Cross tabulations of responses to $q12$ according to age and gender, and proposed orthogonal contrasts.

In this analysis, $k = 3$ (neither agree nor disagree) was used as a baseline. The other four categories where $k = 1, 2, 4, 5$, $\eta_{i,k}$ can be modelled in terms of an age effect $\beta_{a,k}$, a gender effect $\beta_{g,k}$ and an interaction effect between age and gender $\beta_{ag,k}$ as follows:

Group 1: $\eta_{1k} = \beta_{0,k} + \beta_{a_1,k} + \beta_{a_2,k} - \beta_{a_3,k} + \beta_{g,k} + \beta_{a_1 g,k} + \beta_{a_2 g,k} - \beta_{a_3 g,k}$

Group 2: $\eta_{2k} = \beta_{0,k} + \beta_{a_1,k} - \beta_{a_2,k} - \beta_{a_3,k} + \beta_{g,k} + \beta_{a_1 g,k} - \beta_{a_2 g,k} - \beta_{a_3 g,k}$

Group 3: $\eta_{3k} = \beta_{0,k} + \beta_{a_1,k} + 2\beta_{a_3,k} + \beta_{g,k} + \beta_{a_1 g,k} + 2\beta_{a_3 g,k}$

Group 4: $\eta_{4k} = \beta_{0,k} - \beta_{a_1,k} + \beta_{a_4,k} - \beta_{a_5,k} + \beta_{g,k} - \beta_{a_1 g,k} + \beta_{a_4 g,k} - \beta_{a_5 g,k}$

Group 5: $\eta_{5k} = \beta_{0,k} - \beta_{a_1,k} - \beta_{a_4,k} - \beta_{a_5,k} + \beta_{g,k} - \beta_{a_1 g,k} - \beta_{a_4 g,k} - \beta_{a_5 g,k}$

Group 6: $\eta_{6k} = \beta_{0,k} - \beta_{a_1,k} + 2\beta_{a_5,k} + \beta_{g,k} - \beta_{a_1 g,k} + 2\beta_{a_5 g,k}$

Group 7: $\eta_{7k} = \beta_{0,k} + \beta_{a_1,k} + \beta_{a_2,k} - \beta_{a_3,k} - \beta_{g,k} + \beta_{a_1 g,k} + \beta_{a_2 g,k} - \beta_{a_3 g,k}$

Group 8: $\eta_{8k} = \beta_{0,k} + \beta_{a_1,k} - \beta_{a_2,k} - \beta_{a_3,k} - \beta_{g,k} + \beta_{a_1 g,k} - \beta_{a_2 g,k} - \beta_{a_3 g,k}$

Group 9: $\eta_{9k} = \beta_{0,k} + \beta_{a_1,k} + 2\beta_{a_3,k} - \beta_{g,k} + \beta_{a_1 g,k} + 2\beta_{a_3 g,k}$

Group 10: $\eta_{10k} = \beta_{0,k} - \beta_{a_1,k} + \beta_{a_4,k} - \beta_{a_5,k} - \beta_{g,k} - \beta_{a_1g,k} + \beta_{a_4g,k} - \beta_{a_5g,k}$

Group 11: $\eta_{11k} = \beta_{0,k} - \beta_{a_1,k} - \beta_{a_4,k} - \beta_{a_5,k} - \beta_{g,k} - \beta_{a_1g,k} - \beta_{a_4g,k} - \beta_{a_5g,k}$

Group 12: $\eta_{12k} = \beta_{0,k} - \beta_{a_1,k} + 2\beta_{a_5,k} - \beta_{g,k} - \beta_{a_1g,k} + 2\beta_{a_5g,k}$

The covariate a_1 takes the values (1, 1, 1, -1, -1, and -1) in the first six male groups, similarly for female groups. This list of contrasts can be extended to all 12 groups by taking $\beta_{0,3} = \beta_{a,3} = \beta_{g,3} = \beta_{ag,3} = 0$, which in turn gives $\eta_{1,3} = \eta_{2,3} = \cdots = \eta_{12,3} = 0$. In this way, the parameters are made relative to the baseline response category $k = 3$ (neither agree nor disagree).

Apart from the baseline values $\beta_{0,k}$, the coefficients of β_g in η_g are orthogonal contrasts. Briefly, using such contrasts gives some useful structure upon which to develop the prior distribution for the η_g. The contrasts have a useful property if a particular type of prior is taken for the non-baseline β_g: the same prior is used for each $\beta_{0,k}\ (k \neq 3)$, as well as the same prior for each $\beta_{a,k}\ (k \neq 3)$, but possibly different from that for the $\beta_{0,k}$. Similarly, for $\beta_{g,k}\ (k \neq 3)$ and for $\beta_{ag,k}\ (k \neq 3)$, and these blocks are independent of one another.

More time could be spent looking at how to construct the various elements of the prior distribution in detail. However, for clarity, a fairly simple example is used here. There was no evidence that inferences were sensitive to changes in details of the prior. The prior distribution is taken to have independent components, for $k = 1, 2, 4, 5$ as follows:

$\beta_{0,k} | \mu_0 \sim N(\mu_0, 1.0)$, $\quad \mu_0 \sim N(0, 1.0)$

$\beta_{a_1,k} | \mu_{a_1} \sim N(\mu_{a_1}, 1.0)$, $\quad \mu_{a_1} \sim N(0, 0.1)$

$\beta_{a_2,k} | \mu_{a_2} \sim N(\mu_{a_2}, 1.0)$, $\quad \mu_{a_2} \sim N(0, 0.1)$

$\beta_{a_3,k} | \mu_{a_3} \sim N(\mu_{a_3}, 1.0)$, $\quad \mu_{a_3} \sim N(0, 0.1)$

$\beta_{a_4,k} | \mu_{a_4} \sim N(\mu_{a_4}, 1.0)$, $\quad \mu_{a_4} \sim N(0, 0.1)$

$\beta_{a_5,k} | \mu_{a_5} \sim N(\mu_{a_5}, 1.0)$, $\quad \mu_{a_5} \sim N(0, 0.1)$

$\beta_{g,k} | \mu_g \sim N(\mu_g, 1.0)$, $\quad \mu_g \sim N(0, 0.1)$

$\beta_{a_1g,k} | \mu_{a_1g} \sim N(\mu_{a_1g}, 1.0)$, $\quad \mu_{a_1g} \sim N(0, 0.05)$

$$\beta_{a_2g,k}|\mu_{a_2g}\sim N\left(\mu_{a_2g},1.0\right), \qquad \mu_{a_2g}\sim N\left(0,0.05\right)$$

$$\beta_{a_3g,k}|\mu_{a_3g}\sim N\left(\mu_{a_3g},1.0\right), \qquad \mu_{a_3g}\sim N\left(0,0.05\right)$$

$$\beta_{a_4g,k}|\mu_{a_4g}\sim N\left(\mu_{a_4g},1.0\right), \qquad \mu_{a_4g}\sim N\left(0,0.05\right)$$

$$\beta_{a_5g,k}|\mu_{a_5g}\sim N\left(\mu_{a_5g},1.0\right), \qquad \mu_{a_5g}\sim N\left(0,0.05\right)$$

Next, plots of the posterior probability functions were produced by extracting the sampled values first and then using the standard R plot functions to produce the plots. In order to explain these, an example is presented and describes plots in detail. Group 3, male, 35-44 years old, "agree strongly" was chosen and a plot for $\eta_{3,1}$ was produced. The posterior means and standard deviations are presented Table 8.4 where $\eta_{i,3}$, which is fixed at zero for all i , has been removed, and the resulting graphs are shown in Figure 8.3 and Figure 8.4.

Parameter	Mean	SD	Parameter	Mean	SD
Group 1 (Male, 18-24)			**Group 2** (Male, 25-34)		
$\eta_{1,1}$	-0.76	0.55	$\eta_{2,1}$	-1.95	0.49
$\eta_{1,2}$	-0.36	0.51	$\eta_{2,2}$	0.57	0.25
$\eta_{1,4}$	-0.66	0.59	$\eta_{2,4}$	-0.51	0.32
$\eta_{1,5}$	-1.31	0.68	$\eta_{2,5}$	-1.48	0.43
Group 3 (Male, 35-44)			**Group 4** (Male, 45-54)		
$\eta_{3,1}$	-1.23	0.33	$\eta_{4,1}$	-1.40	0.36
$\eta_{3,2}$	0.35	0.21	$\eta_{4,2}$	0.93	0.20
$\eta_{3,4}$	-0.69	0.27	$\eta_{4,4}$	0.01	0.24
$\eta_{3,5}$	-2.24	0.47	$\eta_{4,5}$	-1.54	0.38
Group 5 (Male, 55-64)			**Group 6** (Male, 65+)		
$\eta_{5,1}$	-1.64	0.37	$\eta_{6,1}$	-1.49	0.33
$\eta_{5,2}$	0.70	0.20	$\eta_{6,2}$	0.81	0.17
$\eta_{5,4}$	-0.38	0.25	$\eta_{6,4}$	-0.37	0.23
$\eta_{5,5}$	-1.39	0.34	$\eta_{6,5}$	-1.83	0.38
Group 7 (Female, 18-24)			**Group 8** (Female, 25-34)		
$\eta_{7,1}$	-1.68	0.62	$\eta_{8,1}$	-1.18	0.35
$\eta_{7,2}$	0.27	0.40	$\eta_{8,2}$	0.78	0.21
$\eta_{7,4}$	-0.65	0.48	$\eta_{8,4}$	-0.84	0.311
$\eta_{7,5}$	-2.30	0.68	$\eta_{8,5}$	-2.26	0.49
Group 9 (Female, 35-44)			**Group 10** (Female, 45-54)		
$\eta_{9,1}$	-1.56	0.31	$\eta_{10,1}$	-1.74	0.36
$\eta_{9,2}$	0.57	0.17	$\eta_{10,2}$	0.66	0.18
$\eta_{9,4}$	-1.51	0.31	$\eta_{10,4}$	-0.95	0.27
$\eta_{9,5}$	-2.32	0.42	$\eta_{10,5}$	-2.81	0.50
Group 11 (Female, 55-64)			**Group 12** (Female, 65+)		
$\eta_{11,1}$	-1.50	0.37	$\eta_{12,1}$	-1.07	0.40
$\eta_{11,2}$	0.66	0.21	$\eta_{12,2}$	1.15	0.24
$\eta_{11,4}$	-0.72	0.29	$\eta_{12,4}$	-0.03	0.29
$\eta_{11,5}$	-2.24	0.47	$\eta_{12,5}$	-2.29	0.60

Note: SD = standard deviation.

Table 8.4: Posterior means and standard deviations.

Figure 8.3 has two distributions, the flat red curve is the prior distribution and the sharper normal curve is the posterior. With reference to Table 8.4, it is apparent that the value of $\eta_{3,1}$ of the posterior distribution is negative -0.90 and SD=0.24. This finding suggests that members of this group were less likely to respond “agree strongly” than to respond “neither agree nor disagree”, when they were asked to reduce how much they used their cars for the sake of the environment.

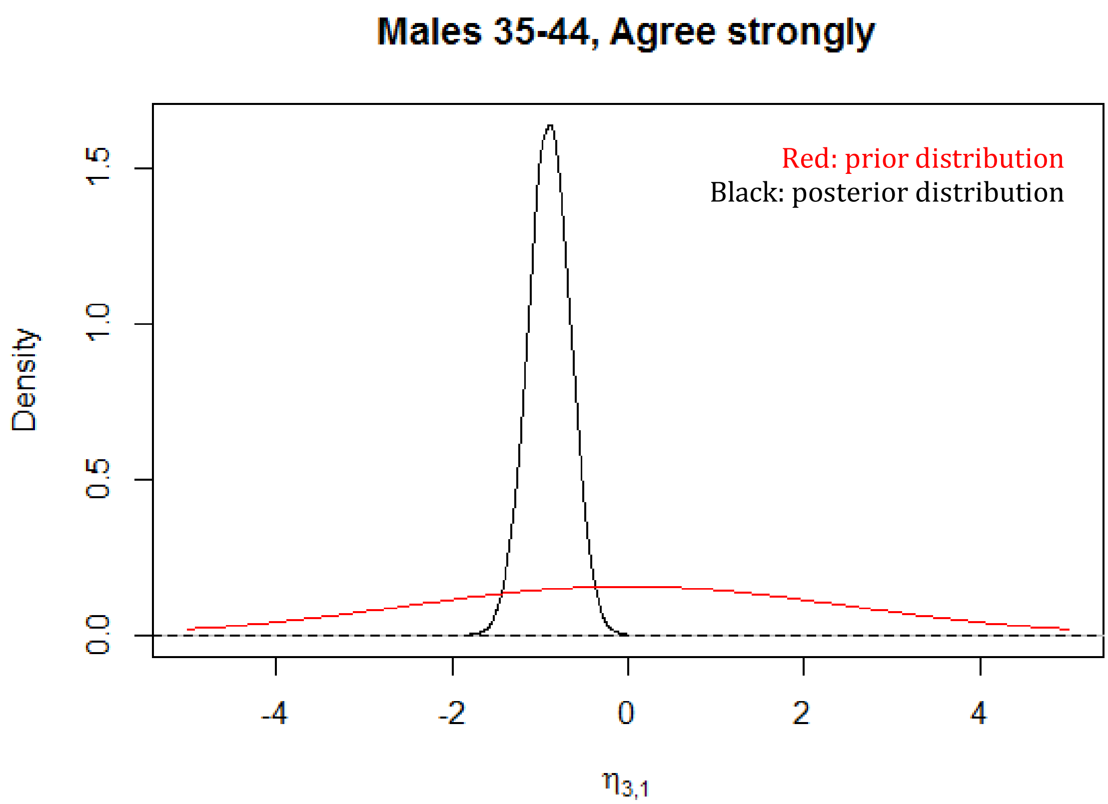


Figure 8.3: Probability density functions for $\eta_{3,1}$, the log odds for "agree strongly" for males aged 35-44.

Subsequently, the same group was chosen to compare their level of perceptions for the same question, $q12$ instead of relative to "neither agree nor disagree", but instead $\eta_{3,2}$ ("agree") is compared to $\eta_{3,4}$("disagree"). The resulting graph is shown in Figure 8.4. It can be clearly seen that a person in this group (male, 35-44 years old) was more likely to respond "agree" than "disagree", when they were asked to reduce the amount of car use for the sake of the environment.

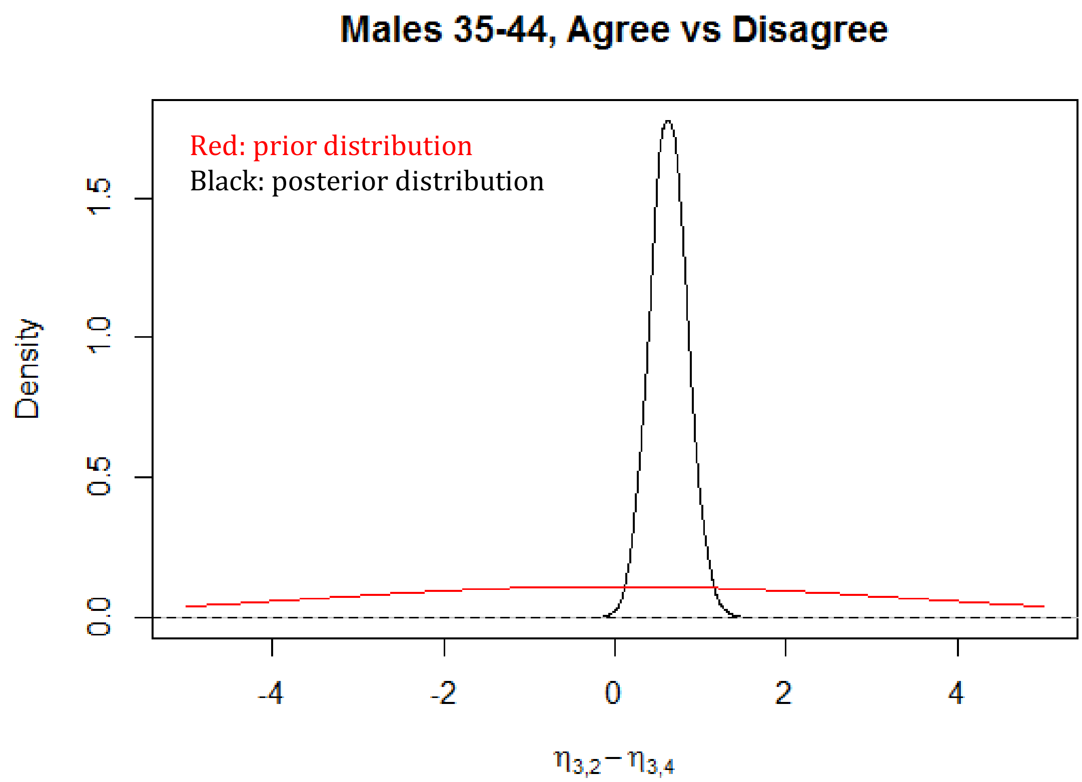


Figure 8.4: Probability density functions for $\eta_{3,2} - \eta_{3,4}$, the log odds for "agree" compared to "disagree" for males aged 35-44.

It is worth noting that most of the posterior means were negatives in Table 8.2 except for $\eta_{i,2}$. This seems to suggest that "agree" tends to be a popular choice for q12, whereas $\eta_{i,2}$ is sometimes positive, but did not always emerge when the interaction of age and gender effect was considered in the models, as shown in Table 8.4.

In log linear modelling, the questions have to be dealt with one question at a time. Therefore, in the next section, MPM model is developed to consider all responses to all attitudinal variables in a single model.

8.3 Developing a Multivariate Probit Model with Bayesian Inference for Categorical Data

A multivariate probit model was proposed for use in this study because a probit model is a model for ordinal responses which will allow for consideration of the responses to several questions in a single model and the responses can be correlated (Congdon, 2005). The effect of responses to explanatory variables such as gender and age can also be related in this model.

8.3.1 Methodology

This model allows for the investigation of relationships between the responses of a car user to the different questions, where each response is a category on an ordinal scale. The basic idea is that, associated with each individual respondent i, there is a vector random variable $\underline{Z}_i$ with a multivariate normal distribution, as shown in equation 8.10.

$$\underline{Z}_i \sim N \quad \underline{\mu}_i, V \qquad (8.10)$$

Here, $\underline{Z}_i$ is normally distributed with a mean vector $\underline{\mu}_i$ and covariance matrix V. So:

$$\underline{Z}_i = \quad Z_{i,1}, \dots, Z_{i,Q} \quad ^{T} \qquad (8.11)$$

where $Z_{i,q}$ is a univariate normal random variable associated with individual i and question q, for $q = 1, \dots, Q$. Let the number of categories be C, then the number of thresholds or cut-off points will be $C - 1$ $(c_{q,1}, \dots, c_{q,C-1})$.

Let the category selected by respondent i in question q be $Y_{i,q}$. Then:

$$
\begin{aligned}
Y_{i,q} &= 1 && \Leftrightarrow && Z_{i,q} \le c_{q,1}, \\
Y_{i,q} &= 2 && \Leftrightarrow \; c_{q,1} && < Z_{i,q} \le c_{q,2}, \\
&\dots \\
Y_{i,q} &= C-1 && \Leftrightarrow \; c_{q,C-2} && < Z_{i,q} \le c_{q,C-1}, \\
Y_{i,q} &= C && \Leftrightarrow \; c_{q,C-1} && < Z_{i,q}
\end{aligned}
\tag{8.12}
$$

The cut-off points are parameters of the model. However, since the mean and variance of $Z_{i,q}$ are also model parameters, without loss of generality, two cut-off points were fixed, to avoid over-parameterising the model.

Given that V is an inverse Wishart[2] prior distribution, by choosing the multivariate normal distribution, modelling correlations is made relatively easy and there does not seem to be an obvious disadvantage to this choice compared to other distributions. In fact, the effect on the model of choosing one kind of distribution rather than another is probably slight.

Thus, for $k = 1, \dots, C-1$,

$$\Pr\left(Y_{i,q} \le k\right) = \Pr\left(Z_{i,q} \le c_{q,k}\right) = \Phi\left(\frac{c_{q,k} - \mu_{i,q}}{\sigma_q}\right) \tag{8.13}$$

where:

$\Phi =$ the standard normal cumulative distribution function (cdf),

$\underline{\mu}_i = \left(\mu_{i,1}, \dots, \mu_{i,Q}\right)^T$,

$\sigma_q^2 =$ the marginal variance of $Z_{i,q}$.

[2] A probability distribution defined on real-valued positive-definite matrices. In Bayesian statistics, it is used as the conjugate prior for the covariance matrix of a multivariate normal distribution.

More generally, the element in row r and column q of V is $v_{r,q}$ and $\sigma_q^2 = v_{q,q}$. The mean $\mu_{i,q}$ is related to the explanatory variables of the linear model, as shown in equation 8.14.

$$\mu_{i,q} = \beta_{0,q} + \sum_{j=1}^{J} \beta_{j,q} x_{j,i} \tag{8.14}$$

where $x_{j,i}$ is the value of covariate j for subject i. Thus Pr $Y_{i,q} \leq k$ is related to the covariate values by the sigmoid function defined in equation 8.13. Other sigmoid link functions, such as a logistic function or a complementary log-log function, could be used, but the standard normal cdf is convenient because the multivariate normal distribution makes modelling correlations relatively easy and the differences between the shapes of these functions are not great.

8.3.2 Application of the model using the BSA dataset

Question q has C_q ordered categories for its response which have to be consistent across all data. The attitudinal variables, depending on the questions can have 3, 4 or 5 categories. In order to overcome this inconsistency, categories were combined as follows: strongly agree and agree, strongly disagree and disagree, a very serious problem and a serious problem. Therefore, the three categories in this study considered ($C_q = C = 3$) were labelled as 0, 1 and 2. By doing so, there were no unknown threshold parameters.

Five socio-demographic variables and 14 attitudinal variables from the BSA dataset were used to illustrate the fitting of a multivariate probit model to investigate car users' attitudes, as presented in Table 8.5 and Table 8.6 respectively. The proportion of age, employment status, household size and car ownership according to gender, as well as 14 attitudinal questions selected for this model with the details of response options, can be seen in Chapter 5 (see Table 5.5).

No.	Variable
1	Age
2	Gender
3	Number of people living in the household, including respondent
4	Number of cars regularly owned and used in the household
5	Employment status

Table 8.5: Socio-demographic variables

No.	Variable
q1	How serious a problem for you is congestion on motorways?
q2	How serious a problem for you is traffic congestion in towns and cities?
q3	How serious a problem for you are exhaust fumes from traffic in towns and cities?
q4	Next time I buy a car, I would be willing to buy a car with lower CO2 emissions.
q5	I am willing to reduce the amount I travel by car (To help reduce the impact of CC).
q6	View on climate change and causes.
q7	Many of the short journeys that I now make by car I could just as easily walk.
q8	Many of the short journeys that I now make by car I could just as easily go by bus.
q9	Many of the short journeys that I now make by car I could just as easily cycle.
q10	For the sake of the environment, car users should pay higher taxes.
q11	People should be allowed to use their cars as much as they like, even it is cause damage to the environment.
q12	For the sake of the environment, everyone should reduce how much they use cars.
q13	There is no point in reducing my car use to help the environment unless others do the same.
q14	People who drive cars that are better for the environment should pay less to use roads.

Table 8.6: Attitudinal variables

In the next section, the model obtained from the data analysis is described.

8.3.3 Model for latent variables

The next step was to construct a model for the latent variables. A model for C probabilities has been built with C responses. However, these must sum to 1, so there are really only $C-1$ free parameters. The normal distribution $Z_{i,q} \sim N(\mu_{i,q}, {\sigma^2}_q)$ requires two parameters, $\mu_{i,q}$ and ${\sigma^2}_q$, and $C-1$ cut-off points. Therefore, to avoid over-parameterising, without loss of generality, two cut-off points were fixed (-1 and 1).

In addition, unlike the analysis in Section 8.2, the latent normal random variables $Z_{i,1}, \ldots, Z_{i,Q}$ were allowed to be dependent on each other and their correlations were treated as unknown. It is more convenient to treat a covariance matrix as an unknown parameter rather than a correlation matrix, since we can give the covariance matrix an inverse Wishart prior distribution. Therefore, it is preferable to treat ${\sigma^2}_q$ as unknown. Then $\mu_{i,q}$ is allowed to be dependent on covariates so that $\mu_{i,q}$ depends on the five covariates for each respondent *i* in a linear model. These are: age, gender, household size, employment status and car ownership, as shown in the equation 8.15.

$$\mu_{i,q} = \beta_{0,q} + \beta_{ActAge,q}x_{ActAge,i} + \beta_{Gender,q}x_{Gender,i} + \sum_{j=1}^{4}\beta_{HH,q,j}\delta_{HH,j,i} + \sum_{j=1}^{4}\beta_{Car,q,j}\delta_{Car,j,i} + \beta_{inwork,q}x_{inwork,i} + \beta_{employee,q}x_{employee,i} + \beta_{fulltime,q}x_{fulltime,i} + \sum_{j=1}^{4}\beta_{noemp,q,j}\delta_{noemp,j,i} \quad (8.15)$$

where $x_{ActAge,i}$ is the age in years of respondent *i* minus 50 (the approximate average age of the respondents). The advantages of centring by subtracting 50 are that it will tend to reduce the correlations in the posterior distribution and therefore improve the mixing of the Markov Chain Monte Carlo (MCMC) sampler. Also, it means that the intercept parameter β_0 corresponds to a realistic case, a person aged 50. Without this "centring", the intercept would correspond to a person aged 0 and it is difficult to interpret such a parameter or to have sensible prior beliefs about how a new-born could answer the questions.

The coding system that was used is as follows:

$x_{Gender,i}$ = -1 if the respondent is a male and 1 if the respondent is a female,

$inwork_i$ = 1 for a respondent in work and -1 otherwise,

$employee_i$ = 1 for employees, -1 for self-employees and 0 otherwise,

$fulltime_i$ = 1 if the respondent is in full-time employment, -1 if the respondent is a part-time employee and 0 otherwise,

$\delta_{HH,j,i}$ = 1 if household category for respondent *i* is *j* and 0 otherwise,

$\delta_{Car,j,i}$ = 1 if car category for respondent *i* is *j* and 0 otherwise,

$\delta_{noemp,j,i}$ = 1 if no employment category for respondent *i* is *j* and 0 otherwise.

The category labels for household size h, car ownership c, and no employment status u (those who are not in work) of respondent *i* are defined in Table 8.7.

Characteristics	Interval	Category label
Household size, h	One	h =1
	Two	h =2
	Three	h =3
	Four or more	h =4
Car ownership, c	One	c =1
	Two	c =2
	Three	c =3
	Four or more	c =4
No employment, u	Unemployed / waiting to take up work	u =1
	Looking after the home / retired	u =2
	In FT education / Other	u =3
	In work	u =4

Table 8.7: Category label of latent variables

In order not to over-parameterize the model, the following constraint was set in equation 8.16.

$$\sum_{h=1}^{4} \beta_{HH,q,h} = \sum_{c=1}^{4} \beta_{Car,q,c} = \sum_{n=1}^{4} \beta_{noemp,q,u} = 0 \tag{8.16}$$

Prior distributions were chosen as follows:

$\beta_{ActAge,q} \sim N(0,1)$,
$\beta_{Gender,q} \sim N(0,1)$,
$\beta_{HH,q,h} \sim N(0,1)$ for $h = 1, \dots, 4$,
$\beta_{Car,q,c} \sim N(0,1)$ for $c = 1, \dots, 4$,
$\beta_{noemp,q,u} \sim N(0,1)$ for $u = 1, \dots, 4$.
(8.17)

8.3.4 Model fitting

For the computation of posterior summaries, the Markov Chain Monte Carlo (MCMC) method was used to draw samples from the posterior distribution of the model parameters. Specifically, the software rjags was used, which is an interface in R (R core team, 2015) to JAGS (Plummer, 2016). Data augmentation (Tanner and Wong, 1987), was used with the latent variables $\underline{Z}_i$ treated as auxiliary data.

The parameters for this model consist of β_{ActAge} (age), β_{Gender} (gender), β_{HH} (household size), β_{Car} (car ownership), and 4 categories of employment status (β_{inwork}, $\beta_{employee}$, $\beta_{fulltime}$, and β_{noemp}). The precision matrix of $\underline{Z}$ is $W = V^{-1}$ and V is the covariance matrix of $\underline{Z}$. The values -1 and 1 were chosen as the fixed cut-off points, without loss of generality. Summaries of the posterior distribution were found using rjags (Plummer, 2003). Two parallel MCMC chains were used. A burn-in[3] of 2000 iterations was used and samples were collected from 10,000 further iterations, and so summaries are based on 20,000 samples from two parallel chains.

Convergence was checked with trace plots and shown to be satisfactory, as illustrated in Figure 8.5 which shows examples of trace plots and density curves. The other trace plots

[3] A colloquial term that describes the practice of throwing away some iterations at the beginning of an MCMC run.

and density curves for all of the parameters and rjags model specification can be found in Appendix B.

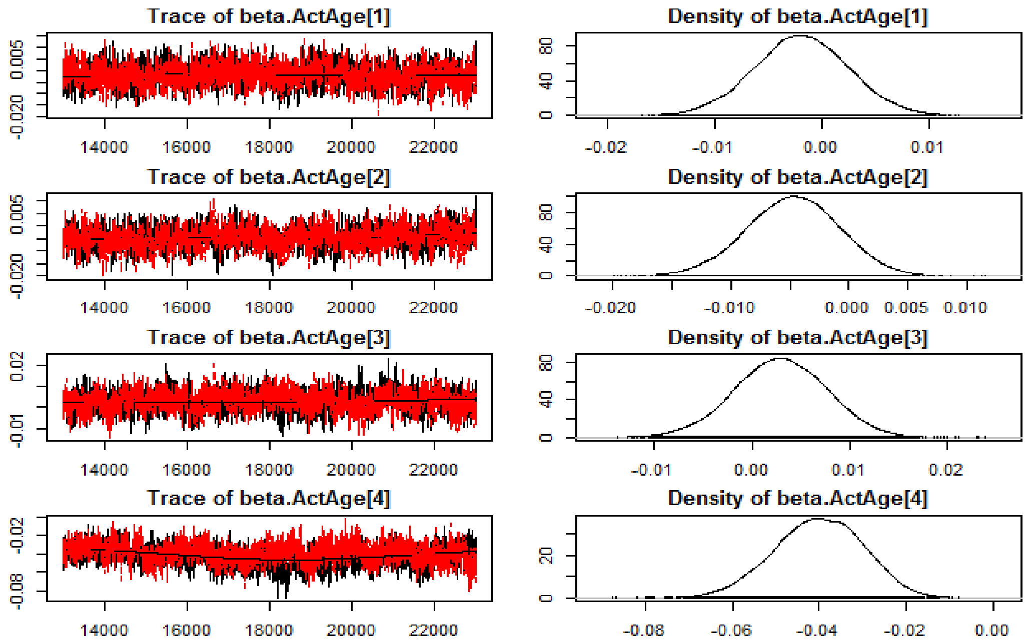


Figure 8.5: Trace plots and posterior density curves for the parameter of actual age in two parallel chains

The parameter trace plots were created to make sure that the sampler was mixing and moving well around the posterior distribution, and that the samples had "converged" so that they may be taken as a representative sample from the posterior distribution. It can be seen that the samples obtained from the two parallel chains have essentially the same properties and that mixing is good with little autocorrelation in the samples. Summaries of the means and standard deviations of the posterior for the coefficients of the socio-demographic variables are shown in Table 8.8. This table has 14 sections, one for each question such that q=1,..., 14 respectively corresponding to questions 1 to 14.

P	Mean	SD	P	Mean	SD	P	Mean	SD	P	Mean	SD
Question 1						Question 2					
β_{Car} 1,1	0.22	0.11	β_{ActAge} 1	0.00	0.00	β_{Car} 2,1	0.03	0.10	β_{ActAge} 2	0.00	0.00
β_{Car} 1,2	-0.15	0.11	β_{Gender} 1	0.17	0.05	β_{Car} 2,2	-0.06	0.10	β_{Gender} 2	0.13	0.05
β_{Car} 1,3	-0.61	0.16	$\beta_{employee}$ 1	-0.14	0.09	β_{Car} 2,3	-0.33	0.15	$\beta_{employee}$ 2	-0.06	0.08
β_{Car} 1,4	0.54	0.23	$\beta_{fulltime}$ 1	-0.10	0.08	β_{Car} 2,4	0.35	0.21	$\beta_{fulltime}$ 2	0.00	0.07
β_{HH} 1,1	-0.00	0.11	β_{inwork} 1	-0.06	0.18	β_{HH} 2,1	0.00	0.10	β_{inwork} 2	-0.06	0.07
β_{HH} 1,2	-0.00	0.08	β_{noemp} 1,1	-0.06	0.18	β_{HH} 2,2	-0.03	0.07	β_{noemp} 2,1	-0.16	0.16
β_{HH} 1,3	0.02	0.11	β_{noemp} 1,2	0.04	0.14	β_{HH} 2,3	0.00	0.09	β_{noemp} 2,2	0.12	0.12
β_{HH} 1,4	-0.02	0.10	β_{noemp} 1,3	0.02	0.17	β_{HH} 2,4	0.02	0.09	β_{noemp} 2,3	0.04	0.15
Question 3						Question 4					
β_{Car} 3,1	-0.22	0.12	β_{ActAge} 3	0.00	0.00	β_{Car} 4,1	-0.60	0.25	β_{ActAge} 4	-0.04	0.01
β_{Car} 3,2	0.00	0.11	β_{Gender} 3	-0.13	0.06	β_{Car} 4,2	-0.43	0.25	β_{Gender} 4	-0.58	0.13
β_{Car} 3,3	-0.29	0.17	$\beta_{employee}$ 3	-0.06	0.09	β_{Car} 4,3	0.30	0.36	$\beta_{employee}$ 4	-0.08	0.22
β_{Car} 3,4	0.50	0.24	$\beta_{fulltime}$ 3	-0.07	0.08	β_{Car} 4,4	0.73	0.47	$\beta_{fulltime}$ 4	0.37	0.22
β_{HH} 3,1	0.19	0.11	β_{inwork} 3	-0.13	0.08	β_{HH} 4,1	0.41	0.25	β_{inwork} 4	-0.23	0.18
β_{HH} 3,2	-0.04	0.08	β_{noemp} 3,1	0.12	0.19	β_{HH} 4,2	0.34	0.20	β_{noemp} 4,1	0.78	0.39
β_{HH} 3,3	0.09	0.11	β_{noemp} 3,2	-0.18	0.14	β_{HH} 4,3	-0.52	0.26	β_{noemp} 4,2	0.10	0.33
β_{HH} 3,4	-0.24	0.10	β_{noemp} 3,3	0.07	0.17	β_{HH} 4,4	-0.23	0.24	β_{noemp} 4,3	-0.88	0.39
Question 5						Question 6					
β_{Car} 5,1	-0.52	0.27	β_{ActAge} 5	0.01	0.01	β_{Car} 6,1	-0.01	0.19	β_{ActAge} 6	-0.01	0.01
β_{Car} 5,2	0.03	0.26	β_{Gender} 5	-0.24	0.14	β_{Car} 6,2	0.17	0.19	β_{Gender} 6	0.35	0.09
β_{Car} 5,3	0.34	0.37	$\beta_{employee}$ 5	-0.02	0.23	β_{Car} 6,3	-0.15	0.26	$\beta_{employee}$ 6	0.02	0.15
β_{Car} 5,4	0.15	0.53	$\beta_{fulltime}$ 5	0.40	0.21	β_{Car} 6,4	-0.01	0.39	$\beta_{fulltime}$ 6	0.12	0.14
β_{HH} 5,1	0.57	0.25	β_{inwork} 5	0.15	0.19	β_{HH} 6,1	-0.08	0.19	β_{inwork} 6	0.15	0.14
β_{HH} 5,2	-0.16	0.21	β_{noemp} 5,1	0.06	0.41	β_{HH} 6,2	0.07	0.14	β_{noemp} 6,1	-0.19	0.31
β_{HH} 5,3	-0.21	0.28	β_{noemp} 5,2	-0.12	0.33	β_{HH} 6,3	0.01	0.18	β_{noemp} 6,2	-0.19	0.23
β_{HH} 5,4	-0.21	0.25	β_{noemp} 5,3	0.06	0.40	β_{HH} 6,4	0.01	0.18	β_{noemp} 6,3	0.37	0.28
Question 7						Question 8					
β_{Car} 7,1	-0.33	0.31	β_{ActAge} 7	0.05	0.01	β_{Car} 8,1	-1.14	0.33	β_{ActAge} 8	0.00	0.01
β_{Car} 7,2	-0.28	0.31	β_{Gender} 7	0.21	0.17	β_{Car} 8,2	-0.13	0.31	β_{Gender} 8	0.01	0.18
β_{Car} 7,3	0.33	0.42	$\beta_{employee}$ 7	-0.34	0.28	β_{Car} 8,3	0.97	0.45	$\beta_{employee}$ 8	-0.16	0.29
β_{Car} 7,4	0.28	0.57	$\beta_{fulltime}$ 7	-0.23	0.24	β_{Car} 8,4	0.29	0.58	$\beta_{fulltime}$ 8	0.09	0.25
β_{HH} 7,1	0.51	0.32	β_{inwork} 7	-0.54	0.23	β_{HH} 8,1	0.77	0.32	β_{inwork} 8	0.37	0.24
β_{HH} 7,2	-0.37	0.26	β_{noemp} 7,1	-0.45	0.49	β_{HH} 8,2	-0.31	0.25	β_{noemp} 8,1	0.53	0.51
β_{HH} 7,3	-0.37	0.26	β_{noemp} 7,2	-0.42	0.39	β_{HH} 8,3	-0.66	0.32	β_{noemp} 8,2	-0.49	0.42
β_{HH} 7,4	-0.06	0.30	β_{noemp} 7,3	0.87	0.46	β_{HH} 8,4	0.20	0.32	β_{noemp} 8,3	-0.04	0.49
Question 9						Question 10					
β_{Car} 91	0.73	0.34	β_{ActAge} 9	0.10	0.02	β_{Car} 10,1	-0.07	0.23	β_{ActAge} 10	0.01	0.01
β_{Car} 9,2	0.33	0.34	β_{Gender} 9	1.24	0.19	β_{Car} 10,2	-0.13	0.21	β_{Gender} 10	-0.05	0.10
β_{Car} 9,3	-0.53	0.51	$\beta_{employee}$ 9	-0.54	0.29	β_{Car} 10,3	-0.28	0.28	$\beta_{employee}$ 10	-0.10	0.17
β_{Car} 9,4	-0.53	0.64	$\beta_{fulltime}$ 9	-0.12	0.27	β_{Car} 10,4	0.48	0.45	$\beta_{fulltime}$ 10	0.07	0.15
β_{HH} 9,1	-0.18	0.35	β_{inwork} 9	-0.79	0.25	β_{HH} 10,1	0.50	0.21	β_{inwork} 10	-0.01	0.15
β_{HH} 9,2	-0.68	0.28	β_{noemp} 9,1	-1.09	0.55	β_{HH} 10,2	0.00	0.15	β_{noemp} 10,1	0.95	0.37
β_{HH} 9,3	0.33	0.36	β_{noemp} 9,2	-0.24	0.43	β_{HH} 10,3	-0.59	0.20	β_{noemp} 10,2	-0.31	0.27
β_{HH} 9,4	0.53	0.34	β_{noemp} 9,3	1.33	0.53	β_{HH} 10,4	0.09	0.19	β_{noemp} 10,3	-0.64	0.33

Continued on the next page

Table 8.8 (continued)											
P	Mean	SD	P	Mean	SD	P	Mean	SD	P	Mean	SD
Question 11						Question 12					
β_{Car} 11,1	0.05	0.13	β_{ActAge} 11	0.02	0.01	β_{Car} 12,1	-0.10	0.17	β_{ActAge} 12	-0.01	0.01
β_{Car} 11,2	-0.07	0.12	β_{Gender} 11	0.31	0.06	β_{Car} 12,2	0.05	0.16	β_{Gender} 12	-0.22	0.09
β_{Car} 11,3	0.21	0.18	$\beta_{employee}$ 11	0.03	0.10	β_{Car} 12,3	0.19	0.24	$\beta_{employee}$ 12	0.37	0.14
β_{Car} 11,4	-0.20	0.25	$\beta_{fulltime}$ 11	0.08	0.09	β_{Car} 12,4	-0.14	0.32	$\beta_{fulltime}$ 12	0.24	0.13
β_{HH} 11,1	-0.25	0.12	β_{inwork} 11	0.10	0.09	β_{HH} 12,1	0.32	0.16	β_{inwork} 12	-0.27	0.12
β_{HH} 11,2	-0.07	0.09	β_{noemp} 11,1	0.09	0.21	β_{HH} 12,2	0.10	0.13	β_{noemp} 12,1	-0.23	0.27
β_{HH} 11,3	0.19	0.12	β_{noemp} 11,2	-0.15	0.16	β_{HH} 12,3	-0.24	0.17	β_{noemp} 12,2	0.04	0.21
β_{HH} 11,4	0.13	0.11	β_{noemp} 11,3	0.06	0.19	β_{HH} 12,4	-0.17	0.15	β_{noemp} 12,3	0.18	0.25
Question 13						Question 14					
β_{Car} 13,1	0.24	0.21	β_{ActAge} 13	-0.02	0.01	β_{Car} 14,1	-0.20	0.22	β_{ActAge} 14	-0.02	0.01
β_{Car} 13,2	-0.14	0.21	β_{Gender} 13	0.19	0.11	β_{Car} 14,2	-0.46	0.21	β_{Gender} 14	-0.15	0.12
β_{Car} 13,3	-0.07	0.31	$\beta_{employee}$ 13	-0.43	0.18	β_{Car} 14,3	0.00	0.29	$\beta_{employee}$ 14	0.25	0.19
β_{Car} 13,4	-0.02	0.42	$\beta_{fulltime}$ 13	0.13	0.15	β_{Car} 14,4	0.66	0.43	$\beta_{fulltime}$ 14	0.14	0.17
β_{HH} 13,1	0.13	0.22	β_{inwork} 13	0.04	0.15	β_{HH} 14,1	0.23	0.22	β_{inwork} 14	-0.54	0.16
β_{HH} 13,2	0.02	0.16	β_{noemp} 13,1	0.31	0.35	β_{HH} 14,2	-0.09	0.17	β_{noemp} 14,1	0.54	0.33
β_{HH} 13,3	-0.50	0.22	β_{noemp} 13,2	-0.63	0.26	β_{HH} 14,3	-0.56	0.24	β_{noemp} 14,2	-0.57	0.26
β_{HH} 13,4	0.35	0.19	β_{noemp} 13,3	0.32	0.32	β_{HH} 14,4	0.42	0.20	β_{noemp} 14,3	0.03	0.31

* P: Parameter, SD: Standard deviation

Table 8.8: Posterior means and standard deviations for coefficients of socio-demographic variables. Each prior mean was 0 and each prior standard deviation was 1.

8.3.5 Results

This section describes the development of the multivariate probit model using all of the attitudinal responses. Using the model, posterior distributions for the effects of the covariates on the responses can be computed. The posterior means and standard deviations are detailed in Table 8.8 for questions, q=1,..., 14. It is interesting to note that some of the responses show age and gender effects, such as for questions q = 3, 4, 7, 9, 11, 13 and 14.

The posterior distributions for β_{ActAge} and β_{Gender} for $q = 7, 8$, and 9 are shown in Figure 8.6. Figure 8.6 (a) shows that almost all of the posterior probability is in the region where $\beta_{ActAge} > 0$, strongly indicating an age effect on responses to $q7$. This indicates that older adults were more likely to disagree with switching from using a car to walking for short journeys. This group also disagreed with changing their travel behaviour from cars to public transport, such as buses as well as to cycling. The reasons for such attitudes and behaviour have been discussed elsewhere (Fuller *et al.*, 2013; Schepers and Heinen, 2013) where safety and health issues often feature.

For q9, "Many of the short journeys less than 2 miles that I now make by car, I could just as easily cycle", Figure 8.6 (e) and (f) show that both $\beta_{ActAge} > 0$ and $\beta_{Gender,9} > 0$. These values indicate the effect of age and gender on responses to $q9$ and suggests that older adults (above 50 years) who are males are more likely to disagree with switching from using a car to cycling for a short journey.

In terms of gender, the evidence suggests that females will be less likely to switch mode from cars to cycling for short journeys. However, switching to walking and going by bus were more or less equally acceptable for both males and females.

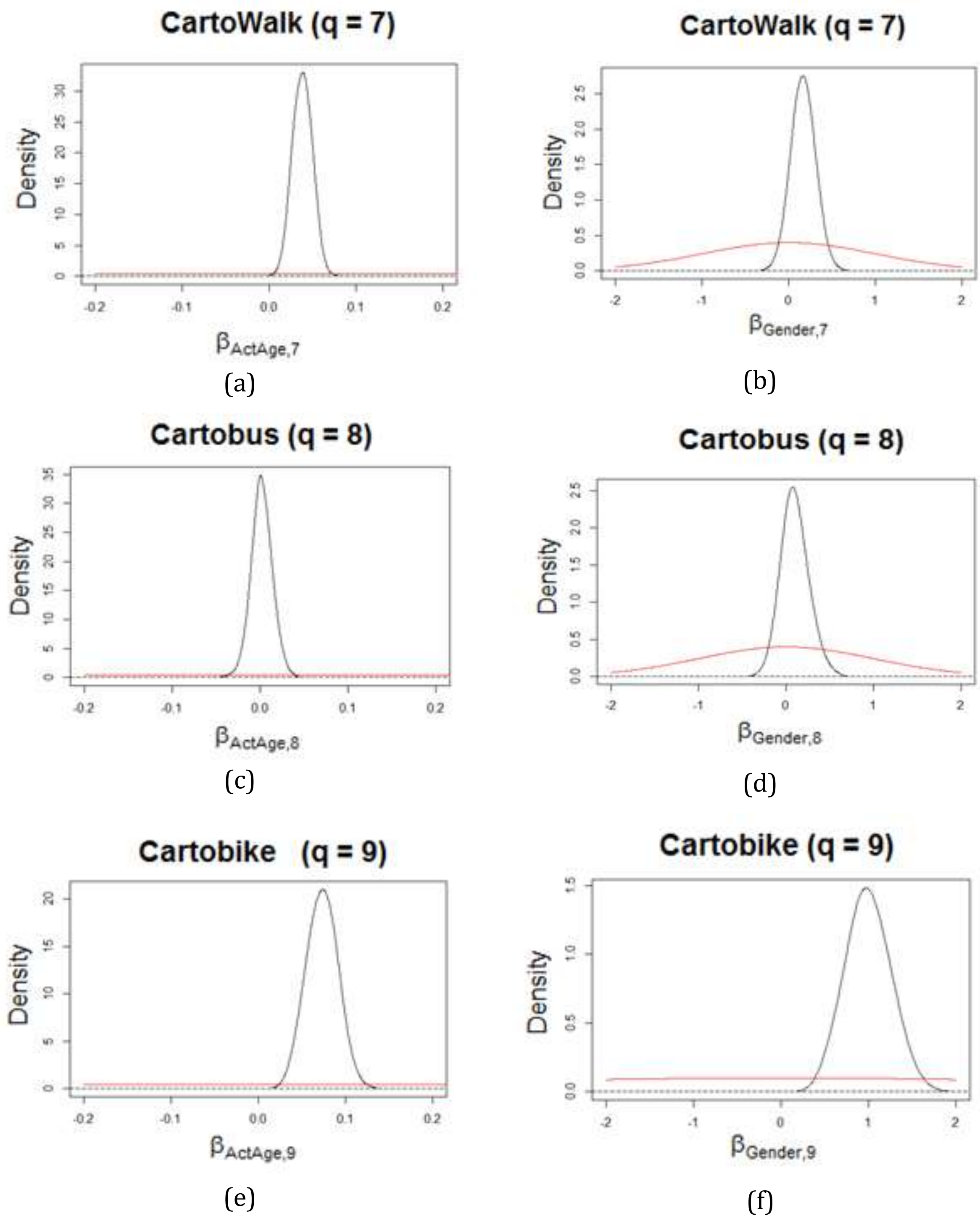


Figure 8.6: Posterior (black) and prior (red) distributions of q = 7, 8, 9 according to β_{ActAge} and β_{Gender}.

Posterior distributions can be computed for other quantities in order to explore what the posterior distributions tell us about questions of interest. Figure 8.7 shows the prior and posterior distributions for the difference in effect between car-ownership groups 1 (one car) and 4 (four or more cars) on q6 (views on climate change and causes). There is little evidence of a difference, but this is to be expected since there were very few respondents in group 4.

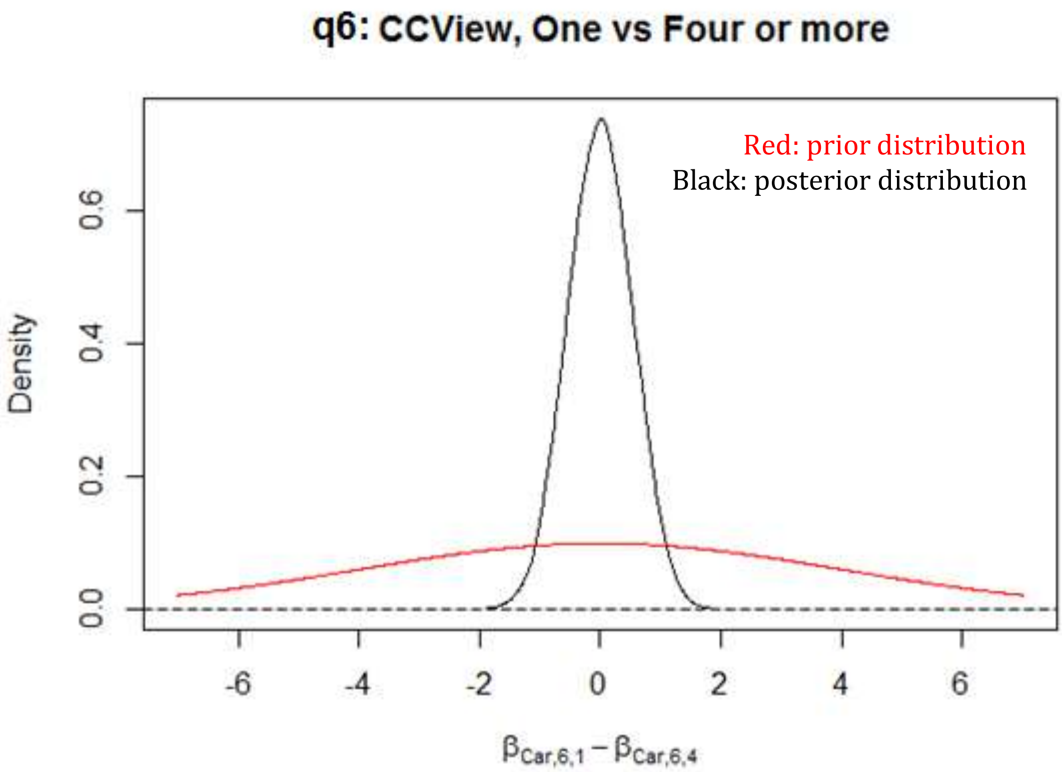


Figure 8.7: Probability density function for $\beta_{Car,6,1} - \beta_{Car,6,4}$, belonging to group with car ownership "one" compared to "four or more" for variable "CCView"

Similarly, Figure 8.8 refers to the difference between car-ownership groups 1 (one car) and 4 (four or more cars) with respect to $q4$ (Next time I buy a car, I would be willing to buy a car with lower CO_2 emissions). Here the evidence suggests a negative difference, proposing a greater tendency to agree by group 1 (one car owning household) compared to group 4 (four or more car owning household) to buy a low emission vehicle in the next purchase.

When $q4$, the willingness to buy a car with lower CO_2 emissions in the future between $\beta_{Car,4,1}$(one car ownership) with $\beta_{Car,4,4}$ (four or more car ownership) is compared, the posterior distribution shows that the difference is almost certainly negative (see Figure 8.8). These results suggest that respondents who have one car in the household are more likely to answer by agreeing to take this action in the future than those who have four or more cars. Moreover, the statistics show that over 70% of respondents (n=1509) were willing to reduce the impact of climate change and environmental problems. Not only were they willing to reduce the amount of travel by car but were also willing to reduce how much they used cars for the sake of the environment.

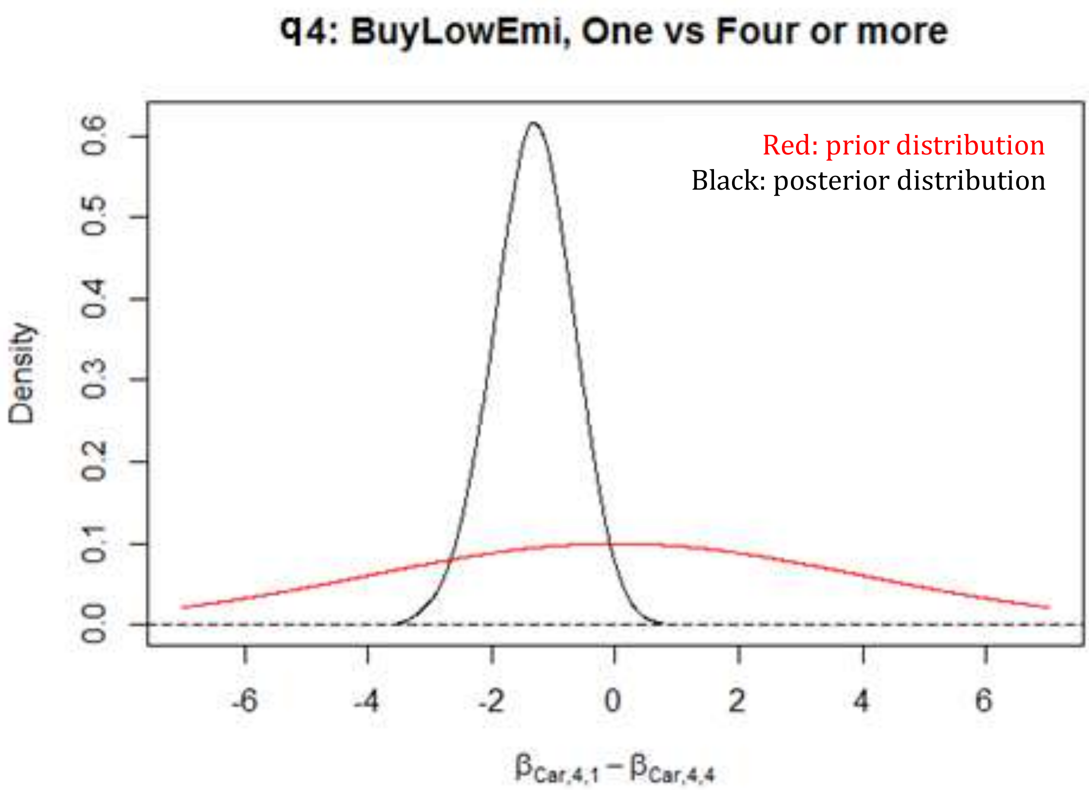


Figure 8.8: Probability density function for $\beta_{Car,4,1} - \beta_{Car,4,4}$, comparing the groups with car ownership "one" and "four or more" for variable "BuyLowEmi".

In terms of the employment status variable, Figure 8.9 refers to the difference between respondents who were in full-time employment and respondents who were in no employment (u = 2), either they were looking after the home or in retirement, with respect to $q6$ (view of climate change and causes). The results indicate a positive difference, suggesting a larger trend to believe that climate change is taking place and is, at least partly, a result of human actions in the group of respondents in full-time employment compared to respondents not in employment.

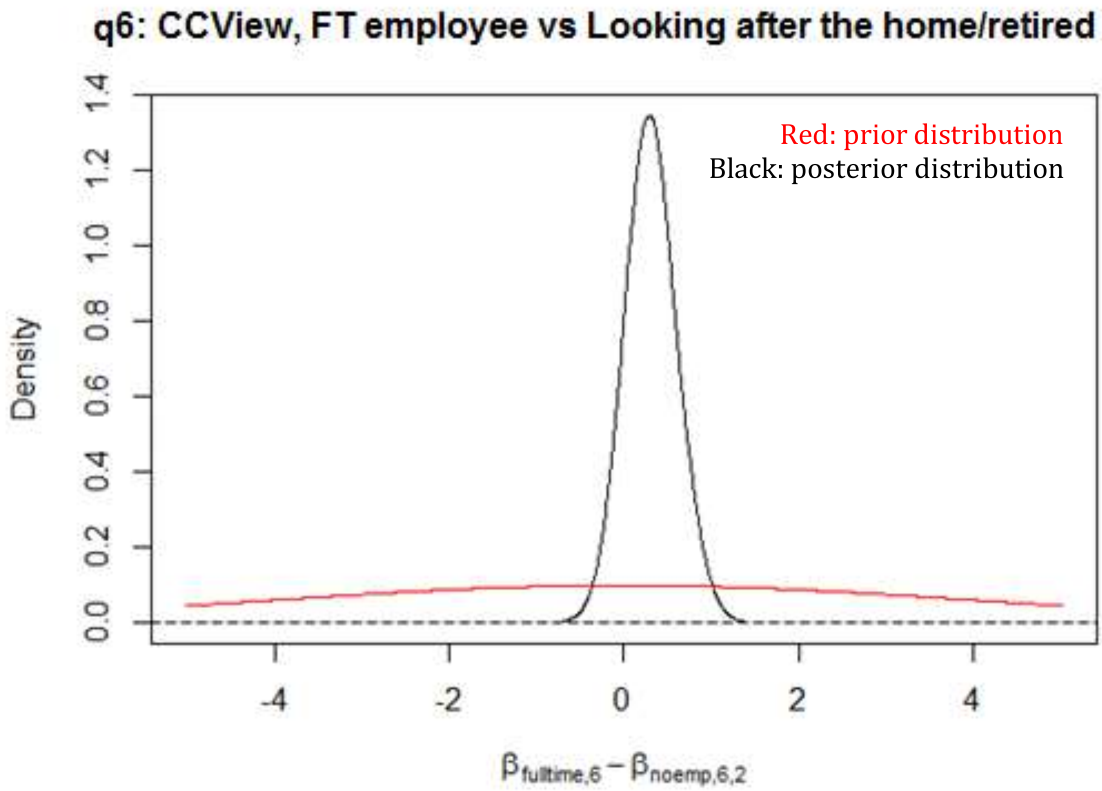


Figure 8.9: Probability density function for $\beta_{fulltime,6} - \beta_{noemp,6,2}$, "full-time employee" compared to "non-employee (u = 2)" for variable "CCView".

In a comparison between two different questions, Figure 8.10 refers to the difference between responses to $q12$ "For the sake of the environment everyone should reduce how much they use cars" and $q13$ "There is no point in reducing my car use to help the environment unless others do the same" in relation to age. The results suggest a positive difference, suggesting a larger tendency to disagree among older respondents toward $q12$ compared to $q13$.

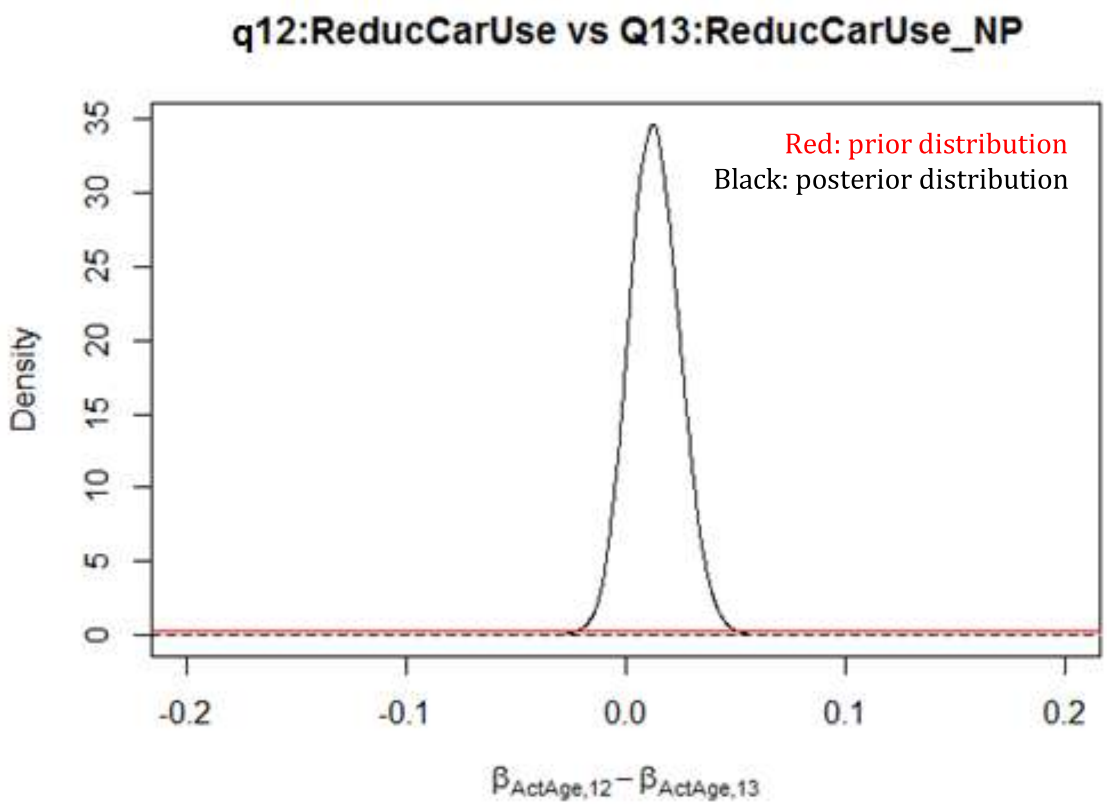


Figure 8.10: Probability density function for $\beta_{ActAge,12} - \beta_{ActAge,13}$, "q12: ReducCarUse" compared to "q13: ReducCarUse_NP" for variable "ActAge".

In order to explore further what the posterior distribution indicates in terms of the relationship of the responses to covariates, the changes in the proportion of respondents giving a particular response was examined when the value of a covariate was changed while holding other covariate values constant. The population of respondents sharing a particular covariate profile was considered. The posterior mean of the proportion of such people who would respond "agree" to a particular question is the same as the posterior predictive probability that a randomly chosen person from this group would respond "agree". For example, to investigate the relationship with age, the other covariates can be fixed, for example a person who is male, in work, in full time employement, has 1 car and household size of 1. Then, the summaries of the posterior distribution were computed for the "agree" proportion for such people aged 20, 25, 30, ..., 70. The model for $\mu_{i,q}$ contains no interaction terms so, although the overall level depends on the choice of values for the other covariates, the dependence on age does not. The posterior median (blue solid line) against age, along with the 2.5% and 97.5% points (red dashed lines) giving a 95% credible interval, were constructed for each question.

For each question, the next step was to explore the effect on age using a) linear, b) quadratic and c) cubic relationships with the linear predictor of the probability of that particular group agreeing with changing from 20-70 years. Figure a) in Appendix C clearly shows that, for each question, the relationship with age is quite smooth monotonically increasing, decreasing or remaining constant. This is because the model has a simple linear dependence of the transformed probability on age.

Probability against age demonstrates an increasing trend from younger to older respondents to questions 2, 4, 12, 13, and 14. These attitudes are towards the level of traffic congestion problem in towns and cities, the willingness to buy a car with lower CO_2 emissions in the future, the environmental awareness towards reducing car use, perceptions of "no point in reducing car use unless others do the same", and "people who drive cars which are better for the environment should pay less to use roads". Whilst responses to $q7$ (CartoWalk), $q9$ (CartoBike) and $q11$ (AllowCarUse) revealed a decreasing trend, the pattern is constant for the other questions.

When the age effect was allowed to be quadratic Figure b) in Appendix C, some of the questions exhibited distinct curves, for example $q2$ (Cong_cities), $q6$ (CCView) and $q7$ (CartoWalk) where the probability of agreeing increases from younger to middle aged, and then decreases to older aged. For congested cities it may be associated with the fact that the younger and older groups have lower levels of disposable income and, therefore, travel fewer urban miles. Next, the age effect was allowed to be cubic Figure c) (Appendix C). Some of the questions show flexibility, for example $q9$ (CartoBike).

Obviously, the shape of the curve changes between linear, quadratic and cubic in a systematic way. However, considering the 95% credible interval, questions 1, 2, 3, 5, 8 and 10 display no important differences between linear, quadratic and cubic. The graph illustrates a clear difference for $q9$ (Many of the short journeys that I now make by car I could just as easily cycle). The linear has a similar shape to the cubic suggesting the younger ages are more likely to agree to mode shift compared to the older group. However, by observing the quadratic plot, it is the middle-aged cohorts that tend to be more likely to agree compared with both younger and older adults.

However, the shapes for linear and cubic tend to provide a similar interpretation with the younger generation are more (or less depending on the question) likely to agree than the older group. Consistently, the quadratic has a shape that reflects the middle-aged group systematically more or less likely to agree with the questions.

Overarching conclusions of this sensitivity testing show that, in the majority of questions, the addition of quadratic and cubic terms makes little difference. However, when the choice of relationship is important it does have a significant influence on the interpretation of the results. Further research is needed to fully understand the age effect.

8.4 Conclusions

An analysis of the relationship between attitudes to travel and travel behaviour in relation to environmental issues was reported in Chapter 7, where 5 clusters were observed within 3 factors obtained using MLR. This chapter subsequently presented a wider and more in-depth analysis of the categorical variables using log-linear and multivariate probit models, aiming to characterise the target groups according to mathematical models shown to be significant and to quantify their potential to shift modes from private to sustainable transport.

Overall, this chapter demonstrates how the novel approach presented here assists considerably in understanding the perceptions of individuals regarding environmental issues and their potential to switch modes. The results of the study have demonstrated that fitting a MPM using Bayesian inference is both a practical and effective way to analyse ordinal survey data.

While the models could be fitted using frequentist inferential methods, such as maximum likelihood estimation, this would not be straightforward and would require special algorithms such as expectation-maximisation (EM) because of the presence of the latent random variables which characterize individuals (Dempster *et al*., 1977). In contrast, by using Bayesian software such as rjags, the need for complicated coding is avoided.

Furthermore, frequentist methods such as maximum likelihood estimation lead to point estimates of model parameters and do not allow the computation of quantities such as predictive probabilities which allow for all sources of uncertainty, including both random differences, or sampling variation, between individuals as well as uncertainty in the values of model parameters. The Bayesian approach, with the help of modern software such as rjags, makes this relatively straightforward. The approach makes possible the computation of predictive probabilities for any combination of responses to the fourteen questions by a hypothetical future respondent with a given set of sociodemographic characteristics.

By using software such as rjags, the need for complicated coding is avoided. This is because the model includes correlations between responses and makes possible the computation of predictive probabilities for any combination of responses to the fourteen questions by a hypothetical future respondent with a given set of socio-demographic characteristics.

This study successfully characterized sections of the population who increasingly demonstrated a higher propensity to take action to adjust their mode choice or activity to address environmental issues. The results indicate that significant differences exist among different age groups and that gender, employment status, household number, and car ownership are influential and should also be taken into account.

Taking the seventh to ninth questions, the majority of respondents had not considered changing mode and there is surprisingly slight variation from this attitude assessed against socio-demographic indicators: non-workers and younger respondents (aged 18-24 years) together with older groups (aged 50+ years) appear to be least likely to have considered a change. This is also the case for those in the full-time employment group. Investigation in terms of travel characteristics does, however, indicate some variation from the routine.

There is some suggestion that susceptibility to change appears to be linked to age. Those in younger and older categories are the least likely to be susceptible to change, whilst those in their middle-age are the most susceptible. Females in full-time employment are undifferentiated from males in full-time employment in terms of their susceptibility to change. Mode shift potential to buses was similar for males and females.

It should be noted that, in this study, the females (<50 years) were found to be more aware and to show more willingness to switch modes for short journeys of less than 2 miles. This finding is contrary to one previous study (Curtis and Headicar, 1997), which suggested that young travellers, who are more likely to be males in their 30s making short commuting journeys of 5 miles or less, are a minority group susceptible to mode change. The results in this chapter indicate that some people think that they may need to change their behaviour to become more sustainable, whilst others do not, and this stance varies according to their socio-demographic background. The next chapter, therefore, moves on to discuss the conclusions and recommendations of the study.

Chapter 9 Discussions, Conclusions and Recommendations

9.1 Introduction

This thesis has set out the aims, objectives and methodology developed to analyse the BSA dataset from 2011-2014 comprehensively. Several statistical approaches were adopted to provide different perspectives on the data.

In this chapter, an overview and discussions of the results emerging from the descriptive analysis, factor analysis, multiple correspondence analysis, cluster analysis, multinomial logistic regression and Bayesian inference approach is given in Section 9.2 followed by Sections 9.3, 9.4, 9.5, 9.6 and 9.7 which address main findings, secondary findings, limitation, policy implications and, finally, directions for future research are suggested.

9.2 Overview of Results

Sustainable mobility targets, the question "What are the characteristics of the population who may switch from private transport to sustainable modes?", have been discussed in many studies (Stradling *et al.*, 2000; Hull, 2008; Schneider, 2013; Okushima, 2015; Zacharias and X.Li, 2016). General concerns include environmental and climate change issues related to transportation which have attracted worldwide attention. Although some studies attempt to discuss and suggest the best solutions (Nilsson and Küller, 2000; Anable, 2005; Steg and Gifford, 2005; Anable *et al.*, 2006; Barr *et al.*, 2011; Banister *et al.*, 2012; Mikiki and Papaioannou, 2012; Susilo *et al.*, 2012; Kamruzzaman *et al.*, 2016), the need for further research is commonly recognised, in particular the need to develop a more reliable method and more statistically sound evidence to support findings (Hickman and Banister, 2007; Rayner *et al.*, 2008; Santos *et al.*, 2010; Thomas and Walker, 2015).

Therefore, the research presented in this thesis aims to provide an insight into which target groups of car users are more likely to switch from private transport to sustainable modes so as to provide scientific evidence to LAs decision and policy makers responsible for the design of marketing strategies and new green transportation schemes. In the UK, there are mandatory targets set at 67% reduction in CO_2 by 2050 over 2010 levels and current research (DfT, 2018) shows that the trend of CO_2 and other GHG emissions has been predicted to increase over time. Therefore, it is argued that this knowledge could play an essential role in targeting sustainable policy initiatives in the future and making better use of public money.

The aim of this research was inspired by the findings of three studies:

1) Research into targeted travel awareness campaigns, exploring which car commuters are likely to be the best targets for promoting non-car modes for the journey to work, has found that males in their 30s were more likely to change their travel behaviour (Curtis and Headicar, 1997).

2) There is evidence to suggest that, in terms of attitudes toward and behaviours in relation to the environment, almost all respondents were aware of environmental issues, but their views did not necessarily 'match' their travel behaviour (Susilo *et al.*, 2012) and;

3) and some indication that combining factor and cluster analyses of data on attitudes towards certain modes of transport and attitudes towards the environment and sustainability, on the other hand, enriches explanatory models for individual travel behaviour and delivers helpful additional information for potential policy and planning measures (Prillwitz and Barr, 2011).

In this study, the perceptions of and attitudes towards travel, and travel behaviour of respondents were investigated using BSA datasets from 2011 to 2014. A sizeable proportion of respondents did not drive or own a vehicle and, given that this research focussed on car travel, this represented redundancy in the datasets. Therefore, such data were removed as it is important to consider only the attitudes of the respondents' cohort who were drivers or passengers in a car to meet the study objectives. The data from respondents were grouped into three types; namely, socio-demographic characteristics, travel behaviour, and attitudinal data based on the completed questionnaire only. However, cluster analysis was conducted to assign respondents into groups based on demographic characteristics and travel behaviour, whilst factor analysis was conducted to reduce the number of descriptors by identifying interrelated variables using attitudinal variables. Relationships between the attitudinal factors obtained from EFA for different clusters of travellers were investigated using multinomial logistic regression analysis. Multivariate probit modelling, using Bayesian inference, was then conducted to model all 14 attitudinal variables, allowing the investigation of relationships between the responses of individuals to different questions, where each response is considered to be a category on an ordinal scale.

Traditionally in the investigation of travel behaviour, socio-demographic characteristics have been relied upon as correlates with behaviour. Similarly, attitudes, preferences and beliefs have been found to be dependent on such characteristics as gender and age. However, in order to prove or disprove the hypothesis that any changes in attitudes and differences in travel behaviour could simply be attributed to personal characteristics, it is necessary to investigate the demographic structure of samples along with the analytical method developed. This represents the originality in this research.

This thesis therefore contributes to the sustainable transport planning field in three respects:

1. ***Analysis of travel behaviour conducted from the impacts of travel attitudinal variables***. This research presented a set of findings concerning travel behaviour with respect to travel attitudinal variables. Five groups were clustered based on their socio-demographic and travel behaviour patterns. Incorporating these clusters into the analysis of travel behaviour significantly increased the explanatory influence of the travel mode choice as well as the future trends when the demographic characteristics change.

2. ***Methods for analysing travel attitudes and behaviour were investigated***. This research used multinomial logistic regression to investigate the relationships between key attitudinal factors among respondents clustered into several distinct travel groups to explain their behaviour; and Bayesian inference was used to measure predictive probabilities for any combination of responses to all

attitudinal variables by a hypothetical future respondent with a given set of socio-demographic characteristics in a more comprehensive way.

3. ***Recommendations on the propensity of groups of individuals switching from private transport to sustainable alternatives***. Building on a critical review of previous work on travel behaviour and attitudes toward environmental issues, in-depth analysis characterised target groups providing evidence to LA to better inform, more cost effective and new developments in sustainable transport planning policy.

9.3 Main Findings

The main findings that can be drawn from the study are shown in Figure 9.1 mapped into the chapters (and therefore the analysis technique) from which the results emerged. The research question posed in this thesis was:

1. *What are the attitudinal factors that characterise the uses of sustainable transport modes?*
2. *What are the key socio-demographic variables that affect travel mode choices and decisions within the cohort of car users?*
3. *Which car users' groups should be targeted in campaigns that promote the uptake of non-car transport alternatives?*
4. *Can key factors be used to derive a model to predict the likely uptake of sustainable modes?*

PRELIMENARY DATA ANALYSIS

Ch. 5

Descriptive Analysis (DA)

1. Car users in the BSA dataset who participated in the questionnaires survey and interviews are representative of all car users in the sample (2011–2014) – proven by χ^2 test.
2. Younger-aged (18-24 years old) car users were the lowest proportion of participants in this study (3.2%) as expected.
3. Normality test showed that in all cases the distributions of responses were not normally distributed at 95% statistical confidence. Non-parametric tests were used throughout the analysis.
4. Car users acknowledged traffic congestion on motorways, in towns and cities as a serious problem.
5. Car users were aware of the importance of reducing CO_2 emissions and very keen to buy a lower emissions vehicle in the future.
6. 35%, 43% and 56% car users were reluctant to change behaviour from car to walking, cycling or use local buses respectively.
7. Charging taxes is definitely not something that car users were in favour of.

MAIN DATA ANALYSIS

Ch. 6

Factor Analysis (FA)

1. With 14 attitudinal variables, 3 factors and 4 factors emerged from PAF and PCA respectively.
2. Provided that correlation and covariance between factors is considered to represent how strongly two factors are related PAF performs better than PCA.
3. In-depth analysis of the characteristics of the factors revealed interrelationships between travel choices, awareness of the environment and demographics.
4. The 3 factors; namely, attitudes to transport and environment, traffic awareness, and modal shift potential were then used in MLR.

Ch. 7

MCA, HCA and MLR

1. Two cohorts emerged from MCA graphical visualisation was labelled as "Socio-demographics" and "Male and female car dependency" for dimensions 1 and 2, respectively.
2. Graphical representation of the MCA analysis enhanced the overall view of the dataset, including how individuals are distributed among the categories, according to responses to each question.
3. 5 clusters emerged from HCA: C1 (M, 65+, retired, 2 in HH and 1 car), C2 (M, 35-44yr, FT, 2 in HH and 2 cars), C3 (M, 45-54yr, FT, 2 in HH and 1 car), C4 (F, 45-54yr, FT, ≥4 in HH and 2 cars) and C5 (F, 45-54yr, looking after the home, ≥4 in HH and 1 car).
4. Cluster 2 represent groups of respondents who have moderate consumption of travel and dependency on sustainable transport.
5. MLR analysis identified that the respondents in Clusters 2 were statistically significant to "modal shift potential".

Ch. 8

Bayesian Inference

1. Fitting a multivariate probit model using Bayesian inference is a practical and effective way to analyse ordinal survey data.
2. Log-linear models was developed dealing with one question at a time. MPM model consider all responses to all attitudinal variables in a single model was then developed.
3. The investigation of whether relationships were linear or polynomial suggested that a cubic was more appropriate.
4. Susceptibility to change appears to be linked to age. Younger and older groups are the least likely to be susceptible to change, whilst those in their middle-age are the most susceptible.
5. Older adults were more likely to disagree for mode shift to walk, using PT, and cycling.
6. Mode shift potential to buses was similar for males and females.
7. One car owning household (group 1) had greater tendency to agree to buy a low emission vehicles compared to four or more car owning household (group 4).
8. Respondents in full-time employment as opposed to unemployed tend to disbelieve that climate change is taking place, partly, a result of human actions.

Figure 9.1: Results diagram

The key messages answering the research questions posed in Chapter 1 were as follows:

1. The descriptive analysis of the BSA dataset from 2011 to 2014 provided evidence of high proportions of people using their cars predominantly in daily travel activities. Through the analysis of travel attitude factors, the details regarding how people travel, how they actually behave, and their levels of environmental awareness and willingness to take action for the sake of the environment were identified. The three key factors characterising the attitudes of car users that emerged from the factor analysis were: F1, attitudes to transport and the environment; F2, traffic awareness; and F3, modal shift potential.

2. The findings from cluster analysis demonstrate that, whilst the majority of respondents are strongly car-orientated, a statistically significant minority who were susceptible to change were found in Cluster 2. These were middle-aged male (35-44 years old) in full-time employment. Cluster 2 represents respondents who have moderate consumption of travel and dependency on sustainable transport and were found to have higher environmental awareness compared to other clustered groups of respondents. Therefore, this cohort was identified as appropriate car users who could be targeted with initiatives to promote sustainable travel.

3. The investigation of the relationships between attitudes to travel and travel behaviour in relation to environmental issues, conducted using MLR analysis, found that respondents in Cluster 2 in 2011 showed no significant causal link with "modal shift potential" (F3). However, by 2013 this cluster had begun to accept the need for action to switch travel modes from private to sustainable modes for short journeys of 2 miles or less. Consistent with cluster analysis, this group was estimated to be the most prone to take action to help reduce the impact of travel on climate change and were the most likely to be willing to switch travel modes.

4. By using Bayesian inference, 14 attitudinal and latent variables were successfully derived in a multivariate probit model. The prediction results from the model indicate that significant differences exist among different age and gender groups.

However, employment status, household size, and car ownership were also found to have little influence on perceptions, attitudes and travel behaviour.

9.4 Secondary Findings

Along with the main findings above, this study also provided a more in-depth understanding and secondary findings were as follows:

1. The originality of the research conducted in this thesis was identified from the literature review which highlighted the need for further research to deliver more detailed evidence concerning sustainable mobility targets and, more specifically, in modelling attitudes to travel and travel behaviour in relation to environmental issues.

2. After considering several data sources, the BSA dataset was identified as the most suitable to conduct this study given that it held data on socio-demographics, travel behaviour, and attitudinal information. Thus, the BSA presented the opportunity to investigate interrelationships between respondents' perceptions and attitudes and opinions towards environmental issues and climate change.

3. The Bayesian multivariate probit model was found to have advantages over other methods documented in the literature, as it is provides a natural and principle way of combining prior information with data.

4. Significant positive correlations were found that suggested car users find congestion on both motorways and in towns and cities a serious problem. Also, traffic congestion in towns and cities was demonstrated to be statistically significantly linked to the problem of exhaust fumes from traffic in towns and cities. These confirm the existence of relationships between variants aspects of traffic awareness.

5. Statistical evidence showed that many of the short journeys made by car could just as easily be made by walking, cycling or bus. This affirmative relationship suggests that modal shift would be acceptable.

6. The respondents considered cars to be the largest transport contributors to UK climate change overall, followed by aeroplanes, vans and lorries, buses and coaches, ships and ferries, and trains, whilst motorbikes were assumed to be the mode least responsible.

7. The majority of the respondents claimed to believe that climate change is happening, and 81.7% of the population were convinced that climate change is linked to human activity. However, it is unclear whether the concern for climate change is currently rising or falling, due to the way the data was collected.

8. Even though the respondents reported that they were willing to take action in response to climate change and environmental problems, an inconsistency existed with their actual behaviour.

9. The results suggest that car dependency, which the government is hoping to tackle, is particularly strong among males and females in full-time employment and those aged 35-54 years old, whilst the results indicate that retired males over the age of 65 years never used public transport such as buses and trains, and never cycled. Hence they can be considered reluctant sustainable mode users. Given that only car users were studied in this thesis, it should be noted that potentially these represent senior citizens with a member of the household with a disability, but this is only conjecture.

10. The one car-owning group displayed significant differences compared to the four or more car-owning group to the extent that they exhibited a willingness to buy a car with lower CO_2 emissions, felt responsible for the environmental effects of their car use and perceived behavioural control over using alternatives.

11. It seems reasonable to suggest from the evidence that a significant minority of the population express a desire to lead a 'greener' lifestyle, and some seem prepared to take action that lessens their impact on the environment.

9.5 Limitations of the Study

The research reported here has drawn upon data collected from the British Social Attitudes (BSA) survey. Caution must be exercised when applying these results in practice. It should be noted that the study was conducted in Great Britain, and therefore the characteristics of the respondents are likely to be different from those in other countries. The issues considered of importance in this study, for example, may be less important in other countries or vice versa, and issues not highlighted here may be of importance in other studies. However, the methodological approach for the systematic analysis of the data is easily transferable. Therefore, scope exists for adopting this analytic procedure and methodology, and repeating the research in other countries on different timescales and with a wider range of variables. Also caution should be exercised given the potential under-representation of younger-aged respondents (18-24 years old) in this study.

Due to the limitations of the BSA data being third party, this study does not consider location (origin–destination). Analysis of the secondary data concerned with attitudinal variables, however, has enabled a detailed understanding of the travel patterns and the relevant socio-demographic backgrounds of respondents (in this case age, gender, household size and car ownership).

Another limitation of this study is that the interviewer did not directly observe the respondents' opinions. Therefore, very little genuine understanding can be gained of how people deal with the complexities of their knowledge of climate change, or how they might deal with it in different information environments.

Difficulties have arisen when attempting to compare the perceptions of car users and non-car users in this study. As mentioned before, only car users were considered in this study. Therefore, differences in attitudes between drivers and non-drivers in this study were not investigated even though they may have very different views on congestion and environmental problems. The lack of knowledge of the perceptions and attitudes of non-drivers in the sample means that further caution is urged regarding the generalisability of these findings.

9.6 Policy Implications of the Study

This study aimed to offer recommendations for evidence-based policy in the future. However, in applying the recommendations made in this section in practice, great care should be taken, as the study has certain limitations as discussed in section 9.5. A large proportion of respondents reported that they would be willing to buy cars with lower CO_2 emissions in the future in order to help reduce the impact of climate change. This is a potential market for car companies to produce such cars and encourage more people to use cars with lower CO_2 emissions for the sake of the environment. From the MPM, a greater tendency to agree to take this action was found among respondents with one car per household compared to respondents with four or more cars per household. This finding challenges car companies to be innovative in their manufacture in response to current issues concerning the environment.

Another aspect that needs more attention from LA is to seriously consider alternatives to higher taxes (road users charging) for private transport and instead lower ticket prices for public transport and other sustainable incentive schemes. Eriksson *et al.* (2008) and Redman *et al.* (2013) present some evidence to suggest that travellers expect to pay less for tickets to encourage them to use train services more. Similarly, the research reported in this thesis found that a small increment in taxes is less acceptable. 73% of those who were interviewed disagreed that for the sake of the environment car users should pay higher taxes. On the other hand, more than half (56%) of respondents agreed that there is no point in reducing their car use to help the environment unless others do the same. Taken together, these statements suggest that there is an association between attitudes towards environmental issues and transport policies.

A positive relationship was found in this study between agreement that people should be allowed to use their cars as much as they like, even if it causes damage to the environment, and agreement that there is no point in reducing car use to help the environment unless others do the same. Those who have negative perceptions on the environmental issue tended also to reinforce the same negative opinions as others and yet they deny the importance of reducing car use for the sake of the environment.

Age and gender were found to be the most influential attributes of modal shift potential. Middle-aged and mature males in full-time employment have the highest propensity to change to sustainable mode and elderly males in retirement and young people in no employment had the lowest, as expected. Since all of the respondents sampled were car users, the analysis concentrated on assessing their susceptibility to change mode, as from a policy perspective it is this group who need to be targeted most to maximise chances of success. The results that emerged are of value to those with responsibility for managing and marketing travel awareness campaigns. Therefore, when the concept of the log-linear relationship is used in estimating categorical variables, target segmentation of car drivers based on their age and gender has to be considered. Based on the evidence obtained from this study, older groups were found to rely more on cars to travel, and therefore this challenges LAs to work towards understanding and accommodating their needs in areas such as improving accessibility and comfort, and thus to encourage older people to use public transport. It is important to note that policy needs to be appropriately aligned with the different characteristics of travellers to affect travel behaviour choices, as well as to create realistic incentives.

From the data collected, it can be recommended that policy should emphasise methods of making walking and cycling both easier and safer so that it can be more conveniently tailored to household routines. Most interventions to increase cycling, in particular, focus on infrastructure, whilst improvements in infrastructure for walking and cycling may include the removal of barriers and parked vehicles on pavements, better maintenance of pavements and the repair of uneven paving stones. The provision of dedicated cycle routes that are segregated from traffic is also important, and may make walking and cycling more acceptable for all household members. However, such strategies alone are unlikely to be sufficient to effect major change. Much more problematically, it can be argued that policies to increase walking and cycling do not require transport solutions but, rather, need more fundamental changes in society and urban structures that allow more flexibility in how and when people travel, so that walking and cycling can be more easily fitted into household routines. Furthermore, the LAs and policymakers might gain valuable information from this study concerning whether to invest more money to provide more facilities to encourage the positive use of sustainable transport or to invest in public transport services.

Finally, implicit in much of the above discussion is the suggestion that greater rates of walking and cycling will only be achieved when car use becomes significantly more costly and less convenient. Therefore, an integrated approach which initiates traffic capacity reducing policies simultaneously with incentives addressing and overcoming household (and other) constraints.

9.7 Future Research

Whilst this research has successfully contributed new knowledge, as with all research, there remain opportunities for further investigation, and potential areas would include modification to the questionnaire and changes in the methods of data collections with further analysis to demonstrate transferability of the method to obtain more robust results.

This study should be repeated using the collection of primary data rather than using secondary data collected by a third party. This should be achieved through questionnaires or face-to-face interviews. However, this would require more manpower and, as a result, increased costs and time consumed processing the data. A combination of direct observation and face-to-face interviews may allow a useful comparative assessment of car and non-car users. Furthermore, insights of respondents provided would have enhanced the understanding of the interaction between travel behaviour, perceptions, and attitudes towards switching travel modes.

It might also be useful to conduct further investigation into the possibility of targeted studies of campaigns aimed at mode shift to sustainable transport in collaboration with employers, since some institutional change is also required. A large group of respondents in the present study claimed that work circumstances were a reason why they could not change mode, but even here there may be a role for policies in addressing the responsibilities of employers in aiding a shift from car dependency. The findings would suggest reviewing the potential for altering work hours and existing car subsidies (company cars) and reconsidering the need for a car for business purposes every day. The extent to which travel behaviour is affected by this measurement remains difficult to measure and will benefit from further research.

In the factor analysis carried out in this study, the value of Cronbach's alpha (calculated to consider the internal consistency of the grouped statements) was 0.16 for the first factor. This is very low to be considered reliable for the factors identified in the PAF. Peterson (1994) indicated that acceptable alpha scores range from 0.5 for a preliminary analysis to 0.9 for applied research. The possible reasons behind the small value include a large number of questions, weak relationships between them, or heterogeneous constructs (Loewenthal, 2001; Johnson and Wichern, 2007). Therefore, it is recommended that more advanced statistical methods and more datasets with supplementary questions should be considered in further analysis of such data in order to obtain a better model fit.

The multivariate probit model presented in this study used $C = 3$, which combined the responses into categories of 0, 1, and 2, for example by combining "Agree" with "Strongly agree" and "Disagree" with "Strongly disagree", so that there were no unknown threshold parameters. In future work, this restriction should be relaxed and allow $C > 3$. Because the model includes correlations between responses, the computation of predictive probabilities is made possible for any combination of responses to the fourteen questions by a hypothetical future respondent with a given set of socio-demographic characteristics. In future work this could be exploited to investigate the characteristics of people most likely to give particular kinds of responses to the whole collection of questions rather than a subset as presented in this thesis.

Furthermore, the application of these methods (to investigate attitudes and travel behaviour patterns and whether or not similarities in attitudes and behaviour exist) can also be applied to similar research in developing countries, where traffic congestion and environmental problems are often greater compared to the UK. Although the results presented in this study refer to the population of people in Great Britain, they illustrate patterns of behaviour that can be investigated for other countries. However, future research should include income as one of the key socio-demographic variables considered, since information on income would help to establish a greater degree of understanding of how disposable income influences travel choices or car ownership since this has been out of the scope of this study.

Future research could further the understanding of why attitudes do not always translate into action and to discover fundamental understanding of the feelings that need to be considered and addressed in order to change attitudes and influence behaviour in preparation for more targeted, community-level sustainability campaigns. In addition, barriers to change differ for different travel behaviours, for different segments of the population, as well as how these variables interact along with their dynamic feedback effects. These all remain important priorities for further research.

All in all, it can be concluded that none of the approaches that could be used in such research can claim absolute superiority. Instead, their various application have specific advantages and disadvantages, which suggests that range of applications could be applied, after careful consideration of their appropriateness, to diverse aspects of the planning and design of sustainable mobility measures.

Appendices

Appendix A

Model specification for log-linear

(a) Without age and gender effects

```
library(rjags)
q1data<-read.table("twelvegroups.txt",header = TRUE)
y<-cbind(q1data[,1:5])
n<-rowSums(y)
q1jagsdata<-list(n=n,y=y)
q1jags<-jags.model("q1bug.txt",data=q1jagsdata,n.chains=2)
update(q1jags,1000)
q1samples<-coda.samples(q1jags,c("eta"),10000)
summary(q1samples)
q1samplesout<-as.matrix(q1samples,iters=TRUE)

eta3.2<-q1samplesout[,16]
x<-seq(-2.5,2.5,0.05)
priordens<-dnorm(x,0,2.5)
plot(density(eta3.2,adj=2.5),xlim=c(-2.5,2.5),main="Males 35-44, Agree Strongly",
xlab=expression (eta["3,2"]))
lines(x,priordens,col="red")
abline(h=0,lty=2)

eta3.2<-q1samplesout[,16]
eta3.4<-q1samplesout[,40]
diff<-eta3.2-eta3.4
newpriordens<-dnorm(x,0,sqrt(12.5))
plot(density(diff,adj=1.5),xlim=c(-2.5,2.5),main="Females 35-44, Agree vs Disagree",
xlab=expression(eta["3,2"]-eta["3,4"]))
lines(x,newpriordens,col="red")
abline(h=0,lty=2)
```

```
model
{
 for (i in 1:12)
    {y[i,1:5]~dmulti(p[i,],n[i])
     for (k in 1:5)
        {p[i,k]<-phi[i,k]/sum(phi[i,])
         phi[i,k]<-exp(eta[i,k])
         }
     eta[i,3]<-0.0
     for (k in 1:2)
        {eta[i,k]~dnorm(mu[k],10)
         }
     for (k in 4:5)
        {eta[i,k]~dnorm(mu[k],10)
```

```
        }
    }
  for (k in 1:2)
    {mu[k]~dnorm(0,0.5)
    }
  for (k in 4:5)
    {mu[k]~dnorm(0,0.2)
    }
}
```

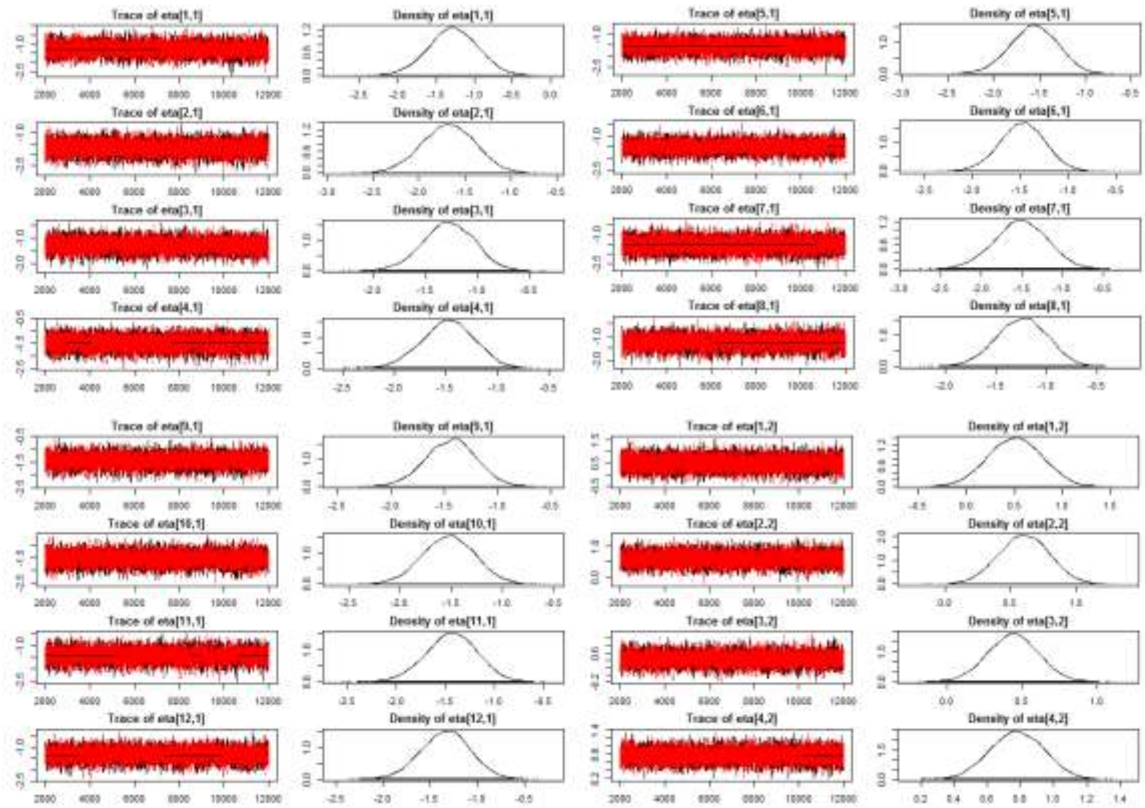
Model summary

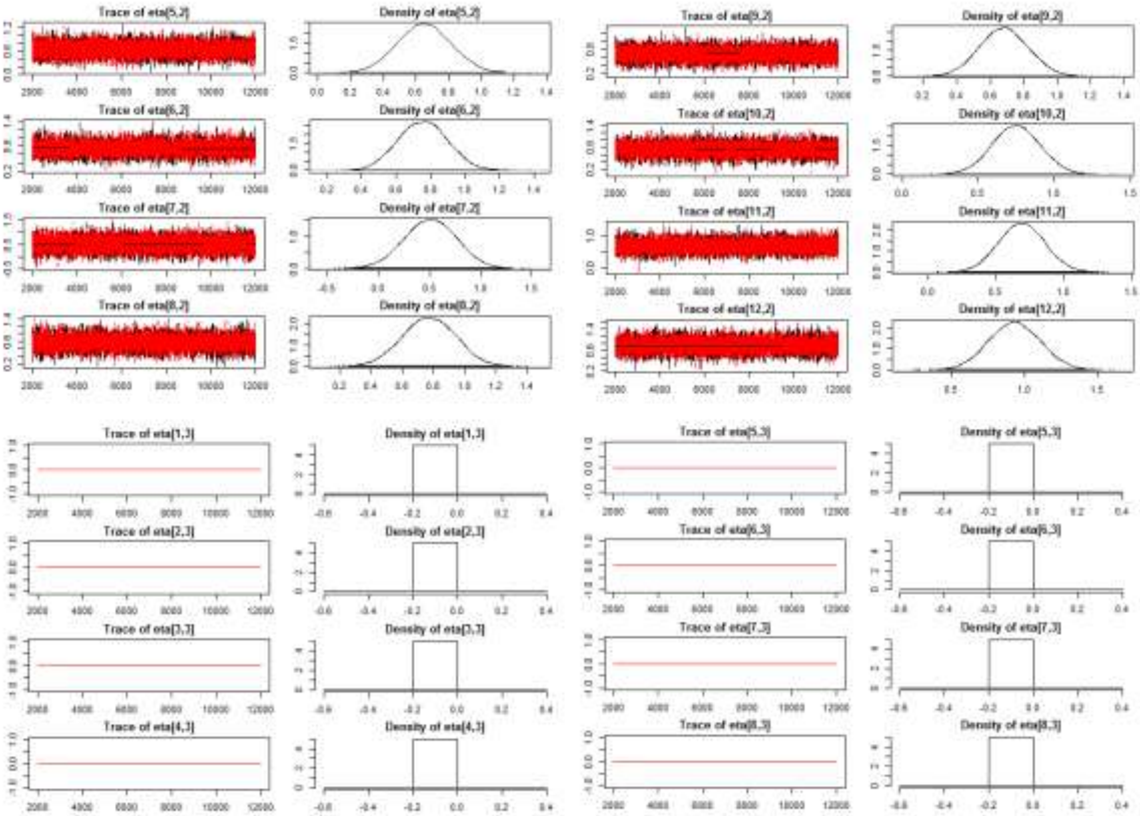
```
Iterations = 2001:12000
Thinning interval = 1
Number of chains = 2
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

              Mean     SD Naive SE Time-series SE
eta[1,1]  -1.2810 0.3233 0.002286       0.003676
eta[2,1]  -1.6804 0.2900 0.002050       0.003370
eta[3,1]  -1.2776 0.2443 0.001727       0.002926
eta[4,1]  -1.4786 0.2559 0.001809       0.003006
eta[5,1]  -1.5612 0.2568 0.001816       0.002936
eta[6,1]  -1.4792 0.2371 0.001676       0.002689
eta[7,1]  -1.5106 0.3231 0.002284       0.003831
eta[8,1]  -1.2441 0.2526 0.001786       0.002965
eta[9,1]  -1.4390 0.2362 0.001670       0.002735
eta[10,1] -1.5258 0.2540 0.001796       0.002936
eta[11,1] -1.4354 0.2615 0.001849       0.003012
eta[12,1] -1.3372 0.2650 0.001874       0.003107
eta[1,2]   0.5079 0.2789 0.001972       0.002741
eta[2,2]   0.6006 0.1922 0.001359       0.002141
eta[3,2]   0.4391 0.1686 0.001192       0.001957
eta[4,2]   0.7782 0.1614 0.001141       0.001934
eta[5,2]   0.6502 0.1612 0.001140       0.001846
eta[6,2]   0.7549 0.1440 0.001018       0.001649
eta[7,2]   0.5064 0.2540 0.001796       0.002660
eta[8,2]   0.7724 0.1707 0.001207       0.001885
eta[9,2]   0.6822 0.1446 0.001023       0.001570
eta[10,2]  0.7481 0.1565 0.001106       0.001782
eta[11,2]  0.6914 0.1689 0.001194       0.001837
eta[12,2]  0.9318 0.1758 0.001243       0.002000
eta[1,3]   0.0000 0.0000 0.000000       0.000000
eta[2,3]   0.0000 0.0000 0.000000       0.000000
eta[3,3]   0.0000 0.0000 0.000000       0.000000
eta[4,3]   0.0000 0.0000 0.000000       0.000000
eta[5,3]   0.0000 0.0000 0.000000       0.000000
eta[6,3]   0.0000 0.0000 0.000000       0.000000
eta[7,3]   0.0000 0.0000 0.000000       0.000000
eta[8,3]   0.0000 0.0000 0.000000       0.000000
eta[9,3]   0.0000 0.0000 0.000000       0.000000
eta[10,3]  0.0000 0.0000 0.000000       0.000000
eta[11,3]  0.0000 0.0000 0.000000       0.000000
eta[12,3]  0.0000 0.0000 0.000000       0.000000
eta[1,4]  -0.6607 0.3072 0.002173       0.003343
eta[2,4]  -0.5110 0.2345 0.001658       0.002515
eta[3,4]  -0.6068 0.2082 0.001472       0.002281
eta[4,4]  -0.2192 0.1936 0.001369       0.002294
eta[5,4]  -0.4863 0.1991 0.001408       0.002314
eta[6,4]  -0.4609 0.1845 0.001305       0.002051
eta[7,4]  -0.4952 0.2877 0.002034       0.003032
eta[8,4]  -0.7596 0.2275 0.001609       0.002479
eta[9,4]  -1.0911 0.2166 0.001532       0.002368
eta[10,4] -0.7719 0.2123 0.001501       0.002390
```

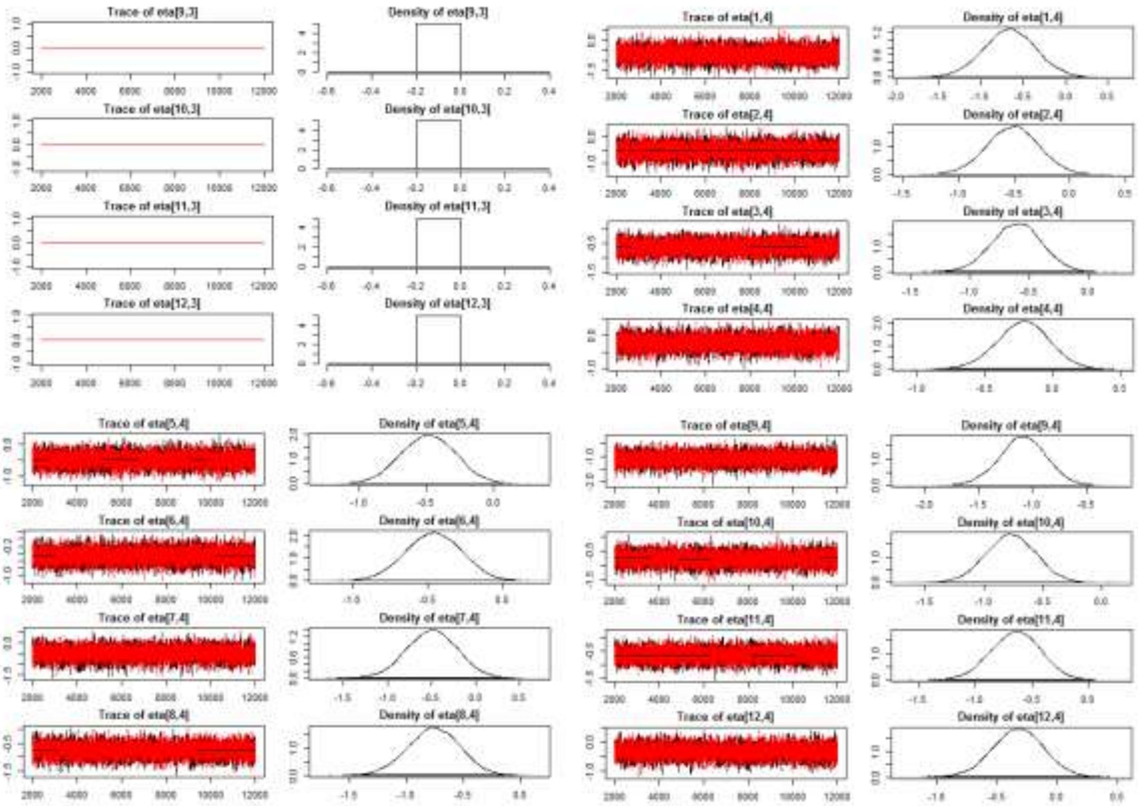
```
eta[11,4] -0.6423 0.2175 0.001538       0.002322
eta[12,4] -0.3247 0.2134 0.001509       0.002343
eta[1,5]  -1.8363 0.3422 0.002419       0.004567
eta[2,5]  -1.7615 0.2959 0.002093       0.003774
eta[3,5]  -2.0446 0.2882 0.002038       0.003740
eta[4,5]  -1.8458 0.2804 0.001983       0.003692
eta[5,5]  -1.6827 0.2691 0.001903       0.003488
eta[6,5]  -1.9140 0.2665 0.001885       0.003434
eta[7,5]  -1.9679 0.3399 0.002404       0.004709
eta[8,5]  -2.0624 0.2952 0.002087       0.003819
eta[9,5]  -2.0301 0.2721 0.001924       0.003559
eta[10,5] -2.2105 0.2917 0.002063       0.003685
eta[11,5] -2.0129 0.2932 0.002073       0.003929
eta[12,5] -2.0950 0.3061 0.002165       0.004147

2. Quantiles for each variable:

                2.5%     25%     50%      75%    97.5%
eta[1,1]  -1.9241777 -1.4946 -1.2810 -1.06121 -0.65061
eta[2,1]  -2.2561861 -1.8765 -1.6770 -1.47935 -1.12501
eta[3,1]  -1.7740718 -1.4397 -1.2752 -1.10729 -0.81128
eta[4,1]  -1.9955321 -1.6448 -1.4740 -1.30439 -0.99960
eta[5,1]  -2.0762145 -1.7321 -1.5576 -1.38418 -1.07281
eta[6,1]  -1.9589791 -1.6364 -1.4758 -1.31816 -1.02123
eta[7,1]  -2.1605865 -1.7238 -1.5092 -1.28977 -0.88875
eta[8,1]  -1.7511694 -1.4136 -1.2388 -1.07121 -0.76470
eta[9,1]  -1.9128027 -1.5970 -1.4320 -1.28015 -0.98567
eta[10,1] -2.0319766 -1.6980 -1.5237 -1.35194 -1.03084
eta[11,1] -1.9612149 -1.6066 -1.4310 -1.25870 -0.93780
eta[12,1] -1.8709396 -1.5112 -1.3313 -1.15874 -0.82787
eta[1,2]  -0.0364442  0.3175  0.5089  0.69948  1.04571
eta[2,2]   0.2238032  0.4713  0.6017  0.73123  0.97332
eta[3,2]   0.1081940  0.3262  0.4406  0.55166  0.76677
eta[4,2]   0.4654442  0.6678  0.7782  0.89059  1.08906
eta[5,2]   0.3386798  0.5398  0.6498  0.75806  0.96523
eta[6,2]   0.4781467  0.6563  0.7542  0.85124  1.03879
eta[7,2]   0.0006438  0.3359  0.5079  0.67989  0.99665
eta[8,2]   0.4364378  0.6572  0.7723  0.88901  1.10390
eta[9,2]   0.4007402  0.5843  0.6812  0.77937  0.96744
eta[10,2]  0.4407520  0.6425  0.7479  0.85303  1.05527
eta[11,2]  0.3636411  0.5769  0.6908  0.80496  1.02454
eta[12,2]  0.5893068  0.8125  0.9318  1.05128  1.27752
eta[1,3]   0.0000000  0.0000  0.0000  0.00000  0.00000
eta[2,3]   0.0000000  0.0000  0.0000  0.00000  0.00000
eta[3,3]   0.0000000  0.0000  0.0000  0.00000  0.00000
eta[4,3]   0.0000000  0.0000  0.0000  0.00000  0.00000
eta[5,3]   0.0000000  0.0000  0.0000  0.00000  0.00000
eta[6,3]   0.0000000  0.0000  0.0000  0.00000  0.00000
eta[7,3]   0.0000000  0.0000  0.0000  0.00000  0.00000
eta[8,3]   0.0000000  0.0000  0.0000  0.00000  0.00000
eta[9,3]   0.0000000  0.0000  0.0000  0.00000  0.00000
eta[10,3]  0.0000000  0.0000  0.0000  0.00000  0.00000
eta[11,3]  0.0000000  0.0000  0.0000  0.00000  0.00000
eta[12,3]  0.0000000  0.0000  0.0000  0.00000  0.00000
eta[1,4]  -1.2703038 -0.8646 -0.6601 -0.45250 -0.05857
eta[2,4]  -0.9763449 -0.6682 -0.5080 -0.35473 -0.05424
eta[3,4]  -1.0229209 -0.7443 -0.6037 -0.46669 -0.20269
eta[4,4]  -0.6103321 -0.3459 -0.2155 -0.08878  0.15571
eta[5,4]  -0.8771465 -0.6229 -0.4859 -0.34956 -0.10131
eta[6,4]  -0.8250718 -0.5862 -0.4601 -0.33482 -0.10585
eta[7,4]  -1.0655631 -0.6879 -0.4929 -0.30162  0.06565
eta[8,4]  -1.2137106 -0.9110 -0.7567 -0.60548 -0.31924
eta[9,4]  -1.5259731 -1.2314 -1.0886 -0.94569 -0.67437
eta[10,4] -1.1914746 -0.9142 -0.7720 -0.62870 -0.35202
```

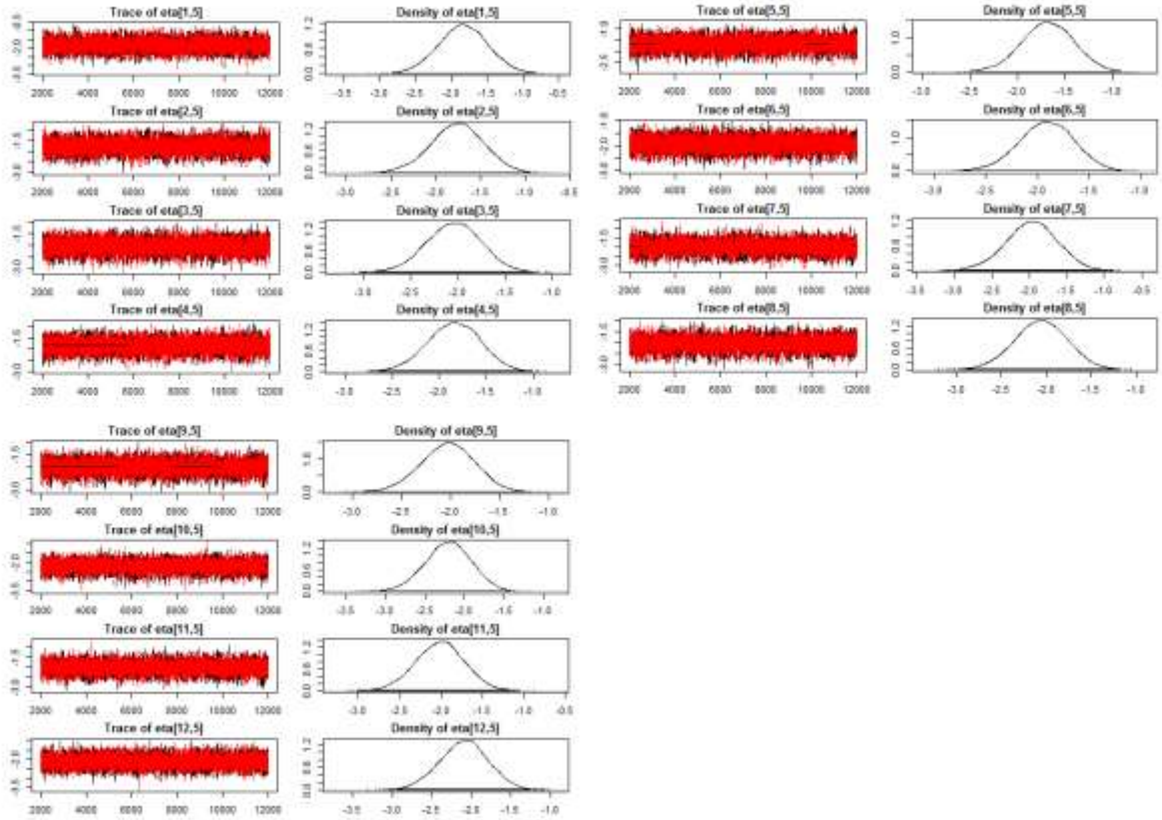
```
eta[11,4] -1.0755641 -0.7884 -0.6386 -0.49344 -0.22757
eta[12,4] -0.7530117 -0.4669 -0.3235 -0.18144  0.09294
eta[1,5]  -2.5107450 -2.0636 -1.8372 -1.60666 -1.17172
eta[2,5]  -2.3658790 -1.9570 -1.7569 -1.56312 -1.19372
eta[3,5]  -2.6210493 -2.2370 -2.0387 -1.84877 -1.48946
eta[4,5]  -2.4115569 -2.0334 -1.8408 -1.65129 -1.30987
eta[5,5]  -2.2219964 -1.8642 -1.6819 -1.49682 -1.16458
eta[6,5]  -2.4571593 -2.0889 -1.9094 -1.73131 -1.40932
eta[7,5]  -2.6604770 -2.1921 -1.9638 -1.74002 -1.32022
eta[8,5]  -2.6532379 -2.2584 -2.0605 -1.86216 -1.49194
eta[9,5]  -2.5820634 -2.2098 -2.0240 -1.84510 -1.51111
eta[10,5] -2.7981214 -2.4007 -2.2045 -2.01234 -1.65672
eta[11,5] -2.5934763 -2.2090 -2.0093 -1.81809 -1.44576
eta[12,5] -2.7076018 -2.2982 -2.0890 -1.88955 -1.50193
```

Trace plots and density curves for log-linear without age-gender effect




Trace of eta[5,2]
Density of eta[5,2]
Trace of eta[9,2]
Density of eta[9,2]
Trace of eta[6,2]
Density of eta[6,2]
Trace of eta[10,2]
Density of eta[10,2]
Trace of eta[7,2]
Density of eta[7,2]
Trace of eta[11,2]
Density of eta[11,2]
Trace of eta[8,2]
Density of eta[8,2]
Trace of eta[12,2]
Density of eta[12,2]
Trace of eta[1,3]
Density of eta[1,3]
Trace of eta[5,3]
Density of eta[5,3]
Trace of eta[2,3]
Density of eta[2,3]
Trace of eta[6,3]
Density of eta[6,3]
Trace of eta[3,3]
Density of eta[3,3]
Trace of eta[7,3]
Density of eta[7,3]
Trace of eta[4,3]
Density of eta[4,3]
Trace of eta[8,3]
Density of eta[8,3]


Trace of eta[9,3]
Density of eta[9,3]
Trace of eta[1,4]
Density of eta[1,4]
Trace of eta[10,3]
Density of eta[10,3]
Trace of eta[2,4]
Density of eta[2,4]
Trace of eta[11,3]
Density of eta[11,3]
Trace of eta[3,4]
Density of eta[3,4]
Trace of eta[12,3]
Density of eta[12,3]
Trace of eta[4,4]
Density of eta[4,4]
Trace of eta[5,4]
Density of eta[5,4]
Trace of eta[9,4]
Density of eta[9,4]
Trace of eta[6,4]
Density of eta[6,4]
Trace of eta[10,4]
Density of eta[10,4]
Trace of eta[7,4]
Density of eta[7,4]
Trace of eta[11,4]
Density of eta[11,4]
Trace of eta[8,4]
Density of eta[8,4]
Trace of eta[12,4]
Density of eta[12,4]



(b) With age and gender effects

```
model
{
for (i in 1:12)
        {y[i,1:5]~dmulti(p[i,],n[i])
        for (k in 1:5)
        {p[i,k]<-phi[i,k]/sum(phi[i,])
        phi[i,k]<-exp(eta[i,k])
                }
eta[i,3]<-0
for (k in 1:2) {
eta[i,k]<-beta0[k]+betaa1[k]*a1[i]+betaa2[k]*a2[i]+betaa3[k]*a3[i]+betaa4[k]*a4[i]+
betaa5[k]*a5[i]+betag[k]*gender[i]+betaa1g[k]*a1[i]*gender[i]+betaa2g[k]*a2[i]*gender[i]+
betaa3g[k]*a3[i]*gender[i]+betaa4g[k]*a4[i]*gender[i]+betaa5g[k]*a5[i]*gender[i]

eta[i,k+3]<-
beta0[k+3]+betaa1[k+3]*a1[i]+betaa2[k+3]*a2[i]+betaa3[k+3]*a3[i]+betaa4[k+3]*a4[i]+
betaa5[k+3]*a5[i]+betag[k+3]*gender[i]+betaa1g[k+3]*a1[i]*gender[i]+betaa2g[k+3]*a2[i]*gen
der[i]+
betaa3g[k+3]*a3[i]*gender[i]+betaa4g[k+3]*a4[i]*gender[i]+betaa5g[k+3]*a5[i]*gender[i]
}
}
for (k in 1:2) {
  beta0[k]~dnorm(mu0,1)
  beta0[k+3]~dnorm(mu0,1)
}

for (k in 1:2) {
```

```
betaa1[k]~dnorm(mua1,10)
betaa1[k+3]~dnorm(mua1,10)
betaa2[k]~dnorm(mua2,10)
betaa2[k+3]~dnorm(mua2,10)
betaa3[k]~dnorm(mua3,10)
betaa3[k+3]~dnorm(mua3,10)
betaa4[k]~dnorm(mua4,10)
betaa4[k+3]~dnorm(mua4,10)
betaa5[k]~dnorm(mua5,10)
betaa5[k+3]~dnorm(mua5,10)
betag[k]~dnorm(mug,10)
betag[k+3]~dnorm(mug,10)
betaa1g[k]~dnorm(mua1g,20)
betaa1g[k+3]~dnorm(mua1g,20)
betaa2g[k]~dnorm(mua2g,20)
betaa2g[k+3]~dnorm(mua2g,20)
betaa3g[k]~dnorm(mua3g,20)
betaa3g[k+3]~dnorm(mua3g,20)
betaa4g[k]~dnorm(mua4g,20)
betaa4g[k+3]~dnorm(mua4g,20)
betaa5g[k]~dnorm(mua5g,20)
betaa5g[k+3]~dnorm(mua5g,20)
}
beta0[3]<-0
betaa1[3]<-0
betaa2[3]<-0
betaa3[3]<-0
betaa4[3]<-0
betaa5[3]<-0
betag[3]<-0
betaa1g[3]<-0
betaa2g[3]<-0
betaa3g[3]<-0
betaa4g[3]<-0
betaa5g[3]<-0
mu0~dnorm(0,1)
mua1~dnorm(0,10)
mua2~dnorm(0,10)
mua3~dnorm(0,10)
mua4~dnorm(0,10)
mua5~dnorm(0,10)
mug~dnorm(0,10)
mua1g~dnorm(0,20)
mua2g~dnorm(0,20)
mua3g~dnorm(0,20)
mua4g~dnorm(0,20)
mua5g~dnorm(0,20)
}
```

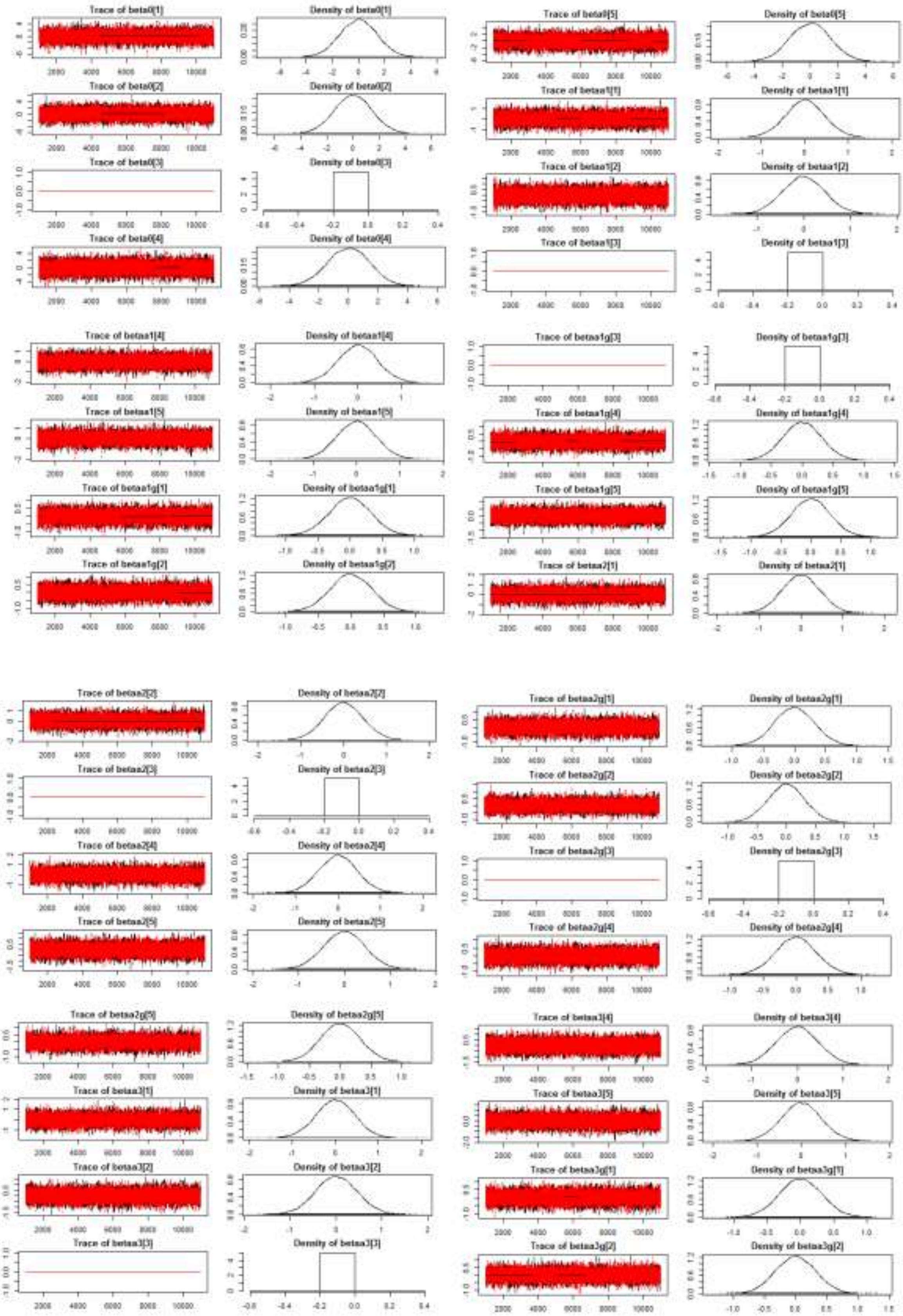
Model summary

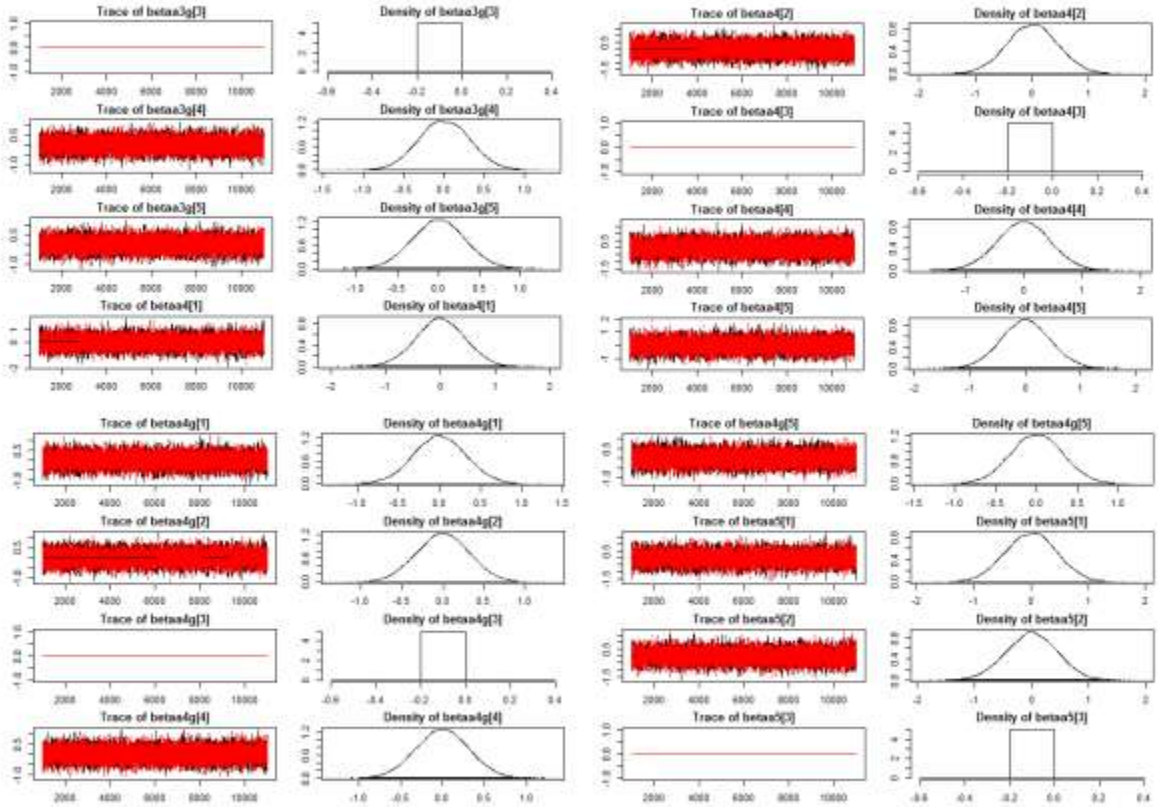
```
Iterations = 2001:12000
Thinning interval = 1
Number of chains = 2
Sample size per chain = 10000
```

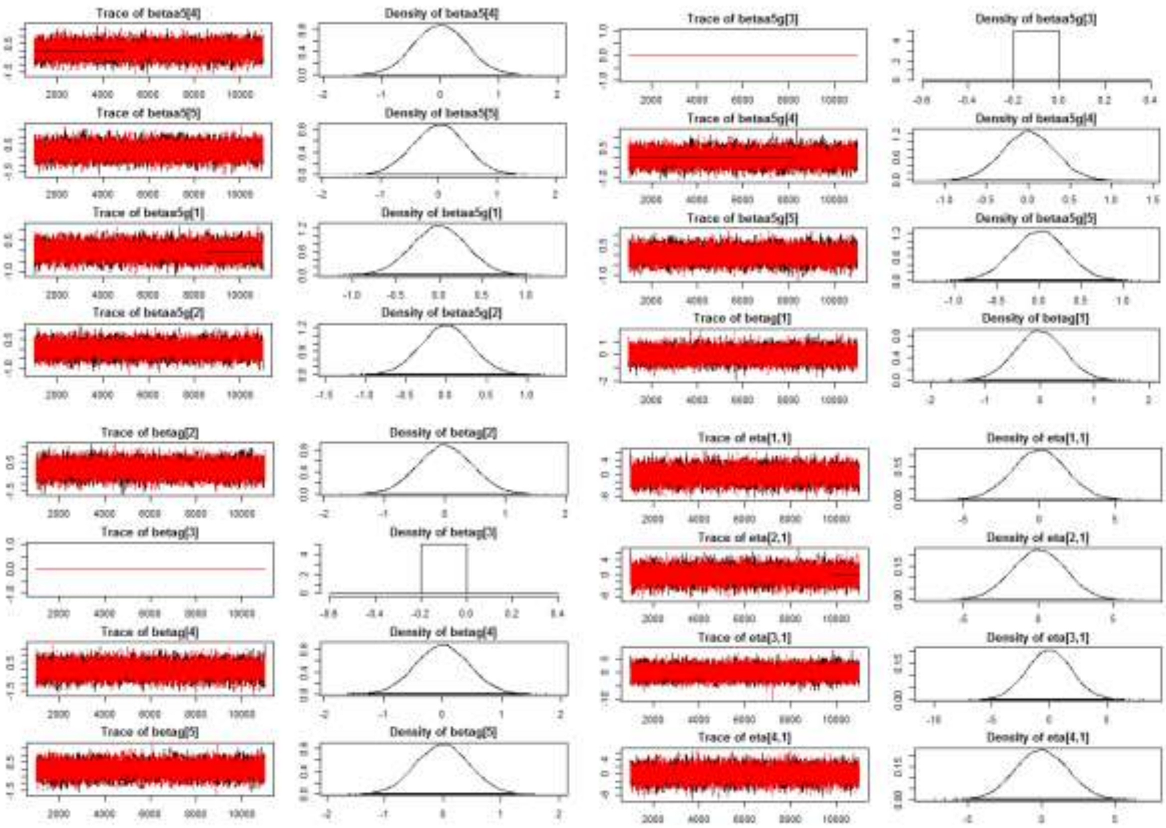
```
1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

              Mean     SD Naive SE Time-series SE
eta[1,1]  -1.16548 0.6553 0.004634       0.014010
eta[2,1]  -1.94551 0.4922 0.003481       0.007243
eta[3,1]  -1.23469 0.3303 0.002336       0.003451
eta[4,1]  -1.40121 0.3643 0.002576       0.004374
eta[5,1]  -1.64010 0.3711 0.002624       0.004335
eta[6,1]  -1.48596 0.3253 0.002300       0.003234
eta[7,1]  -1.68340 0.6154 0.004352       0.011296
eta[8,1]  -1.18168 0.3455 0.002443       0.003845
eta[9,1]  -1.56465 0.3111 0.002200       0.002853
eta[10,1] -1.73599 0.3550 0.002510       0.003912
eta[11,1] -1.50299 0.3658 0.002586       0.004603
eta[12,1] -1.07328 0.3984 0.002817       0.005271
eta[1,2]   0.36105 0.5065 0.003582       0.013932
eta[2,2]   0.57491 0.2481 0.001755       0.003508
eta[3,2]   0.34832 0.2085 0.001474       0.002609
eta[4,2]   0.93078 0.2046 0.001447       0.002887
eta[5,2]   0.70309 0.1981 0.001400       0.002640
eta[6,2]   0.81054 0.1734 0.001226       0.002044
eta[7,2]   0.27366 0.3953 0.002795       0.008494
eta[8,2]   0.77964 0.2123 0.001501       0.002526
eta[9,2]   0.57316 0.1667 0.001179       0.001728
eta[10,2]  0.66340 0.1824 0.001289       0.002131
eta[11,2]  0.66324 0.2071 0.001464       0.002915
eta[12,2]  1.15231 0.2385 0.001686       0.003865
eta[1,3]   0.00000 0.0000 0.000000       0.000000
eta[2,3]   0.00000 0.0000 0.000000       0.000000
eta[3,3]   0.00000 0.0000 0.000000       0.000000
eta[4,3]   0.00000 0.0000 0.000000       0.000000
eta[5,3]   0.00000 0.0000 0.000000       0.000000
eta[6,3]   0.00000 0.0000 0.000000       0.000000
eta[7,3]   0.00000 0.0000 0.000000       0.000000
eta[8,3]   0.00000 0.0000 0.000000       0.000000
eta[9,3]   0.00000 0.0000 0.000000       0.000000
eta[10,3]  0.00000 0.0000 0.000000       0.000000
eta[11,3]  0.00000 0.0000 0.000000       0.000000
eta[12,3]  0.00000 0.0000 0.000000       0.000000
eta[1,4]  -0.65994 0.5935 0.004197       0.014634
eta[2,4]  -0.51042 0.3182 0.002250       0.004167
eta[3,4]  -0.69287 0.2741 0.001939       0.003032
eta[4,4]   0.01432 0.2433 0.001720       0.003161
eta[5,4]  -0.38459 0.2523 0.001784       0.003073
eta[6,4]  -0.37436 0.2268 0.001603       0.002531
eta[7,4]  -0.65312 0.4826 0.003413       0.009303
eta[8,4]  -0.83601 0.3098 0.002190       0.003636
eta[9,4]  -1.51115 0.3068 0.002169       0.003518
eta[10,4] -0.95057 0.2725 0.001927       0.003261
eta[11,4] -0.72435 0.2891 0.002044       0.004120
eta[12,4] -0.03421 0.2929 0.002071       0.004394
eta[1,5]  -1.30600 0.6754 0.004776       0.014289
eta[2,5]  -1.47722 0.4265 0.003016       0.005068
eta[3,5]  -2.23987 0.4710 0.003331       0.005083
eta[4,5]  -1.53721 0.3791 0.002681       0.004259
eta[5,5]  -1.39135 0.3413 0.002414       0.003622
eta[6,5]  -1.83269 0.3754 0.002654       0.003573
eta[7,5]  -2.30125 0.6818 0.004821       0.011777
eta[8,5]  -2.26253 0.4871 0.003445       0.005636
eta[9,5]  -2.32280 0.4161 0.002942       0.003910
eta[10,5] -2.80735 0.4989 0.003528       0.006172
eta[11,5] -2.23699 0.4652 0.003289       0.005796
eta[12,5] -2.28735 0.5963 0.004216       0.008584

2. Quantiles for each variable:

              2.5%       25%      50%     75%     97.5%
eta[1,1]  -2.47231 -1.601713 -1.15320 -0.7161  0.092343
eta[2,1]  -2.97291 -2.259818 -1.92786 -1.6070 -1.035777
eta[3,1]  -1.91291 -1.448613 -1.22518 -1.0108 -0.613053
eta[4,1]  -2.13287 -1.640465 -1.39335 -1.1543 -0.712511
```

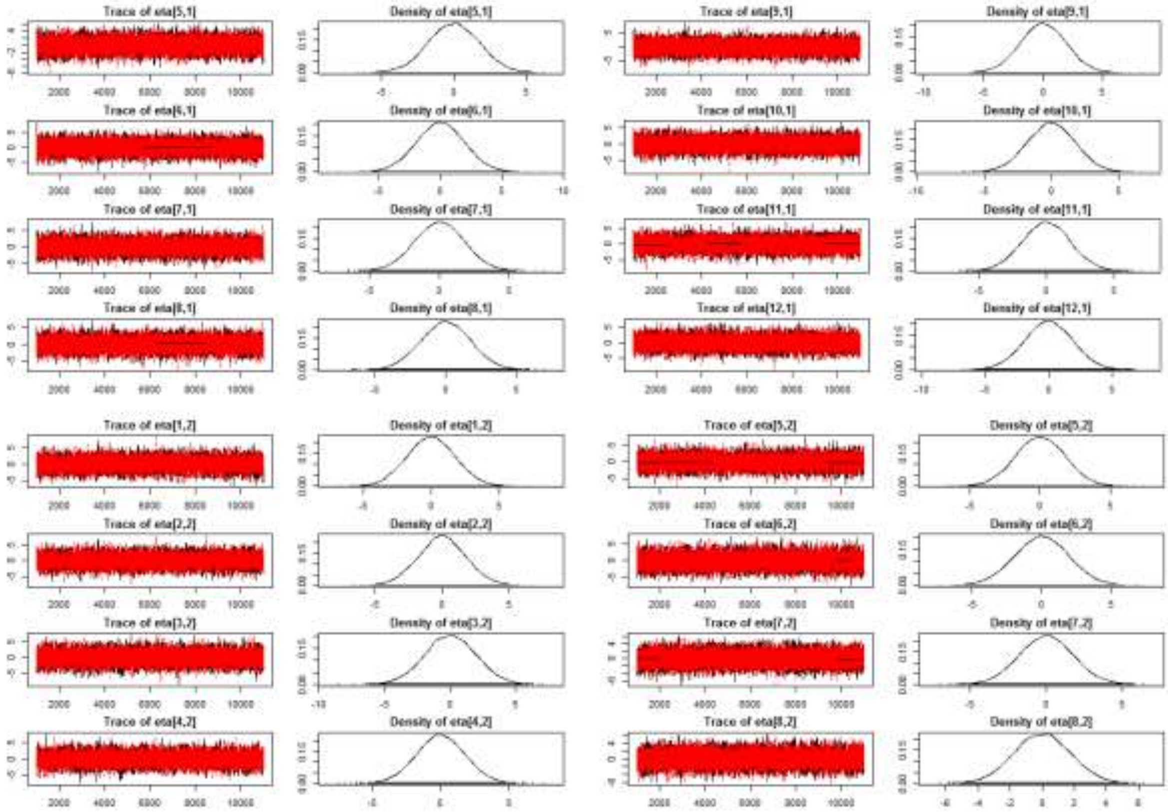
eta[5,1]	-2.39312	-1.883633	-1.63092	-1.3846	-0.938136
eta[6,1]	-2.15074	-1.698990	-1.47702	-1.2613	-0.876847
eta[7,1]	-2.94036	-2.087166	-1.66058	-1.2572	-0.538260
eta[8,1]	-1.88029	-1.410948	-1.17215	-0.9446	-0.523784
eta[9,1]	-2.19898	-1.767932	-1.55592	-1.3520	-0.973893
eta[10,1]	-2.46681	-1.967521	-1.72172	-1.4906	-1.082698
eta[11,1]	-2.24391	-1.741957	-1.49580	-1.2500	-0.825542
eta[12,1]	-1.88627	-1.331800	-1.06122	-0.8044	-0.320995
eta[1,2]	-0.63665	0.022646	0.35095	0.7018	1.363401
eta[2,2]	0.09046	0.407454	0.57404	0.7403	1.064507
eta[3,2]	-0.05689	0.205880	0.34670	0.4893	0.760450
eta[4,2]	0.54754	0.790131	0.92677	1.0644	1.343290
eta[5,2]	0.32704	0.564898	0.69977	0.8343	1.101365
eta[6,2]	0.47440	0.692081	0.80966	0.9270	1.154854
eta[7,2]	-0.50086	0.007997	0.26967	0.5366	1.063150
eta[8,2]	0.36962	0.634979	0.77796	0.9208	1.204131
eta[9,2]	0.24547	0.462034	0.57223	0.6845	0.903971
eta[10,2]	0.31181	0.539577	0.66309	0.7862	1.020912
eta[11,2]	0.26254	0.524441	0.66090	0.8011	1.077106
eta[12,2]	0.69531	0.989045	1.14566	1.3120	1.635587
eta[1,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[2,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[3,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[4,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[5,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[6,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[7,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[8,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[9,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[10,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[11,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[12,3]	0.00000	0.000000	0.00000	0.0000	0.000000
eta[1,4]	-1.84022	-1.059482	-0.65382	-0.2582	0.481529
eta[2,4]	-1.15019	-0.719819	-0.50743	-0.2934	0.103005
eta[3,4]	-1.24518	-0.873616	-0.68669	-0.5079	-0.171912
eta[4,4]	-0.45605	-0.151088	0.01327	0.1756	0.502887
eta[5,4]	-0.88225	-0.551646	-0.38499	-0.2154	0.104430
eta[6,4]	-0.82262	-0.527503	-0.37461	-0.2226	0.069840
eta[7,4]	-1.61704	-0.978689	-0.65004	-0.3240	0.277148
eta[8,4]	-1.45766	-1.040828	-0.83115	-0.6264	-0.245895
eta[9,4]	-2.13037	-1.711929	-1.50184	-1.3039	-0.929313
eta[10,4]	-1.49822	-1.128315	-0.94526	-0.7661	-0.431478
eta[11,4]	-1.29704	-0.919924	-0.71937	-0.5278	-0.160346
eta[12,4]	-0.61718	-0.230538	-0.03339	0.1657	0.531120
eta[1,5]	-2.66875	-1.758649	-1.29466	-0.8411	-0.008841
eta[2,5]	-2.35278	-1.758268	-1.45933	-1.1811	-0.688208
eta[3,5]	-3.22616	-2.543455	-2.21752	-1.9118	-1.380105
eta[4,5]	-2.31597	-1.786380	-1.52641	-1.2770	-0.834258
eta[5,5]	-2.09563	-1.615873	-1.38028	-1.1564	-0.746054
eta[6,5]	-2.60430	-2.078414	-1.81769	-1.5707	-1.138450
eta[7,5]	-3.69915	-2.749025	-2.28286	-1.8290	-1.037896
eta[8,5]	-3.26893	-2.578628	-2.24667	-1.9186	-1.370867
eta[9,5]	-3.19651	-2.590463	-2.30214	-2.0349	-1.567415
eta[10,5]	-3.84667	-3.134095	-2.78968	-2.4593	-1.880092
eta[11,5]	-3.20016	-2.534340	-2.21477	-1.9187	-1.374634
eta[12,5]	-3.55094	-2.666580	-2.25522	-1.8682	-1.214440

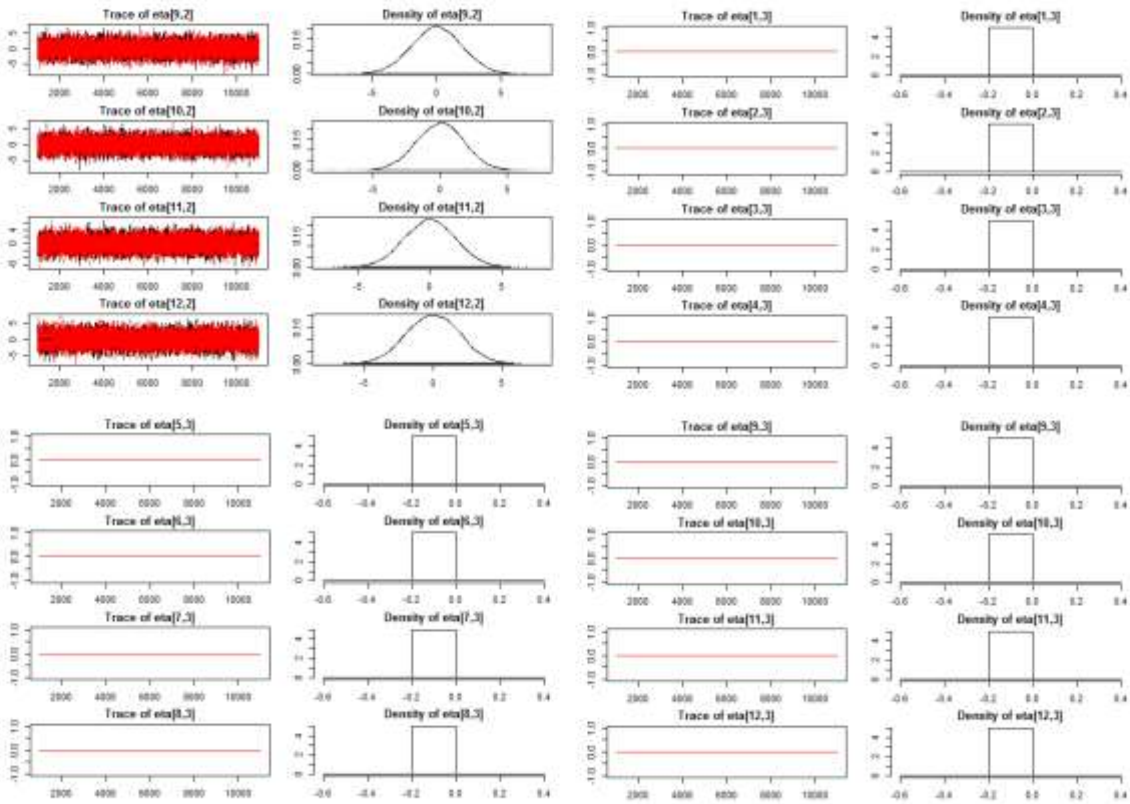
Trace plots and density curves for log-linear with age-gender effect

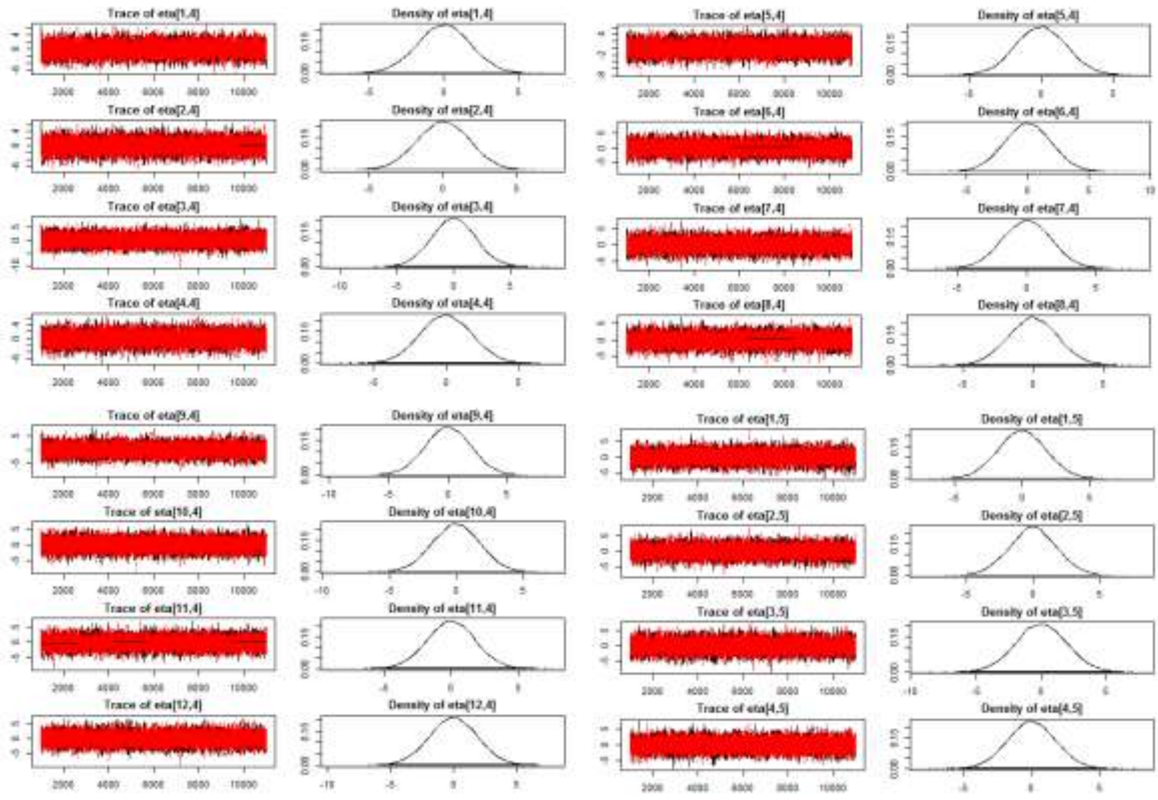


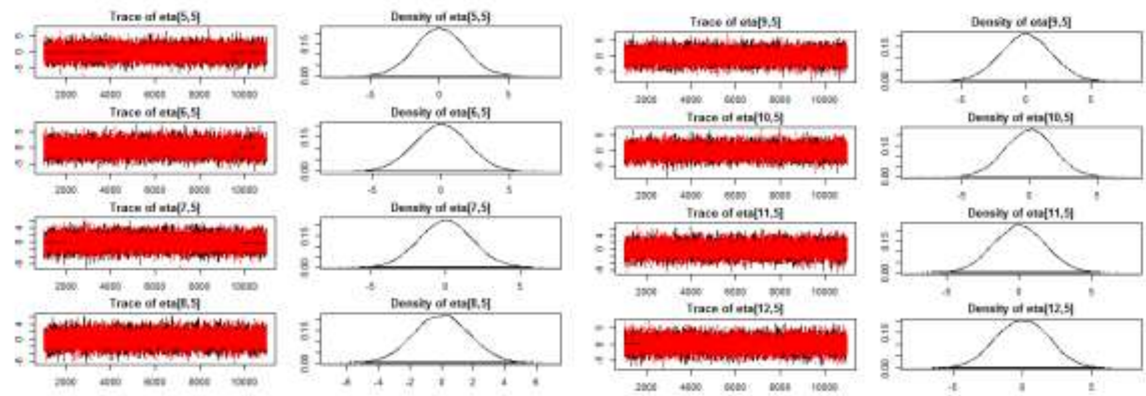

Trace of betaa3g[3]
Density of betaa3g[3]
Trace of betaa4[2]
Density of betaa4[2]
Trace of betaa3g[4]
Density of betaa3g[4]
Trace of betaa4[3]
Density of betaa4[3]
Trace of betaa3g[5]
Density of betaa3g[5]
Trace of betaa4[4]
Density of betaa4[4]
Trace of betaa4[1]
Density of betaa4[1]
Trace of betaa4[5]
Density of betaa4[5]
Trace of betaa4g[1]
Density of betaa4g[1]
Trace of betaa4g[5]
Density of betaa4g[5]
Trace of betaa4g[2]
Density of betaa4g[2]
Trace of betaa5[1]
Density of betaa5[1]
Trace of betaa4g[3]
Density of betaa4g[3]
Trace of betaa5[2]
Density of betaa5[2]
Trace of betaa4g[4]
Density of betaa4g[4]
Trace of betaa5[3]
Density of betaa5[3]


Trace of betaa5[4]
Density of betaa5[4]
Trace of betaa5g[3]
Density of betaa5g[3]
Trace of betaa5[5]
Density of betaa5[5]
Trace of betaa5g[4]
Density of betaa5g[4]
Trace of betaa5g[1]
Density of betaa5g[1]
Trace of betaa5g[5]
Density of betaa5g[5]
Trace of betaa5g[2]
Density of betaa5g[2]
Trace of betag[1]
Density of betag[1]
Trace of betag[2]
Density of betag[2]
Trace of eta[1,1]
Density of eta[1,1]
Trace of betag[3]
Density of betag[3]
Trace of eta[2,1]
Density of eta[2,1]
Trace of betag[4]
Density of betag[4]
Trace of eta[3,1]
Density of eta[3,1]
Trace of betag[5]
Density of betag[5]
Trace of eta[4,1]
Density of eta[4,1]


Trace of eta[5,1]
Density of eta[5,1]
Trace of eta[9,1]
Density of eta[9,1]
Trace of eta[6,1]
Density of eta[6,1]
Trace of eta[10,1]
Density of eta[10,1]
Trace of eta[7,1]
Density of eta[7,1]
Trace of eta[11,1]
Density of eta[11,1]
Trace of eta[8,1]
Density of eta[8,1]
Trace of eta[12,1]
Density of eta[12,1]
Trace of eta[1,2]
Density of eta[1,2]
Trace of eta[5,2]
Density of eta[5,2]
Trace of eta[2,2]
Density of eta[2,2]
Trace of eta[6,2]
Density of eta[6,2]
Trace of eta[3,2]
Density of eta[3,2]
Trace of eta[7,2]
Density of eta[7,2]
Trace of eta[4,2]
Density of eta[4,2]
Trace of eta[8,2]
Density of eta[8,2]


Trace of eta[9,2]
Density of eta[9,2]
Trace of eta[1,3]
Density of eta[1,3]
Trace of eta[10,2]
Density of eta[10,2]
Trace of eta[2,3]
Density of eta[2,3]
Trace of eta[11,2]
Density of eta[11,2]
Trace of eta[3,3]
Density of eta[3,3]
Trace of eta[12,2]
Density of eta[12,2]
Trace of eta[4,3]
Density of eta[4,3]
Trace of eta[5,3]
Density of eta[5,3]
Trace of eta[9,3]
Density of eta[9,3]
Trace of eta[6,3]
Density of eta[6,3]
Trace of eta[10,3]
Density of eta[10,3]
Trace of eta[7,3]
Density of eta[7,3]
Trace of eta[11,3]
Density of eta[11,3]
Trace of eta[8,3]
Density of eta[8,3]
Trace of eta[12,3]
Density of eta[12,3]





Content of data for log-linear model with and without age and gender effects

y1	y2	y3	y4	y5	n	a1	a2	a3	a4	a5	gender
3	7	5	2	2	19	1	1	-1	0	0	1
2	42	24	15	6	89	1	-1	-1	0	0	1
12	55	39	20	4	130	1	0	2	0	0	1
8	80	31	33	7	159	-1	0	0	1	-1	1
7	74	37	25	10	153	-1	0	0	-1	-1	1
11	107	48	33	8	207	-1	0	0	0	2	1
1	12	10	6	1	30	1	1	-1	0	0	-1
11	68	31	13	3	126	1	-1	-1	0	0	-1
12	99	56	12	6	185	1	0	2	0	0	-1
8	86	45	17	2	158	-1	0	0	1	-1	-1
8	66	34	17	4	129	-1	0	0	-1	-1	-1
8	70	22	22	2	124	-1	0	0	0	2	-1

Appendix B

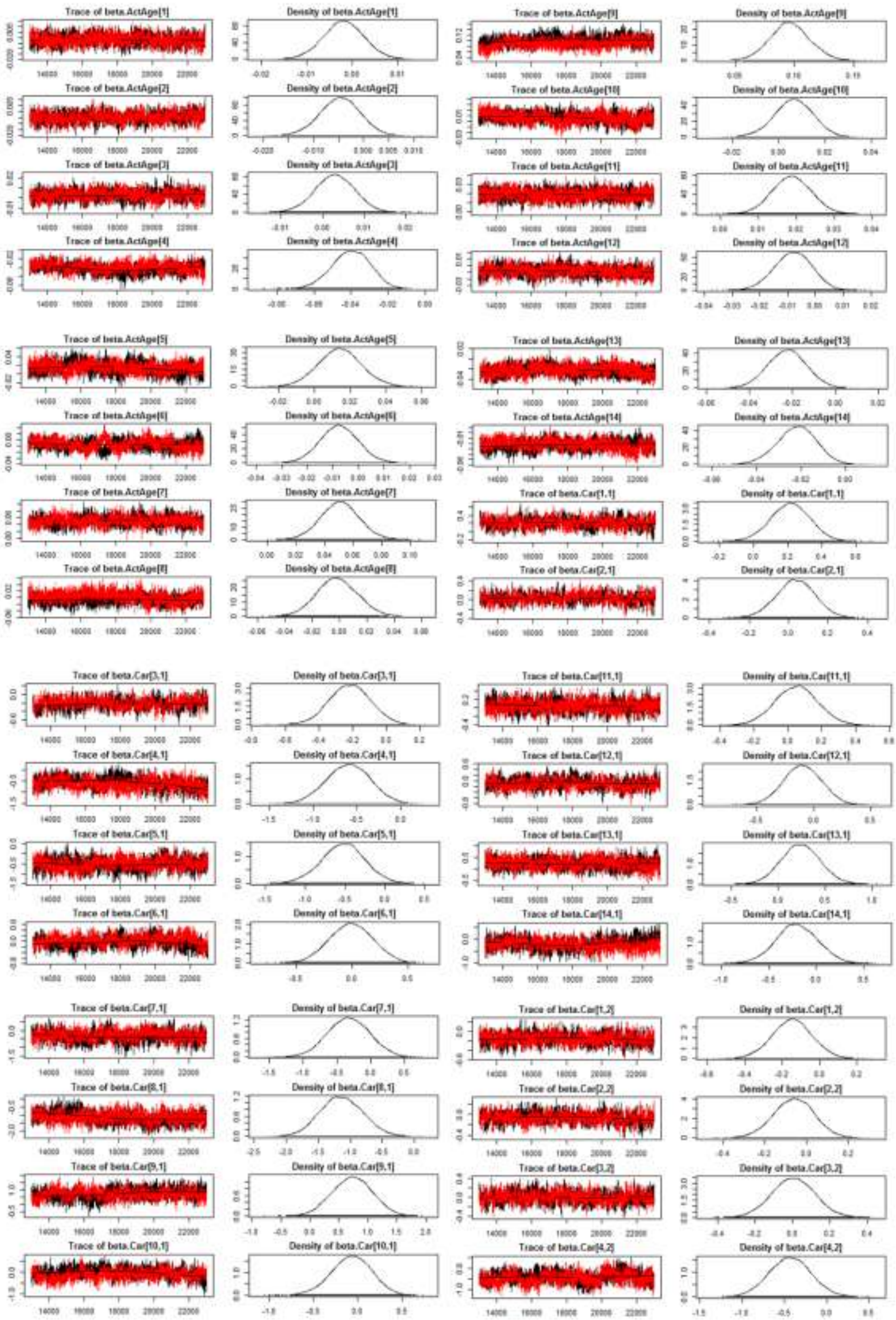
```
library(rjags)
mydata<-read.csv(file.choose(),header=TRUE)

y<-with (mydata , cbind (Cong_MWs, Cong_cities, Exhaustfumes, BuyLowEmi, ReducTravCar,
CCView, CartoWalk, CartoBus, CartoBike, Tax_CarUse, AllowCarUse, ReducCarUse,
ReducCarUse_NP, CarBetterPayLess))

ActAge<-mydata$ActAge
Gender<-mydata$Gender
HH<-mydata$HH
inwork<-mydata$inwork
employee<-mydata$employee
fulltime<-mydata$fulltime
noemp<-mydata$noemp
Car<-mydata$Car
cut<-c(-1,1)

R<-rep(1,14)
R<-diag(R)

orddat<-list(y=y, ActAge=ActAge, Gender=Gender, HH=HH, inwork=inwork,
employee=employee, fulltime=fulltime, noemp=noemp, Car=Car, cut=cut, R=R)

zinits<-y-1
zinits1<-2*zinits
zinits2<-3*zinits
zinits<-list(list(z=zinits1),list(z=zinits2))

ordjags<-jags.model("probitmodel.txt",data=orddat,inits=zinits,n.chains=2)
update(ordjags,2000)
ordsamples<-coda.samples(ordjags,c("beta.ActAge", "beta.Gender", "beta.HH", "beta.inwork",
"beta.employee", "beta.fulltime", "beta.noemp", "beta.Car"), 10000)
summary(ordsamples)
```

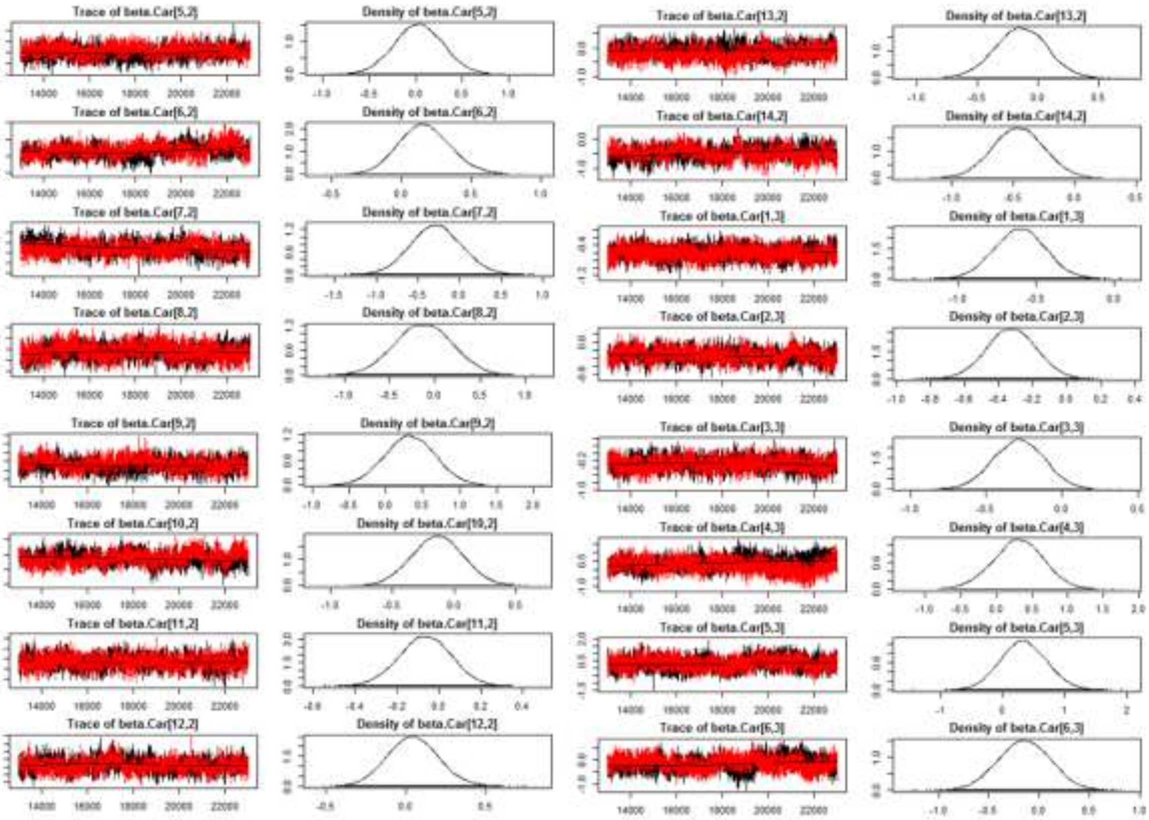
Multivariate Probit model

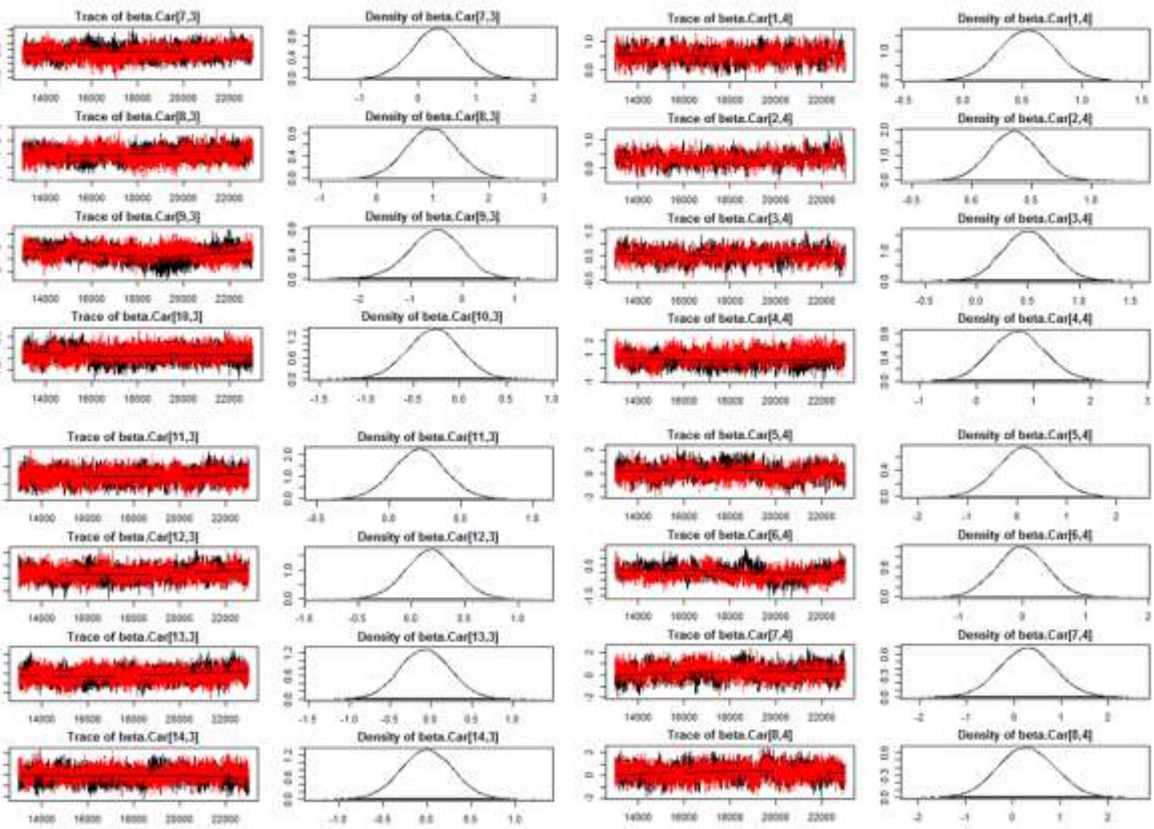
```
var z[1509,14];
model
{
 for (i in 1:1509)
   {
    z[i,]~dmnorm(mu[i,],Omega)
 for (j in 1:14)
   {
    y[i,j]~dinterval(z[i,j],cut)
    mu[i,j]<-beta0[j]+beta.ActAge[j]*(ActAge[i]-50)+beta.Gender[j]*(2*Gender[i]-
3)+beta.HH[j,HH[i]]+beta.Car[j,Car[i]]+beta.inwork[j]*inwork[i]+beta.employee[j]*employee[i]+
beta.fulltime[j]*fulltime[i]+beta.noemp[j,noemp[i]]
   }
   }
 Omega~dwish(R,20)
```

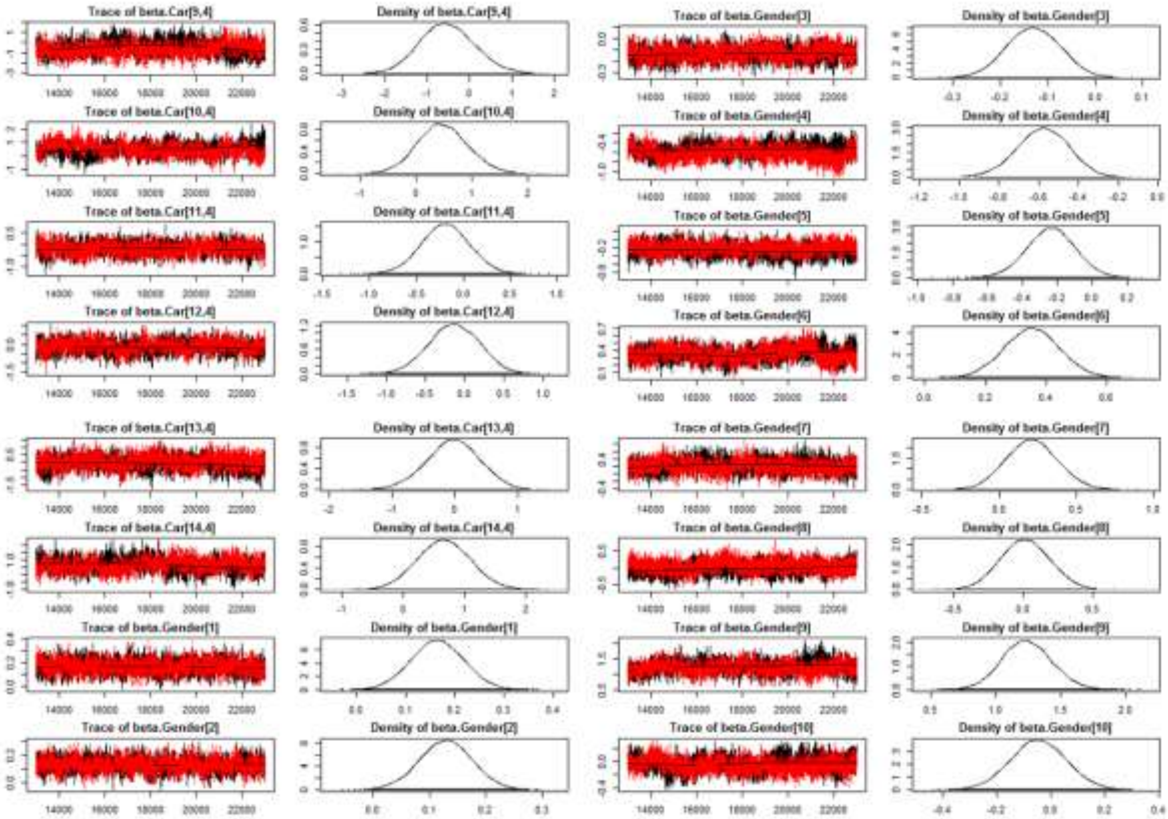
```
 for (j in 1:14)
{
beta0[j]~dnorm(0,1)
beta.ActAge[j]~dnorm(0,1)
beta.Gender[j]~dnorm(0,1)
beta.employee[j]~dnorm(0,0.0625)
beta.inwork[j]~dnorm(0,0.0625)
beta.fulltime[j]~dnorm(0,0.0625)

  beta.HH[j,1]<-3*U.HH[j,1]
  beta.HH[j,2]<-2*U.HH[j,2]-U.HH[j,1]
  beta.HH[j,3]<-  U.HH[j,3]-U.HH[j,2]-U.HH[j,1]
  beta.HH[j,4]<- -U.HH[j,3]-U.HH[j,2]-U.HH[j,1]
  k.HH[j,1]<-9
  k.HH[j,2]<-9/2
  k.HH[j,3]<-9/6
  for (k in 1:3)
  {U.HH[j,k]~dnorm(0,k.HH[j,k])

  }

  beta.Car[j,1]<-3*U.Car[j,1]
  beta.Car[j,2]<-2*U.Car[j,2]-U.Car[j,1]
  beta.Car[j,3]<-  U.Car[j,3]-U.Car[j,2]-U.Car[j,1]
  beta.Car[j,4]<- -U.Car[j,3]-U.Car[j,2]-U.Car[j,1]
  k.Car[j,1]<-9
  k.Car[j,2]<-9/2
  k.Car[j,3]<-9/6
  for (k in 1:3)
  {U.Car[j,k]~dnorm(0,k.Car[j,k])

  }

  beta.noemp[j,1]<-2*U.noemp[j,1]
  beta.noemp[j,2]<-U.noemp[j,2]-U.noemp[j,1]
  beta.noemp[j,3]<--U.noemp[j,2]-U.noemp[j,1]
  k.noemp[j,1]<-4
  k.noemp[j,2]<-4/3
  for (k in 1:2)
  {U.noemp[j,k]~dnorm(0,k.noemp[j,k])


  beta.noemp[j,4]<-0

  }
}
```

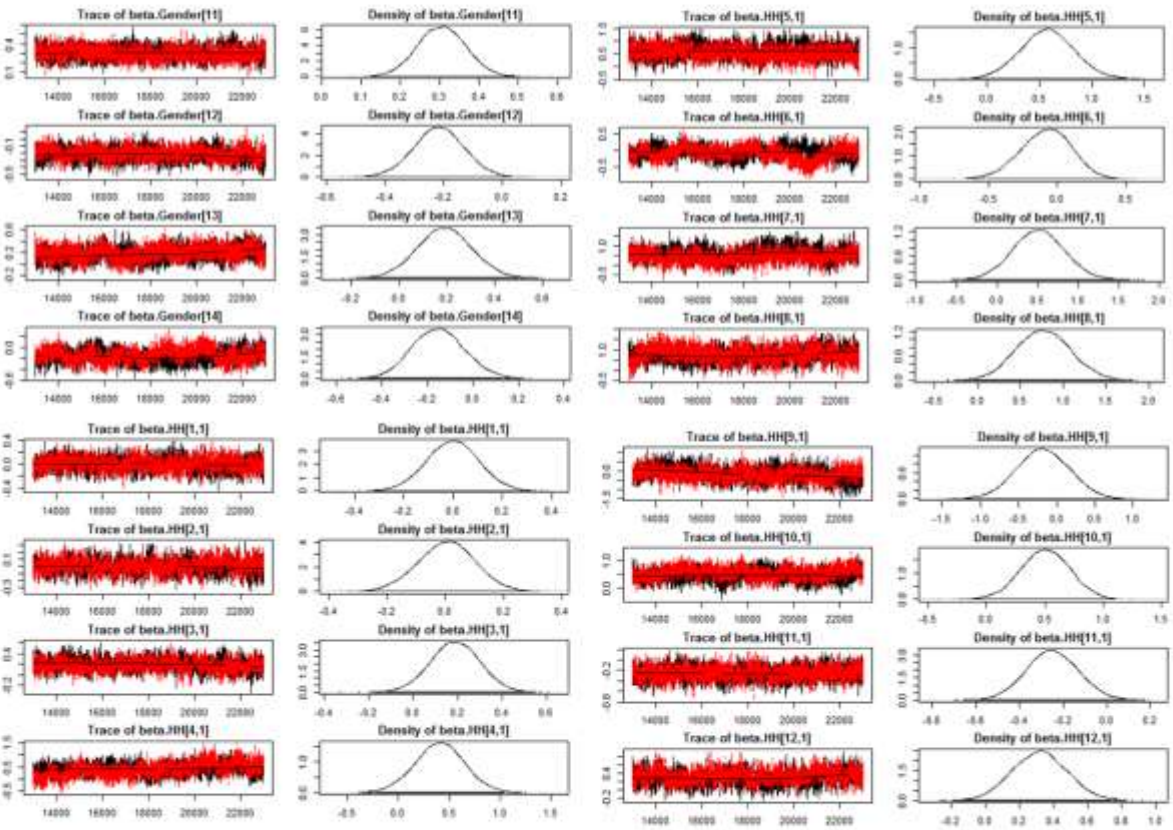
Trace plots and density curves for latent variables

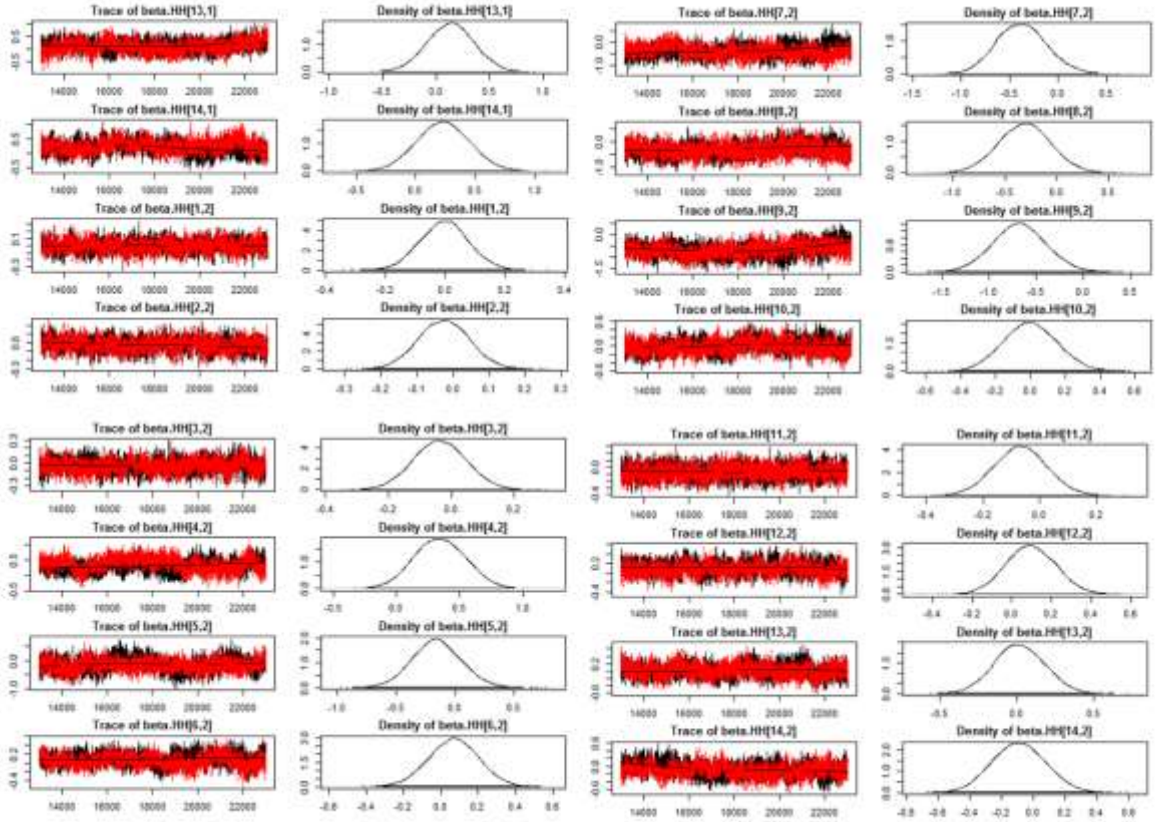


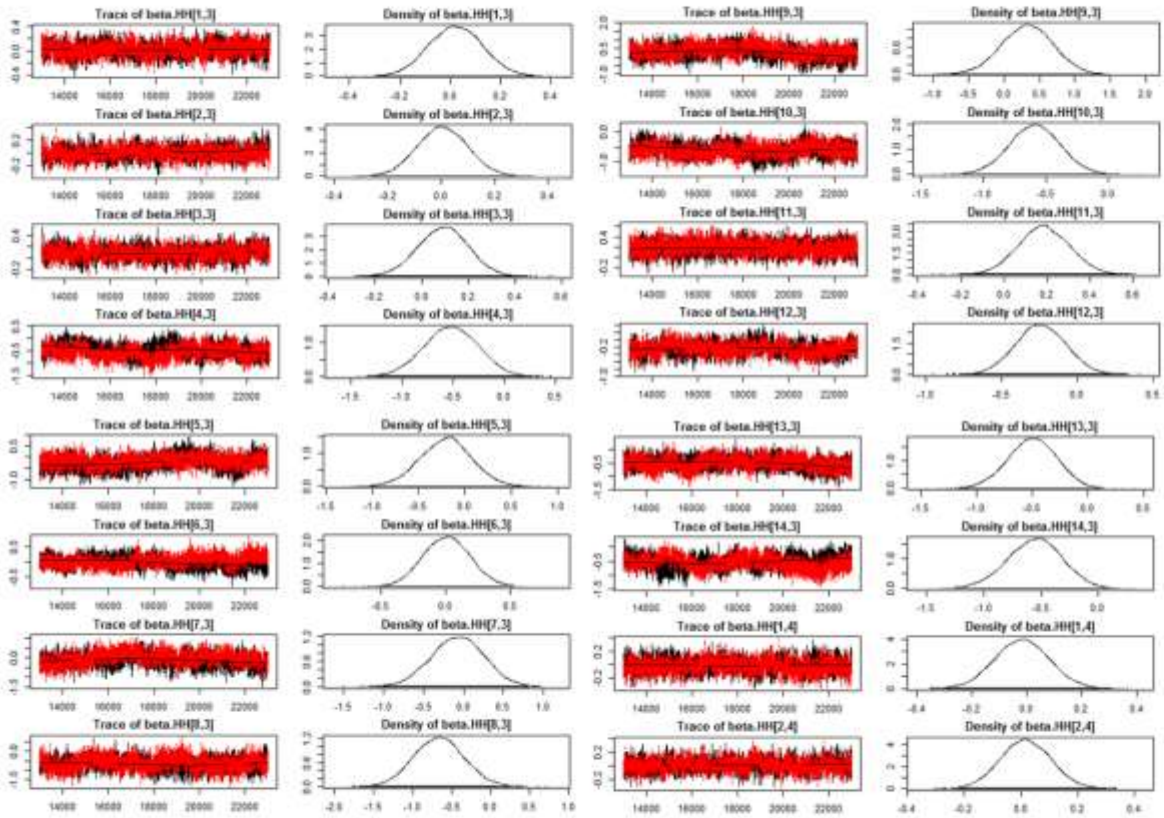

Trace of beta.Car[5,2]
Density of beta.Car[5,2]
Trace of beta.Car[13,2]
Density of beta.Car[13,2]
Trace of beta.Car[6,2]
Density of beta.Car[6,2]
Trace of beta.Car[14,2]
Density of beta.Car[14,2]
Trace of beta.Car[7,2]
Density of beta.Car[7,2]
Trace of beta.Car[1,3]
Density of beta.Car[1,3]
Trace of beta.Car[8,2]
Density of beta.Car[8,2]
Trace of beta.Car[2,3]
Density of beta.Car[2,3]
Trace of beta.Car[9,2]
Density of beta.Car[9,2]
Trace of beta.Car[3,3]
Density of beta.Car[3,3]
Trace of beta.Car[10,2]
Density of beta.Car[10,2]
Trace of beta.Car[4,3]
Density of beta.Car[4,3]
Trace of beta.Car[11,2]
Density of beta.Car[11,2]
Trace of beta.Car[5,3]
Density of beta.Car[5,3]
Trace of beta.Car[12,2]
Density of beta.Car[12,2]
Trace of beta.Car[6,3]
Density of beta.Car[6,3]

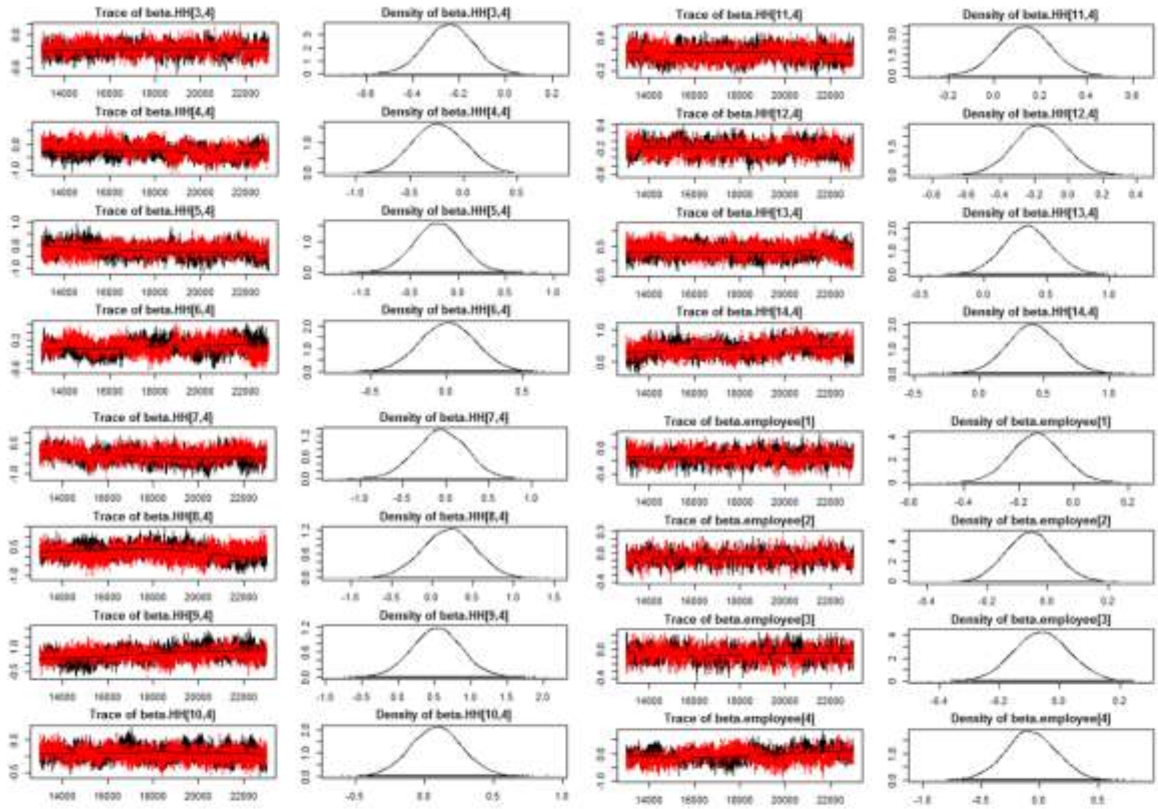

Trace of beta.Car[7,3]
Density of beta.Car[7,3]
Trace of beta.Car[1,4]
Density of beta.Car[1,4]
Trace of beta.Car[8,3]
Density of beta.Car[8,3]
Trace of beta.Car[2,4]
Density of beta.Car[2,4]
Trace of beta.Car[9,3]
Density of beta.Car[9,3]
Trace of beta.Car[3,4]
Density of beta.Car[3,4]
Trace of beta.Car[10,3]
Density of beta.Car[10,3]
Trace of beta.Car[4,4]
Density of beta.Car[4,4]
Trace of beta.Car[11,3]
Density of beta.Car[11,3]
Trace of beta.Car[5,4]
Density of beta.Car[5,4]
Trace of beta.Car[12,3]
Density of beta.Car[12,3]
Trace of beta.Car[6,4]
Density of beta.Car[6,4]
Trace of beta.Car[13,3]
Density of beta.Car[13,3]
Trace of beta.Car[7,4]
Density of beta.Car[7,4]
Trace of beta.Car[14,3]
Density of beta.Car[14,3]
Trace of beta.Car[8,4]
Density of beta.Car[8,4]

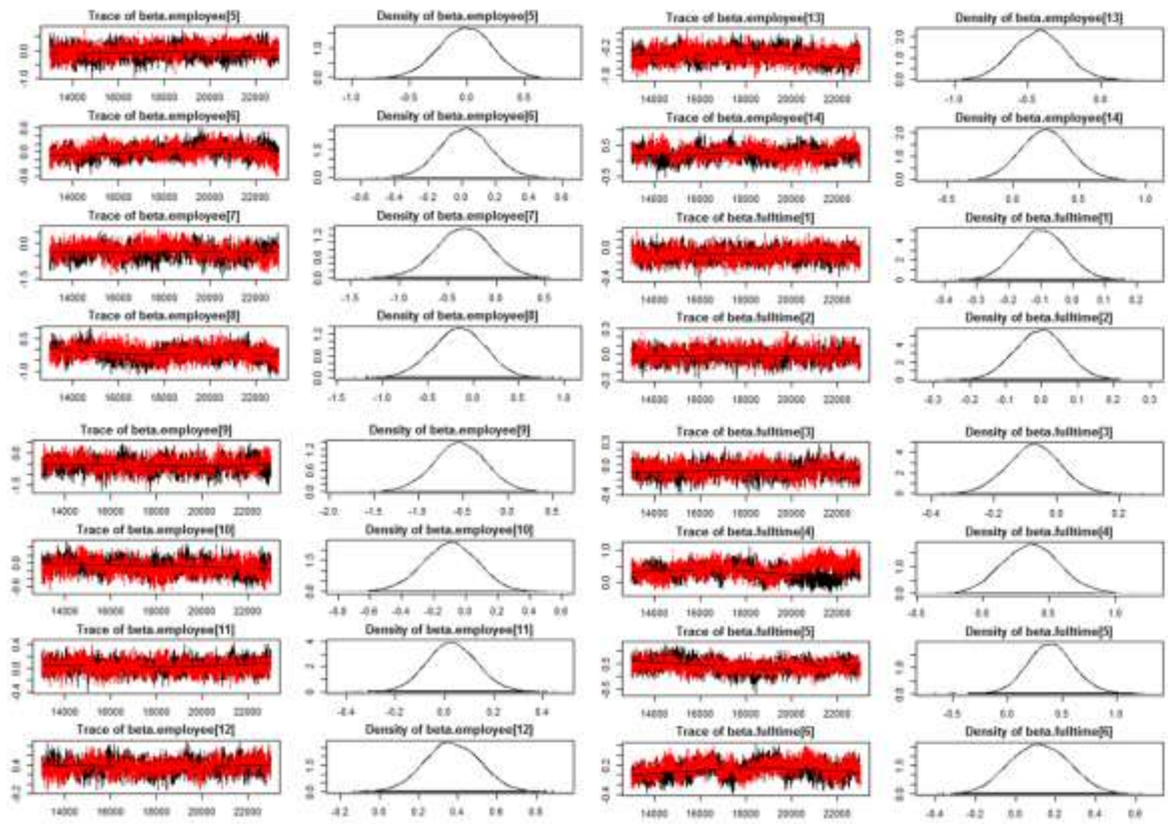

Trace of beta.Car[9,4]
Density of beta.Car[9,4]
Trace of beta.Gender[3]
Density of beta.Gender[3]
Trace of beta.Car[10,4]
Density of beta.Car[10,4]
Trace of beta.Gender[4]
Density of beta.Gender[4]
Trace of beta.Car[11,4]
Density of beta.Car[11,4]
Trace of beta.Gender[5]
Density of beta.Gender[5]
Trace of beta.Car[12,4]
Density of beta.Car[12,4]
Trace of beta.Gender[6]
Density of beta.Gender[6]
Trace of beta.Car[13,4]
Density of beta.Car[13,4]
Trace of beta.Gender[7]
Density of beta.Gender[7]
Trace of beta.Car[14,4]
Density of beta.Car[14,4]
Trace of beta.Gender[8]
Density of beta.Gender[8]
Trace of beta.Gender[1]
Density of beta.Gender[1]
Trace of beta.Gender[9]
Density of beta.Gender[9]
Trace of beta.Gender[2]
Density of beta.Gender[2]
Trace of beta.Gender[10]
Density of beta.Gender[10]

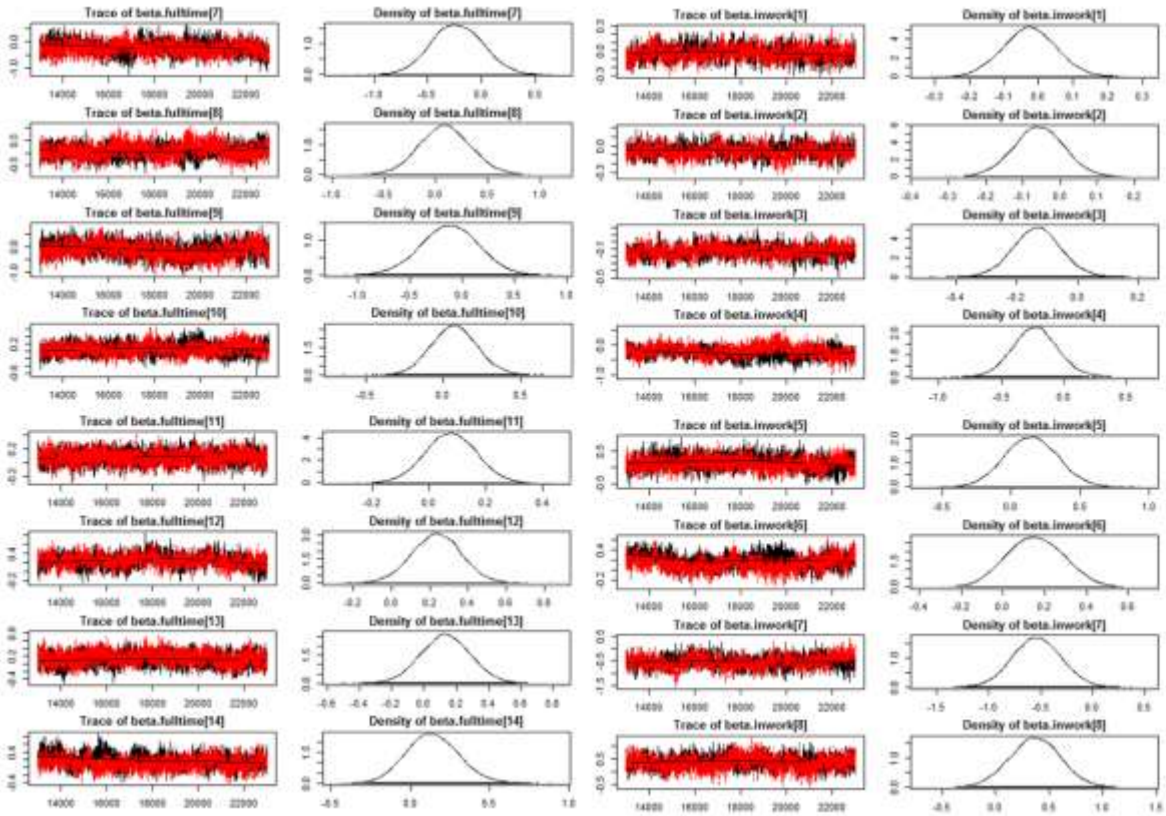

Trace of beta.Gender[11]
Density of beta.Gender[11]
Trace of beta.HH[5,1]
Density of beta.HH[5,1]
Trace of beta.Gender[12]
Density of beta.Gender[12]
Trace of beta.HH[6,1]
Density of beta.HH[6,1]
Trace of beta.Gender[13]
Density of beta.Gender[13]
Trace of beta.HH[7,1]
Density of beta.HH[7,1]
Trace of beta.Gender[14]
Density of beta.Gender[14]
Trace of beta.HH[8,1]
Density of beta.HH[8,1]
Trace of beta.HH[1,1]
Density of beta.HH[1,1]
Trace of beta.HH[9,1]
Density of beta.HH[9,1]
Trace of beta.HH[2,1]
Density of beta.HH[2,1]
Trace of beta.HH[10,1]
Density of beta.HH[10,1]
Trace of beta.HH[3,1]
Density of beta.HH[3,1]
Trace of beta.HH[11,1]
Density of beta.HH[11,1]
Trace of beta.HH[4,1]
Density of beta.HH[4,1]
Trace of beta.HH[12,1]
Density of beta.HH[12,1]

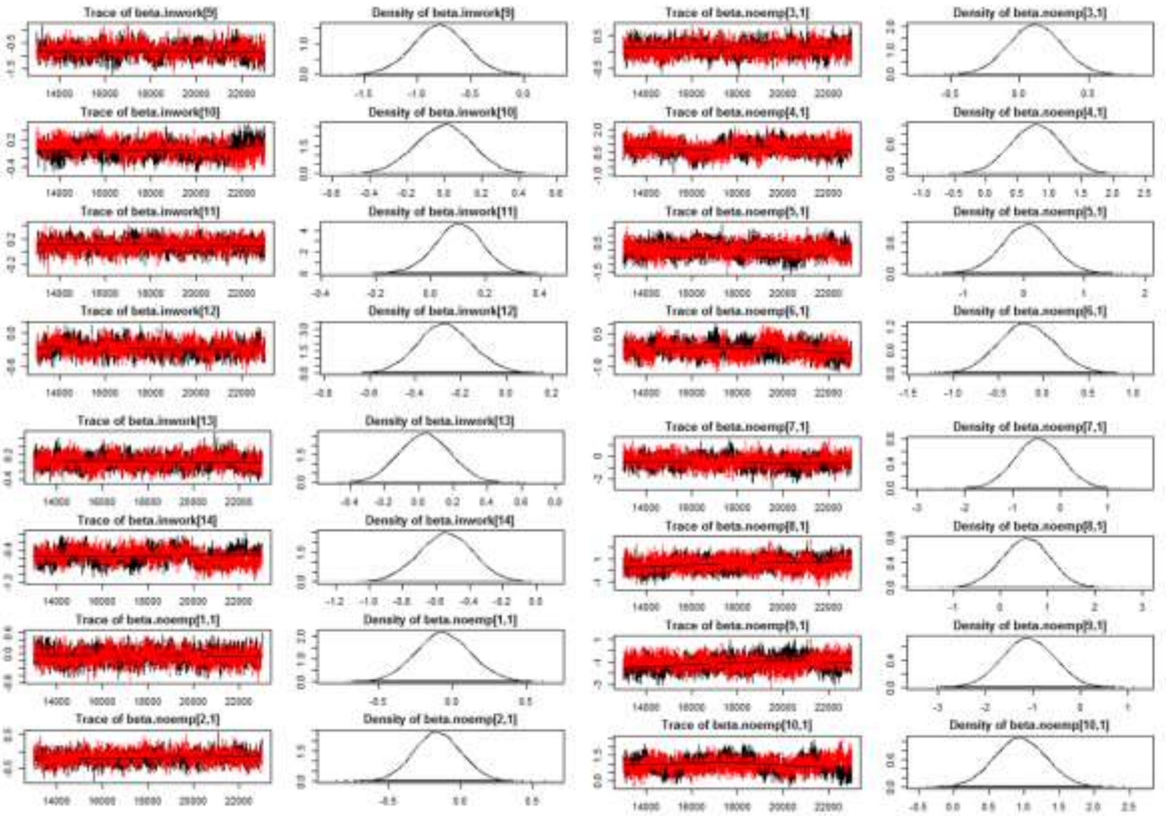

Trace of beta.HH[13,1]
Density of beta.HH[13,1]
Trace of beta.HH[7,2]
Density of beta.HH[7,2]
Trace of beta.HH[14,1]
Density of beta.HH[14,1]
Trace of beta.HH[8,2]
Density of beta.HH[8,2]
Trace of beta.HH[1,2]
Density of beta.HH[1,2]
Trace of beta.HH[9,2]
Density of beta.HH[9,2]
Trace of beta.HH[2,2]
Density of beta.HH[2,2]
Trace of beta.HH[10,2]
Density of beta.HH[10,2]
Trace of beta.HH[3,2]
Density of beta.HH[3,2]
Trace of beta.HH[11,2]
Density of beta.HH[11,2]
Trace of beta.HH[4,2]
Density of beta.HH[4,2]
Trace of beta.HH[12,2]
Density of beta.HH[12,2]
Trace of beta.HH[5,2]
Density of beta.HH[5,2]
Trace of beta.HH[13,2]
Density of beta.HH[13,2]
Trace of beta.HH[6,2]
Density of beta.HH[6,2]
Trace of beta.HH[14,2]
Density of beta.HH[14,2]

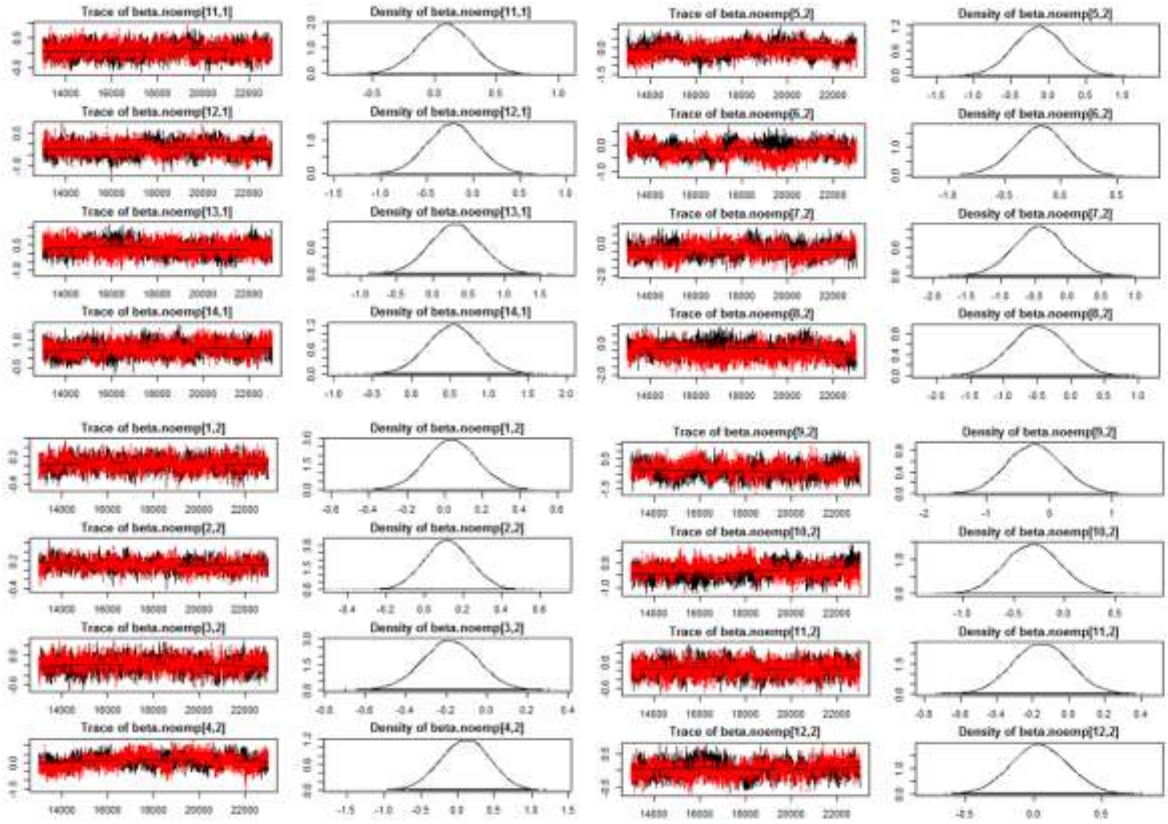

Trace of beta.HH[1,3]
Density of beta.HH[1,3]
Trace of beta.HH[9,3]
Density of beta.HH[9,3]
Trace of beta.HH[2,3]
Density of beta.HH[2,3]
Trace of beta.HH[10,3]
Density of beta.HH[10,3]
Trace of beta.HH[3,3]
Density of beta.HH[3,3]
Trace of beta.HH[11,3]
Density of beta.HH[11,3]
Trace of beta.HH[4,3]
Density of beta.HH[4,3]
Trace of beta.HH[12,3]
Density of beta.HH[12,3]
Trace of beta.HH[5,3]
Density of beta.HH[5,3]
Trace of beta.HH[13,3]
Density of beta.HH[13,3]
Trace of beta.HH[6,3]
Density of beta.HH[6,3]
Trace of beta.HH[14,3]
Density of beta.HH[14,3]
Trace of beta.HH[7,3]
Density of beta.HH[7,3]
Trace of beta.HH[1,4]
Density of beta.HH[1,4]
Trace of beta.HH[8,3]
Density of beta.HH[8,3]
Trace of beta.HH[2,4]
Density of beta.HH[2,4]


Trace of beta.HH[3,4]
Density of beta.HH[3,4]
Trace of beta.HH[11,4]
Density of beta.HH[11,4]
Trace of beta.HH[4,4]
Density of beta.HH[4,4]
Trace of beta.HH[12,4]
Density of beta.HH[12,4]
Trace of beta.HH[5,4]
Density of beta.HH[5,4]
Trace of beta.HH[13,4]
Density of beta.HH[13,4]
Trace of beta.HH[6,4]
Density of beta.HH[6,4]
Trace of beta.HH[14,4]
Density of beta.HH[14,4]
Trace of beta.HH[7,4]
Density of beta.HH[7,4]
Trace of beta.employee[1]
Density of beta.employee[1]
Trace of beta.HH[8,4]
Density of beta.HH[8,4]
Trace of beta.employee[2]
Density of beta.employee[2]
Trace of beta.HH[9,4]
Density of beta.HH[9,4]
Trace of beta.employee[3]
Density of beta.employee[3]
Trace of beta.HH[10,4]
Density of beta.HH[10,4]
Trace of beta.employee[4]
Density of beta.employee[4]

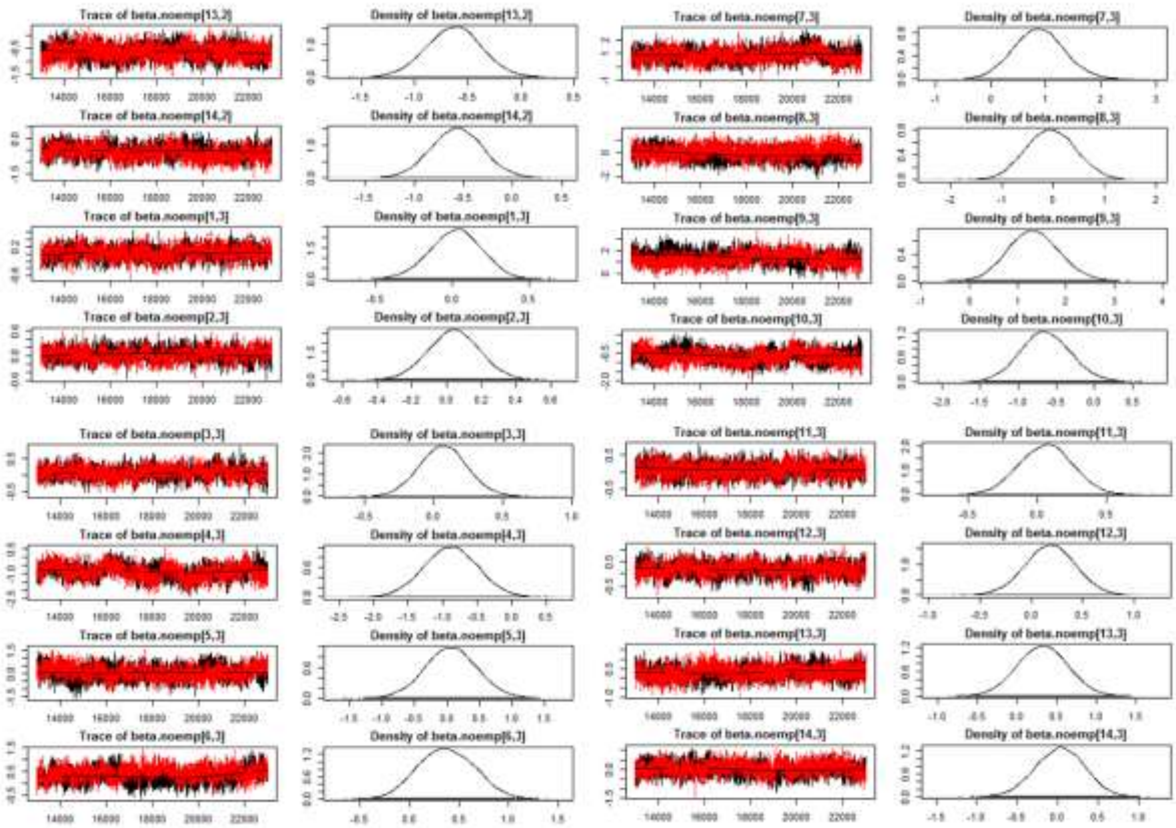

Trace of beta.employee[5]
Density of beta.employee[5]
Trace of beta.employee[13]
Density of beta.employee[13]
Trace of beta.employee[6]
Density of beta.employee[6]
Trace of beta.employee[14]
Density of beta.employee[14]
Trace of beta.employee[7]
Density of beta.employee[7]
Trace of beta.fulltime[1]
Density of beta.fulltime[1]
Trace of beta.employee[8]
Density of beta.employee[8]
Trace of beta.fulltime[2]
Density of beta.fulltime[2]
Trace of beta.employee[9]
Density of beta.employee[9]
Trace of beta.fulltime[3]
Density of beta.fulltime[3]
Trace of beta.employee[10]
Density of beta.employee[10]
Trace of beta.fulltime[4]
Density of beta.fulltime[4]
Trace of beta.employee[11]
Density of beta.employee[11]
Trace of beta.fulltime[5]
Density of beta.fulltime[5]
Trace of beta.employee[12]
Density of beta.employee[12]
Trace of beta.fulltime[6]
Density of beta.fulltime[6]


Trace of beta.fulltime[7]
Density of beta.fulltime[7]
Trace of beta.inwork[1]
Density of beta.inwork[1]
Trace of beta.fulltime[8]
Density of beta.fulltime[8]
Trace of beta.inwork[2]
Density of beta.inwork[2]
Trace of beta.fulltime[9]
Density of beta.fulltime[9]
Trace of beta.inwork[3]
Density of beta.inwork[3]
Trace of beta.fulltime[10]
Density of beta.fulltime[10]
Trace of beta.inwork[4]
Density of beta.inwork[4]
Trace of beta.fulltime[11]
Density of beta.fulltime[11]
Trace of beta.inwork[5]
Density of beta.inwork[5]
Trace of beta.fulltime[12]
Density of beta.fulltime[12]
Trace of beta.inwork[6]
Density of beta.inwork[6]
Trace of beta.fulltime[13]
Density of beta.fulltime[13]
Trace of beta.inwork[7]
Density of beta.inwork[7]
Trace of beta.fulltime[14]
Density of beta.fulltime[14]
Trace of beta.inwork[8]
Density of beta.inwork[8]

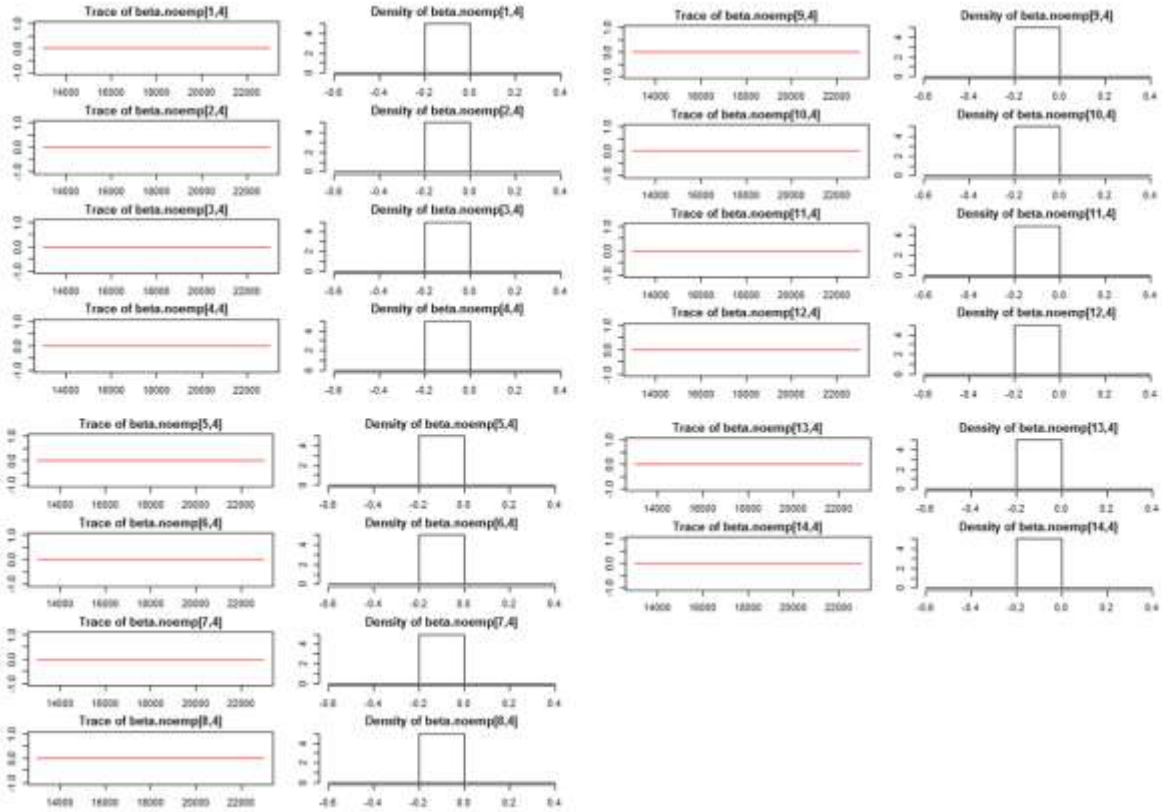

Trace of beta.inwork[9]
Density of beta.inwork[9]
Trace of beta.noemp[3,1]
Density of beta.noemp[3,1]
Trace of beta.inwork[10]
Density of beta.inwork[10]
Trace of beta.noemp[4,1]
Density of beta.noemp[4,1]
Trace of beta.inwork[11]
Density of beta.inwork[11]
Trace of beta.noemp[5,1]
Density of beta.noemp[5,1]
Trace of beta.inwork[12]
Density of beta.inwork[12]
Trace of beta.noemp[6,1]
Density of beta.noemp[6,1]
Trace of beta.inwork[13]
Density of beta.inwork[13]
Trace of beta.noemp[7,1]
Density of beta.noemp[7,1]
Trace of beta.inwork[14]
Density of beta.inwork[14]
Trace of beta.noemp[8,1]
Density of beta.noemp[8,1]
Trace of beta.noemp[1,1]
Density of beta.noemp[1,1]
Trace of beta.noemp[9,1]
Density of beta.noemp[9,1]
Trace of beta.noemp[2,1]
Density of beta.noemp[2,1]
Trace of beta.noemp[10,1]
Density of beta.noemp[10,1]



Trace of beta.noemp[11,1]
Density of beta.noemp[11,1]
Trace of beta.noemp[5,2]
Density of beta.noemp[5,2]
Trace of beta.noemp[12,1]
Density of beta.noemp[12,1]
Trace of beta.noemp[6,2]
Density of beta.noemp[6,2]
Trace of beta.noemp[13,1]
Density of beta.noemp[13,1]
Trace of beta.noemp[7,2]
Density of beta.noemp[7,2]
Trace of beta.noemp[14,1]
Density of beta.noemp[14,1]
Trace of beta.noemp[8,2]
Density of beta.noemp[8,2]
Trace of beta.noemp[1,2]
Density of beta.noemp[1,2]
Trace of beta.noemp[9,2]
Density of beta.noemp[9,2]
Trace of beta.noemp[2,2]
Density of beta.noemp[2,2]
Trace of beta.noemp[10,2]
Density of beta.noemp[10,2]
Trace of beta.noemp[3,2]
Density of beta.noemp[3,2]
Trace of beta.noemp[11,2]
Density of beta.noemp[11,2]
Trace of beta.noemp[4,2]
Density of beta.noemp[4,2]
Trace of beta.noemp[12,2]
Density of beta.noemp[12,2]



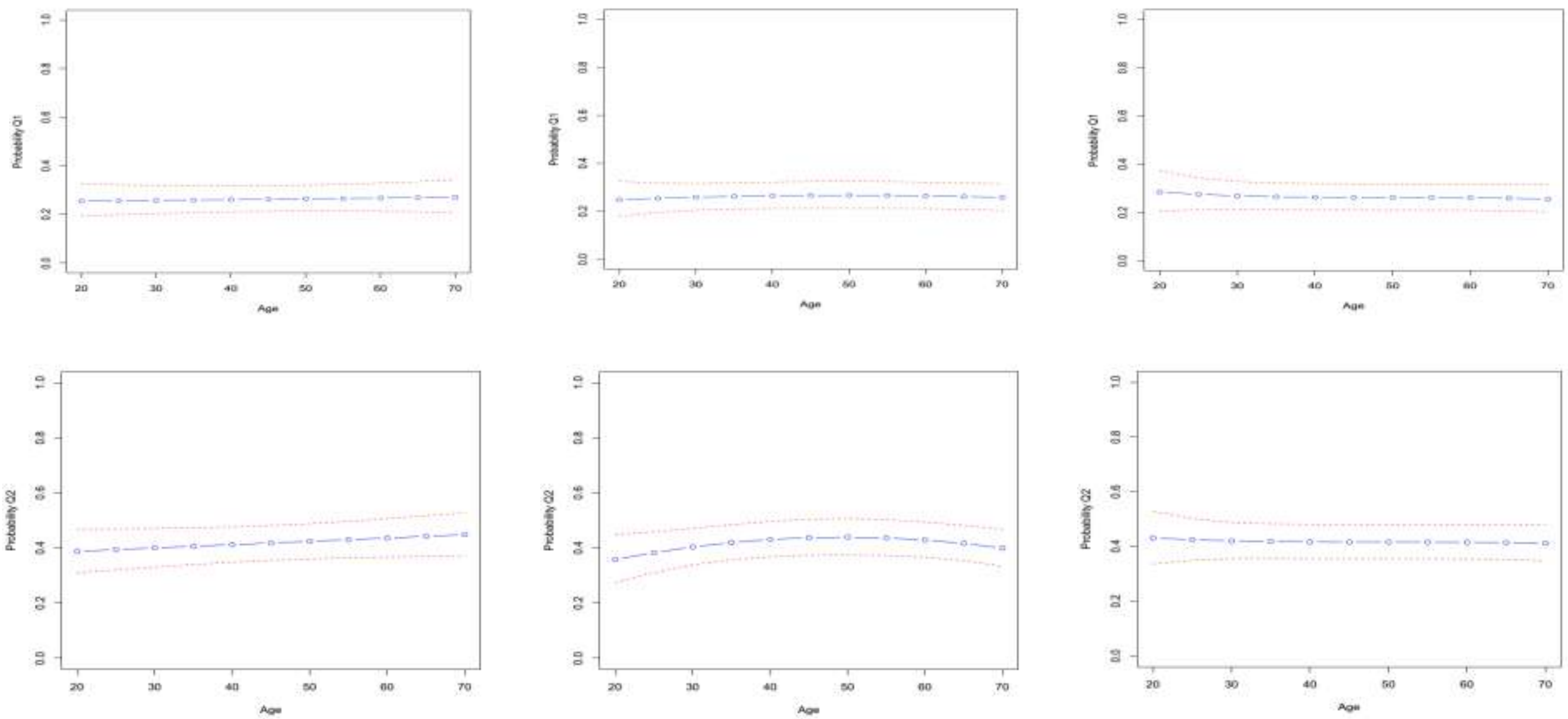
Trace of beta.noemp[13,2]
Density of beta.noemp[13,2]
Trace of beta.noemp[7,3]
Density of beta.noemp[7,3]
Trace of beta.noemp[14,2]
Density of beta.noemp[14,2]
Trace of beta.noemp[8,3]
Density of beta.noemp[8,3]
Trace of beta.noemp[1,3]
Density of beta.noemp[1,3]
Trace of beta.noemp[9,3]
Density of beta.noemp[9,3]
Trace of beta.noemp[2,3]
Density of beta.noemp[2,3]
Trace of beta.noemp[10,3]
Density of beta.noemp[10,3]
Trace of beta.noemp[3,3]
Density of beta.noemp[3,3]
Trace of beta.noemp[11,3]
Density of beta.noemp[11,3]
Trace of beta.noemp[4,3]
Density of beta.noemp[4,3]
Trace of beta.noemp[12,3]
Density of beta.noemp[12,3]
Trace of beta.noemp[5,3]
Density of beta.noemp[5,3]
Trace of beta.noemp[13,3]
Density of beta.noemp[13,3]
Trace of beta.noemp[6,3]
Density of beta.noemp[6,3]
Trace of beta.noemp[14,3]
Density of beta.noemp[14,3]

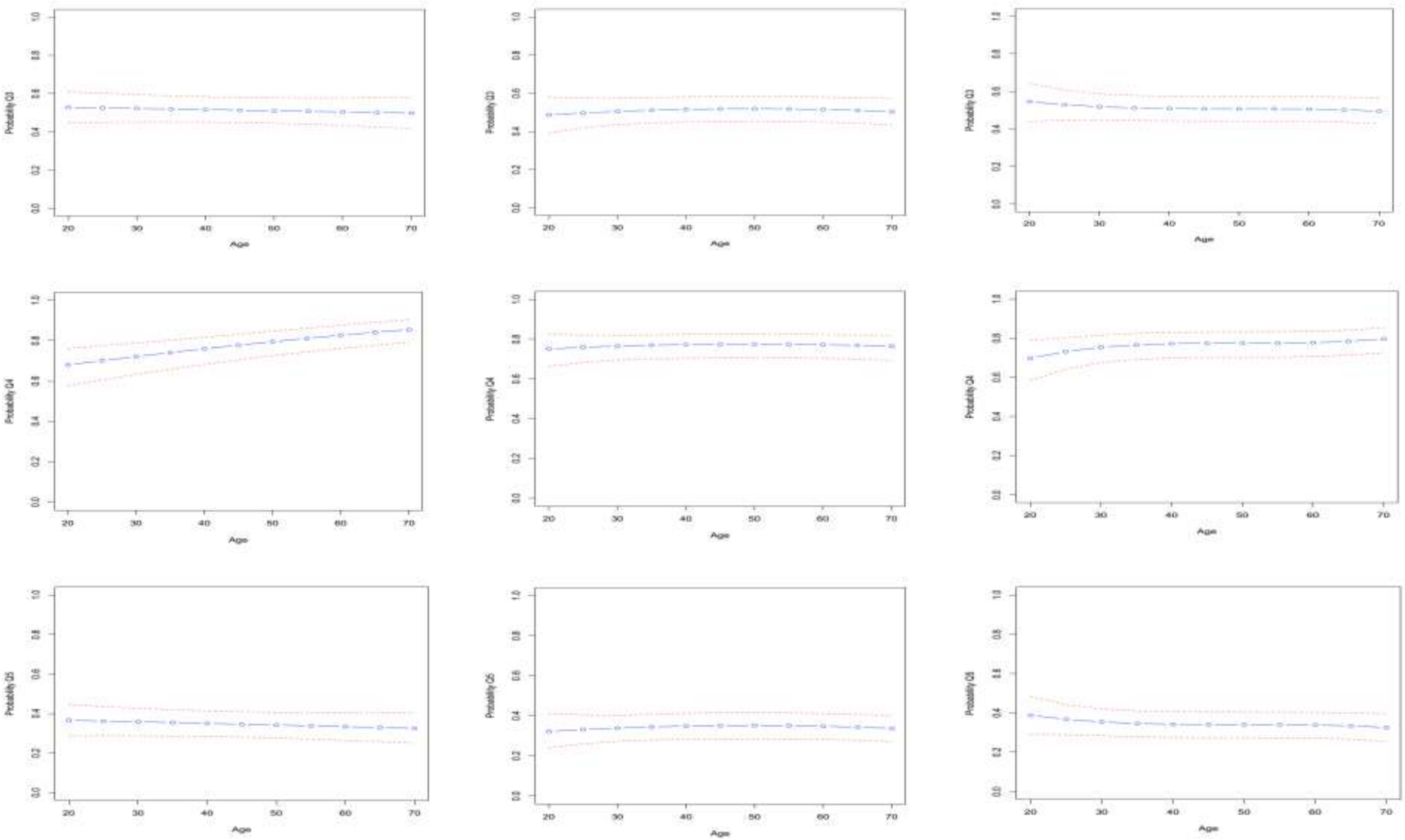

Trace of beta.noemp[1,4]
Density of beta.noemp[1,4]
Trace of beta.noemp[2,4]
Density of beta.noemp[2,4]
Trace of beta.noemp[3,4]
Density of beta.noemp[3,4]
Trace of beta.noemp[4,4]
Density of beta.noemp[4,4]
Trace of beta.noemp[5,4]
Density of beta.noemp[5,4]
Trace of beta.noemp[6,4]
Density of beta.noemp[6,4]
Trace of beta.noemp[7,4]
Density of beta.noemp[7,4]
Trace of beta.noemp[8,4]
Density of beta.noemp[8,4]
Trace of beta.noemp[9,4]
Density of beta.noemp[9,4]
Trace of beta.noemp[10,4]
Density of beta.noemp[10,4]
Trace of beta.noemp[11,4]
Density of beta.noemp[11,4]
Trace of beta.noemp[12,4]
Density of beta.noemp[12,4]
Trace of beta.noemp[13,4]
Density of beta.noemp[13,4]
Trace of beta.noemp[14,4]
Density of beta.noemp[14,4]

Appendix C

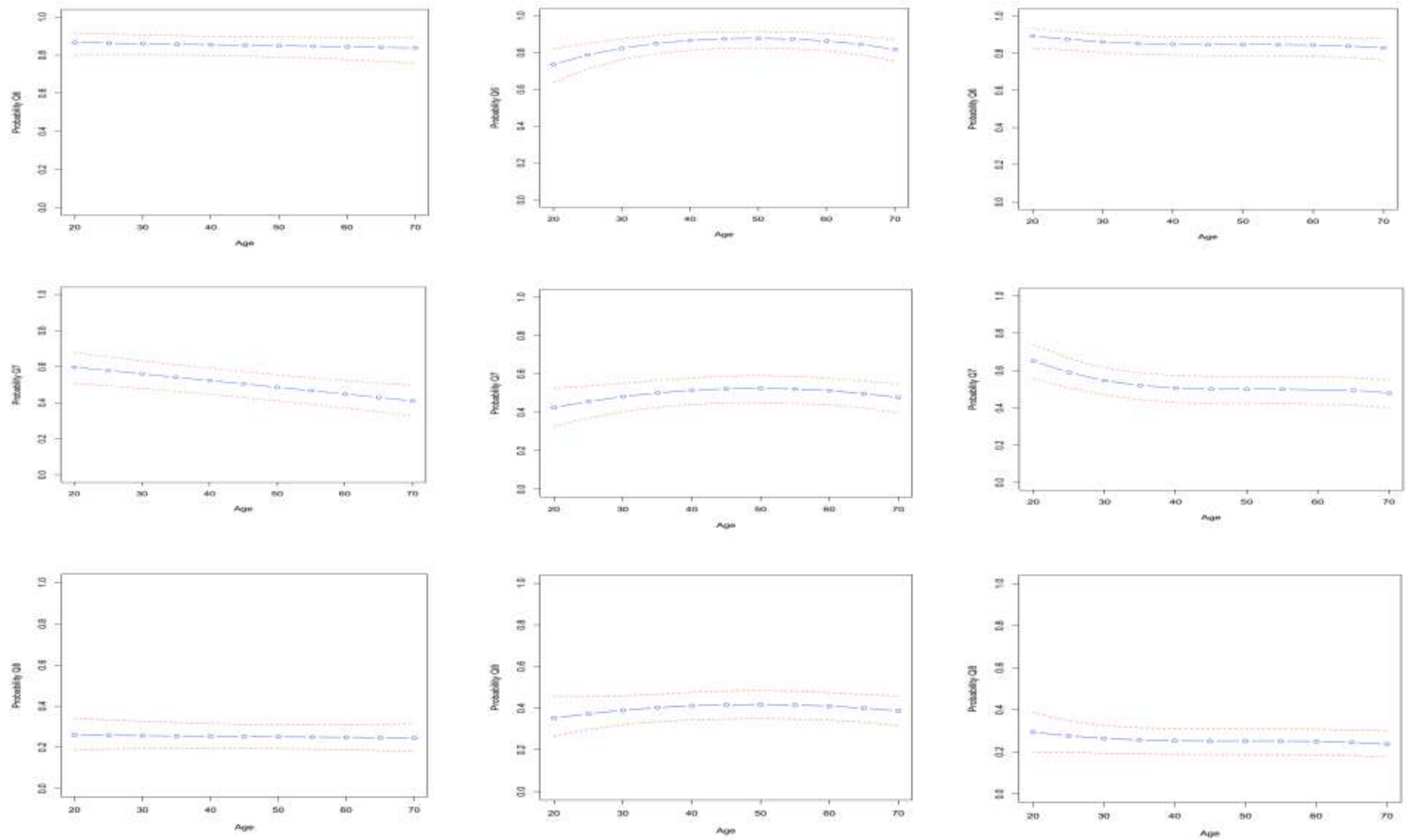
Probability of “agree” against age: a) linear, b) quadratic and c) cubic. Blue-posterior median, red-2.5% and 97.5% points.

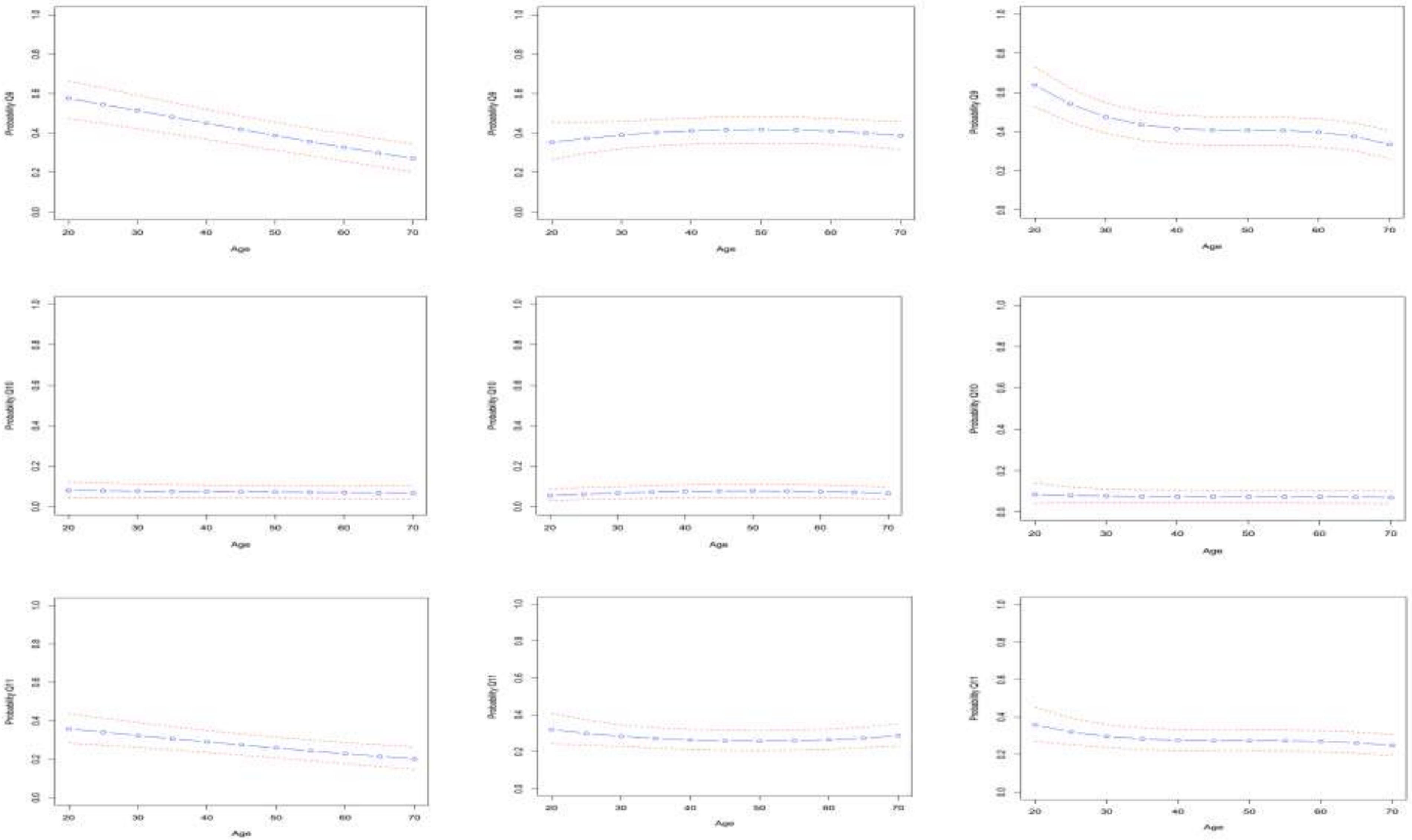
a) Linear b) Quadratic c) Cubic

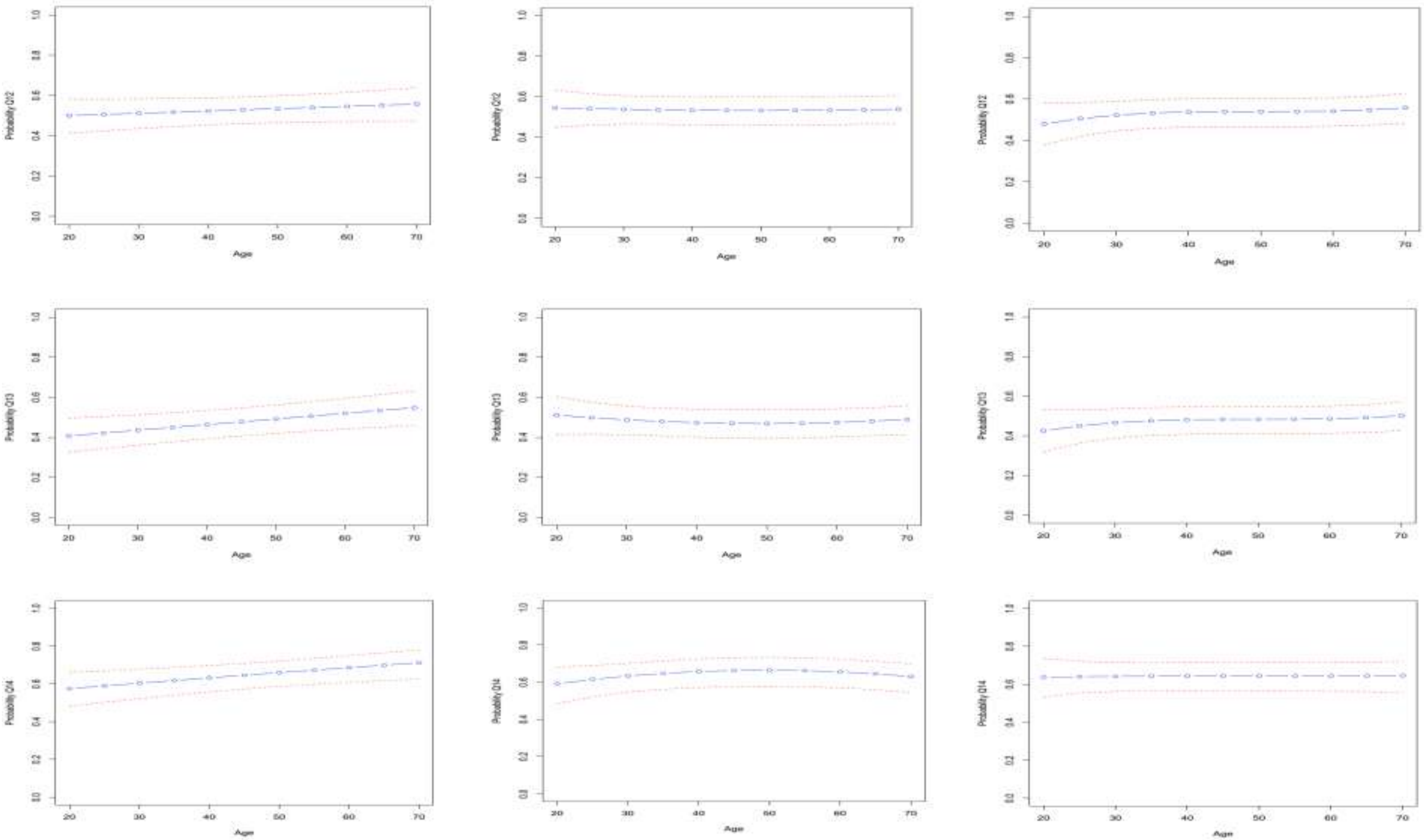




Probability Q3
Probability Q4
Probability Q5
Age


Probability Q6
Probability Q7
Probability Q8
Age


Probability Q9
Probability Q10
Probability Q11
Age



References

1. Abdi, H. (2007) 'Multiple Correspondence Analysis', *Encyclopedia of Measurement and Statistics*, 95(2), pp. 116-128.

2. Ajzen, I. (1987) 'Attitudes, Traits and Actions: Dispositional Prediction of Behavior in Personality and Social Psychology', *Advances in Experimental Social Psychology*, 20, pp. 1-63.

3. Alison Park, C.B., Elizabeth Clery, J.C.a. and Phillips, M. (2013) *BritishSocial Attitudes: the 30th Report*. London: NatCen Social Research.

4. Anable, J. (2005) ''Complacent Car Addicts' or 'Aspiring Environmentalists'? Identifying Travel Behaviour Segments Using Attitude Theory', *Transport Policy*, 12(1), pp. 65–78.

5. Anable, J. and 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 Part A: Policy and Practice*, 39(2-3), pp. 163-181.

6. Anable, J., Lane, B. and Kelay, T. (2006) *An Evidence Base Review of Public Attitudes to Climate Change and Transport Behaviour*. London: Department for Transport.

7. Andrews, G. (2012) *Just the ticket? Understanding the wide-ranging benefits of England's concessionary fares policy*. University of the West of England.

8. Aristodemou, K. (2014) 'New Regression Methods for Measures of Central Tendency', *Ph.D Thesis*, Brunel University.

9. Ayele, D., Zewotir, T. and Mwambi, H. (2014) 'Multiple correspondence analysis as a tool for analysis of large health surveys in African settings', *African Health Sciences*, 14(4), pp. 1036-1045.

10. Bamberg, S., Fujii, S., Friman, M. and Gärling, T. (2011) 'Behaviour theory and soft transport policy measures', *Transport Policy*, 18(1), pp. 228-235.

11. Banister, D. (2008) 'The Sustainable Mobility Paradigm', *Transport Policy*, 15(2), pp. 73-80.

12. Banister, D., Schwanen, T. and Anable, J. (2012) 'Introduction to the Special Section on Theoretical Perspectives on Climate Change Mitigation in Transport', *Transport Geography*, 24, pp. 467-470.

13. Barr, S., Gilg, A. and Shaw, G. (2011) 'Helping People Make Better Choices: Exploring the Behaviour Change Agenda for Environmental Sustainability', *Applied Geography*, 31, pp. 712-720.

14. Bartlett, M.S. (1950) 'Test of significance in factor analysis', *British Journal of Statistical Psychology*, 3(2), pp. 77-85.

15. Batool, Z. and Carsten, O. (2017) 'Self-reported dimensions of aberrant behaviours among drivers in Pakistan', *Transportation Research Part F: Traffic Psychology and Behaviour*, 47, pp. 176-186.

16. Becker, H., Loder, A., Schmid, B. and Axhausen, K.W. (2017) 'Modeling car-sharing membership as a mobility tool: A multivariate Probit approach with latent variables', *Travel Behaviour and Society*, 8, pp. 26-36.

17. Beirão, G. and Cabral, S.J. (2007) 'Understanding Attitudes Towards Public Transport and Private Car: A Qualitative Study', *Transport Policy*, 14(6), pp. 478-489.

18. Ben-Akiva, M.E. and Lerman, S.R. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA, USA: MIT Press.

19. Berger, J.O. (1985) *Statistical Decision Theory and Bayesian Analysis*. Second edn. New York: Springer-Verlag.

20. Bergstad, C.J., Gamble, A., Hagman, O., Polk, M., Gärling, T. and Olsson, L.E. (2011) 'Affective–symbolic and instrumental–independence psychological motives mediating effects of socio-demographic variables on daily car use', *Journal of Transport Geography*, 19(1), pp. 33-38.

21. Bernardo, J.M. and Smith, A.F.M. (1994) *Bayesian Theory*. New York: John Wiley & Sons.

22. Bertolini, L. and Martin, D. (2003) 'Mobility Environments and Network Cities', *Journal of Urban Design*, 8(1), pp. 27-43.

23. Bertram, D. (2006) *Likert Scales are the Meaning of Life*.

24. Bharadwaj, S., Ballare, S., Rohit and Chandel, M.K. (2017) 'Impact of congestion on greenhouse gas emissions for road transport in Mumbai metropolitan region', *Transportation Research Procedia*, 25, pp. 3538-3551.

25. Billot-Grasset, A., Amoros, E. and Hours, M. (2016) 'How Cyclist Behavior Affects Bicycle Accident Configurations?', *Transportation Research Part F: Traffic Psychology and Behaviour*, 41 (Part B), pp. 261-276.

26. Boardman, B. (2005) *Home Truths: A Low-Carbon Strategy to Reduce UK Housing Emissions by 80% by 2050*. London: Friends of the Earth.

27. Boarnet, M. and Crane, R. (2001a) 'The influence of land use on travel behavior: specification and estimation strategies', *Transportation Research Part A: Policy and Practice*, 35(9), pp. 823-845.

28. Boarnet, M.G. and Crane, R. (2001b) *Travel by Design: The Influence of Urban Form on Travel*. Oxford University Press.

29. Bolstad, W.M. and Curran, J.M. (2016) *Introduction to Bayesian Statistics*. 3rd edn. John Wiley & Sons.

30. Bösehans, G. and Walker, I. (2016) ''Daily Drags' and 'Wannabe Walkers' – Identifying dissatisfied public transport users who might travel more actively and sustainably', *Journal of Transport & Health*, 3(3), pp. 395-403.

31. Brady, J. and O'Mahony, M. (2016) 'Modelling charging profiles of electric vehicles based on real-world electric vehicle charging data', *Sustainable Cities and Society*, 26, pp. 203-216.

32. Briand, A.-S., Côme, E., Trépanier, M. and Oukhellou, L. (2017) 'Analyzing year-to-year changes in public transport passenger behaviour using smart card data', *Transportation Research Part C: Emerging Technologies*, 79, pp. 274-289.

33. Burton, E., Jenks, M. and Williams, K. (2013) *Achieving Sustainable Urban Form*. Taylor & Francis.

34. Cairns, S., Sloman, L., Newson, C., Anable, J., Kirkbride, A. and Goodwin, P. (2008) 'Smarter Choices: Assessing the Potential to Achieve Traffic Reduction Using 'Soft Measures'', *Transport Reviews*, 28(5), pp. 593-618.

35. Campbell, A.R., Ryley, T. and Thring, R. (2012) 'Identifying the early adopters of alternative fuel vehicles: A case study of Birmingham, United Kingdom', *Transportation Research Part A: Policy and Practice*, 46(8), pp. 1318-1327.

36. Cantrell, C.D. (2000) *Modern Mathematical Methods for Physicists and Engineers*. Cambridge University Press.

37. Cao, X., Mokhtarian, P.L. and Handy, S.L. (2009) 'The relationship between the built environment and nonwork travel: A case study of Northern California', *Transportation Research Part A: Policy and Practice*, 43(5), pp. 548-559.

38. Carlin, B.P. and Louis, T.A. (2000) *Bayes and Empirical Bayes Methods for Data Analysis*. 2nd. edn. London: Chapman & Hall.

39. Carrasco, J.-A. and Lucas, K. (2015) 'Workshop Synthesis: Measuring Attitudes; Quantitative and Qualitative Methods', *Transportation Research Procedia*, 11(165-171).

40. Cass, N. and Faulconbridge, J. (2016) 'Commuting Practices: New Insights into Modal Shift From Theories of Social Practice', *Transport Policy*, 45, pp. 1-14.

41. Cattell, R.B. (1996) 'The Scree Test for the Number of Factors', *Multivariate Behavioral Research* 1(2), pp. 245-276.

42. Chapman, L. (2007) 'Transport and Climate Change: A Review', *Journal of Transport Geography*, 15, pp. 354-367.

43. Chevalier, A. and Lantz, F. (2013) 'Personal Car, Public Transport and Other Alternatives? Predicting Potential Modal Shifts from Multinomial Logit Models and Bootstrap Confidence Intervals', *International Journal of Automotive Technology and Management*, 15(2), pp. 149-170.

44. Choo, S. and Mokhtarian, P. (2008) 'How Do People Respond to Congestion Mitigation Policies? A Multivariate Probit Model of The Individual Consideration of Three Travel-Related Strategy Bundles', *Transportation*, 35(2), pp. 145-163.

45. Clark, B., Lyons, G. and Chatterjee, K. (2016) 'Understanding the process that gives rise to household car ownership level changes', *Journal of Transport Geography*, 55, pp. 110-120.

46. Clifton, K. (2013) 'Collecting Qualitative and Quantitative Data on the Social Context of Travel Behaviour', in *Transport Survey Methods: Best Practice for Decision Making*. J. Zmud, M. Lee-Gosselin, M. Munizaga, and J. Carrasco edn. Bingley, UK: Emerald Press, pp. 441-448.

47. Clifton, K.J. and Handy, S.L. (2003) 'Qualitative Methods In Travel Behaviour Research', in *Transport Survey Quality and Innovation*. Peter R. Stopher edn. P. Jones, pp. 283 - 302.

48. Cohen-Blankshtain, G. and Rotem-Mindali, O. (2016) ' Key research themes on ICT and sustainable urban mobility', International Journal of Sustainable Transportation, 10, pp. 9-17.

49. Congdon, P. (2005) *Bayesian Models for Categorical Data*. Hoboken, New Jersey, United States: Wiley.

50. Congdon, P. (2006) *Bayesian Statistical Modelling*. 2nd. edn. Wiley-Blackwell.

51. Costa, P.S., Santos, N.C., Cunha, P., Cotter, J. and Sousa, N. (2013) 'The use of multiple correspondence analysis to explore associations between categories of qualitative variables in healthy ageing', *Journal of Aging Research*, pp. 1-12.

52. Costello, A.B. and Osborne, J.W. (2005) 'Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis', *Practical Assessment, Research, and Evaluation*, 10(7), pp. 1-9.

53. Cramer, D. (1998) *Fundamental Statistics for Social Research: Step-by-Step Calculations and Computer Techniques Using SPSS for Windows*. NY: Routledge New York.

54. Cramer, D. (2004) *The SAGE Dictionary of Statistics: A Practical Resource for Students in the Social Sciences*. SAGE Publications Inc.

55. Cronbach, L.J. (1951) 'Coefficient Alpha and the Internal Structure of Tests', *Psychometrika*, 16(3), pp. 297-334.

56. Currie, G. and Stanley, J. (2008) 'Investigating Links between Social Capital and Public Transport', *Transport Reviews*, 28(4), pp. 529-547.

57. Curtis, C. and Headicar, P. (1997) 'Targeting travel awareness campaigns: Which individuals are more likely to switch from car to other transport for the journey to work?', *Transport Policy*, 4(1), pp. 57-65.

58. Das, S. and Sun, X. (2015) ' Factor association with multiple correspondence analysis in vehicle-pedestrian crashes', *Transportation Research Record: Journal of the Transportation Research Board.* , 2519, pp. 95-103.

59. Dempster, A.P., Laird, N.M. and Rubin, D,B. (1977) 'Maximum likelihood from incomplete data via the EM algorithm', *Journal of the Royal Statistical Society*, Series B, 39, 1-38

60. Davison, L.J. and Knowles, R.D. (2006) 'Bus Quality Partnerships, Modal Shift and Traffic Decongestion', *Journal of Transport Geography*, 14(3), pp. 177-194.

61. DCLG (2012) *National Planning Policy Framework*. London: Department for Communities and Local Government.

62. DECC (2016) *Final UK Greenhouse Gas Emissions National Statistics*. London: Department of Energy and Climate Change.

63. DETR (2005) *Encouraging Walking: Advice to Local Authorities* London: Department of Environment, Transport, and the Regions.

64. DfT (2004) *Walking and Cycling: An Action Plan*. London: Department for Transport.

65. DfT (2009) *Low Carbon Transport: A Greener Future*. London: Department for Transport.

66. DfT (2010) *Sustainable Travel Guide*. London: Department for Transport.

67. DfT (2011) *Creating Growth, Cutting Carbon - Making Sustainable Local Transport Happen*. London: Department for Transport.

68. DfT (2014) *Transport Statistics Great Britain 2014*. London: Department for Transport.

69. DfT (2015) *Road Traffic Forecasts 2015*. London: Department for Transport.

70. DfT (2017) *Transport Statistics Great Britain: 2017*. London: Department for Transport.

71. DfT (2018) *Improving Air Quality: Reducing Emissions from Non-Road Mobile Machinery*. London.

72. Dickinson, J.E., Kingham, S., Copsey, S., Pearlman Hougie and J., D. (2003) 'Employer Travel Plans, Cycling and Gender: Will Travel Plan Measures Improve the Outlook for Cycling to Work in the UK?', *Transportation Research Part D: Transport and Environment*, 8(1), pp. 53-67.

73. Dissanayake, D. and Morikawa, T. (2001) 'A Combined RP/SP Nested Logit Model To Investigate Household Decisions On Vehicle Usage, Mode Choice and Trip-chaining in Developing Countries', *Journal of the Eastern Asia society for Transportation Studies*, 4(2).

74. Dissanayake, D. and Morikawa, T. (2010) 'Investigating household vehicle ownership, mode choice and trip sharing decisions using a combined revealed preference/stated preference Nested Logit model: case study in Bangkok Metropolitan Region', *Journal of Transport Geography*, 18(3), pp. 402-410.

75. Doane, D.P. and Seward, L.E. (2011) 'Measuring Skewness: A Forgotten Statistic?', *Journal of Statistics Education*, 19(2).

76. DoH (2010) *Active Travel Strategy*. London: Department of Health.

77. Domarchi, C., Tudela, A. and Gonzalez, A. (2008) 'Effect of Attitudes, Habit and Affective Appraisal on Mode Choice: An Application to University Workers', *Transportation*, 35, pp. 585–599.

78. Domencich, T. and McFadden, D. (1975) *Urban Travel Demand: A Behavioral Analysis*. Oxford: North-Holland.

79. DREDF (2014) *FTA Research Report: Federal Transit Administration*. Berkeley, CA: Disability Rights Education & Defense Fund

80. Eriksson, L., Friman, M. and Gärling, T. (2008) 'Stated Reasons for Reducing Work-Commute by Car', *Transportation Research Part F: Traffic Psychology and Behaviour*, 11(6), pp. 427-433.

81. Everitt, B.S., Landau, S., Leese, M. and Stahl, D. (2011) *Cluster Analysis*. John Wiley & Sons, Ltd.

82. Farber, S., Bartholomew, K., Li, X., Páez, A. and Nurul Habib, K.M. (2014) 'Assessing social equity in distance based transit fares using a model of travel behavior', *Transportation Research Part A: Policy and Practice*, 67, pp. 291-303.

83. Fatmi, M.R. and Habib, M.A. (2017) 'Modelling mode switch associated with the change of residential location', *Travel Behaviour and Society*, 9, pp. 21-28.

84. Ferdous, N., Eluru, N., Bhat, C.R. and Meloni, I. (2010) 'A multivariate ordered-response model system for adults' weekday activity episode generation by activity purpose and social context', *Transportation Research Part B: Methodological*, 44(8), pp. 922-943.

85. Field, A.P. (2009) *Discovering Statistics Using SPSS*. 3rd edn. London: SAGE publications.

86. Franco, G.D. (2015) 'Multiple correspondence analysis: one only or several techniques?', *International Journal of Methodology*, 50(3), pp. 1299-1315.

87. Fujii, S. and Garling, T. (2003) 'Application of Attitude Theory for Improved Predictive Accuracy of Stated Preference Methods in Travel Demand Analysis', *Transportation Research Part A*, 37(4), pp. 389-402.

88. Fuller, D., Gauvin, L., Kestens, Y., Morency, P. and Drouin, L. (2013) 'The potential modal shift and health benefits of implementing a public bicycle share program in Montreal, Canada', *International Journal of Behavioral Nutrition and Physical Activity*, 10, pp. 1-6.

89. Garling, T., Gillholm, R. and Garling, A. (1998) 'Reintroducing Attitude Theory in Travel Behavior Research: The Validity of an Interactive Interview Procedure to Predict Car Use', *Transportation*, 25(2), pp. 129- 146.

90. Gärling, T. and Schuitema, G. (2007) 'Travel Demand Management Targeting Reduced Private Car Use: Effectiveness, Public Acceptability and Political Feasibility', *Journal of Social Issues*, 63(1), pp. 139-153.

91. Gary Haq, John Whitelegg, Steve Cinderby & Anne Owen (2008) The use of personalised social marketing to foster voluntary behavioural change for sustainable travel and lifestyles, *Local Environment*, 13(7), pp. 549-569.

92. Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2013) *Bayesian Data Analysis*. 3rd. edn. Taylor & Francis.

93. Geng, J., Long, R., Chen, H., Yue, T., Li, W. and Li, Q. (2017) 'Exploring Multiple Motivations on Urban Residents' Travel Mode Choices: An Empirical Study from Jiangsu Province in China.', *Sustainability* 9(136).

94. Geraghty, D. and O'Mahony, M. (2015) 'Urban Noise Analysis Using Multinomial Logistic Regression', *Journal of Transportation Engineering*, 142(6).

95. Giannopoulos, G.A. (2004) 'The application of information and communication technologies in transport', *European Journal of Operational Research*, 152(2), pp. 302-320.

96. Gim, T.-H.T. (2017) 'Investigating travel utility elements in association with travel time and mode choice: the case of Seoul, South Korea', *Transportation Planning and Technology*, 40(6), pp. 641-660.

97. Goh, M. (2002) 'Congestion management and electronic road pricing in Singapore', *Journal of Transport Geography*, 10, pp. 29-38.

98. Goldman, T. and Gorham, R. (2006) 'Sustainable Urban Transport: Four Innovative Directions', *Technology in Society*, 28(1-2), pp. 261-273.

99. Gorsuch, R.L. (1983) 'Factor Analysis', *Work Study*, 42(1), pp. 10-11.

100. Graham-Rowe, E., Skippon, S., Gardner, B. and Abraham, C. (2011) 'Can We Reduce Car Use and, if so, How? A Review of Available Evidence', *Transportation Research Part A: Policy and Practice*, 45(5), pp. 401-418.

101. Greenacre, M. and Blasius, J. (2006) *Multiple Correspondence Analysis and Related Methods*. London: Chapman & Hall/CRC.

102. Greene, W.H. (2003) *Econometric Analysis* 5th edn. New York: Prentice-Hall Press.

103. Grison, E., Gyselinck, V. and Burkhardt, J.-M. (2015) 'Exploring Factors Related to Users' Experience of Public Transport Route Choice: Influence of Context and User Profiles', *Cognition, Technology & Work*, 18(2), pp. 287-301.

104. Grosvenor, T. (2000) 'Qualitative Research in the Transport Sector, Resource Paper for the Workshop on Qualitative or Quantitative Methods', *Proceedings of an International Conference on Transport Survey and Innovation*. Grainau, Germany, May 24-30, 1997. Transportation Research E-Circular.

105. Gwenaelle, F., Marine, C., Olivier, F., Francois, L.P., Jerome, B. and Sebastien, L. (2010) *Enquirer: A Package Dedicated to Questionnaires*. R Package Version 0.10. [Online]. Available at: https://CRAN.R-project.org/package=EnQuireR.

106. Habibian, M. and Kermanshah, M. (2011) 'Exploring the Role of Transportation Demand Management Policies' Interactions', *Scientia Iranica*, 18(5), pp. 1037-1044.

107. Habibian, M. and Kermanshah, M. (2013) 'Coping with Congestion: Understanding the Role of Simultaneous Transportation Demand Management Policies on Commuters', *Transport Policy*, 30, pp. 229-237.

108. Hair, J.F., Black, B., Babin, B., Anderson, E.R. and Tatham, R.L. (2006) *Multivariate Data Analysis* 6th edn. Upper Saddle River, N.J. : Pearson Prentice Hall.

109. Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2010) *Multivariate Data Analysis: A Global Perspective*. 7th edn. London: Pearson Education.

110. Hamersma, M., Heinen, E., Tillema, T. and Arts, J. (2017) 'New highway development in the Netherlands: A residents' perspective', *Transportation Research Part D: Transport and Environment*, 51, pp. 326-339.

111. Hanson, S. and Schwab, M. (1986) 'Describing Disaggregate Flows: Individual and Household Activity Patterns', *The Geography of Urban Transportation*, pp. 154-178.

112. Harris, I., Wang, Y. and Wang, H. (2015) 'ICT in multimodal transport and technological trends: Unleashing potential for the future', *International Journal of Production Economics*, 159, pp. 88-103.

113. Heinen, E. and Handy, S. (2012) 'Similarities in Attitudes and Norms and the Effect on Bicycle Commuting: Evidence from the Bicycle Cities Davis and Delft', *International Journal of Sustainable Transportation* 6(5), pp. 257-281.

114. Heinen, E., Maat, K. and Wee, B.v. (2011) 'The Role of Attitudes Toward Characteristics of Bicycle Commuting on the Choice to Cycle to Work Over Various Distances', *Transportation Research Part D: Transport and Environment*, 16(2), pp. 102-109.

115. Hensher, D.A. (2007) 'Sustainable public transport systems: Moving towards a value for money and network-based approach and away from blind commitment', *Transport Policy*, 14(1), pp. 98-102.

116. Hensher, D.A. (2008) 'Climate Change, Enhanced Greenhouse Gas Emissions and Passenger Transport – What Can We Do to Make a Difference?', *Transportation Research Part D: Transport and Environment*, 13(2), pp. 95-111.

117. Hickman, R. and Banister, D. (2007) 'Looking Over the Horizon: Transport and Reduced CO_2 Emissions in the UK by 2030', *Transport Policy*, 14(5), pp. 377-387.

118. Hickman, R., D.Bonilla, Givoni, M. and Banister, D. (2015) *International Handbook on Transport and Development*. Edward Elgar Publishing.

119. Ho, C., Mulley, C., Tsai, C.-H., Ison, S. and Wiblin, S. (2017) 'Area-Wide Travel Plans—Targeting Strategies for Greater Participation in Green Travel Initiatives: A Case Study of Rouse Hill Town Centre, NSW Australia', *Transportation*, 44(2), pp. 325–352.

120. Hotelling, H. (1933) 'Analysis of a complex of statistical variables into principal components', *Journal of Education Psychology*, 24(6), p. 417.

121. Howarth, C. (2006) 'How Social Representations of Attitudes Have Informed Attitude Theories: The Consensual and the Reified', *Theory & Psychology*, 16(5), pp. 691-714.

122. Hull, A. (2008) 'Policy integration: What will it take to achieve more sustainable transport solutions in cities?', *Transport Policy*, 15(2), pp. 94-103.

123. Hunecke, M., Haustein, S., Böhler, S. and Grischkat, S. (2010) 'Attitude-Based Target Groups to Reduce the Ecological Impact of Daily Mobility Behavior.', *Environment Behaviour*, 42, pp. 3-43.

124. IBM (2013) 'IBM SPSS Statistics for Windows', in Armonk, NY: IBM Corporation.

125. Jakobsson, C. (2004) 'Accuracy of household planning of car use: comparing prospective to actual car logs', *Transportation Research Part F: Traffic Psychology and Behaviour*, 7(1), pp. 31-42.

126. Jalayer, M. and Zhou, H. (2017) 'A Multiple Correspondence Analysis of At-Fault Motorcycle-Involved Crashes In Alabama', *Journal of Advanced Transportation*, 50, pp. 2089-2099.

127. Jensen, M. (1999) 'Passion and Heart in Transport-A Sociological Analysis on Transport Behaviour', *Transport Policy*, 6(1), pp. 19-33.

128. Jin, X., Hossan, S., Asgari, H. and Shams, K. (2018) 'Incorporating attitudinal aspects in roadway pricing analysis', *Transport Policy*, 62, pp. 38-47.

129. Jochem, P., Rothengatter, W. and Schade, W. (2016) 'Climate change and transport', *Transportation Research Part D: Transport and Environment*, 45, pp. 1-3.

130. Johnson, R.A. and Wichern, D.W. (2007) 'Applied Multivariate Corresponding Analysis', in *Applied Multivariate Statistical Analysis*. 6th edn. Upper Saddle River, NJ, USA: Prentice-Hall.

131. Juho, H., Koivisto, J. and Sarsa, H. (2014) 'Does Gamification Work-A Literature Review of Empirical Studies on Gamification', *47th Hawaii International Conference on System Sciences (HICSS)*. Manoa, HI, USA.

132. Kaiser, H.F. (1960) 'The application of electronic computers to factor analysis', *Educational and Psychological Measurement*, 20, pp. 141-151.

133. Kamruzzaman, M., Shatu, F., Hine, J. and GavinTurrell (2015) 'Commuting Mode Choice Intransi Toriented Development: Disentangling The Effects of Competitive Neighbourhoods,Trave Lattitudes, and Self-Selection', *Transport Policy*, 42, pp. 187–196.

134. Kamruzzaman, M., Washington, S., Baker, D., Brown, W., Giles-Corti, B. and Turrell, G. (2016) 'Built Environment Impacts on Walking for Transport in Brisbane, Australia', *Transportation*, 43(1), pp. 53-77.

135. Kandt, J., Rode, P., Hoffmann, C., Graff, A. and Smith, D. (2015) 'Gauging interventions for sustainable travel: A comparative study of travel attitudes in Berlin and London', *Transportation Research Part A: Policy and Practice*, 80, pp. 35-48.

136. Kingham, S., Dickinson, J. and Copsey, S. (2001) 'Travelling to Work: Will People Move Out of Their Cars', *Transport Policy*, 8(2), pp. 151-160.

137. Kinnear, P. and Gray, C. (2000) *SPSS for Windows Made Simple*. 1st edn. Abingdon: Psychology Press.

138. Klieštik, T., Kočišová, K. and Mišanková, M. (2015) 'Logit and Probit Model used For Prediction of Financial Health of Company', *Procedia Economics and Finance*, 23, pp. 850 - 855.

139. Koetse, M.J. and Rietveld, P. (2009) 'The impact of climate change and weather on transport: An overview of empirical findings', *Transportation Research Part D: Transport and Environment*, 14, pp. 205–221.

140. Kollmuss, A. and Agyeman, J. (2002) 'Mind the Gap: Why do People Act Environmentally and What are the Barriers to Pro-Environmental Behavior?', *Environmental Education Research*, 8(3), pp. 239-260.

141. Kruschke, J.K. (2015) *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. 2nd. edn. Academic Press.

142. Kumar, R. (2011) *Research Methodology: A Step-By-Step Guide for Beginners*. 3rd edn. SAGE Publication.

143. Kwak, C. and Clayton-Matthews, A. (2002) 'Multinomial Logistic Regression', *Nursing Research*, 51(6), pp. 404-410.

144. Le Quéré, C., Andrew, R.M., Canadell, J.G. and Sitch, S. (2016) 'Global carbon budget 2016', *Earth System Science Data*, 8, pp. 1-145.

145. Lee, P.M. (2012) *Bayesian Statistics: An Introduction*. 4th. edn. John Wiley & Sons.

146. Lehman, A., O'Rourke, N., Hatcher, L. and Stepanski, E.J. (2005) *JMP For Basic Univariate And Multivariate Statistics: A Step-by-step Guide*. Cary, NC: SAS Institute.

147. LeRoux, B. and Rouanetm, H. (2010) *Multiple Correspondence Analysis*. United State: SAGE Publications.

148. Li, Z., Wang, W., Yang, C. and Ragland, D.R. (2013) 'Bicycle Commuting Market Analysis using Attitudinal Market Segmentation Approach', *Transport Policy*, 47, pp. 56-68.

149. Liberti, L., Lavor, C., Maculan, N. and Mucherino, A. (2014) 'Euclidean Distance Geometry and Applications', *Society for Industrial and Applied Mathematics Review*, 56(1), pp. 3-69.

150. Likert, R. (1932) 'A Technique for the Measurement of Attittudes', *Archives of Psychology*, 22(142), pp. 1-55.

151. Linardakis, M. and Dellaportas, P. (2003) 'Assessment of Athens's metro passenger behaviour via a multiranked probit model', *Applied Statistics*, 52(2), pp. 185-200.

152. Line, T., Chatterjee, K. and Lyons, G. (2010) 'The Travel Behaviour Intentions of Young People in the Context of Climate Change', *Journal of Transport Geography*, 18(2), pp. 238–246.

153. Line, T., Chatterjee, K. and Lyons, G. (2012) 'Applying Behavioural Theories to Studying the Influence of Climate Change on Young People's Future Travel

Intentions', *Transportation Research Part D: Transport and Environment*, 17(3), pp. 270-276.

154. Litman, T. (2003) 'The Online TDM Encyclopedia: Mobility Management Information Gateway', *Transport Policy*, 10(3), pp. 245-249.

155. Litman, T. (2009) 'Sustainable Transportation and TDM', *Online TDM Encyclopedia. Victoria Transport Policy Institue*, 1(4), pp. 9-16.

156. Litman, T. and Fitzroy, S. (2009) 'Evaluating Mobility Management Traffic Safety Impacts', *Injury Prevention*, 15(6).

157. Liu, C., Susilo, Y.O. and Karlström, A. (2016) 'Estimating changes in transport CO2 emissions due to changes in weather and climate in Sweden', *Transportation Research Part D: Transport and Environment*, 49, pp. 172–187.

158. Loewenthal, K. (2001) *An Introduction to Psychological Tests and Scales*. 2nd. edn. New York, USA: Routledge.

159. Lomas, K.J., Bell, M.C., Firth, S.K., Gaston, K.J., Goodman, P., Leake, J.R., Namdeo, A., Rylatt, M., Allinson, D., Davies, Z.G., Edmondson, J., Galatioto, F., Guo, L., Hill, G., Irvine, K., Taylor, S.C. and Tiwary, A. (2010) 'The carbon footprint of UK cities : 4M: Measurement, modelling, mapping and management', in Gossop, C. (ed.) *Sustainable city / developing world*. The Hague, Netherlands: ISOCARP.

160. Long, J.S. (2012) *Regression models for nominal and ordinal outcomes*. In Best and Wolf edn. Thousand Oaks, CA: Sage Publications.

161. Lucas, K. (2013) 'Qualitative Methods in Transport Research: The 'Action Research' Approach ', *Transport Survey Methods: Best Practice for Decision Making*, pp. 427 - 440.

162. Mackett, R.L. (2000) 'How to Reduce the Number of Short Trips by Car', *European Transport Conference*. Cambridge, UK. pp. 87-98.

163. Marsden, G. and Rye, T. (2010) 'The Governance of Transport and Climate Change', *Journal of Transport Geography*, 18, pp. 669–678.

164. Martin, A.D., Quinn, K.M. and Park, J.H. (2011) 'MCMC pack: Markov Chain Monte Carlo in R', *Journal of Statistical Software*, 42(9), pp. 1-21.

165. Mikiki, F. and Papaioannou, P. (2012) 'Investigating Pro-Environmental and Active Travel Behaviour for Successful Sustainable Travel Promotion', *Procedia-Social and Behavioural Sciences*, 48, pp. 1424-1433.

166. Milioti, C.P., Karlaftis, M.G. and Akkogiounoglou, E. (2015) 'Traveler perceptions and airline choice: A multivariate probit approach', *Journal of Air Transport Management*, 49, pp. 46-52.

167. Miralles-Guasch, C., Melo, M.M. and Marquet, O. (2016) 'A Gender Analysis of Everyday Mobility in Urban and Rural Territories: from Challenges to Sustainability', *Gender, Place & Culture*, 23(3), pp. 398-417.

168. Molin, E., Mokhtarian, P. and Kroesen, M. (2016) 'Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers', *Transportation Research Part A: Policy and Practice*, 83, pp. 14-29.

169. Mondschein, A. and Parkany, E. (2017) 'Hitting the sweet spot: Variability in commute lengths and vehicle emissions across a diverse state', *Transportation Research Part D: Transport and Environment*, 54, pp. 348-359.

170. Morey, E. (2006) 'Using Angler Characteristics and Attitudinal Data to Identify Environmental Preference Classes: A Latent-Class Model', *Environmental and Resource Economics*, 34(1), pp. 91-115.

171. Muradov, N. (2013) 'Decarbonization at crossroads: the cessation of the positive historical trend or a temporary detour?', *Energy Environment Science*, 6, pp. 1060-1073.

172. Murtagh, F. (2007) 'Multiple correspondence analysis and related methods', *Psychometrika*, 72(2), pp. 275-277.

173. Musti, S., Kortum, K. and Kockelman, K.M. (2011) 'Household energy use and travel: Opportunities for behavioral change', *Transportation Research Part D: Transport and Environment*, 16(1), pp. 49-56.

174. Myers, J.L. and Well, A.D. (2003) *Research Design and Statistical Analysis*. 2nd. edn. Lawrence Erlbaum.

175. Nilsson, M. and Küller, R. (2000) 'Travel Behaviour and Environmental Concern', *Transportation Research Part D: Transport and Environment*, 5(3), pp. 211–234.

176. Nkurunziza, A., Zuidgeest, M., Brussel, M. and Maarseveen, M. (2012) 'Examining the Potential for Modal Change: Motivators and Barriers for Bicycle Commuting in Dar-es-Salaam', *Transport Policy*, 24, pp. 249–259.

177. Noland, R.B. (2001) 'Relationships between highway capacity and induced vehicle travel', *Transportation Research Part A*, 35(47-72).

178. Noland, R.B. and Polak, J.W. (2002) 'Travel Time Variability: A Review of Theoretical and Empirical Issues', *Transport Reviews*, 22(1), pp. 39-54.

179. Nordlund, A.M. and Garvill, J. (2003) 'Effects of Values, Problem Awareness, and Personal Norm on Willingness to Reduce Personal Car Use', *Journal of Environmental Psychology*, 23, pp. 339-347.

180. Nurdden, A., Rahmat, R.A.O.K. and Ismail, A. (2007) 'Effect of Transportation Policies on Modal Shift from Private Car to Public Transport in Malaysia', *Journal of Applied Sciences*, 7, pp. 1014-1018.

181. Nutsugbodo, R.Y., Amenumey, E.K. and Mensah, C.A. (2018) 'Public transport mode preferences of international tourists in Ghana: Implications for transport planning', *Travel Behaviour and Society*, 11, pp. 1-8.

182. O'Flaherty, C.A. (1997) 'Traffic planning strategies', in *Traffic Planning and Traffic Engineering*. Oxford, UK: Taylor & Francis.

183. Okushima, M. (2015) 'Simulating Social Influences on Sustainable Mobility Shifts for Heterogeneous Agents', *Transportation*, 42(5), pp. 827-855.

184. Onunwor, E., Miller, I. and Frank-Ito, D. (2014) *Essentials of Industrial Mathematics*. 2nd edn. Morrisville, North Carolina: Lulu.

185. Orloff, J. and Bloom, J. (2014) *Comparison of frequentist and Bayesian inference*. MIT.

186. Ortúzar, J.d.D. and Willumsen, L. (2011) *Modelling Transport*. 4th edn. London: John Wiley & Son.

187. Ory, D. and Mokhtarian, P. (2009) 'Modeling the Structural Relationships Among Short-Distance Travel Amounts, Perceptions, Affections, and Desires', *Transportation Research Part A: Policy and Practice*, 43(1), pp. 26-43.

188. Osborne, J.W. and Castello, A.B. (2005) 'Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis', *Practical Assessment, Research & Evaluation*, 10(7), pp. 1–9.

189. Outwater, M.L., Castleberry, S., Shiftan, Y., Ben-Akiva, M., Zhou, Y.S. and Kuppam, A. (2003) 'Attitudinal market segmentation approach to mode choice and ridership forecasting – structural equation modeling', *Transportation Research Record*, 1854, pp. 32-42.

190. Papadimitriou, E., Lassarre, S. and Yannis, G. (2017) 'Human factors of pedestrian walking and crossing behaviour', *Transportation Research Procedia*, 25, pp. 2002-2015.

191. Parkany, E., Gallagher, R. and Viveiros, P. (2004) 'Are Attitudes Important in Travel Choice?', *Transport Research Record*, 1894.

192. Pearson, K. (1901) 'On Lines and Planes of Closest Fit to Systems of Points in Space', *Philosophical Magazine Series 6*, 2(11), pp. 559–572.

193. Pendyala, R. and Bricka, S. (2006) 'Collection and Analysis of Behavioural Process Data: Challenges and Opportunities', in *Travel Survey Methods: Quality and Future Directions.* In P.R. Stopher and C. Stecher edn. Oxford, UK: Elsevier, pp. 511-530.

194. Peterson, R.A. (1994) 'A Meta-Analysis of Cronbach's Coefficient Alpha', *Journal of Consumer Research*, pp. 381-391.

195. Plummer, M. (2003) *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling*.

196. Plummer, M. (2016) *rjags: Bayesian Graphical Models using MCMC*. R Package Version 4-6. [Online]. Available at: http://mcmc-jags.sourceforge.net.

197. Posavac, E.J. (2015) *Program Evaluation: Methods and Case Studies*. 8th edn. Routledge.

198. Prillwitz, J. and Barr, S. (2011) 'Moving Towards Sustainability? Mobility Styles, Attitudes and Individual Travel Behaviour', *Journal of Transport Geography*, 19(6), pp. 1590-1600.

199. Raubenheimer, J. (2004) 'An item selection procedure to maximise scale reliability and validity', *Journal of Industrial Psychology*, 30(4), pp. 59-64.

200. Rayner, T., Russel, D. and Lorenzoni, I. (2008) ''It's Demand, Stupid': The Failure and Future of Integrating Climate Change Concerns into UK Transport Policy', *British Politics*, 3, pp. 373–389.

201. Razali, N.M. and Wah, Y.B. (2011) 'Power comparisons of Shapiro-Wilk , Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests', *Journal of Statistical Modeling and Analytics*, 2(1), pp. 21-33.

202. Redman, L., Friman, M., Gärling, T. and Hartig, T. (2013) 'Quality Attributes of Public Transport that Attract Car Users: A Research Review', *Transport Policy*, 25, pp. 119-127.

203. Rietveld, P. and Stough, R.R. (2005) *Barriers to Sustainable Transport: Institutions, Regulation and Sustainability*. London: Routledge.

204. Robert, C.P. (2001) *The Bayesian Choice*. 2nd. edn. New York: Springer-Verlag.

205. Rodriguez Cote, J.D. and Diana, M. (2017) 'Exploring the benefits of a traveller clustering approach based on multimodality attitudes and behaviours', *Transportation Research Procedia*, 25, pp. 2552-2565.

206. Ruhe, A.-K., Wager, J., Hirschfeld, G. and Zernikow, B. (2016) 'Household income determines access to specialized pediatric chronic pain treatment in Germany', *BMC Health Services Research*, 16(1), p. 140.

207. Rumsey, D. (2007) *Intermediate Statistics for Dummies*. Indianapolis, Indiana: Wiley Publishing, Inc.

208. Ryley, T. (2001) 'Translating Cycling Policy into Cycling Practice', *World Transport Policy and Practice*, 7, pp. 38-43.

209. Saleh, W. (2007) 'Success and failure of travel demand management: is congestion charging the way forward?', *Special Issue of Journal of Transportation Research Part A*, 41(7), pp. 611–614.

210. Salkind, N.J. (2010) *Encyclopedia of Research Design*. Thousand Oaks: CA: SAGE Publications.

211. Sandelowski, M. (1986) 'The problem of rigor in qualitative research', *Advances in Nursing Science*, 8(3), pp. 27-37.

212. Santos, G., Behrendt, H. and Teytelboym, A. (2010) 'Part II: Policy Instruments for Sustainable Road Transport', *Research in Transportation Economics*, 28(1), pp. 46-91.

213. Sarstedt, M. and Mooi, E. (2014) *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics*. Springer.

214. Schade, J. and Schlag, B. (2003) 'Acceptability of Urban Transport Pricing Strategies', *Transportation Resesearch Part F: Traffic Psychology Behaviour*, 6, pp. 45-61.

215. Scheiner, J. and Holz-Rau, C. (2007) 'Travel Mode Choice: Affected by Objective or Subjective Determinants?', *Transportation: Planning, Policy, Research, Practice*, 34(4), pp. 487-511.

216. Schepers, J.P. and Heinen, E. (2013) 'How does a modal shift from short car trips to cycling affect road safety?', *Accident Analysis & Prevention*, 50, pp. 1118-1127.

217. Schneider, R.J. (2013) 'Theory of Routine Mode Choice Decisions: An Operational Framework to Increase Sustainable Transportation', *Transport Policy*, 25, pp. 128-137.

218. Schwanen, T., Banister, D. and Anable, J. (2011) 'Scientific research about climate change mitigation in transport: A critical review', *Transportation Research Part A: Policy and Practice*, 45(10), pp. 993-1006.

219. Sebastien, L., Josse, J. and Husson, F. (2008) 'Factominer: an R package for multivariate analysis', *Journal of Statistical Software*, 25(1), pp. 1-18.

220. Semanjski, I., Aguirre, A.J.L., Mol, J.D. and Gautama, S. (2016) 'Policy 2.0 Platform for Mobile Sensing and Incentivized Targeted Shifts in Mobility Behavior', *Sensors* 16(7), p. 1035.

221. Shannon, T., Giles-Corti, B., Pikora, T., Bulsara, M., Shilton, T. and Bull, F. (2006) 'Active commuting in a university setting: Assessing commuting habits and potential for modal change', *Transport Policy*, 13(3), pp. 240-253.

222. Shapiro, S.S. and Wilk., M.B. (1965) 'An Analysis of Variance Test for Normality (Complete Samples)', *Biometrika*, 52(3/4), pp. 591–611.

223. Shepherd, S.P., Zhang, X., Emberger, G., Hudson, M., May, A.D. and Paulley, N. (2006) 'Designing Optimal Urban Transport Strategies: The Role of Individual Policy Instruments and the Impact of Financial Constraints', *Transport Policy*, 13(1), pp. 49-65.

224. Shiftan, Y., Outwater, M.L. and Zhou, Y. (2008) 'Transit Market Research using Structural Equation Modeling and Attitudinal Market Segmentation', *Transportation Policy* 15, pp. 186-195.

225. Sillaparcharn, P. (2007) 'Vehicle Ownership and Trip Generation Modelling: A Case Study of Thailand', *IATSS Research*, 31(2), pp. 17-26.

226. Skinner, I., Pridmore, A., Halsey, S., Wilkins, G., Barham, P., Jones, S., Edge, V. and Lucas, K. (2011) *Knowledge Review of the Social and Distributional Impacts of DfT Climate Change Policy Options* London: Department of Transport.

227. SKTP (2016) *Role of Sustainable Transportation to Combat Climate Change* Available at: https://sustainabledevelopment.un.org/topics/climatechange (Accessed: 1).

228. Souza, R.M.C.R. and Carvalho, F.d.A.T. (2004) 'Clustering of interval data based on city–block distances', *Pattern Recognition Letters*, 25(3), pp. 353-365.

229. Spearman, C. (1904) '"General Intelligence," Objectively Determined and Measured', *The American Journal of Psychology*, 5(2), pp. 201-292.

230. Stanley, J.K., Hensher, D.A. and Loader, C. (2011) 'Road Transport and Climate Change: Stepping Off the Greenhouse Gas', *Transportation Research Part A: Policy and Practice*, 45(10), pp. 1020-1030.

231. Stapleton, L., Sorrell, S. and Schwanen, T. (2017) 'Peak car and increasing rebound: A closer look at car travel trends in Great Britain', *Transportation Research Part D: Transport and Environment*, 53, pp. 217-233.

232. Stark, J., Beyer Bartana, I., Fritz, A., Unbehaun, W. and Hössinger, R. (2018) 'The influence of external factors on children's travel mode: A comparison of school trips and non-school trips', *Journal of Transport Geography*, 68, pp. 55-66.

233. Steg, L. (2005) 'Car Use: Lust and Must. Instrumental, Symbolic and Affective Motives for Car Use', *Transportation Research Part A: Policy and Practice*, 39(2-3), pp. 147-162.

234. Steg, L. and Gifford, R. (2005) 'Sustainable transportation and quality of life', *Journal of Transport Geography*, 13(1), pp. 59-69.

235. Steg, L. and Vlek, C. (2009) 'Encouraging pro-environmental behaviour: An integrative review and research agenda', *Journal of Environmental Psychology*, 29, pp. 309–317.

236. Stern, N. 7 (2006) 'Stern Review: The Economics of Climate Change '. WORLD ECONOMICS, pp. 1-2.

237. Stradling, S.G., Meadows, M.L. and Beatty, S. (2000) 'Helping Drivers Out of Their Cars Integrating Transport Policy and Social Psychology for Sustainable Change', *Transport Policy*, 7, pp. 207-215.

238. Sundar Rao, P.S.S. and Richard, J. (2012) *Introduction to Biostatistics and Research Methods*. 5th edn. India: Prentice Hall.

239. Sunkanapalli, S., Pendyala, R.M. and Kuppam, A.R. (2000) 'Dynamic analysis of traveler attitudes and perceptions using panel data', *Transportation Research Record*, 1718, pp. 52-60.

240. Susilo, Y.O. and Cats, O. (2014) 'Exploring Key Determinants of Travel Satisfaction for Multi-Modal Trips by Different Traveler Groups', *Transportation Research Part A: Policy and Practice*, 67, pp. 366-380.

241. Susilo, Y.O., Williams, K., Lindsay, M. and Dair, C. (2012) 'The Influence of Individuals' Environmental Attitudes and Urban Design Features on Their Travel Patterns in Sustainable Neighborhoods in the UK', *Transportation Research Part D: Transport and Environment*, 17(3), pp. 190-200.

242. Tabachnick, B.G. and Fidell, L.S. (2013) *Using Multivariate Statistics*. 6th edn. Boston: Pearson Education.

243. Talhouk, A., Doucet, A. and Murphy, K. (2010) 'Efficient Bayesian Inference for Multivariate Probit Models With Sparse Inverse Correlation Matrices', *Journal of Computational and Graphical Statistics* 21(3), pp. 739-757.

244. Tanner, M.A. and Wong, W.H. (1987) 'The Calculation of Posterior Distributions by Data Augmentation', *Journal of the American Statistical Association*, 82, pp. 528-540.

245. Taylor, B. and Brook, L. (1998) 'Public Attitudes to Transport Issues: Findings from the British Social Attitudes Surveys', in Banister, D. (ed.) *Transport Policy and the Environment*. pp. 72-96.

246. Tertoolen, G., Kreveld, D. and Verstraten, B. (1998) 'Psychological Resistance Against Attempts To Reduce Private Car Use', *Transportation Research Part A: Policy and Practice*, 32, pp. 171-181.

247. Thigpen, C.G., Driller, B.K. and Handy, S.L. (2015) 'Using a Stages of Change Approach to Explore Opportunities for Increasing Bicycle Commuting', *Transportation Research Part D: Transport and Environment*, 39, pp. 44-55.

248. Thomas, G.O. and Walker, I. (2015) 'Users of Different Travel Modes Differ in Journey Satisfaction and Habit Strength but not Environmental Worldviews: A Large-Scale Survey of Drivers, Walkers, Bicyclists and Bus Users Commuting to a UK University', *Transportation Research Part F*, 34, pp. 86–93.

249. Thurstone, L.L. (1931) 'Multiple Factor Analysis', *Psychological Review*, 38(5), pp. 406-427.

250. Tight, M.R., Bristow, A.L., Pridmore, A. and May, A.D. (2005) 'What is a Sustainable Level of CO_2 Emissions from Transport Activity in the UK in 2050?', *Transport Policy*, 12, pp. 235-244.

251. Truong, L.T. and Somenahalli, S.V. (2015) 'Exploring frequency of public transport use among older adults: a study in Adelaide, Australia', *Travel Behaviour and Society*, 2(3), pp. 148-155.

252. Urwin, K. and Jordan, A. (2008) 'Does Public Policy Support or Undermine Climate Change Adaptation? Exploring Policy Interplay Across Different Scales of Governance', *Global Environmental Change*, 18, pp. 180–191.

253. Van Acker, V. and Witlox, F. (2010) 'Car ownership as a mediating variable in car travel behaviour research using a structural equation modelling approach to identify its dual relationship', *Journal of Transport Geography*, 18(1), pp. 65-74.

254. Van, H.T., Choocharukul, K. and Fujii, S. (2014) 'The effect of attitudes toward cars and public transportation on behavioral intention in commuting mode choice-A comparison across six Asian countries', *Transportation Research Part A: Policy and Practice*, 69, pp. 36-44.

255. Van Wee, B. and Holwerd, H. (2002) 'Preferences for Modes, Residential Location and Travel Behaviour: The Relevance for Land-Use Impacts on Mobility', *European Journal of Transport and Infrastructure Research*, 2, pp. 305-316.

256. Verplanken, B. and Orbell, S. (2003) 'Reflections on Past Behavior: A Self-Report Index of Habit Strength', *Journal of Applied Social Psychology*, 33(6), pp. 1313-1330.

257. Veterník, M. and Gogola, M. (2017) 'Examining of correlation between demographic development of population and their travel behaviour', *Procedia Engineering*, 192, pp. 929 – 934

258. Victor, D.J. and Ponnuswamy, S. (2012) *Urban Transportation: Planning, Operation and Management*. Tata McGraw-Hill Education.

259. Wakefield, J. (2013) 'Frequentist Inference', in *Bayesian and Frequentist Regression Methods*. New York, NY: Springer New York, pp. 27-83.

260. Wang, C. (2016) 'The Impact of Car Ownership and Public Transport Usage on Cancer Screening Coverage: Empirical Evidence Using a Spatial Analysis in England', *Journal of Transport Geography*, 56, pp. 15-22.

261. Wang, S. (2015) 'The Function of Individual Factors on Travel Behaviour: Comparative Studies on Perth and Shanghai', *State of Australian Cities Conference*, SOAC conference proceedings.

262. Wang, X. and Kockleman, K.M. (2009) 'Bayesian Inference for Ordered Response Data With Dynamic Spatial-Ordered Probit Model', *Journal of Regional Science*, 49(5), pp. 887-913.

263. Waterson, B.J., Rajbhandari, B. and Hounsell, N.B. (2003) 'Simulating the Impacts of Strong Bus Priority Measures', *Journal of Transportation Engineering*, 129(6), pp. 642–647.

264. Wen, C.-H. and Chen, W.-Y. (2011) 'Using Multiple Correspondence Cluster Analysis to Map the Competitive Position of Airlines', *Journal of Air Transport Management*, 17, pp. 302-304.

265. Wu, W. and Hong, J. (2017) 'Does public transit improvement affect commuting behavior in Beijing, China? A spatial multilevel approach', *Transportation Research Part D: Transport and Environment*, 52, pp. 471-479.

266. Wuerzer, T. and Mason, S.G. (2015) 'Cycling willingness: Investigating distance as a dependent variable in cycling behavior among college students', *Applied Geography*, 60, pp. 95-106.

267. Xenias, D. and Whitmarsh, L. (2013) 'Dimensions and Determinants of Expert and Public Attitudes to Sustainable Transport Policies and Technologies', *Transportation Research Part A: Policy and Practice*, 48, pp. 75–85.

268. Xiong, C. and Zhang, L. (2017) 'Dynamic travel mode searching and switching analysis considering hidden model preference and behavioral decision processes', *Transportation*, 44(3), pp. 511-532.

269. Xu, X., Wong, S. and Choi, K. (2016) 'Quasi-Induced Exposure Method For Pedestrian Safety At Signalized Intersections', *Journal of Transportation Safety & Security*, 8(2), pp. 129-147.

270. Yamamoto, T. (2009) 'Comparative analysis of household car, motorcycle and bicycle ownership between Osaka metropolitan area, Japan and Kuala Lumpur, Malaysia', *Transportation*, 36(3), pp. 351-366.

271. Yang, M., Wu, J., Rasouli, S., Cirillo, C. and Li, D. (2017) 'Exploring the impact of residential relocation on modal shift in commute trips: Evidence from a quasi-longitudinal analysis', *Transport Policy*, 59, pp. 142-152.

272. Yim, O. and Ramdeen, K.T. (2015) 'Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data', *The Quantitative Methods for Psychology*, 11(1), pp. 8-21.

273. Yong, A.G. and Pearce, S. (2013) 'A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis ', *Tutorials in Quantitative Methods for Psychology*, 9(2), pp. 79-94.

274. Zacharias, J. and X.Li (2016) 'Shifting from Metro to Sustainable Modes for Short-Distance Travel', *Transportation Research Record: Journal of the Transportation Research Board*, 2541, pp. 38-45.

275. Zandvliet, R., Dijst, M. and Bertolini, L. (2006) 'Destination choice and the identity of places: A disaggregated analysis for different types of visitor population environment in the Netherlands', *Journal of Transport Geography*, 14(6), pp. 451-462.

276. Zhao, C., Nielsen, T.A.S., Olafsson, A.S., Carstensen, T.A. and Fertner, C. (2018) 'Cycling environmental perception in Beijing – A study of residents' attitudes towards future cycling and car purchasing', *Transport Policy*.

277. Zhou, J. (2012) 'Sustainable transportation in the US: A review of proposals, policies, and programs since 2000', *Frontiers of Architectural Research*, 1(2), pp. 150-165.