Preference Elicitation and Preference Uncertainty

An Application to Noise Valuation

Kaushali Dave

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
Institute for Transport Studies

November, 2011

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publication has been included. The contribution of the candidate and other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The jointly authored paper:

Dave K., Toner J. and Chen H. (2009), 'Examining the Effect of Attribute Representation on Preference Uncertainty', *International Choice Modelling Conference*, Harrogate

contains elements from Chapters 2, 3, 5, 6 and 7. I hereby declare that I had made the major intellectual contribution to the paper. The co-authors of the paper were my PhD supervisors.

Shri Krishnarpanamastu

Acknowledgements

Foremost, I would like to thank my supervisors Jeremy Toner and Haibo Chen. Thanks to Jeremy for teaching me many things, for his immense patience, faith and support and to Haibo for always encouraging me to stretch my horizons and evoking my interest in many different techniques of analysis. I am also thankful to Prof. Bengt Kriström for getting me interested in preference uncertainty. During different stages of the PhD, I benefitted from interesting discussions and comments from Pedro Abrantes, Richard Batley, Peter Bonsall, Phani Kumar Chintakayala, Andrew Daly, Xu Hao, Stephane Hess, Hui Lu, Nigel Tapley and Mark Wardman. Thanks to Richard Batley for getting me interested in higher model forms, for the many opportunities to present my work and receive critical feedback. Thanks to Andrew Daly and Stephane Hess for their invaluable insight on the workings of ALOGIT and BIOGEME, to Nigel Tapley for his help in getting me started with BIOGEME and to Anthony Fowkes for giving me confidence when it was much needed. Thanks are also due to Richard Rigby for getting BIOGEME on Linux up and running. Thanks to Pedro Abrantes for kindly translating the questionnaire to Portuguese, to Nusrat Walid for all her help in getting the questionnaires ready and to Steve Peacock for the important Risk Assessment information. I am very grateful to Filipa Mateus and her gang of friends for successfully conducting the interviews.

I am thankful to the many friends I made in ITS over the years for the good times and to Zahara and Rizwana for their support during a difficult period. I would also like to thank the Marie Curie Fellowship for the generous funding. I am indebted to my family, for all their love and support. Thanks to my father for giving me wings and to my mother for giving me roots. Thanks to Motabhai for his love and the Bhagavad Gita, without which I would not have managed as well when the going got rough and to my mother for the Sundarkand and the Chandi Path. Thanks to Ba, Keshavdas, Haridas and Madhavi for their ever strong support. I am also thankful to Zahid for his friendship, faith, guidance and his help during the survey, to Xun for her friendship and many hilarious times and to Nidhi, Shilpa, Srija, and Vikram for their friendship and immense faith in my abilities. Thanks to Kamal for the Dhammapada and all the encouragement, to Jasuben for her love and to Amrish Pandeyji and his family for their love and kindness. Finally, most importantly, thanks to God, for everything.

Abstract

The valuation of environmental impacts through Choice Experiments (CE) has been increasing applied in order to estimate the cost of environmental externalities. While this valuation technique offers several advantages over other methods, a crucial problem lies in representing the attributes in a manner that can be easily understood by the respondents. Another problem associated with this valuation technique is the assumption that respondents have known and consistent preferences. This thesis relaxes the restraint by allowing respondents to indicate their level of preference certainty. The effect of different attribute representation techniques especially in context of traffic noise is also examined in relation to the level of preference certainty, while the effect of preference elicitation methods on certainty levels is also scrutinised.

Several CE surveys were conducted to evaluate the impact of traffic noise under a residential setting. In order to examine the effects of attribute representation method on the respondents, two different surveys were undertaken using the location and the linguistic representation techniques. This has been carried out in conjunction with three different methods of preference elicitation: the binary choice, one stage Likert and two stage Likert methods. Thus for each of the attribute representation methods, different preference elicitation techniques have been employed.

The main purpose of the analyses has been to examine the variation in error structure and the need for error flexibility due to the different preference elicitation and representation techniques. The results reveal that these components of choice design significantly affect respondents' decision making and subsequent valuation. Moreover, different methods of representation also influence the level and cause of preference uncertainty as well the decision process.

Contents

Acknowledgements .... **iv**

Abstract .... **v**

Figures .... **ix**

Tables .... **x**

1 INTRODUCTION **1**

1.1 Noise Externality and Valuation 1

1.2 The Valuation Framework 2

1.3 Thesis Outline 3

2 NOISE VALUATION **5**

2.1 Introduction 5

2.2 Noise and Noise Valuation 6

2.3 Exposure Based Methods 13

2.3.1 Contingent Valuation 14

2.3.2 Choice Experiments 16

2.3.3 Summary 23

2.4 Annoyance Based Methods 29

2.4.1 Contingent Valuation 32

2.4.2 Choice Experiments 35

2.4.3 Summary 37

2.5 Other Attributes 39

2.6 Research Implications and Research Hypotheses 40

2.7 Conclusions 42

3 PREFERENCE UNCERTAINTY **43**

3.1 Introduction 43

3.2 The Psychology of Choice 44

3.3 Preference Uncertainty in SP 47

3.3.1 Causes of Preference Uncertainty 47

3.3.2 Methods of Preference Elicitation 61

3.4 Effect on Valuation and Forecasting 79

3.4.1 Uncertainty Representations 87

3.5 Research Implications and Research Hypothesis 89

3.6 Conclusions 91

4 ANALYTICAL METHODS 92
4.1 Introduction 92
4.2 The Stated Preference Method 92
4.3 Logit Methods 94
4.3.1 Random Utility Theory 95
4.3.2 Types of Choice Model 97
4.4 Fuzzy Logic 110
4.4.1 Fuzzy Logic Theory 110
4.4.2 Applications of Fuzzy Logic 116
4.5 Conclusions and Implications..119
5 DATA COLLECTION AND DESCRIPTION 121
5.1 Introduction 121
5.2 Survey Background and Previous Study 121
5.3 Experimental Design and Simulation 123
5.3.1 The Location Method 127
5.3.2 The Linguistic Method 132
5.4 Data Collection and Description..141
5.4.1 Survey Objectives 141
5.4.2 Survey Method 142
5.4.3 Questionnaire format..145
5.5 Descriptive Statistics 154
5.5.1 The Location Method 155
5.5.2 The Linguistic Method 162
5.5.3 Comparative Statistics and Sample Representativeness 168
5.6 Summary..174
6 LOGIT DATA ANALYSIS – I..175
6.1 Introduction 175
6.2 Research Hypotheses and Model Structure 175
6.1.1 Research hypothesis 176
6.2.1 Model Structure and Rationale..176
6.3 Model Results..182
6.3.1 Pooled Binary-Multinomial Logit Model 182
6.3.2 Segmentation and Interaction Analyses 200
6.4 Attribute Valuation..203

6.5 Comparison with Previous Study 209
6.6 Summary and Conclusions 213
7 LOGIT DATA ANALYSIS – II 217
7.1 Introduction 217
7.2 Research Hypotheses 217
7.3 Model Structure and Rationale 219
7.4 Flexible Error Structures 221
7.4.1 The Nested Logit (NL) model 222
7.4.2 The Error Components Logit (ECL) model 228
7.5 The Ordered Logit Model 235
7.6 Effect on Valuation 240
7.7 Summary and Conclusions 245
8 FUZZY LOGIC ANALYSIS 249
8.1 Introduction 249
8.2 Fuzzy Inference System with Location Representation Data 250
8.2.1 Binary Choice Data 251
8.3 The Linguistic Method 263
8.3.1 Binary Choice Data 263
8.3.2 Five Point Likert Data 271
8.4 Discussion and Conclusions 273
9 CONCLUSIONS 275
9.1 Thesis Summary 275
9.2 Research Objectives 276
9.3 Main Research Findings 280
9.4 Research Implications 283
9.5 Research Limitations 284
9.6 Conclusions and Future Research 286
10 REFERENCES 290
11 APPENDICES 310
11.1 APPENDIX A 311
11.2 APPENDIX B 321
11.3 APPENDIX C 323
11.4 APPENDIX D 325

Figures

Figure 2.1 Noise Levels in Lisbon in year 2000. Source: http://lisboaverde.cm-lisboa.pt 7
Figure 4.1Nesting Structure for the preference levels........ 100
Figure 4.2 Preference Levels with Ordered Logit Analysis........ 108
Figure 4.3 Classical and Fuzzy set representations........ 111
Figure 4.4 Graphical Representation of Membership functions 113
Figure 4.5 The fuzzy inference system 114
Figure 5.1 Map of the survey area surrounded by three main traffic roads - Av.Gen. Norton de Matos, Eixo Norte-Sul and Av.Padre Cruz........ 122
Figure 5.2 Survey Phases and Experiments 125
Figure 5.3 Gender, Age and Income characteristics across location and linguistic representation methods........ 168
Figure 5.4 Education and Occupation characteristics across location and linguistic representation methods........ 169
Figure 5.5 Noise characteristics and perceptions of respondents across the two phases of the survey 170
Figure 5.6 Preference levels obtained across the two elicitation and representation methods 171
Figure 8.1 Membership function shapes for ‘view’, ‘noise’ and ‘sunlight’........ 253
Figure 8.2 Output membership functions........ 254
Figure 8.3 Membership functions for Gumbel distributed error input........ 260
Figure 8.4 Membership function shape and range for normal distributed error 261
Figure 8.5 Gumbel error input for linguistic representation 268
Figure 8.6 Output membership function shapes for five point Likert data........ 271

Tables

Table 2.1 Overview of noise valuation SP studies, taken from Navrud (2002) 11
Table 2.2 Summary of studies with exposure based method of noise representation 26
Table 2.3 Summary of noise valuation studies using annoyance based method 38
Table 5.1 Apartment location description.. 128
Table 5.2 Experimental design levels for view, noise and sunlight....................... 129
Table 5.3 Simulation results for location method with different input and coefficient values 131
Table 5.4 Average perception rating value obtained for each attribute at each level from the Location method.. 133
Table 5.5 Average perception category values obtained for each attribute for each category from the Location method 133
Table 5.6 Membership function level for each category of each attribute 134
Table 5.7 Category counts for attributes in the location representation method ... 135
Table 5.8 Dummy levels of attributes under linguistic method............................ 137
Table 5.9 Attribute category values used in simulation for the linguistic method 138
Table 5.10 Simulation results for linguistic method with different input and coefficient values 140
Table 5.11 Attribute rating exercise for the location method 147
Table 5.12 Example of a five point Likert scale choice scenario with the location method.. 147
Table 5.13 Example of a two stage Likert scale choice scenario with location representation... 148
Table 5.14 Attribute rating exercise for phase 2 153
Table 5.15 Socio-economic characteristics for location method - aggregate data of all experiments 156
Table 5.16 Percentage of respondents with different socio-economic characteristics across different preference elicitation experiments with Location method 158
Table 5.17 Day and night noise source and disturbance level of respondents with location method.. 160
Table 5.18 Ranking of noise causes of respondents in Location method 161

Table 5.19 Preference levels elicited across the two Likert methods 161
Table 5.20 Socio-economic characteristics of respondents with the linguistic method...163
Table 5.21 Percentage of respondents with different socio-economic characteristics across different preference elicitation experiments with the linguistic method.....164
Table 5.22 Day and night noise source and disturbance level of respondents with the linguistic method 166
Table 5.23 Ranking of noise causes of respondents with the linguistic method....167
Table 5.24 Ranking of noise causes of respondents in Linguistic method 167
Table 5.25 Socio-economic characteristics across different studies and experimental phases 172
Table 6.1 Attribute dummy levels with location method......................................178
Table 6.2 Attribute dummy levels with linguistic method....................................178
Table 6.3 Average ratings for attribute levels across different representation methods 179
Table 6.4 Binary-Likert pooled MNL model for Location and Linguistic data with specific availability conditions and panel effects, with each variable level common across the elicitation methods 188
Table 6.5 Panel variance across different model specifications............................197
Table 6.6 BL-MNL pooled panel model with noise interaction 201
Table 6.7 Willingness to pay (in Euro and 2008 prices) for each attribute across the different representation and data input methods 204
Table 6.8 Willingness to pay (in Euro and 2008 prices) for each attribute level using the average ratings method..207
Table 6.9 Comparison of BL results from current and 2002 studies 210
Table 6.10 WTP values for 2002 and current study..211
Table 7.1 NL results for pooled data across different representation and data input methods 224
Table 7.2 LR test for NL model versus MNL model 228
Table 7.3 ECL model with heteroskedastic error variances across different model specifications..230
Table 7.4 LR test for ECL model versus MNL model..234
Table 7.5 OL results for one stage Likert data..236
Table 7.6 OL results for two stage Likert data..238

Table 7.7 Willingness to pay (in Euro and 2008 prices) for each attribute across the different representation and data input methods from the NL model 241

Table 7.8 Willingness to pay (in Euro and 2008 prices) for each attribute across the different representation and data input methods from the ECL model 242

Table 7.9 Valuation (in Euro and 2008 prices) across the different representation and data input methods from the one and two stage OL models 243

Table 7.10 Model results summary and application recommendation 248

Table 8.1 Membership function values for attributes with location representation method 252

Table 8.2 Summary of results for binary choice, location representation method in percentage correct predicted as well as RMSE 262

Table 8.3 Attribute input levels for linguistic difference design 263

Table 8.4 Summary of results for binary choice, linguistic representation method for percentage correct predicted and RMSE 269

1 INTRODUCTION

1.1 Noise Externality and Valuation

With increasing environmental impacts arising from transportation, transport externalities have been gaining significant importance. The effect of air pollution on materials and human health, the economic costs of congestion and human impact of noise pollution are some effects considered in cost benefit and environmental impact assessments. However, failure to incorporate these externalities into project appraisals can result in a reduction in social and economic welfare.

In order to thus account for environmental externalities, it becomes imperative to incorporate these effects into monetary or other comparative quantitative terms. Several techniques have thus been developed to estimate the effects of transport externalities. While the effect of air pollution on materials and human can be estimated by traditional techniques such as dose-response method, noise impacts have commonly been evaluated using avertive cost behaviour and hedonic pricing (Bickel *et al*., 2003; Navrud, 2002). In case of the classic methods of noise valuation, both avertive cost and hedonic pricing seek to estimate the implicit price of the environmental good, through expenditures incurred to protect oneself from the environmental change or through observed effects on the land market. However the technique of directly eliciting the impacts of noise on an individual’s utility along with his/her valuation has been increasingly gaining attention. Among these are stated preference techniques which either seek to elicit respondents’ willingness to pay for change in noise level (through contingent valuation) or elicit preferences for a set of various characteristics, including noise (through choice experiments). The latter method allows direct valuation of attributes with lesser strategic distortions caused by the respondent; however, in case of environmental amenities in general, and noise in particular, the optimal representation format of the attributes for both the methods, poses a particular challenge to researchers. The application of choice experiments to evaluate noise, while being relatively uncommon, allows increased scope for experimentation with various representation techniques; thus promising an interesting research arena.

1.2 The Valuation Framework

Eliciting individual preferences for economic valuation requires certain behavioural assumptions in accordance with the economic theory. While these assumptions can seem perfectly reasonable in the economic context, methods modelling human decisions are prone to individual idiosyncrasies and anomalies. An interesting facet of research arises not only through differences in individual tastes and preferences but also through respondents' differing abilities to understand the valuation question. Another aspect affecting decision-making and valuation, particularly in the context of choice experiments, arises from the different methods to represent the attributes in question. Thus several factors affect respondents' understanding of the experiment, potentially causing some effect on the valuation of attributes.

Besides the problem associated with variation in tastes and understanding, the theoretical foundation of choice experiments assume the certain formation, knowledge and affirmation of choices, thus resulting in elicitation of fixed preferences. While preferences could very well be certain, allowance needs to be made for respondents to state any levels of doubts in the decision-making. Moreover, the characteristics of the experiment and the level of choice certainty can be closely related.

Though much attention is generally paid to the effect of various factors on choice, any understanding gained of the effect on the decision-making process can also shed light on the relative influence of the different factors. An application of a heuristic (rules) based approach thus promises further insight on methods employed during the choice formation.

Following the empirical limitation of the noise valuation literature as it stands so far and relating it to test the assumptions of the choice experiment theory, the research aim can thus be stated as follows:

Research Aim: to examine the effects of different attribute representation and preference elicitation methods on respondents' choice set[1] understanding and level of preference uncertainty

The aim will be followed by conducting a choice experiment study with different representations of noise and the effect on preference certainty will be tested through varying methods of elicitation.

1.3 Thesis Outline

The thesis is organised in the following manner:

Chapter 2 provides an overview of the noise valuation method commonly applied in the literature as well as the different methods of noise representation in stated preference. The difficulties and problems associated with different representation techniques are outlined along with the need for alternative methods of representation and comparative analysis. The research hypotheses associated with the representation effect is also given in this chapter

Chapter 3 examines the psychological underpinnings of the economic theory of choice. The causes of preference uncertainty in stated preference methods as well as the different methods of preference elicitation conducted in the literature are critically reviewed in this chapter. The different forms of uncertainties will also be outlined along with research implications and the associated research hypotheses

Chapter 4 serves to provide a theoretical framework for the research analysis. The chapter focuses on the methods of stated preference technique, the econometrics of choice modelling as well as the specific models applied in the analysis. The theory of each of the logit models applied in the analysis is explained along with its relation to the preference structure. Additional logit model forms that could be

[1] Choice set effect/understanding relates to the effect of different attribute representation as well as different preference elicitation methods on respondents understanding and decision making. This definition will thus be adhered to in the rest of the thesis.

applied in the analysis is also briefly outlined in this chapter along with the theory and methods of fuzzy logic analysis and its application in choice experiments

Chapter 5 focuses on the survey methodology employed in this research. The background of the study area, the types of different experiments conducted in the survey as well as the methods of level formation are provided in this chapter. This chapter also describes the method employed in the survey while also providing results from pilot analysis along with the questionnaire format. Some descriptive statistics obtained from the survey are also presented in this chapter

Chapter 6 deals with data analysis in context of the effect of attribute representation method on the relative understanding of the attributes. A brief recollection of the research hypothesis as relating to the analysis conducted in this chapter is provided along with the results obtained from the pooled binary-multinomial model. The attribute valuations obtained as well as comparison with previous study conducted in the area are also given in this chapter

Chapter 7 focuses on the analytical effects of preference uncertainty. The results obtained from nested, error components and ordered logit models are provided in this chapter. The relative attribute valuation obtained across the different models as well as an outline of the relation between socio-economic characteristics and preference level are also provided

Chapter 8 provides the results obtained from fuzzy logic analysis of the binary choice data

Chapter 9 provides conclusions, research implications, research limitations and an outline of future research work arising from this thesis

2 NOISE VALUATION

2.1 Introduction

Noise is one of the important transport negative externalities significantly affecting many people. Due to the broader effect on human health and well-being and its associated social costs, several approaches are adopted by researchers to evaluate the economic value of noise. These methods can include valuation through land markets as well as social surveys (stated preference techniques) aiming to elicit respondents' perceptions and valuation. While valuation through land markets constitutes an indirect method with little dependence on conveying noise levels to the respondents, an important aspect of the application of social surveys affecting the valuation is the method of attribute representation. This chapter thus serves two main purposes – 1) it provides a brief background on the methods of noise valuation in the literature, mainly in context of applying the stated preference technique to value noise and 2) it provides a critical review of the methods of attribute representation used in the previous studies.

Two principal methods of noise representation can be identified: exposure based method and annoyance based method. Section 2.2 will provide a background on noise and a summary of noise valuation studies undertaken as cited by Navrud (2002). This will then be followed by analysing the literature using the exposure based method in Section 2.3 while the annoyance literature and studies using that method will be examined in Section 2.4. Section 2.6 will outline the research implications and research hypothesis while Section 2.7 will provide chapter conclusions. Though particular emphasis is placed on noise as it forms the key attribute considered in the thesis, methods of sunlight and view representation and valuation conducted in the literature will also be outlined in Section 2.5.

2.2 Noise and Noise Valuation

Noise can be defined as unwanted sound, the perception of which is highly subjective (Papi and Halleman, 2004; Watkins, 1981). While it is estimated that about 80 million people in Europe suffer from unacceptable levels of noise, about 170 million people live in areas where noise is a considerable cause of annoyance during the daytime (Nijland and van Wee, 2005; Papi and Halleman, 2004). One of the major sources of noise is through transportation where it forms a key externality along with congestion, accident and air pollution (Pratt, 2002).

Noise pollution from traffic is gaining increasing attention in Europe as high noise levels have physiological as well as psychological effects (Tinch, 1996). As per the EEA (1999), about 32% of the EU population is exposed to road noise levels over 55 Ldn dB (A) at house façade. In another study conducted by Nijland and van Wee (2005), the authors estimated that about 20% of the EU population suffer from noise annoyance of which road traffic is the main cause. Moreover, individual perception of different transport noise differs based on the source of the noise. Thus, for a 60 Ldn dB (A), while 15% are highly annoyed by aircraft noise, 10% are highly annoyed by road traffic noise while 5% are highly annoyed by railway noise (Barreiro *et al.*, 2005; EEA, 1999). Though it is generally accepted that noise is an important transport externality affecting Europe, differences of up to 15 dB (A) are found in noise measurements across different European countries based on the national calculation methods which are country specific. These differences are caused due to variation in noise emission calculation across the different countries. Another cause of difference is due to the different implementation of national calculation method and the different interpretation of actual noise levels by the experts. Nonetheless, based on the percentage of population exposed to more than 65 dB (A) of noise, Portugal was found to have the highest percentage with about 30% of the population exposed to road traffic noise (Nijland and van Wee, 2005).

While recent noise levels in Portugal is difficult to obtain, the following map gives overall noise levels in Lisbon for the year 2000. Examining the map it can be observed that most of major road networks show noise levels greater than 55 dB (A). The highlighted region on the map shows an enclosed residential area surrounded by three major road networks which comprises the survey area of this research. Focussing on this enclosed area, it can be observed that some parts within the area have noise levels greater than 70 dB (A).

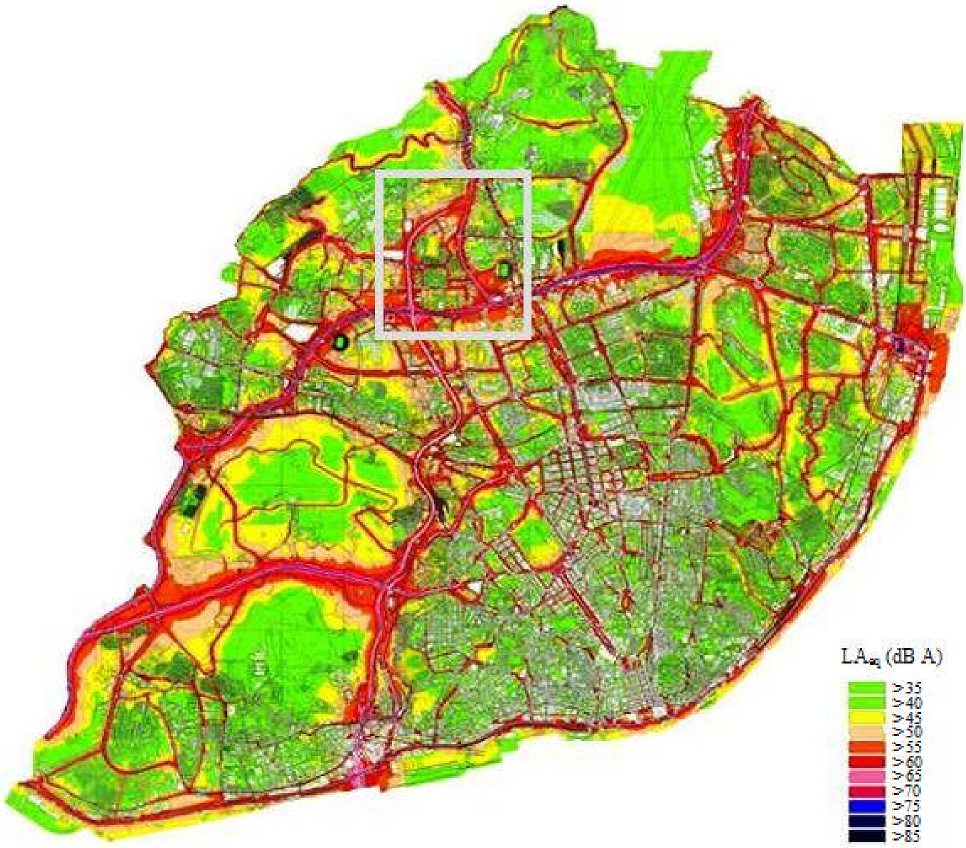


Figure 2.1 Noise Levels in Lisbon in year 2000. Source: http://lisboaverde.cm-lisboa.pt

The method of noise measurement is of particular significance in noise studies as different techniques exist to measure noise. The most common measurement technique is the decibel (dB) scale which is twenty times the logarithm of the ratio of measured sound pressure to a reference pressure of $20N/m^2$. By itself the dB is not an adequate measure of noise as it ranks noise only according to its sound pressure level and does not take in to account the ear's decreasing response at low and high frequencies. Thus an increase of 10 dB can seem like a doubling of sound, due to its logarithmic measure. Most people do not have the equipment or the facilities to hear noise from different decibels and accustom to the significance of the corresponding number (Watkins, 1981). A dB change of noise corresponds roughly to the variation required for a change in perceived intensity. To account for different sensitivity of the human ear to different frequencies, various weighting scales have been devised. For ranking noise, the A weighted decibel dB (A) is considered suitable for *most everyday noise* (Tinch, 1996). The A-weighted scale considers the sensitivity of the human ear, on some form of statistical level based on equivalent sound level (Nijland and van Wee, 2005). However, the way in which noise is measured and the way it is perceived do not always coincide. While the scientific doubling of sound intensity occurs with a 3 dB (A) increase, human perception of doubling only follows after an increase of about 10 dB (A) (Tinch, 1996). Moreover, noise can also be perceived differently based on its source. Thus, in many instances, a 'railway bonus' of 5 dB (A) is applied which implies that the perceived annoyance caused by 60 dB (A) of road noise is similar to the perceived annoyance caused by 65 dB (A) of railway noise (Tinch, 1996).

Due to the difference in the physical measures of noise and the associated perception, higher difficulty is experienced by people in order to understand correct noise level based on the physical measure depiction. Another problem associated in using the physical measure for representation is the correct understanding of the noise level, as respondents need some form of noise source example or actual sound perception in order to understand the physical measure.

In a comprehensive report providing guidelines to the European Commission on the economic valuation of noise, Navrud (2002) listed several techniques for noise valuation along with a review of the different studies which have applied these techniques. The author mainly listed the damage function approach and environmental valuation techniques such as stated and revealed preference as the different methods to evaluate noise impacts. The damage function approach outlined by the author, encompasses several steps which include the description of the noise emission reduction, construction of the noise dispersion model (which estimates the changed exposures to noise at different locations), the exposure response function (ERF) (which links the level of noise and annoyance with effects of noise such as heart disease, subjective sleep quality etc.), the calculation of *overall changes in noise impact* from the ERF and the number of cases of each endpoint (such as number of people highly annoyed by noise per year), the economic value for a unit of each endpoint for ERF computed either through benefit transfer or valuation studies and the economic benefit of noise mitigating measures. While the ERF allows the researcher to compute the economic value of noise through its direct effect on the respondent, the main shortcoming of this method is that health effects arising from chronic exposure to noise can be non-specific and the exact cause of the effect can be difficult to determine (Passchier-Vermeer and Passchier, 2000). Thus, increasing attention has been given to other methods of valuation especially when the health effects do not form the sole aim of the valuation exercise.

The revealed preference technique is applied for noise valuation mainly in the form of the Hedonic Pricing Method (HPM) whose main strength lies in the dependence on actual behaviour in the housing market to explain preferences (Navrud, 2002). HPM comprises of estimating the marginal implicit price of noise followed by the estimation of the demand equation for quiet. While the technique derives its strength in the dependence on actual behaviour, a general shortcoming of the technique lies in the calculation of the implicit price i.e., the Noise Sensitivity Depreciation Index which is extremely sensitive to the modelling technique used as well as the conditions of the local housing market (Staskeviciute and Kaklauskas, 2007). Thus, Navrud (2002) notes that the implicit price is very sensitive to *model specification, the estimation procedure, the functional form*, amount of noise level information respondent had, perception of marginal changes in the physical noise measure used, presence of prefect competition in the housing market, *zero transaction costs and other strict assumptions*, reducing the favourability of this technique. Another technique included in the RP method is the avoidance cost approach which estimates the cost of noise from the amount of expenses that the respondent has incurred in trying to avoid the exposure. The main weakness of this technique as argued by Navrud (2002) is based on the fact that the values obtained can be regarded as a proxy for welfare loss/gain only under certain conditions.

With the stated preference technique, two methods are generally recognised – the Contingent Valuation Method (CVM) and Choice Experiments (CE). A contingent valuation (CV) survey aims to elicit the willingness to pay (WTP) for a specified good or service. The CV questionnaire consists of an introductory section to set the context, description of the good to be valued, the institutional setting through which the good would be provided, method of eliciting the WTP, preference elicitation, debriefing and socio-economic questions (Mitchell and Carson, 1989; Navrud, 2002). While CVM has been vigorously applied in environmental valuation, increasing attention is now paid to the application of CE as it allows valuation of multiple attributes at a time thus masking the real motivation behind the exercise and thereby reducing the risk of strategic bias.

Under both the CVM and CE valuation methods, Navrud (2002) cites several examples which mainly directs towards two main methods of noise representation – the exposure based method and the annoyance based method.

The following SP studies, methods of noise representation and the valuation obtained are cited in Navrud (2002):

Table 2.1 Overview of noise valuation SP studies, taken from Navrud (2002)

Study (Valuation method)	**Site/scenario description/year of study**	**WTP/dB/hh/year in original estimate in national currency**	**WTP/dB/hh/year in Euro (in 2001 price level)**
Road traffic noise			
Pommerehne 1988 (CV)	Basel, 50% reduction in experienced noise level/1988	112 CHF (=75 CHF/month for 8 dB)	99
Soguel 1994a (CV)	Neuchatel, 50% reduction in experience noise level/1993	84-110 CHF (=56-67 CHF/month for 8 dB)	60-71
Saelensminde & Hammer 1994, Saelensminde 1999 (CV and CE)	Oslo and Akershus, 50% reduction in experienced noise level/1993	281-562 NOK (=2250-4500 NOK/year for 8 dB)	47-97
Wibe 1995 (CV)	Sweden national study, elimination of noise annoyance/1995	240 SEK (=200 SEK/month for 10 dB)	28
Vainio 1995, 2001 (CV)	Helsinki, elimination of noise annoyance/1993	33-48 FIM	6-9
Thune-Larsen 1995 (CV and CE)	Oslo and Ullensaker, 50% reduction in experienced noise level/1994	117 NOK (=78 NOK/month for 8 dB)	19
Navrud 1997 (CV)	Norway national study, elimination of noise annoyance/1996	11 NOK (=115 NOK/year for 10 dB)	2
Navrud 2000b (CV)	Oslo, only household exposed to > 55 dB,	152-220 NOK (=1520-2200 NOK/year for 10 dB)	23-32

	elimination of noise annoyance/1999		
Arsenio *et al*. 2000 (CE)	Lisbon, avoiding a doubling of noise level/1999	9,480 PTE (=7900 PTE/month for 10-15 dB)	50
Barreiro *et al*. 2000 (CV)	Pamplona, elimination of noise annoyance/1999	476 ESP (=4765 ESP/year for 10 dB)	2-3
Lambert *et al*. 2001 (CV)	Rhones-Alps region, elimination of noise annoyance/2000	7 Euros (=73 Euros/year for 10 dB)	7
Aircraft Noise			
Pommerehne 1988 (CV)	Basel, 50% reduction in experienced noise level/1988	48 CHF (=32 CHF/month for 8 dB)	43
Thune-Larsen 1995 (CV and CA)	Residents around Oslo airport, 50% reduction in experienced noise level/1994	NOK 1.092 – 5.220 NOK (=91-460 NOK per month and 104-353 NOK/month for 8 dB from CV and CE resp.)	190-959
Faburel 2001 (CV)	Residents around Paris-Orly airport, elimination of noise annoyance/1999	8 Euro (84 Euro/year for 10 dB)	8

Examining the studies listed in the above table, it is evident that the primary method applied for noise valuation is the CV method. Moreover, two main methods can be observed for the valuation process – the exposure based method which is based on noise level reduction and the annoyance based method which focuses on the level of reduction in noise annoyance. Studies from both the techniques will be examined in more detail subsequently in the chapter. It is to be noted that the main form of noise representation cited under the exposure based method by the author, uses percentage reduction. However, as will be seen later, other techniques could also be included in the method.

As the studies using the exposure based method cited in Navrud (2002) uses mainly percentage reduction as the scenario description method, the author is prompted to recommend increased focus on *impacts and level of annoyance* as a description method over the *change in exposure* (thus, implying that the '*50% reduction in noise level*' should be avoided over the '*elimination of noise annoyance*' scenario). However, relative merits and shortcomings of both exposure based and annoyance based techniques will be examined in this chapter.

2.3 Exposure Based Methods

While the specific classification of noise representation techniques into exposure and annoyance based methods have been rarely conducted in the literature, except Navrud (2002) who states that increased focus should be paid on noise representation as change in annoyance rather than change in exposure levels, the different methods used for representation can be generally classified into these categories. The exposure based method encompasses noise representation technique which refers to noise level variation as experienced by the respondent. This can take several forms depending on the method adopted by the researcher. However, the primary focus of this method lies in representing noise as change in the exposure level.

The exposure based noise representation technique has been applied for both CV and CE methods. The following sections will outline some studies which have applied this method in CV as well as CE.

2.3.1 Contingent Valuation

The application of CV to evaluate noise has been relatively common until 2002, as can be seen in the literature cited in Table 2.1. Examining the exposure based method adopted in the studies it can be observed that the most common method of noise representation in this case has been percentage reduction (mostly 50% reduction in experienced noise level).

While an important focus of the Soguel (1996) study was to handle hypothetical and strategic bias in the modelling process, the study estimated the benefits derived from traffic noise reduction in Switzerland by linking noise level to the housing environment. The respondents were asked how much more monthly rent they are willing to pay in order to halve the housing noise level using a monthly starting bid which was then further increased by the interviewer. Using the ordered and weighted least squares methods, the author found that the presence of children, gender, net household income (in case of ordered least squares method), respondent's sensitivity to noise nuisance and their education level had a significant effect on the WTP parameter estimate. While the results obtained from the study showed that the WTP estimate is well dependent on respondent's utility perceived from the independent variables, the method of noise representation remains vague for both the respondent and the policy maker as it is generally not only difficult to imagine the halving of noise but also to derive any policy measures based on this representation form unless physical noise level is also computed. The application of this form of noise representation in case of other studies listed in Table 2.1 pose similar problems in terms of difficulty associated with the respondent's understanding of the noise level.

Using the CV method to estimate economic value of noise reduction in Spain, Barriero *et al*. (2005) asked respondents to indicate their willingness to pay (WTP) for three different methods of noise reduction. In this case, respondents' perception on the type of noise that is more disturbing, the time when it is more disturbing (day or night) and the type of noise more disturbing at each time of the day was elicited. In order to elicit the subjective measurement of noise, respondents were asked to rank the three time periods based on the level of noise experienced while the difference between these levels indicated the level of noise reduction. Using the one and one-half bound method[2] the authors found that the WTP value was higher for greater level of noise reduction. Along with the subjective measurement, the authors also conducted objective measurement of noise and found an increased difference between the level of noise and respondent's perception. Thus while the noise representation method used in this study is the level of noise reduction as perceived by the respondent, this method has obvious shortcoming in that it is completely based on respondent's perception and no measure is elicited in the survey in order to compare the variation obtained between different respondents. Moreover, the author also notes that a variation exists between the 'subjective' and 'objective' measures of noise as the respondents rank a noise reduction between '*a working hour on a working day and a Sunday morning*' to be greater than that between '*a working hour on a working day and 9:30 pm on the same day*' although this is not actually the case when objective noise measures are considered. Moreover, the noise reduction in the former case (compared to a Sunday morning) is valued higher. This finding thus implies that total dependence on subjective perception without a comparable numeric or objective measurement could be dubious.

While CVM method has been commonly applied for environmental valuation, increasing attention is paid on the application of CE as apart from the hypothetical and strategic bias affecting the CV method (Arsenio *et al*., 2000), it is able to value only one good at a time, thus not allowing for the motive of the experiment to be concealed. While the CE method has an advantage over the shortcomings associated with the HPM and CVM, the problem of optimal representation of noise is also prevalent in CE where several techniques have been experimented to effectively communicate noise levels to the respondents.

---

[2] With this method, an upper and a lower bid were chosen for WTP elicitation and one of these was randomly offered to the respondent. When a higher bid was offered but rejected by the respondent, a lower bid was subsequently offered. If the lower bid was initially offered to the respondent which was accepted, the higher bid was then offered. The elicitation stopped when the higher bid was accepted or the lower bid was rejected by the respondent.

2.3.2 Choice Experiments

Saelensminde (1999) applied the CE exercise with attributes such as in-vehicle time, fuel cost, level of noise and air pollution, using different exercises based on different combinations of the attributes. The author provided definitions of air and noise pollution along with other environmental factors in addition to the information on damage and nuisance caused by road traffic. Using demand for alternative fuels as the base and the different fuel costs as the monetary measure, the author employed different exercises based on varying combinations of the attributes. Respondents were informed on the definitions of local air and noise pollution that comprised the study and the attributes air and noise pollution were represented in the CE exercise as a percentage change (20, 40 and 60% reductions compared to the current level).

While the noise variable had a significant value with the correct sign, the author found the value to be higher for lower level of noise reduction compared to larger reductions, inducing the conclusion of declining marginal utility. Using the WTP estimated from these exercises, the author computed the WTP for a 50% reduction which was further multiplied by the total number of households in Oslo to obtain the total WTP. In order to compute the 'WTP per annoyed person', the total WTP was divided by the number of people annoyed by noise, which was captured during the survey through the question on respondent's *experience of various forms of nuisance*. This study was a pioneering application of CE for air and noise pollution in Norway. However, a major shortcoming of the application lies in the representation method of these attributes through percentage change as it is difficult for respondents to understand the level of environmental change based on this representation form. The authors rightly acknowledged this problem however and suggested the use of 'no choice' or 'don't know' alternatives in the choice experiment in order to elicit any preference uncertainty caused by the lack of attribute level understanding.

To assess the benefits of traffic calming measures from speed reduction, noise and community severance, Garrod *et al*. (2002) conducted a SP study on road traffic noise. Three noise levels were incorporated in the choice experiment – 60, 70 and 80 dB which were presented to the respondents through audio recordings of traffic noise. In order to accomplish this method of representation, noise was measured at the pavement and *the nearest residence to the line of traffic*. The authors allowed for interaction term between the distance from the traffic and noise levels showing that the benefits derived from increasing the distance and having a lower noise exposure exceeded the benefits derived from avoiding the increase in the effective speed limit. An evident advantage of this study was the possibility for respondents to hear the noise levels which would have aided their understanding. However, the effect of respondent's socio-physical characteristics such as age or lifestyle as well as the ambient sound levels during the survey could have an effect on the interpretation of the experimental noise level. Moreover, the level of loudness and annoyance could also possibly vary. However, this study offers an interesting alternative to noise representation compared to other exposure based methods and especially the use of percentage reduction.

Wardman and Bristow (2004) applied the SP method to evaluate air and noise pollution under residential choice context in Edinburgh. While the authors employed *proportionate change* as well as the *location method* (where respondents were offered different locations based on the level of air quality) to represent air pollution, noise levels were again represented as percentage change in the SP choice scenario. Though the authors focussed more on the size, sign and level effects, to examine whether gains and losses are valued differently and whether there is a significant effect dependent on the size of the change, they also noted the difference in perception caused by different methods of representation. In case of noise with 50% variation, the authors found that respondents valued increase in noise higher than reduction while in the case of 100% change, a large variation in interpretation was found in the case of 100% improvement 'twice as good' than in 100% deterioration 'twice as bad' prompting the authors to conclude that respondents did not interpret this form of representation in a way that was intended. The paper thus further emphasised that the percentage reduction method of noise representation is not very conducive for respondent's understanding of the attribute.

While Saelensminde (1999) conducted the CE application on noise with a relatively new form of representation thus providing insight on its effect, the results showed that the method employed posed a significant challenge on respondents' understanding. The application of CE to value air and noise quality in residential context by Wardman and Bristow (2004) formed one of the earliest applications in this field. Moreover, the authors used two different methods of representation for air quality along with different forms of representation in terms of gains and losses. However, in this case too, the representation of noise as a percentage change proved to be a shortcoming as a clear variation in respondents' understanding in the case of improvement and deterioration was not observed.

While Wardman and Bristow (2004) applied the *location method* only for air pollution, this method in residential CE survey to evaluate noise from road traffic in Portugal can be found in Arsenio (2002) who represented noise as a *perceived stimuli* for different apartment locations. Thus in this case, respondents' noise perception for the specific apartment locations was considered in the modelling process. The CE was held in the context of residential choice with different characteristics of the apartment based on view, noise, sunlight and housing service charge levels. With the location method applied in these studies, the author asked respondents to provide a numeric rating (from 0-100) for their perception on view, noise and sunlight attributes. Along with the location method for road traffic noise, the author also measured the physical noise level which was incorporated in the modelling process. Using binary and mixed logit models, Arsenio (2002) found that models based on respondent's perceptions statistically outperformed those based on physical noise measure. Moreover, Arsenio *et al.* (2002) reveal that models based on indoor noise measurement outperformed model on outdoor noise measurement, implying that outdoor noise measure is not a true proxy of respondents' noise perception. Apart from the relative significance of the location method compared to the physical noise measurement, it was also stated that households at the quieter façade had a higher marginal value for quiet. The application of this representation method for noise valuation was one of the first examples of using spatial references to convey noise levels. This method thus provided a unique approach to noise representation.

Carlsson *et al.* (2004) evaluated aircraft noise in Sweden using a residential postal CE survey based on the aircraft frequency at different times of the day and different days of the week. Two separate CE surveys were conducted based on increase and decrease in take-off as well as for the weekday and weekend. In case of the increased aircraft frequency version, the authors used compensation as the monetary measure while the payment was used in case of noise decrease. Using a random parameters logit model with normal distribution and accounting for preference heterogeneity and household characteristics, the authors found that respondents were likely to choose the opt-out choice if no one from the household flew from the airport, in the increased take-off version while less likely to do so in case of decreased take-off. Moreover, households with a detached house chose 'current situation' in fear of changes. While the main objective of the paper was to evaluate the value of aircraft noise based on aircraft movement at different times of the day and different days of the week, the authors also laid emphasis on the 'opt-out' alternative indicating that most residents were satisfied with the current state of the airport. However, a shortcoming associated with the study lay in the different factors considered to capture preference heterogeneity. While the authors considered the socio-economic characteristics as indicative of preferences and the usage of the airport in the household, factors such as the number of hours spent at home, the presence of children or the elderly and the perceived level of annoyance and attitudes towards noise, which would have been important factors in choice, were not considered in the study. Nonetheless, this study employed a different approach to noise representation through the use of, what will be termed subsequently, as the 'proxy method'.

In a study conducted to evaluate aircraft noise from Manchester and Lyon airports, Bristow and Wardman (2006a) applied the stated choice (SC) as well as the priority ranking (PR) methods in a residential survey. The PR method comprised of various quality of life attributes along with aircraft noise (which was specified as aircraft movement during day and evening time) which the respondents were asked to rank in order of preference in terms of improvement from the current state. The quality of life variables included 'traffic noise at home' which was represented as 'extremely noisy', 'very noisy' ... 'not at all noisy'.

Using the SC method which comprised of eight binary choice scenarios, the authors conducted the survey by characterising the noise level using the ‘proxy method’ which was based on the type of aircraft and the frequency of movement. In the first SP exercise (PR), noise was represented as total aircraft movement across the day and evenings while in the other case it was represented as aircraft movements (based on different aircraft types) across different time periods. Developing logit models with these two representations as well as with physical noise measure, the authors found that higher perceived noisiness increased the WTP in case of Manchester airport while in case of Lyon, people living in quiet areas had a higher WTP to reduce noise. In case of the PR method, models were formed based on aircraft movement as well as the Leq measure for day and evening time. The authors found that most of the PR models improved by substituting aircraft movements by Leq measure.

This study applied the ‘proxy’ approach in the SC experiment, already mentioned in Carlsson *et al.* (2004) albeit with slight variations. While the variations to the method were an interesting extension to the previous method, several shortcomings can be observed in the noise representation technique for the SC method. Using the SC method, the authors note that though the current level of noise is experienced by individuals, the other levels of aircraft movement and noise are only experienced to the extent that they may occur in other time periods. While this has been cited as a shortcoming by the authors in the paper, this in itself is not a major problem as long as the levels are realistic. A more significant problem lies in using different aircraft types as it might not be possible for respondents to know the variation in noise level based on the aircraft type without that information provided by the interviewer.

In another paper reporting the comparison of SC and PR methods to evaluate aircraft noise, Wardman and Bristow (2008) represented noise as the number of different aircraft movement in an hour for specific time periods for the SC exercise. The PR method consisted of noise represented as different aircraft movement every hour for evening and day time, along with other quality of life variables or as categorical semantic levels such as: extremely noisy, very noisy, moderate, slightly noisy and not at all noisy. In order to examine the different methods of aircraft noise representations applied in the PR method, the authors compared the monetary values obtained under the aircraft movement scenario as well as for the different categorical values, concluding that *the two different means of presenting aircraft noise exhibit an encouraging degree of similarity*. While the authors examined the applicability of the semantic method of noise representation by comparing the values obtained from the aircraft movements model, they failed to examine which method is more suitable for noise representation based on the relative ease in respondent's understanding. However, the results from this study do indicate that the linguistic representation of noise can provide logical valuation estimates. Comparing the valuations obtained from the SC and the PR methods, the authors found valuation variations ranging from ratios of 1.5 to 3.0 for Lyon and Manchester which when examined across a range of reasons was thought to arise from strategic bias caused in the SC exercise.

In Bristow and Wardman (2006b), application of the 'proxy method' was again cited with Bucharest airport along with Manchester and Lyon. In this case, the authors applied the method to conduct 'within period' (WP) and 'between period' (BP) SP experiments. In case of the 'within period' experiment, trade-off was required for different aircraft movements within a given time period while in the case of the 'between period' experiment, more emphasis was laid on the flight frequency for the whole range of time periods across the two options. Reporting results from the same study as outlined in Bristow and Wardman (2006a), several similarities can be found in the two; however, an important point of shortcoming in Bristow and Wardman (2006b), lies in the application of the BP method where the CE employs eight different time periods with flight frequency in each scenario, with a potential to cause increased task complexity. This caused the authors to adopt a 'priority evaluator' type question where current flight, deterioration and improvements were offered to the respondents with different weekly tax to elicit 25 rankings on what they most prefer. Moreover, in addition to some lexicographic choices with the BP experiment, respondents were unable to rank the alternatives in logical order, implying that this is a difficult method for respondents to understand. Using the ordered logit model for the BP method and the standard logit method for the WP exercise, high degrees of similarity were found between the *relative valuations by time period* across the WP and BP results in Manchester while lower similarity was found in case of Lyon. However, though the authors mention the difference in the modelling method employed across the two representation methods, the different model structures are not explicitly specified.

To evaluate aircraft noise in the context of airport relocation, Thanos *et al*. (2006) conducted several residential SC studies around two different airports in Greece. As the study context involved opening of the Eleftherios Venizelos airport and closure of the Hellenikon airport, the authors employed 'noise as of now' and 'not subjected to aircraft noise' as the means of noise representation in the SC experiment. Pooling the data across three different SC experiments and developing different logit models with income and socio-economic characteristics, the authors found that sensitivity to aircraft noise varies for the different airports. Moreover, differences in sensitivity was also observed in case of socio-economic characteristics and the level of reported noise annoyance, with respondents who considered *noise annoyance in their neighbourhood extreme,* had a greater sensitivity to the 'no aircraft noise' variable. This study adopted an extreme variation scenario with 'noise' or 'no noise' as the alternatives. While this approach is suitable for the specific case-study due to its context, the application of this method could not be relevant in other studies where such representation of noise is either unrealistic or infeasible. Incorporating different socio-economic and attitudinal characteristics in the SC model, the authors found that these factors affect noise valuation and the willingness to pay amount. However, this induced the authors to conclude that these factors influence aircraft noise annoyance (rather than aircraft noise valuation), thus causing confusion between aircraft noise annoyance and aircraft noise valuation.

In a unique application of SC methods to evaluate rail noise, Nunes and Travisi (2007) conducted a residential CE survey to assess rail noise annoyance in Italy. The key attributes employed in the SC exercise included noise dB reduction, height of the trackside barrier, investments in trains and tracks technology and the cost of the noise mitigation programme. While reduction in the dB (A) levels were explained to the respondents using visual aids, the main method of representation along with dB (A) in the SC experiment consisted of the specification of the distance of the house from the railway. Thus, for example, in case of a 9-11 dB noise reduction, the level was described as '10 times increase in distance of the residence from the railway'. The authors found that rail noise was ranked more important by the respondents than traffic noise but less important than air pollution. Preferences of respondents for different noise abatement program was found to be *sensitive to the type of instrument, noise abatement target* and the price associated with each policy. Moreover, the authors found that the intermediate level of noise reduction (12-14 dB (A)) was preferred more than the other levels. While this is one of the first studies applying SC to evaluate rail noise abatement, the representation of the noise variable can be difficult for respondents to understand as the noise level is characterised by dB (A) and different levels of distance of the residence from the railway, both of which could be difficult to imagine. However, the authors have attempted to alleviate some of these effects by providing visual aids in relation to the noise levels, prior to the SC experiment.

2.3.3 Summary

The above literature reviewed has provided an overview of the different types of noise representation techniques that can be applied in case of the exposure based method. The different techniques that can be listed range from percentage reduction, auditory examples, reference to spatial noise variation to extreme classification such as 'noise' or 'no noise' scenarios. Examining the studies conducted by Soguel (1996), Saelensminde (1999) and Wardman and Bristow (2004), it is evident that representing noise in terms of percentage reduction or halving poses considerable problem in respondents to understand the levels. The method of noise representation adopted by Garrod *et al*. (2002) through the auditory description of the levels has a relative advantage over percentage reduction description as it allows respondents to better understand the levels. However, problems could arise with this technique when some variation exists in the ambient noise levels for the different respondents during the experiment as well as from the different hearing capabilities of the respondents.

The application of 'noise' and 'no noise' representation as conducted by Thanos *et al.* (2006) has revealed that this extreme form of representation can be successfully applied for noise representation, however, only in unique cases where it is possible to totally eliminate the noise source. In most noise valuation studies though, this situation is not generally realistic. Bristow and Wardman (2006a) and Bristow and Wardman (2006b) represented noise using the aircraft movement at different times of the day (the ‘proxy method’), which can be adequately used to represent noise, though the usefulness of its applicability can be larger in cases of aircraft and rail noise where every traffic movement can be distinctly noticed. In case of road traffic however, the ‘proxy method’ can only be applied where a substantial and marked difference in traffic movement or traffic type (large motor vehicles vs. small motor vehicles) exists between different days and/or time periods. Where a marked variation in traffic movement is not available, this method cannot provide sufficient insight on the level of road traffic noise.

A shift from the percentage reduction as well as auditory exposure of noise level to spatial variations in noise can be observed in Arsenio (2002) and Arsenio *et al.* (2002) who used the ‘location method’ to represent the noise levels. While this method has the disadvantage that respondents might not be entirely familiar with the noise levels of the referred locations, the application of this technique on residential noise exposure from road traffic proved to be quite acceptable. A slight variation of this spatial reference method can be seen in Nunes and Travisi (2007) who specified the noise levels as distance from the noise source along with the dB reduction. Though the authors provided an alternative description of the levels using pictorial representation, this method on its own can pose some difficulty on the respondents to understand the levels. Based on the monetary values obtained from the aircraft movement as well as the different categorical values of aircraft noise using the PR method, Wardman and Bristow (2008) concluded that representation of noise in linguistic terms can provide logical valuation estimates.

Assessing the different methods adopted to represent noise, it can be concluded that some techniques such as the presentation of ‘noise’ and ‘no noise’ scenarios is suitable when total elimination of a specific noise is a realistic option. Noise representation using the ‘location’ method in case of residential choice is also a relatively more comprehensible representation method than specifying noise levels in terms of distance from the noise source or through percentage/proportionate change. Moreover, representing noise in linguistic terms has also been seen to provide logical valuation while representing aircraft noise in terms of aircraft movement can be an adequate approach when classification of different aircraft types is either not required or is clearly understood by the respondents. Thus, depending on the type of noise source and the context, certain representations of noise can be more suitable than others. However, as the economic cost of noise is largely related to the potential harmful effects of the externality on health and productivity, interests is also shown on representing noise based on the level of annoyance caused. This is especially so where a damage function approach is required (Navrud, 2002). The next section outlines the theory and rationale of the annoyance based method along with some examples of its application and the associated shortcomings.

The following overview can be provided of the reviewed literature using the exposure based method of representation:

Table 2.2 Summary of studies with exposure based method of noise representation

Author	**Study Context**	**Data type**	**Attributes**	**Representation Method for Noise**	**Conducing variables**	**Monetary measure and WTP**
Soguel (1996)	Residential survey	CVM	Traffic noise	Halving of noise level	N/A	Monthly housing rent Mean WTP SFr 67/month/hh with OLS and SFr 56/month/hh with WLS
Barriero *et al*. (2005)	Residential study	CVM	Urban noise	Level of noise reduction as perceived by the respondents	Noise experienced at different time periods	City taxes Mean WTP 26-29 Euro/household (hh)/year
Saelensmi-nde (1999)	Travel choice	CE	In-vehicle time, fuel cost, noise pollution, air pollution	Percentage reductions	N/A	Fuel cost NOK 45-90/percentage change/year/hh
Garrod et al. (2002)	Residential survey, traffic calming measure	CE	Effective speed limit, noise level from traffic, waiting time for pedestrians to cross the road, appearance of traffic calming scheme and annual cost to household	Auditory measure (pre-recorded traffic noise) for 60, 70 and 80 dB	N/A	. WTP is mapped on different ranges for effective speed limit, noise level in dB + distance of house from main road
Wardman and Bristow (2004)	Residential choice	CE and CV	Travel accessibility, environmental quality (air and noise quality), local	Proportionate/ percentage change	100% improvement/deteriora-tion for CE	Council tax 50% variation in noise valued at £3.15-4.65/week/hh

			council tax		50% improvement for CV	with SP and £1.48-2.55/week/hh for CVM
Arsenio (2002)	Residential choice	CE	View, sunlight, noise and housing service charge	Location method	Numeric rating for noise perception at different locations	Housing service charge 1.95 Euro/month/hh for unit change in perceived ratings
Carlsson *et al.* (2004)	Aircraft landings and take-off at Bromma Airport	CE	Noise, monetary payment/compensa-tion	Proxy method	Aircraft movements at different times of the day	Payment/compensa-tion Marginal values of 4.16-18.22 SEK per hourly take-off/landing using point estimates
Bristow and Wardman (2006a)	Residential survey, aircraft noise	CE	Noise, council tax	Proxy method	Type of aircraft and frequency of movement	Council tax WTP for removal of 1 aircraft movement/one dB(A)/hr: Euro1.09 (day), 0.41 (even.) for Man. improve. and 1.01(day) & 1.45 (even.) for Lyon
Thanos et al. (2006)	Airport relocation, residential survey	CE	Aircraft noise at home, public transport travel time, tram/light rail service, traffic	Aircraft noise 'as of now' or 'as was subjected' versus 'no noise'	N/A	Local tax 11.85 Euro/month for EV airport, 13.23 Euro/month for H5 model and 8.22

			congestion, use of airport land, local tax			Euro/month for H6 model [3]
Nunes and Travisi (2007)	Rail noise abatement	CE	Noise reduction, height of track side barrier, investment in trains and tracks and cost of noise mitigation programme	dB reduction accompanied by proxy method	Increase in distance between residence and railway	Cost of noise mitigation programme – most favoured alternative – Euro 74.7 – 622.8/hh/year
Wardman and Bristow (2008)	Aircraft noise, residential survey	CE and PR	Noise, council tax (CE) Various 'quality of life' indicators incl. aircraft noise and council tax for PR	Proxy method for CE Linguistic category and proxy method for PR	Type of aircraft and frequency of movement	WTP ranging from Euro 1.79-29.68 for total movements Euro 0.73-24.54 for 'not at all noisy' level

[3] Model H5 is Hellenikon airport SC experiment with five variables (without the airport land-use variable) while Model H6 is Hellenikon airport SC experiment with six variables (including the airport land use variable).

2.4 Annoyance Based Methods

Annoyance based methods in valuation studies identify the level of annoyance caused from noise as the end-point of the valuation (Navrud, 2002). Annoyance level forms a component on the exposure response function which considers loss in human well-being caused from the noise as an important component in the damage analysis. While the noise exposure based method can be sufficiently applied when a single source of noise is identified and chosen to be evaluated, the shift to annoyance based method is required when there are multiple sources of noise (such as aircraft and traffic noise) and reducing one dominant source of noise level will have little effect on the level of annoyance experienced. Moreover, using the annoyance based method has the relative advantage of using the same value for different noise sources while in the case of the exposure based method, adjustments would be needed based on the source of the noise (Navrud, 2002). Despite these apparent advantages of the annoyance based method, definition of the annoyance level is of significant importance in the application of the technique.

Methods to estimate the correct scale and wording of noise annoyance have been a crucial factor in applying this method of representation. Fields *et al*. (2001) conducted a study to identify the correct scale and wording for noise annoyance question in community noise surveys across nine different languages in order to develop good quality survey questions. The main rationale for the study arose from the recognition that comparable questions in socio-acoustical surveys are of significant importance as it is vital to know whether respondents across different locations exposed to the same level of noise give sufficiently similar responses in terms of stated annoyance and whether relatively standard methods are used to measure the level of noise exposure. In order to develop noise reaction measure which allows for international comparison within and between languages, a reliable measure of general reaction to noise and an interval based measurement scale suitable for all questionnaire administration modes, the International Commission of the Biological Effects of Noise (ICBEN) conducted a series of studies and workshops where the type of question (open or close ended), description of the reaction (annoyed, bothered, disturbed), reference to acoustical environment, specification of residential conditions, the type of answer scale and the wording of the question were examined.

In order to develop the necessary noise reaction questions, the study adopted several steps such as the usage of standard wording, examination of the wording by experts, translation and back translation to English and the use of empirical studies for language specific verbal modifiers. In order to obtain a comparable annoyance measure, the authors recommended that each survey should contain a linguistic and a numeric scale question to measure annoyance reactions; these two questions on noise annoyance should be asked to all respondents in order to obtain a reliable estimate; the full scale should be presented, exactly as worded to all respondents; the annoyance questions should be asked early in the questionnaire; appropriate instructions should be included if the questions are perceived as repetitive and written instructions should be prepared for the interviewers in order to respond to ‘I don’t understand’ or to urge respondents to answer all questions.

The authors provided the guidelines for the noise reaction questions on the rationale that short, direct closed rating questions with only negative or neutral reactions are easier to understand than a bipolar scale which could confuse respondents. Due to the varying strengths of numeric as well as linguistic scales, both were recommended in the surveys. In case of the linguistic scale, a five point was preferred while an 11 point (0-10) was recommended for the numeric scale. The intensity score findings from the annoyance modifier study supported that the ‘very’ and ‘extremely’ points of the five point linguistic scale can be combined to define ‘percentage highly annoyed’. Based on the intensity score of different words, the level ‘highly’ is closely associated with ‘very’ and thus can be used to imply high level of noise annoyance (Fields *et al.*, 2001).

Based on the above study and the recommendations provided to conduct noise annoyance survey, it is evident that several factors need to be considered when using this method for noise valuation. While the study focussed on the formation of appropriate questions for noise annoyance survey, the primary focus has been on the type of scale and the wording to be used. Another factor however, that affects the application of annoyance based techniques in noise valuation is the relation of the annoyance level to the physical sound level.

Miedema and Oudshoorn (2001) note that though the linkage between annoyance level and the associated noise exposure level is vital, the uncertainty on the exact relationship between exposure and the response in the population is crucial for environmental policy. In order to address this problem, the authors provide a confidence interval along with the exposure response curve where the distribution of the annoyance response is modelled as a function of the noise exposure. Thus using the normally distributed variable for the noise annoyance level for an individual (obtained on a scale from 0 to 100), the authors estimate the probability that the annoyance level of a randomly selected individual from a randomly selected study is greater than a specific annoyance level. Moreover, the *predictability of the annoyance of a general population exposed to a certain noise level* is given by the width of the confidence interval for that noise level for the related noise and annoyance measure. In addition to the confidence interval for each exposure response curve, the EC (2002) also estimates the exact corresponding proportion of percentage annoyed and percentage highly annoyed for different levels of noise exposure (Lden) for aircraft, rail as well as road traffic noise for a given population.

While some attempt has been made to link the noise exposure level to annoyance, the commonly applied method of relating annoyance with higher sound intensity is examined and challenged in Fujii *et al*. (2002). The study notes that while annoyance is generally described as a *feeling of displeasure of adverse reaction generated by noise* and related to loudness and sound level, for sound sources from widely differing acoustical properties, this relationship may no longer apply. Thus, annoyance cannot be predicted by sound intensity alone and the authors list several studies that have examined the effect of other factors such as frequency distribution, tonality, temporal fluctuation and impulsivity on annoyance. Conducting various physical experiments, the authors found that the variance of the sound pressure level (SPL) had a much higher effect on annoyance than the range of the SPL. Thus, it was found that sounds with strong tonal components were perceived to be more annoying along with those with a clear pitch. While high pitch sounds tested in the experiment were found to have weak pitch sensation, the authors found that low pitch sounds are tonal, indicating that the effects of pitch annoyance should be considered along with the pitch strength.

In a study linking loudness perception with noise indices in Chile, Sommerhoff *et al*. (2006) asked the respondents to give a loudness perception inside their dwelling caused by road traffic noise. The following verbal scale was used in the survey: Not at all Loud, Slightly Loud, Moderately Loud, Very Loud and Extremely Loud which was combined with the objective measure of noise such that the authors created a matrix of noise level and perception inside and outside the dwelling. Using relevant conversion equations and defining loudness in terms of sound intensity while annoyance as the level of sound effect, the authors found that the percentage of people 'highly annoyed' by noise was slightly higher for 'extremely loud' noise perception. This finding revealed that noise hearing sensitivity is one of the variables that explain loudness classification difference. Moreover, the number of people highly annoyed by noise was found to slightly greater than the number obtained in the 'extremely loud' category of loudness perception.

The above studies have attempted to form a uniform measurement scale for annoyance as well as relate annoyance with different acoustical properties and noise exposure level; some studies are also found in the valuation literature where respondent's subjective annoyance levels forms a component in the valuation exercise. The following sections will outline these studies from the CVM as well as the CE literature.

2.4.1 Contingent Valuation

Bjorner (2004) used CVM to evaluate the WTP for removing noise annoyance. Using a CV questionnaire which sought to elicit respondent's level of noise annoyance from road traffic using a payment card followed by an open ended WTP question, the author stated that as the valuation scenario is closely associated with the annoyance level experienced by the respondents, it is important to determine the exposure-response relationship between annoyance and noise level in order to apply the WTP results. Data on noise exposure was obtained by the author from the Environmental Protection Agency of Copenhagen. Models were estimated by the author by including noise in the index function in a linear form and as an additional term as noise squared as an explanatory variable. The author found that noise annoyance was related to age, respondent's stated sensitivity to noise and the orientation of bedroom facing the main traffic road. However, as the latter two variables can be closely related to noise annoyance, a model was built without including these variables.

Comparing the predicted probabilities obtained from the linear and quadratic models it was observed that both models gave reasonable predictions of the observed shares of annoyance levels though the quadratic model was slightly better in predicting the observed average shares. The author found that the share of positive bid increased with the level of annoyance, while the share of positive bids reached its peak for moderately annoyed. When included with the annoyance levels in the model, noise was not found to be significant; however, without it, noise became significant with the expected positive sign. The model fit thus suggested that the annoyance level dummies provided a better description of the data. With the specific context, the study thus found that the annoyance level acts as a better explanatory variable than the noise level and can also act as a proxy for noise. Moreover, the author also combined the exposure-annoyance relation with the valuation exercise in order to estimate the value of a dB reduction. Conducting this exercise, the author found that with the chosen exposure-annoyance relationship, the value of noise reduction depended on the initial level of noise. While the study provided reasonable results in terms of WTP values for different annoyance levels, it also highlighted the importance of defining a correct functional form to combine the exposure-annoyance relation with the WTP responses.

Though some SP studies elicit respondents' noise annoyance at different times of the day (Arsenio, 2002) few examples can be found where noise is actually valued in this manner. Martin *et al*. (2006) conducted a survey asking respondents to indicate their degree of annoyance from different traffic sources. The physical measurement of noise was carried out at different time periods during the day and night as well as across different months. In order to relate the annoyance level to the physical measure of noise, different functions were computed for the annoyance level with the noise level in Ldn. The authors found that the level of annoyance corresponded well to the level of noise, especially for 'highly annoyed' and Ldn above 65 dB (A). However, in terms of noise valuation, the authors adopted a series of CV like WTP questions to elicit respondents' preferences for a house in a quieter place, expenditure to improve sound insulation in the dwelling and preference for a 'noise free' neighbourhood. While this study estimated a link between the level of annoyance and the physical measure of noise, the annoyance representation method was not used to capture the WTP estimate. Moreover, the description of 'noise free' neighbourhood can be rather vague and seemingly impractical representation in many instances to elicit the WTP.

Fosgerau and Bjorner (2006) conducted a study to develop a joint model for noise annoyance and willingness to pay. While the previous application of CVM by Bjorner (2004) estimated the WTP conditional on annoyance which is further linked to noise exposure through the noise exposure-annoyance relationship, this method of computing the expected WTP for reduction in noise exposure has a potential for endogeneity bias. The authors thus applied a joint model of annoyance and log (WTP) to address this issue, without estimating the relationship between annoyance and noise. In this case, annoyance was treated as an endogenous variable and estimated along with the WTP. The stated annoyance was an ordinal variable with five categories while annoyance was described as an ordered probit model with latent variable A^* which was a function of the independent variables and a random component. The respondents were asked to indicate their level of noise annoyance on a five point scale while the WTP was elicited using an open-ended question that provided a continuous variable censored at zero.

The authors incorporated noise, noise squared, age, age squared and gender as independent variables in the ordered probit model while for the regression, the variables were noise, noise squared, log (income) and dummies for high and medium education. In the first model, the authors estimated an ordered probit model for annoyance and a linear log (WTP) regression separately. In the next model, the probit and regression models were combined by the coefficient of noise squared. The results indicated that the correlation introduced improved the model fit substantially, indicating that the correlation of the error terms is quite significant in this case. By including the stated annoyance level as an endogenous variable in the estimation of the expected WTP, the authors found that the standard errors of the expected marginal WTP in this case was reduced by 3-10% which implied that the model efficiency can be improved by specifying the model in this form. As the main purpose of this study was to identify methods of improving model efficiency using the information available from the respondents, alternative model forms were thus experimented, with the finding that the annoyance level can be jointly modelled as an endogenous variable and used to estimate the expected WTP.

2.4.2 Choice Experiments

While the representation of noise as an annoyance measure is relatively more commonly found in contingent valuation, fewer applications of it are found with the CE. However, in an application of the annoyance level in CE, Li *et al*. (2009) examined the characteristics that affect individual's preferences for a residential apartment by examining the effect of apartment orientation, travel time to nearby public transportation facility, annoyance level and monthly rent/management fee on respondent's choice. Three levels for annoyance were employed in the survey: *one annoyance level lower than the current level at home*, *same annoyance level* and *one annoyance level higher than the current level at home*. The orientation characteristic was defined as best and worst orientation. In order to estimate the WTP for different annoyance level experienced by the respondents as well as to form an annoyance-dB relationship, the authors provided 10 seconds to the respondents to listen to the background noise before asking them to provide their level of annoyance on a 10 point scale (0 indicating 'not noticeable' and 10 indicating 'extremely annoyed'). The perceived noise annoyance at home was utilised to form a segmentation model of the CE utility function.

Using an ordered logit model, the authors sought to correlate the annoyance at a particular noise level with the objectively measured dB levels. In accordance with the results obtained from the above noise annoyance studies, the authors found a substantial difference between annoyance and physical dB levels of noise as the probability of getting annoyed at a particular level varied based on the initial objective noise level. Moreover, the authors found that this relationship also differed based on the respondent’s stated level of annoyance at home. In this case, it was found that the highly annoyed group was more sensitive to noise levels compared to the moderately and slightly annoyed groups. While the authors applied the annoyance based method in the CE, the method of its representation however, is quite complex as ‘one annoyance level lower than the current’ can be difficult for respondents to understand. However, this study highlights the problem associated with the application of annoyance based method in stated preference and especially in choice experiments. While the exposure based method allows the researcher to represent noise in myriad different forms, the possibility is significantly reduced when incorporating noise as the change in experienced/perceived annoyance level.

In a SP study to estimate the willingness to pay for reduction in noise levels in a residential location, Galilea and Ortuzar (2005) encountered problems in respondents’ understanding of the physical measure of noise (dB) during the focus group study. A 10-point rating scale was thus adopted where grade one represented noise level ‘as in the countryside’ and grade 10 represented ‘unbearable noise’. Using this representation method, the authors calculated the subjective value of reducing noise outside the dwellings. In order to relate the subjective levels of noise to the objective measurement, the authors developed a multiple regression method with the subjective noise level as the dependent variable while awareness of the noise level and the importance given to the noise attribute as the independent variables. With the aid of the regression model, the authors aimed to transform the estimated noise level parameter by multiplying it with the coefficient for the objective decibel measure. While this technique formed a pioneering approach to link the subjective and the objective measures of noise, the validity of the method has not been externally tested. Moreover, the term ‘unbearable noise’ can be rather vague, subjective and prone to wide variation across respondents. Another shortcoming of the study was the noise range offered to the respondents. While the authors sought to relate the subjective noise values to the decibel scale measurement, they found that though the noise measure was quite high, the noise range was not sufficiently wide resulting in the failure of the linear regression to give a reasonable fit. A multiple regression considering awareness of the noise level as well as the importance given to it had to be developed to improve the model fit thus indicating the effect of other factors such as awareness of noise level to link the relation between the subjective and the objective measures of noise.

2.4.3 Summary

Examining the above studies, several points can be noted. First, it is extremely important that a uniform method of defining noise annoyance as well as eliciting respondents' annoyance level should be used as it is imperative that these measures should be comparable across different studies. The Fields *et al.* (2001) study has attempted to provide a guideline on how noise annoyance should be elicited from the respondents such that valid comparisons can be made while the Fujii *et al.* (2002) study has revealed that noise annoyance is not only related to the sound pressure level but also to the tonality and the pitch of the sound. Some attempts have also been made by Miedema and Oudshoorn (2001) as well as in EC (2002) to relate annoyance level with noise exposure level.

The contingent valuation studies with noise annoyance method have shown that while incorporating the annoyance level gives a better model fit than the noise level in the regression model, the examples of annoyance level used in the survey such as 'unbearable noise' or 'noise free' neighbourhood can be either extremely subjective or highly unrealistic. While adequate method to obtain valuation for different annoyance levels is one concern of the studies, another factor outlined is the exposure-annoyance relationship and estimating WTP for variation in physical noise level when valuation for the annoyance level is obtained. CE studies applying annoyance based method of noise representation, while fewer than those applying the exposure based method, have revealed that the presentation of the annoyance levels in the choice scenario poses a significant challenge in its application, which could explicate the causes for lesser applications of the method in CE.

The following table summarises the methods and findings of the valuation studies using the annoyance based method:

Table 2.3 Summary of noise valuation studies using annoyance based method

Author	**Study Context**	**Data type**	**Attributes**	**Representation Method for Noise**	**Conducing variables**	**Monetary measure**
Bjorner (2004)	Socio-acoustic + CV	CV	Road traffic noise	Noise annoyance level	Noise reducing surface on street	Varied based on the level of annoyance; 12-20 Euro/dB/hh/year
Fosgerau and Bjorner (2006)	Socio-acoustic + CV	CV	Traffic noise	Noise annoyance level	Noise reducing surface on street	Variable based on the level of annoyance
Martin *et al*. (2006)	Socio-acoustic + CV	CV	Traffic noise	WTP to reduce 'noise level'	N/A	7.22 Euro/person/year
Galilea and Ortuzar (2005)	Residential location choice	CE	Travel time to work, monthly rent, sun orientation, subjective noise level	Subjective noise level on a 10 point scale	N/A	\$2.12-\$4.10/dB(A)/month for best model
Li *et al*. (2009)	Residential survey	CE	Flat orientation, travel time to nearby transport facilities, annoyance level and monthly rent/management fee	Annoyance level	N/A	HK\$61.6/unit annoyance level/month

2.5 Other Attributes

Though the literature review has focussed mainly on the methods of noise representation and valuation, the current study following from Arsenio (2002), comprises of view, noise, sunlight and housing service charge as the main attributes in the CE. Arsenio (2002) selected view and sunlight as additional attributes in the experimental design as the information of these attributes is provided by the housing agency, thus resulting in respondents being accustomed to these attributes in the residential decision-making context. Moreover, the inclusion of these attributes in the residential choice experiment would mask the main objective of the study (i.e., noise valuation) and hence were selected in the choice task. Though these attributes are not generally valued in the residential context using the SP method, Galilea and Ortuzar (2005) incorporated sunlight in the CE which was represented as perceived best/worst orientation. In case of Li *et al*. (2009), the authors incorporated apartment orientation in the choice set which had best/worst levels. However, the definition of 'orientation' was not well specified in this study as it could be interpreted in terms of view, sunlight or direction. Thus, in case of Galilea and Ortuzar (2005) as well as Li *et al*. (2009) studies, substantial problem can be observed in the representation of the attribute with added problem associated with the definition of the alternative in case of Li *et al*. (2009).

While pictorial representation (Campbell *et al*., 2007) and now, virtual reality experiments (Fiore *et al*., 2009) have been used in landscape amenity studies to aid respondent's understanding, in case of the residential SP study, Arsenio (2002) elicited respondent's knowledge of other flat characteristics which reflected their understanding of the attribute levels. Hence in this case, the need for pictorial representation was diminished. The above studies reveal that while some attempt has been made to incorporate other apartment characteristics in the SP exercise, the valuation of these is still relatively uncommon. Moreover, the comparison of different representation methods for these attributes in the SP context being rare supports the application of alternative representation techniques for these attribute valuations.

2.6 Research Implications and Research Hypotheses

The literature review conducted to examine the methods of noise representation and valuation reveals that while representing noise in physical dB measure is difficult for respondents to understand, studies using percentage change as a representation method has also consistently shown problems associated with respondent's understanding of the attribute. In case of the annoyance level application, variation was found between the objective level of noise and the annoyance level along with the problem of subjective interpretation of annoyance.

Though some examples can be found in the literature which applies alternative methods of representation along with comparison of different noise representation methods, the relative effect of linguistic representation method compared to other techniques of noise representation under the SP method, is still relatively uncommon. The method of attribute representation can be crucial in the choice task as Stone and Schkade (1991) observed varying methods of decision strategies when applying numeric and linguistic representation methods for choice of alternative computer information systems. While the semantic representation of noise levels is relatively common with the annoyance based method, it is less commonly applied in the case of the exposure based method. A comparison of the effects of this form of noise representation with other methods can be an interesting research question within the broader area of choice experiment. While Wardman and Bristow (2008) compare the aircraft movements representation method with the semantic method, this is conducted using the priority rankings technique and hence the examination of alternative forms of noise representation, and especially the linguistic method, on respondents' understanding of the choice task is a relatively unchartered research area.

As road networks in close proximity to residential areas causes greater exposure and annoyance to households, conducting residential valuation surveys is generally closely associated with road traffic noise. Considering the residential area examined by Arsenio (2002), the location offers an interesting site for examining the varying effects of different noise representation techniques as well as to compare temporal variations in the valuations. Moreover, compared to the annoyance based method, the exposure based method offers greater opportunity to examine the effects of alternative forms of representation and hence a residential, exposure based study on road traffic noise offers a good possibility to examine the effect of alternative noise representation forms on valuation as well respondents' preferences.

As the comparison of alternative noise representation forms is still not widely conducted while the valuation of apartment characteristics such as view and sunlight is rare, the inclusion of these attributes and the examination of alternative representation methods forms an interesting research area. In light of the literature reviewed, two implications can be derived: 1) a comparison of different methods of attribute representation offers an interesting research possibility and 2) the effect of linguistic representation of noise on respondents' understanding could be tested within the exposure based method.

Based on the implications for research and integrating the methods of attribute representation to level of choice set understanding and preference uncertainty, the following research hypothesis is formed:

Hypothesis 1: The method of attribute representation affects respondent understanding in terms of the specific attribute understanding and valuation as well as the understanding of the choice set.

This hypothesis in relation to the models developed can thus be decomposed to the following two hypotheses:

Hypothesis 1a: The location and linguistic representation methods can have varying effects on the understanding of the different attributes. In case of noise, the linguistic representation method could be clearer for respondents to understand than the location method.

Hypothesis 1b: The type of choice set is expected to have an effect on the level of respondent's preference certainty. Focussing only on the random effect on preference certainty, it is expected that the linguistic representation method will have a lower need for complex error assumption.

2.7 Conclusions

This chapter focussed on the various methods of noise valuation conducted in the literature. Noise is a significant externality arising from transportation and several studies have been conducted to estimate the economic value of noise. While hedonic pricing method and contingent valuation have been largely applied in the past, increasing attention is now being given to the application of choice experiment in the valuation of this environmental amenity. However, the method of attribute representation presents a peculiar problem for noise valuation as the physical measure of noise is difficult for respondents to understand. Various alternative methods of representation such as percentage changes, the use of annoyance levels as well as the location method have been experimented. However, few studies have sought to compare the effect of alternative representation methods. In case of sunlight and view, significantly fewer valuation studies and comparative effect of representation methods is observed indicating greater need for incorporating these attributes in the residential choice experiment. It is hypothesised that compared with the representation methods generally used for noise valuation under the exposure based method, linguistic representation would be easier for respondents to understand. Hence, the research objectives aim to examine the effect of linguistic representation on choice set comprehension as well as conduct a comparative analysis of different representation methods.

3 PREFERENCE UNCERTAINTY

3.1 Introduction

While the previous chapter outlined the methods of noise valuation and the problems associated with its representation, this chapter will focus on another aspect of interest related to choice experiments. An assumption of surveys aiming to capture respondents' preferences is that those preferences are known and consistent. However, in practice, this is not always the case as preferences can be affected by nescience and inconsistency due to several causes, one of which could be the method of experiment presentation. While the assumption of certain preferences allows for simplification of the experiment and the subsequent analyses, any unidentified uncertainty can have serious implications on the computed valuations. The purpose of this chapter thus, is to examine the causes of preference uncertainty, the methods of eliciting preferences as well as the tools available to analyse the preference data.

Section 3.2 will provide a background of the choice theory along with the perspective and implications from the psychology literature. Section 3.3 will outline the causes of preference uncertainty as found in the SP literature along with the methods of preference elicitation and treatment of preference uncertainty data. Examining the effects of preference uncertainty, preference elicitation and attribute representation methods in Section 3.4, Section 3.4.1 will define different forms of uncertainty representations, thus introducing alternative approaches to preference data analysis. The implications of the literature review to the specific research will be given in Section 3.5 along with the associated research hypothesis based on this aspect of the study while the chapter conclusions will be given in Section 3.6.

3.2 The Psychology of Choice

Choice theory is based on the assumption that individuals are rational and their choices are based on utility maximisation (Train, 2003). The basis of this assumption lies in random utility theory (RUT) (Ben-Akiva and Lerman, 1985) and the maximisation of expected utility (Restle, 1961) which assumes that utility is quantitatively measureable and individuals have clear preferences (von Neumann and Morgenstern, 1944). While RUT is based on utility maximising behaviour where the decision maker can compute certain utility associated with each alternative (Train, 2003), the expected utility theory (EUT) sheds light on individual choice under risk or uncertainty. Under the EUT, the choice of an alternative with several outcomes is dependent on utility and probability of occurrence associated with each outcome. Based on finite number of outcomes and the associated utilities, the individual is then anticipated to choose the alternative with the highest expected utility (Kliendorfer *et al.*, 1993). Under this theory, the individual is thus expected to select the same alternative when faced with the choice scenario repeatedly *ceteris paribus*; however, empirical experiments have shown that people can be inconsistent (Restle, 1961) thus paving way for the significant insight provided by the psychology literature.

Several experiments have been conducted by experimental economists and psychologists to understand individual's decision making and to examine whether the EUT is unequivocally followed during the decision making process. When applied to decision making under risk and choice between different prospects, Kahneman and Tversky (2000) found that the EUT was violated under different experimental situations due to several causes. Besides the gain-loss asymmetry where individuals were found to value gains and losses differently, the authors found that in case of gains, individuals gave higher value to certain outcomes relative to those that were probable. Where losses formed the part of the experiment, the authors found that the preference for certainty resulted in increased risk seeking while the same principle resulted in higher risk aversion in the gains domain. In other experiments, the authors found that individuals focussed on distinctive components of the choice task, ignoring the common components thus resulting in inconsistent preferences.

Besides the failure to follow utility maximisation, other factors affect individual's choice, two of which can be classified as the systematic treatment of preferences and the violation of the invariance principle. The invariance principle can be decomposed to descriptive and procedural invariance. While descriptive invariance state that preference is independent of the method in which the alternatives are described, procedural invariance relates to the independence of the elicitation technique used (Slovic, 2000). However, research has shown that the assumption of both the invariances does not hold. In case of the procedural invariance, preference reversals are known to occur when equivalent methods of preference elicitations are used interchangeably (Slovic, 2000); however, the exact insight on which rationality axiom is violated and why is not always overt (Kliendorfer *et al*., 1993). The concept of descriptive invariance pertains mostly to the framing of the decision problem. Tversky and Kahneman (2000) have shown that individual decision-making is significantly affected by the *language of presentation, the context of choice and the nature of display* thus proving the violation of the descriptive invariance principle.

Besides descriptive and procedural invariance, systematic treatment of preferences has been an important factor of choice inconsistencies. While an important tenet of choice theory is the individual's ability to have well-defined and consistent preferences, this in turn is based on the assumption that the causal tastes are *absolute, relevant, stable, precise, consistent* and *exogenous* (March, 1988). However, it is argued that not only do individuals ignore their preferences during the decision-making process and tastes can possibly be partly endogenous (such that future tastes are dependent on the consequences of the present choices or current tastes could be a function of the consequences of the past choices), but preferences can also be managed to modify future tastes, constructed to allow for flexibility through vagueness, treated strategically, confounded, avoided, changed and suppressed (March, 1988). Individuals are also known to construct different psychological representation of information based on temporal distance. Though economists have assumed exponential discounting, experimental evidence under construal level theory have proved that real-life decisions are based on hyperbolic discounting where the discount rate becomes higher as the outcomes gets closer in time and thus when the time for action arrives, people choose to adjourn it although they consider the action to be beneficial (Leiser *et al*., 2008).

Thus, evidence from psychology and behavioural decision theory has proved that the assumptions of choice theory are often violated and contorted by individuals during the decision-making process. Focussing on the specific application of descriptive invariance and preference awareness, this research aims to examine how different methods of attribute representation affect respondents' choice. The effect on respondents' decision making is examined through the levels of preference strength using different methods of preference elicitation.

Before proceeding to the next section of preference uncertainty in SP, clarity on three main concepts are sought as pertaining to choice theory – uncertainty, ambiguity and risk. Much work has been carried out on uncertainty and risk in relation to subjective expected utility (Fishburn, 1988). While risk denotes possible random outcomes of a certain event with known associated probabilities, uncertainty relates to the state where randomness cannot be expressed in mathematical probabilities or the probabilities themselves are unknown (Acquisti and Grossklags, 2005). In contrast to risk, where the individual is not aware which state will occur (though he knows the certain probabilities), and uncertainty, where the probabilities are unknown, ambiguity relates to the uncertainty associated with the probability of a state's occurrence due to the lack of some relevant information. Thus, in this case, the correct distribution of the probabilities over states is uncertain (Camerer and Weber, 1992). It can be seen that these definitions in relation to the expected utility, relate to the states of outcome (Batley and Ibanez, 2009). Hence, uncertainty, ambiguity and risk in context of expected utility are functions of the probability of the outcome.

In context of random utility and this research in particular, ambiguity and uncertainty will be taken to have the following meanings: ambiguity will represent respondent's level of perceived clarity of the attribute level and representation, while uncertainty relating to preferences will comprise of the effect of randomness on choice as well as the level of choice commitment the respondent is willing to make. Both these types of preference uncertainty (based on level of choice randomness and level of choice commitment) can be a function of several other factors including attribute ambiguity.

3.3 Preference Uncertainty in SP

While accounting for preference uncertainty is relatively uncommon in choice experiments, its importance has been more acknowledged within contingent valuation. Several causes of preference uncertainty have been outlined in the literature for both CV and CE methods while different methods of elicitation have also been applied. This section provides the causes of preference uncertainty as commonly found in the literature along with the different methods of preference elicitation. Section 3.3.1 provides the review of literature on preference uncertainty as found in the CV and CE fields along with some review from the marketing literature while Section 3.3.2 gives an overview of the different elicitation methods used to capture preference uncertainty within SP as well as marketing literature.

3.3.1 Causes of Preference Uncertainty

3.3.1.1 Preference Uncertainty in SP

This section examines studies conducted within contingent valuation and choice experiments that have sought to elicit respondents' preferences. The main emphasis on this section will be to examine the methods of preference elicitation and the analytical techniques employed to treat the preference data. Studies eliciting respondents' choice certainty levels in contingent valuation will be examined first followed by a review of studies conducted within SP in the 1980s which offered respondents a Likert scale. The rationale of these studies as well as the method of analysis will be critically reviewed followed by an examination of studies which have elicited respondents' preference within the choice experiment literature.

The contingent valuation method (CVM) is used to elicit respondents' willingness to pay (WTP) for a given good or service using one of the several methods of preference elicitation. As the standard format of a CV exercise consists of the specification of the good or service, the description of the proposed policy and the WTP elicitation question, the main causes of preference uncertainty cited within the CV literature are related to the uncertainty of the good being valued, suspicions on the proposed policy instrument, hypothetical nature of the survey and the effect of survey instrument and design (Kosenius, 2009; Shaikh *et al*., 2007; Wang, 1997).

More recently, the link between cognitive psychology and preference uncertainty has also been investigated in CV by studying the effect of bid levels and respondents' attitudes towards the good to the level of preference certainty where the studies found that a higher level of respondent's responsibility results in increased probability of higher payment (Akter and Bennett, 2009). In explaining the level of preference certainty, CV studies have mainly focussed on the effect of respondent's characteristics such as education, income, age and gender (Akter *et al*., 2009; Alberini *et al*., 1997; Brouwer, 2009).

The inclusion of a 'don't know' or 'not sure' response in several initial CVM studies arose from the recommendation given by the NOAA panel on CVM[4]. Thus, Wang (1997), Alberini *et al.* (1997) and Akter and Bennett (2010) included the 'don't know/not sure' response or different methods of preference elicitation in accordance with the guidelines given by the NOAA panel. While early CVM studies do not cite any specific causes of preference uncertainty in CVM, they acknowledge the findings from several economics and psychology literature which state different causes of preference uncertainty among decision-makers such as arising from stochastic and unstable preferences, policy uncertainty, exogenous, endogenous and extrinsic uncertainty, inability to evaluate future utility and change of preferences which they recommend to be captured by using an interval WTP measure[5] rather than a point estimate (Hanemann *et al.*, 1996; Hanley and Kriström, 2002). In other cases, researchers have argued that as individuals do not know their

[4] Arrow K., Solow R., Portney P.R., Leamer E.E., Radner R. and Schuman H. (1993), 'Report of the NOAA Panel on Contingent Valuation', *Federal Register* **58**: 4601-4614.
[5] Under this approach, a respondent's true WTP lies somewhere within an interval which bounds the WTP value, which is used as the respondents could be uncertain of their precise WTP value.

true valuation with certainty, this should be captured using the stochastic error term (Li and Mattsson, 1995).

In a study to evaluate the value for the critical habitat for Mexican spotted owl, Loomis and Ekstrand (1997) acknowledged that respondent uncertainty could arise from doubts of the importance of the species in relation to other problems, preference uncertainty due to insufficient thought about the species and uncertainty of other people's vote in the survey. In terms of market products, Alberini *et al.* (2003) noted that lack of prior purchase experience, inability to use brands and models as signals of quality, lack of prior experience with researcher along with lack of recourse for bad decision making and little available information in the scenario on the item to be valued can cause doubts on the credibility of the valuation exercise. Thus, unclear questions or poor guidance during the valuation exercise, the hypothetical nature of the survey, inexperience of the good or the survey method and cognitive inability of the respondents to compare and rank alternatives have all been stated as causes of preference uncertainty, increasing the need to explicitly capture uncertainty (Boman, 2009; Broberg and Brannlund, 2008; Brouwer, 2009). Brouwer (2009) categorises preference uncertainty into: *uncertainty related to own future scenario, uncertainty related to bid price, policy scenario uncertainty and other survey instrument uncertainty* and find that policy scenario uncertainty is the largest source of uncertainty in the study.

While most of the studies have listed the hypothetical nature of the CVM exercise as well as the unfamiliarity of the good to be valued as the main causes of uncertainty, Welsh and Poe (1998) acknowledged that CV respondents could use different decision heuristics based on the elicitation method used and hence the authors compared the effect of different elicitation methods on respondent's uncertainty while Akter and Bennett (2009) regarded respondent uncertainty as a form of *cognitive uncertainty* caused due to the respondent's lack of confidence on their decisions. This was thought to arise in CV *due to error occurring at various stages of the cognitive information process* while ambivalent attitudes can cause dissonance which can further lower the certainty levels. Fraser and Balcolme (2010) noted that people could be better at responding to probabilistic intentions than absolutes as they could be inherently uncertain. As lack of market experience and previous thought on the valuation question along with greater knowledge of the good or service to be valued and lack of understanding of the future consequences from the committed payments have been stated as causing uncertainty (Kingsley, 2008; Loureiro and Loomis, 2008), it was suggested that uncertainty could be reduced from preference learning.

In addition to the uncertainty caused from the good to be valued, questionnaire design and respondent's cognitive inability, Shaikh *et al.* (2007) noted that the value an individual assigns to non-market goods is influenced by both its substitutes and complements while van Kooten *et al.* (2001) argued that the standard WTP estimate conceals underlying preference vagueness and could thus lead to biased outcomes.

Compared to the CV method where mainly one good is valued at a time, CE allows for multiple attribute valuation at a time where two or more alternatives are offered to respondents over several choice scenarios. As the method of attribute valuation and preference elicitation is quite different from the CV method, the questionnaire-based causes of preference uncertainty are also different. In case of the CE method, the difference between the attribute levels across two alternatives is considered to have a significant effect on respondent's choice. Restle (1961) point out two important causes of preference variation due to low differences between the utilities of the alternatives. The author states that when the utilities of alternatives are quite close to one another, random effects can significantly affect choice. This phenomenon is explained by the author through the threshold as well as the normal-curve theory. While the threshold theory implies a margin of error where choices are not consistently made, the normal-curve theory allows for the computation of the difference between the mean of the two alternatives' utilities, under a normal distribution assumption. In both the cases however, inconsistent choices are expected to occur when the utility difference is lower than the *error of judgment* (Restle, 1961). The proponents of the threshold theory in choice experiment apply this technique to imply that when the difference between the alternatives is too little, there is an increase in task complexity experienced by the respondent while when the difference is greater than a particular threshold, choices are easier to make and hence a greater certainty can be obtained (Cantillo *et al*., 2006; de Palma, 1998; Fowkes and Wardman, 1988; Jia *et al*., 2004).

Fowkes and Wardman (1988) acknowledge the threshold effect where an attribute could be ignored when the difference in the attribute levels is too small across the alternatives and especially when other attributes vary. This approach is also adapted by Cantillo *et al.* (2006) who state that perceptible changes above a threshold is expected to cause a reaction within an individual while below it is not. Thus they find that in models with threshold, parameter recovery is better and the model fits better than a misspecified model.

Jia *et al.* (2004) examined the effects of within alternative attribute conflict (defined as *discrepancy among the attributes of an alternative*) as well as the level of attribute difference on preference uncertainty on a rating exercise, using test-retest reliability measure and average response time. In order to obtain an implicit preference uncertainty measure, individual's response time was noted. The explicit measure of preference uncertainty was obtained by eliciting respondents' upper and lower bounds from the re-rating exercise. It was found that with more important attributes, there was increased uncertainty while attribute levels closer to each other showed a higher variance. The authors concluded that *greater attribute conflict leads to larger response errors, wider confidence intervals and longer response times while greater attribute extremity leads to less response errors, smaller confidence intervals and shorter response times*. The authors also noted that between and within alternative conflict[6] can cause feelings of regret and ambivalence respectively. The authors argued that within alternative conflict would result in consumer's uncertainty for an item and thus, attribute conflict and attribute extremity are important factors affecting preference strength.

Though level of attribute difference, attribute conflict and attribute extremity are important factors of choice design affecting respondent's preferences, Brouwer *et*

[6] While between alternative conflict can arise when one alternative is better on some attributes while another alternative is better in other set of attributes, within alternative conflict can arise when a particular alternative has a mixture of desirable and undesirable attribute levels thus resulting in ambivalent feelings towards that alternative (Jia *et al.*, 2004).

al. (2010) and Brown *et al.* (2006) notes that preferences change over the course of the session as they get more precise with experience. Hence, the stochastic error component is expected to decrease with repeated choice. Moreover, self-reported certainty level in this case is expected to increase with repeated choice and decrease as the utility differential decreases (Brouwer *et al.*, 2010).

In another study Oppewal *et al.* (2009) stated that preferences change after receiving new product information, thus prompting towards the concept of preference instability. In an attempt to elicit respondent's strength of preference and incorporate this information in modelling, Swallow *et al.* (2001) argued that dichotomous choice provides only limited information on the underlying preferences and hence the data is *inefficient as a foundation for the econometric exercise*. The strength of preference information was thus elicited to allow analysts to obtain further information and thus improve statistical efficiency. Moreover, the authors argued that respondents would be interested in stating their strength of preference (SOP) and would be less inclined to perceive the SOP indicators as manipulative or misleading. While the use of categorical response scale has been used in early SP studies as a rating exercise, Wardman (1988) acknowledged that this form of response scale provided more information about preferences than discrete choice, though no cause for preference uncertainty was specifically stated. Whelan and Tapley (2006) further reiterated this point that allowing respondents to indicate their strength of preference could provide a richer dataset and enhance the respondent's experience of the choice task.

Other causes of preference uncertainty in CE have been stated as arising from the precise value of the good, the meaning of the words used in the CE and the unfamiliarity with the good to be valued in monetary terms (Olsen *et al.*, 2011). Lundhede *et al.*, (2009) stated that previous studies have identified biases associated with respondent's processing of different attributes and the use of heuristics as the causes of uncertainty. Though not directly related to preference uncertainty, learning and fatigue during the choice experiment is known to affect respondents' choices with learning reducing the error variance and fatigue increasing the error variance (Savage and Waldman, 2008). Swait and Adamowicz

(2001) explored the effect of complexity[7] on utility by making the variance/scale a function of entropy/complexity. In some cases, the authors found a concave relationship between complexity and variance indicating that at smaller and larger levels of complexity the variance was small while it was maximum at intermediate levels of complexity. This supported the authors' hypothesis that at low levels of complexity, the ease of decision-making resulted in higher preference consistency across respondents while at high levels of complexity, preference consistency was obtained due to the similarity between alternatives. However, at intermediate levels of complexity, the effort expended by different respondents could vary, resulting in higher variances.

To summarise, Kosenius (2009) states several causes of preference uncertainty as can be found in the various literature. These include uncertainty on the good being valued as well as suspicions on proposed policy instrument, hypothetical nature of the exercise, difficulty in making trade-offs and respondents' difficulty in understanding of the question, the expression of certain (or uncertain) WTP as a means for accepting (or rejecting) a proposed program, trustworthiness of the authority and the level of information provided in the survey, the effect of survey instrument such as alternatives' utility difference, the bid amount given to the respondents and respondents characteristics such as prior thought on the good and the experience of the good as the main factors of preference uncertainty.

In light of the above-mentioned causes, two main factors for uncertainty can be identified under CVM and CE. The first factor pertains to choice set characteristics while the second factor is based on respondent's characteristics (Sun and van Kooten, 2009).

The examination of framing effects on preference and choice though well acknowledged in the psychology and risk literature (Tversky and Kahnemann, 2000) has been relatively less examined within the CE literature. Though some studies which examine the effect of different noise representation methods on valuation has been covered in the previous chapter, it is important to scrutinise this

[7] Defined as the number of alternatives and attributes as well as the levels of attribute correlation and preference similarity between alternatives (Swait and Adamowicz, 2001).

issue more generally in order to test the effects. In a study based on car driving commuters and non-commuters making choice in terms of travel time and cost, Hess *et al.* (2008) examined the effect of framing the attribute levels as increases or decreases from the reference alternative. By taking differences of the relevant attributes compared to the reference alternative and estimating separate coefficients for increases and decreases, the authors found that preference formation may not be dependent on absolute values of the attributes but difference from respondents' reference points. Based on whether the attribute was referenced positively or negatively, asymmetric preferences were obtained.

Along with the examination of different causes of preference uncertainty, another important factor is the method of eliciting respondent's preferences. Following the causes of consumer uncertainty in marketing literature, the subsequent section outlines the common methods of preference elicitation found in the literature along with the methods of analysing the preference data.

3.3.1.2 Consumer Uncertainty in Marketing Research

Consumer uncertainty has been examined in marketing research where studies have sought to examine the effect of choice and knowledge uncertainty on buyers' information search method. In a study conducted by Urbany *et al.* (1989), the authors examined the effects of these uncertainties on consumers' information search using a follow-up survey. While knowledge uncertainty (KU) was defined as the uncertainty related to the knowledge about the alternatives, choice uncertainty (CU) was defined as the uncertainty associated with buyer's choice. Hence while KU comprised of the uncertainty regarding the presence of certain attribute characteristics, the importance of the specified characteristics and which alternatives had which features, CU comprised of uncertainty associated with which brand, model and store to choose in the buying decision. The uncertainties were elicited by asking the consumer to rate their level of certainty on different factors on a seven point scale (sure/unsure). Based on a factor analytic approach, the authors found that both KU and CU explained about 51.7% of the variance in the measures. Though different combinations of these uncertainties within a consumer revealed different search patterns, choice uncertainty on average appeared to increase search while knowledge uncertainty had a weaker negative effect on search.

A similar study was conducted by Lauraeus-Niinivaara *et al.* (2007) where the effect of knowledge and choice uncertainty was examined on consumer search and buying behaviour. In this case, the authors extended the definition of knowledge uncertainty as that arising from the lack of factual information about different alternatives, the uncertainty about which decision rules to employ as well as the uncertainty associated with methods of acquiring the relevant information to make the choice. Thus, KU followed from the consumer's lack of awareness of the choice set and therefore which alternatives are available and what attributes are of significance and relevance. The uncertainties in this case were elicited by asking the consumers to speak their every move and reasons for choice aloud before an observer. Extending the definition of KU implied that reduced ability to understand and use new information would affect information search. Moreover, information search was likely to be greater when choice sets were more similar, resulting in greater choice uncertainty. Considering actual shopping time, number of brands considered and the number of stores considered as measures of search behaviour, the width of search was based on the number of alternatives considered while the depth of search was described as the number of product attributes that are evaluated. Search effort was characterised as the amount of time spent on search while the size of the consideration set was defined as the number of different alternatives (music records) considered during the experiment.

In this case the authors found that while KU increased the amount of time expended on search, consumers increased their consideration sets (i.e., the number of alternatives and product attributes) when faced with high CU. These studies revealed that consumer choice and knowledge uncertainty affects their search method as well as their buying behaviour. The search extent measures utilised in this study were search time and search range. While choice uncertainty was explained in terms of the number of alternatives considered by the buyers which implied that the number of choice alternatives were not fixed (as buyers were allowed to consider as many records as they wished), this measure has higher relevance in the marketing research. In case of choice experiments however, this cannot be a reasonable method as the number of alternatives are generally fixed in the choice scenario and hence alternative methods of capturing choice uncertainty needs to be devised. Nonetheless, these studies have revealed that KU increases search time while CU affects the number of alternatives and product attributes considered in the buyers' decision making.

Stone and Schkade (1991) applied numeric and linguistic representation methods for choice of alternative computer information systems. The numeric and linguistic representation methods were conducted using an equivalence scale where the following conversion measures were used: 2 (very poor), 4 (poor), 6 (fair), 8 (good) and 10 (excellent). Thus while some choice sets comprised of attributes represented by numeric values, others used the semantic form. Using a computer based survey where respondents were required to click on each attribute box for each alternative to have their values revealed, the authors examined the effects of information representation on search patterns through task complexity and effects on total effort through similarity of alternatives.

The authors recorded the time spent with each choice set along with the alternative selected. In order to assess the effects of different representation methods on decision strategies, the authors examined strategy characteristics as well as processing operations such as the *sequence in which the attribute values are read*, whether respondents compared two alternatives or whether decision was made based on one *item of information*, using MANOVA. The authors found that as task complexity increased, the search index shifted more quickly to attribute based method with numeric representation than with linguistic representation. While alternative based search method was associated with less complex choice tasks, a greater use of this search method with linguistic representation when higher task complexity was involved could imply that the choice task was less complex with linguistic representation than with numeric. However, through this result, the authors were tempted to conclude that *words may have inhibited the adaption of acquisition processes to task demand* causing a lesser move towards attribute based method. The authors found that with the numeric method, the respondents adopted a comparison method while with the linguistic method, the respondents studied each alternative leading the authors to conclude that numeric and linguistic representation methods *can lead to different choice processes*. Thus, this result suggests that varying methods of attribute representation has important effects on respondents' decision process.

In another study, Stone and Schkade (1994) used different attribute scaling methods for different word processing packages. The authors examined the difference between common measurement technique and context relevant measurement technique by comparing the effect of three different types of attribute scaling – the common context independent (CCI) scaling, the unique context independent (UCI) scaling and the unique context related (UCR) scaling. Under the CCI method, the attributes were represented on a single unit of measurement, a 1-10 point rating scale while in the case of the UCI method, the attributes were represented in a rating form albeit the rating scale differed for each alternative. For the UCR method, the attributes had a context dependent representation method (for example, cost in dollars, ease of use in number of keystrokes, wait time in seconds etc.).

Decision accuracy and decision speed were estimated and using MANOVA to analyse the data, the authors found that context relevant scaling resulted in faster responses while commonality of scaling resulted in more accurate decisions. Comparing the type of search methods employed with the common scales as well as the context related scales, it was observed that within alternative search was greater with CCI while within attribute search was greater with UCR. Thus, it was observed that the different scales resulted in different choices and choice processes while variation in representation from CCI to UCR scale, resulted in a significant shift in the direction of the search. From different experiments conducted, the authors found that description invariance assumption was violated even under minor variations in the representation method. These violations were found by *causing systematic differences in choices and choice processes*.

While the authors note that the CCI scale resulted in higher decision accuracy, the method of measuring this criterion of evaluation needs to be further examined. Taking the software ratings given by the Harold Gregory's guidelines as a reference, the respondents were asked to give a quality rating from a scale of 1 (worst) – 8 (best) for each software package. The accuracy of the respondents' ratings was assessed against that provided in the guidelines and thus, the level of decision accuracy was estimated. Though the authors nominate this technique as measuring decision accuracy, it is more related to the level of knowledge respondents have about each software (based especially on the guidelines used as a reference for the analysis) and thus this technique can be a better measure of level of knowledge uncertainty than decision accuracy. Though this is a shortcoming of the method of criterion definition, the research provided insight on the effect of different representation methods on respondents' choice and their choice processes.

In a study to examine temporal effects on the level of preference uncertainty, Salisbury and Feinberg (2008) offered participants to choose snacks for future consumption under simultaneous as well as sequential conditions. While all snack for future consumption were chosen in the current state under the simultaneous condition, the sequential method involved selecting each snack immediately prior to consumption. The authors observed that under the simultaneous condition, participants chose a greater variety of snacks and were less likely to select their preferred snack. The research argued that preference uncertainty was greater when making a choice for future consumption due to higher stochastic effects around the mean attractiveness of the available items. Moreover, choice set characteristics was argued to be a crucial component as the presence of a strongly favoured alternative resulted in low diversification irrespective of the time of consumption while the case of a weakly favoured item resulted in higher diversification.

The effect of choice set characteristics on consumers' preferences has also been examined by Yoon and Simonson (2008) who conducted several experiments to examine the effect of choice context on respondents' choices as well as preference strength. By offering participants choice sets under different conditions (absolute attribute values, context dependence as well as a control group), the authors found that under the asymmetric dominance effect where preferences are based on alternative's absolute attribute values, preferences are more stable with a higher level of confidence than in the case of the compromise effect where preferences are based considering the context. When participants were asked to identify the choices made during the experiment in the prior week, respondents of the asymmetric dominating option were better able to correctly identify their choice. While eliciting choice confidence (defined as the confidence participants had that their chosen alternative was the best option) across the two effects revealed a higher confidence associated with the asymmetric dominating option than under the compromise effect. The authors thus remarked that while preferences are observed to be unstable in some cases, in others they are extremely well-established and resistant to change and hence it is crucial to examine the factors that affect respondents' preference strength as well as the effect of *consumers' understanding of the impact of context on their preferences*.

The above studies have examined the different kinds of uncertainties affecting buyer's product search and purchasing decision as well as the effect of different forms of attribute representation on decision processes. While knowledge and choice uncertainties have been recognised as the main factors affecting buying behaviour, a comparison of numeric and linguistic representation methods showed that attribute based search method was predominant with the numeric method while alternative based search was more adapted with linguistic representation. Moreover, using different types of attribute scaling, it was observed that a significant change in decision process can be caused from slight variations in methods of representation. Thus, these studies revealed that not only are buyers affected by choice uncertainty, the decision invariance assumption is grossly violated under slight variations in the attribute representation method. While the above studies have focussed on the problem of buyer's uncertainty and the effect of representation forms on decision processes in the marketing literature, the following section examines the method of preference elicitation employed mainly in the environmental valuation literature.

As found in the literature reviewed for preference uncertainty in SP as well as consumer uncertainty in marketing research, the following causes of preference uncertainty and differing decision processes can be summarised:

1. In CV, the bid level as well as respondent characteristics such as income and education have been found to affect the level of certainty
2. The hypothetical nature of the survey (especially in CV) has also been known to cause uncertainty
3. The amount of difference between attribute levels across the alternatives in CE affects choice certainty where lower differences results in increased uncertainty
4. Task complexity in CE as measured by within and between alternative conflict has also shown to affect preference certainty
5. Some studies on different forms of attribute representations have shown that respondents adopt different decision processes based on the type and scale of representation
6. Respondents' cognitive ability is found to cause preference uncertainty in some cases while some researchers argue that respondents can be inherently uncertain

While the importance of attribute difference level has been explicitly stated in choice experiments and the effect of respondents' characteristics have been well-studied in CV, studies in SP show sparse examination of the effect of different attribute representation techniques in relation to choice modelling and its effect on respondents' preference uncertainty. In case of environmental and noise valuation especially, where different forms of attribute representation can be possible, not much evidence can be found in the literature within this area, where the effect of different representation forms on preference uncertainty have been examined. This comparative analysis thus forms an interesting research avenue which will be undertaken in this thesis.

3.3.2 Methods of Preference Elicitation

This section examines the methods of preference elicitation used with contingent valuation methods (CVM) as well choice experiment (CE) within the environmental valuation literature. The main focus of this section is to highlight the different methods of eliciting and treating the preference information obtained from the respondents as well as its possible effects on the valuation exercise. The section begins with the methods adapted in the contingent valuation (CV) literature followed by the methods used in choice experiment. While the main focus of eliciting detailed preference information in contingent valuation has been to examine the effect of respondent's preference uncertainty on the WTP estimates, the application of preference scale in CE especially in the 1980s has been less applied for the purpose of valuation, though increasing attention is now being paid to the effect of respondents' varying level of preference certainty. The rationale behind the use of the preference scales in these studies will be critically examined, followed by the application of preference scale in recent CE studies which aim to explicitly treat the preference information obtained from the respondents. The types, causes and methods of eliciting consumer uncertainty from the marketing literature will also be briefly examined later in the section.

3.3.2.1 Preference Elicitation and Treatment in CVM

The elicitation of respondents' preference certainty can take several forms as can be found in the literature. In CV, while some studies have allowed respondents to state their level of certainty by responding 'Yes', 'No' or 'Don't know' to a valuation question or through a status quo alternative (Mitchell and Carson, 1989; Morrison *et al.*, 1996; Wang, 1997; Welsh and Poe, 1998), others have sought to elicit the level of certainty by asking a post-decisional certainty question (Li and Mattsson, 1995). Another method of preference elicitation within the CV framework is by allowing respondents to state their WTP through a bounded range (interval method) or by allowing them to state their level of certainty through a polychotomous choice (PC) or numeric certainty scale (NCS) (Akter *et al.*, 2008; Hanley and Kriström, 2002; Loomis and Ekstrand, 1997; Loureiro and Loomis, 2008). Preference uncertainty has also been measured in CV using reliability studies through test-retest method where the same sample is faced with the same scenario on two different occasions (Brown *et al.*, 2006).

While the post-decisional certainty question can take the form of a debriefing question aimed to elicit respondents' confidence measure (0-100%) for each yes/no answer (Li and Mattsson, 1995) or elicit the level of their preference strength on a numeric certainty scale with a 1-10 point rating (Akter *et al*., 2008), the interval method either offers a bounded range to the respondents or compute this through a bidding process using the respondents' accepted and rejected bid amounts (Hanley and Kriström, 2002). The use of multiple-bounded or polychotomous choice in CV allows respondents to express their level of preference uncertainty over a Likert like scale of – 'Definitely Yes', 'Probably Yes', 'Maybe (Yes/No)/Not Sure/Uncertain', 'Probably No' and 'Definitely No' (Akter *et al*., 2008; Alberini *et al*., 1997).

The analyses of the numeric certainty scale (NCS) as well as the polychotomous choice (PC) in CV have been conducted using several methods. In case of the numeric certainty scale, a cut-off point can be assigned, below which the choice is deemed to be a 'No' response. Different recoding techniques are also applied with the NCS method where the 'Yes' responses are either coded based on the certainty score obtained (while the 'No' responses are taken to be all responses with 0% certainty score) or both the 'Yes' and 'No' responses are coded using the level of respondent's stated certainty (Akter *et al*., 2008). Preference levels from the PC data have been treated by recoding of the preference levels to either 'Yes' or 'No' responses using different methods such as calibrating 'Definitely Yes' and 'Probably/Maybe Yes' as 'Yes' or only considering 'Definitely Yes' as 'Yes' etc. (Akter *et al*., 2008; Alberini *et al*., 1997), through multinomial logit model (Alberini *et al*., 1997) as well as through ordered logit (Swallow *et al*., 2001) and ordered probit (Akter *et al*., 2009; Wang, 1997) models.

In a classroom experiment to evaluate the effect on natural resources from the Grand Canyon Dam, Welsh and Poe (1998) examined different elicitation methods by comparing different techniques such as dichotomous choice (DC), open ended

(OE) question, payment card (where respondents were offered different WTP amounts and asked to indicate their maximum WTP) and multiple bounded dichotomous choice (MBDC) – also known as the polychotomous choice (where respondents were asked to state their level of certainty on a five point scale from Definitely Yes to Definitely No) and observed that the mean and median WTP amounts differ based on the elicitation method used. Taking the DC, OE and payment card response methods as special cases of the MBDC likelihood function, different threshold values were computed for each of these.

Comparing the estimated logit distributions between the different certainty levels of the MBDC format and the other elicitation methods, it was found that respondents who indicated that they are ‘unsure’ of their WTP to a specific amount with the MBDC method, would accept a similar dollar amount threshold in case of the DC method while in the OE and payment card methods, a higher level of certainty was observed. In order to obtain a more reliable estimate of the WTP amount, the authors recommended treating only ‘Definitely Yes’ of the MBDC method as the certain WTP amount while rejecting all other levels of preference certainty. While this method could yield a more stringent WTP estimate, the lower levels of certainty (such as ‘Probably’) could provide useful implications for the welfare estimates. Moreover, to develop a bounded interval for the responses obtained from the DC method, the authors considered that with a ‘Yes’ response to a DC threshold, the WTP amount exceeded the DC threshold while for a ‘No’ response, the WTP was considered to be less than the DC threshold. However, this might not be the case in reality as for a ‘Yes’ response, respondents might be unwilling to accept any further bids exceeding that amount while for a ‘No’ response, some participants could also refuse lower bids (or indeed be genuine zero bidders). This method of applying threshold levels therefore to DC questions can be a contestable technique.

In a study to examine the predator protection policy in Sweden, Broberg and Brannlund (2008) combined the ordinary payment card and the polychotomous question to form the multiple bounded choice format where respondents were offered multiple bids and were asked to state their level of preference for each of the bids. This information was treated using the interval-estimation approach to obtain a less conservative mean and median WTP than that obtained by Welsh and Poe (1998). Using this expansion technique, the authors noted that the treatment of uncertainty using the payment card method was closer to the open-ended interval approach where respondents are free to state the WTP interval rather than a precise value. While in the case of Welsh and Poe (1998) where the uncertainty information is used to shift the WTP estimate, the approach adopted in this study increased bounds of the WTP estimate, thus expanding them. This approach was adopted based on the rationale that with larger WTP bounds, respondents would reveal a higher level of certainty; thus, by expanding the bounds, respondent's WTP would certainly lie within that range. Moreover, the authors noted that for respondents with high uncertainty a wider stated interval would be obtained and thus the expansion method could be used to account for uncertainty.

Boman (2009) offered multiple bids and the polychotomous choice with uncertain responses (also referred as the MBDC format) for each of the bids in order to elicit respondents' uncertainty, in a study to evaluate forest land protection for biodiversity in Sweden. Assuming that subjective numeric probability associated with each verbal category of choice can be used to inform the WTP estimates, the author examined different studies where respondents were asked to provide a numeric probability for each category. Based on the mean numeric probability values obtained from these studies for each of the verbal categories, the author estimated the range of WTP estimate among individuals by allowing the range of the distribution to vary based on their certainty levels (with normal distribution around mean WTP of SEK 70 and WTP range of SEK 288). Thus, respondents who were more certain (assessed by their numeric probability) would have a smaller range of interval than those with lesser level of certainty, which could have effect for benefit estimate. While the study outlined that the unmeasured probability could be associated with the ratio scale which underlies the MBDC response, the paper did not explicitly ask respondents to indicate the numeric probabilities associated with the verbal categories. However, Weijters *et al.* (2010) argue that providing levels of certainty is not commonly conducted through numeric probability and hence numeric probability scale can be more difficult for respondents to understand than verbal categories. However, a combined elicitation of numeric ratings and verbal categories of choice certainty as closely applied by

Boman (2009) could be adopted to better understand respondents' certainty levels as well as conduct other analyses.

Several analytical methods have been applied with CV to treat the preference uncertainty data. In a study to evaluate the Galveston Bay environmental quality, Wang (1997) included 'don't know' category in the referendum bid method where respondents were asked to vote whether they are: for the program, against the program or not sure (don't know) to different bid amounts. The author developed different models based on the treatment of the 'don't know' responses. While in one case constant thresholds were developed, in second case the 'don't know' responses were treated as 'no' and in the third case, these responses were eliminated. Thus three probit models were developed: ordered probit with explicit treatment of the 'don't know' responses, binomial probit with 'don't know' recoded as 'no' and binomial probit with 'don't know' excluded. It was found that in the latter two cases, the WTP value was underestimated. The results thus implied that explicitly treating the 'don't know' response provided a better fit (in terms of better estimation of the mean WTP) than either treating it as a 'no' or eliminating it.

In an application of post-decisional certainty question in CV, Li and Mattsson (1995) conducted a follow-up debriefing question to a yes/no question where respondents were asked to indicate their certainty to the WTP amount on a percentage graphical scale where 0% represented absolutely uncertain and 100% represented absolutely certain (with 5% interval). The follow-up question took the form, '*How certain were you of your answer to the previous question?*' The authors estimated the parameter vectors and standard deviation using the log-linear function as well as the normally distributed conventional estimate with a yes/no response. Comparing the parameter estimates obtained from the two models, the authors found that their method of considering certainty values yielded more accurate estimates. Moreover, they found that not accounting for preference uncertainty resulted in upwardly biased WTP values. Both Wang (1997) and Li and Mattsson (1995) have attempted to capture respondent's preference certainty using different methods with contradictory results obtained from the two studies in terms of WTP estimates, indicating that the method of preference elicitation has an important effect on the model output.

The main purpose of Akter *et al*. (2008) study was to examine whether consistent results are obtained from the different preference elicitation and calibration techniques across different CVM studies in the environmental literature. To that end, the authors compared polychotomous (PC) and numeric certainty scale (NCS) elicitation methods. While the NCS method generally allows respondents to indicate their choice over a numeric scale from 1 to 10 (or a percentage scale), the PC method offers the varied range Likert scale ranging from 'Definitely No' to 'Definitely Yes' for a WTP estimate. For both the NCS and PC data, the authors examined the effect of different recoding methods, observing that the WTP estimate varied significantly based on the recoding method used. Based on the upper and lower bounds of the 95% confidence interval and the mean WTP estimate, the authors estimated the efficiency of the WTP estimate. Contrary to the expectation, it was observed that models that accounted for preference uncertainty gave less efficient welfare estimate than the standard dichotomous choice data irrespective of the elicitation and calibration technique used. While this study aimed to examine the effect of two different elicitation methods in CV, the process of recoding the preference data significantly affected the results, implying that alternative methods of treating the uncertainty data needs to be considered instead of applying the recoding procedure.

In a CVM study conducted by Akter *et al*. (2009) to examine travellers' certainty of paying a carbon travel tax to offset carbon emissions if the tax was voluntary, the authors used the double bounded method (with start bid varying from Euro 5-100 per flight) which was followed by an open ended question to elicit the maximum WTP and a subsequent five point PC response (extremely unlikely, fairly unlikely, unlikely, fairly likely and extremely likely). Comparing the WTP estimates obtained from the double bounded and open ended question without considering uncertainty and the WTP from open ended with uncertainty, it was found that the mean WTP through the open ended question is higher than that obtained from the double bounded question largely due to anchoring of the OE WTP value on value cues given by the double bounded method. Using an ordered probit regression model, it was revealed that there is a significant negative relationship between the start bid and the likelihood of paying the voluntary tax. Thus uncertainty was observed to be higher with higher bid price and value cue in the valuation process.

Adapting a similar analytical approach, Alberini *et al*. (1997) applied the multiple bounded method in CV where respondents were offered a range of bid values and were asked to indicate the level of confidence with which they would pay each WTP amount on a matrix. Different models were fit to the data based on different assumptions and treatment of the response category. It was observed that the mean WTP estimate varied significantly while the coefficient estimates also varied in terms of the size, sign and the significance levels based on the type of model and recoding used. While one model used a separate error variance approach (incorporating heteroskedasticity), the effect of this technique was only observed on the mean maximum WTP estimate and the associated standard error without discussing the levels of the error variances for each of the preference levels. However, examining the results, a higher variance was observed for 'probably' alternatives compared to 'definitely' and 'not sure'. Examining the results obtained from the MNL model which was developed to estimate which characteristics explained the different preference levels, it was observed that the socio-economic characteristics (such as age and income) played an important role in explaining the preferences along with other study specific variables (i.e., *fishing experience variables*). Though the authors did not make any comparisons on the alternative specific constant (ASC) values obtained for each of the preference levels, it can be observed that the ASC value for 'definitely yes' had a high, positive and significant value while plausible values were also obtained for most other preference levels. Though discussion on the ASC values as well as the error variance obtained from the heteroskedastic model could have enhanced the result discussion in the paper, the approach adopted by the authors provided a different perspective to the treatment of preference data compared to commonly applied recoding and ordered probit methods.

3.3.2.2 Preference Elicitation in CE

Likert scales have been applied in SP in the 1980s on what can be termed as SP rating exercise. In one of the early applications of the Likert scale in stated preference to develop the work trip mode choice model, Kocur *et al*. (1982) conducted a survey with different experiments based on the different types of mode.

For each of these experiments reported in the paper, the alternative ‘drive alone’ was used as the base mode while the other alternative was one of the following: ridesharing, walking, bicycling and local bus service. For each of the binary choice experiment, the respondents were asked to indicate their choice on a Likert scale such as in the case of drive alone and car sharing, they were asked to indicate whether they would: always drive alone, probably drive alone, indifferent, probably share a ride, always share a ride. Using the categorical levels of preference as the dependent variables and the experimental attributes as the independent variables, the authors developed a multiple regression model *equivalent to the linear approximation of the logit function*. The sensitivity analysis was conducted using the incremental form of the logit model before the selection of the final model coefficients.

Through adjustments achieved from the calibration model, the final model of coefficients was developed. The study thus attempted to examine whether the functional measurement technique can be applied within the logit framework without adjustment and *whether sufficient variability existed to check the performance of the model*. While the authors developed policy implications and demand for the various modes using the model results, the explanation of the calibration method implied that both the ‘always...’ and ‘probably...’ levels were combined for the coefficient estimation, in which case the use of binary choice with an ‘indifferent’ alternative would have sufficed. Apart from these shortcomings, the rationale of using the scale was not provided by the authors. Thus, it was observed that though the Likert scale is used in the experiment, no treatment or conclusions can be derived for the cases with varying levels of preference certainty.

In another application of Likert scales in SP, Bates and Roberts (1983) offered different mode choice (car and rail) to the respondents. The respondents in this case were asked to provide their preference on a five point scale: Definitely prefer train, Probably prefer train, Indifferent, Probably prefer car and Definitely prefer car and were asked to make their choice based on car time, car cost, train time and train cost as the available variables.

The semantic scale of preferences was converted into a numeric scale where the probabilities were assumed to be at equal intervals. Thus, the scale from definitely prefer train to definitely prefer car was converted to a numeric scale corresponding to the following values: 0.9 (definitely prefer train), 0.7 (on balance prefer train), 0.5 (indifferent), 0.3 (on balance prefer car) and 0.1 (definitely prefer car). This conversion from semantic to numeric scale was carried out in order to examine *how sensitive the results are to the exact assumptions made about the probabilities*. The numeric scale was further modelled using the Berkson-Theil logit transform *log (p/(1-p))*. While the authors applied the Likert scale to elicit respondents' preferences in this case, the exact rationale for using the scale rather than just a binary choice was not stated by the authors. It could be implied that the use of this scale was undertaken more as a rating exercise and the log transformation of the numeric probabilities probably indicated a prevalent modelling technique adopted then.

Louviere and Kocur (1983) offered respondents an 11 point scale (11-almost always to 1-almost never) where the respondents were asked to indicate their proportion of trips for each of the 18 transit alternatives (all offered to respondents on one page), in relation to car use. Separate regression models were estimated in order to understand the trade-off between the different levels of service and the cost attributes. Examining the individual coefficients for each attribute across the different alternatives, the relative preference of the attribute level along with the correlation between different coefficients was computed (the authors found that as the fare rose from $0 to $0.5, the likelihood of transit decreased by 3.5 points on the 11 point scale). The preference rating data obtained was thus used to develop a forecast model. The authors found that using the deterministic forecasting model, the respondents would choose transit if their valuation was greater than six while use auto when this was less than six. While this study applied the 11 point scale in an interesting way to understand the effect of different levels of attributes on the likelihood of using the transit, the method of application implied that one of the alternatives need to be fixed (with consequence of possible fixed attribute values) throughout the experiment, which could not be a realistic possibility in all kinds of surveys. Thus the application of this method could be fairly restricted.

Bates (1984) asked commuters to indicate their preference of coach and train travel by rating each SP replication on a semantic scale of: definitely choose coach, probably choose coach, no preference might use either, probably choose train and definitely choose train. In order to analyse the preference data, it was converted from the semantic to the numeric scale using the method specified in Bates and Roberts (1983). By converting the semantic scale into a linear scale (1-2-3-4-5), the author calculated the average score for each of the 32 replications which revealed that an average score of greater than 3.0 indicated a preference for rail while that of less than 3.0 indicated a preference for coach. Thus, this study examined the scores obtained for the two modes under each replication within a set as well as across the sets to derive the relative preference for one mode of transport over another. While the information obtained from the preference scale was more incorporated into the models as well as in understanding the effects of different attributes on respondents' certainty levels, the implications of varying levels of preferences for valuation and forecasting was not derived in this exercise. However, the study estimated different values of time across different sets and types of the experiment.

In several case studies conducted by the MVA Consultancy (1987) on value of travel time savings, the SP method was applied either through a rating or a ranking exercise. Therefore in order to develop a balanced technique across the binary choice method (which does not elicit respondents varying levels of preferences) and the functional measurement method (which increases task complexity by asking respondents to state their preference level on a scale from 0-100), the authors applied the five point semantic Likert scale. While the methods of converting the semantic scale into a utility scale were exactly as described by Bates and Roberts (1983), this project report was one of the first to explicitly classify the application of the five point semantic scale of preference elicitation as a rating exercise.

Using the SP data obtained from the 1983 survey conducted for the UK Department of Transport where the choice between train and coach was offered to commuters with 16 replications, Wardman (1988) reported the application of the five point preference scale on commuters' mode choice in North Kent. While the main attributes used in the experiment were: main mode in vehicle time, other mode in vehicle time, walk time, wait time and cost, the respondents were asked to indicate their level of preference on the following five point scale: definitely prefer coach, probably prefer coach, no preference, probably prefer train and definitely prefer train. The author noted that the preference scale was offered to the respondents as the categorical scale provided more information about preferences than a simple binary exercise while at the same time offsetting the complexity involved with a continuous rating response scale. In order to analyse the categorical response data, the author assigned numeric values to the categorical responses in order to represent the probability associated with each level, which was then converted to a binary logit model using the linear transformation of the logit.

Using this method, the author found that the estimates obtained for the time coefficient and the money value associated with the alternative specific constant were similar to that obtained from the standard discrete choice analysis. However, the t-ratios obtained from the assumed probabilities showed a higher value, inducing the author to conclude that this method captures more information of respondents' preferences than the standard discrete choice. The author also found that the coefficient estimates varied with the different choice of the probabilities though the estimated relative valuation remained unaffected. While the author did apply the categorical response scale to elicit respondents' preferences, this information was not explicitly treated and hence no insight was gathered on preference uncertainty or the causes for it.

In another application of a five point Likert scale in travel demand forecasting, Wardman (1987) compared the application of the logit transform on the categorical responses as well as the discrete choice model (with elimination of the 'uncertain' responses and combining the 'definitely' and 'probably' levels). Comparing the effects on probabilistic and deterministic forecasting[8], the author noted that with the logit transform which assigned probability values for each response category the linear logit model though based on sensible assumptions of the probability values

[8] The deterministic forecasting method assigns individual to the alternative with the highest utility, given the estimated coefficients prevailing at the situation to be forecast while the probabilistic method calculates the probability of choosing the alternative for each individual given the estimated utility difference across the alternatives *for the situation to be forecast* (Fowkes and Preston, 1991).

could affect the coefficient estimates as this is dependent on the assumptions made. However, the relative utility weights were found to be unaffected to the assumptions made and were similar to that obtained from the discrete choice model. Thus, the author concluded that this method of analysis is more suitable for deterministic forecasting and inter-attribute valuation while in the case of probabilistic forecasting this method could be unreliable as the coefficient estimates obtained is dependent on the probability assumptions made. Moreover, the method would also introduce an additional error due to the arbitrary assignment of the probability values.

Examining the studies conducted in 1980s which allowed respondents to indicate their varying levels of preferences, it can be observed that while Louviere and Kocur (1983) have tried to incorporate the information obtained from the preference scale to understand respondents' choice, other studies have converted the semantic scale to a numeric scale with assigned probability levels. While modelling from the 11 point scale by Louviere and Kocur (1983) have been conducted by treating the scale response as estimates of utility dependent on attribute characteristics, the method of application implied that one of the alternatives need to be fixed. The rationale of eliciting respondents' preference levels in cases of numeric probability transformation has neither been explicitly stated nor did the modelling method of the obtained data justify the use of a preference scale. Thus, elicitation of preferences in this form can be regarded as a rating exercise adapted due to technical/technological limitations of the time. Moreover, the logit transform method adapted for the numeric transform of the semantic categories has been argued to affect coefficient estimates due to the probability assumptions made. The application of the five point Likert scale in the older SP studies has thus been mainly conducted for reasons other than to explicitly model respondents' elicited preference levels.

While later choice experiments were commonly found to elicit preferences either through binary or multinomial choice, the application of Likert scales can take two forms - the commonly used five point Likert scale which forms the one stage Likert method or a post-decisional certainty measure which can be regarded as the two stage Likert method (Albaum, 1997). Both these methods of preference elicitation have found some applications within recent choice experiments (Lundhede *et al*., 2009; Whelan and Tapley, 2006).

In a binary choice between landfill sites, Swallow *et al*. (2001)[9] asked respondents to choose a preferred site and then indicate the levels of their preference on a five point scale ranging from slightly prefer (1) to strongly prefer (5) where only the end-points of the scale were labelled. The preference data obtained from the Likert scale was analysed using ordered logit while binary logit from the dichotomous choice was also developed; these models were then compared through parameter estimates for alternative sub-samples as well as the mean square difference (MSD) between estimated utilities of small and large sample and MSD for WTP estimates. The study revealed that the ordered logit gave a better model fit, tighter standard errors and the implicit ordering of preferences was also maintained (which was judged based on the monotonic relationship exhibited by the threshold values along with the middle parameter not being statistically different from zero while all other parameters were very significantly different from zero). Comparing the MSD for estimated utilities obtained from large and small samples across the two analytical models, it was revealed that the ordered logit model performed better than the binary model. The small sample ordered logit model was found to be a better predictor of the large sample ordered logit than in the case of the binary logit model. The authors noted that allowing for extra preference information obtained from the Likert scale improved the econometric ability to evaluate utility changes. This study thus proved that accounting for different preference strengths resulted in improved model fit and greater efficiency.

Another application of ordered logit analysis can be found in Whelan and Tapley (2006) who applied the five point Likert scale to elicit preferences between tolled tunnel and un-tolled bridge routes. The main aim of the authors in this study was to identify and capture the subjective difference in the Likert scale interpretation across respondents using a Mixed Ordered Logit model. Hence in this case, the threshold parameters of the OL model were allowed to vary following a distribution, along with the parameter estimates. The authors incorporated the

[9] Though the authors in the paper state that they employ CVM, the review of this study is positioned in this section as it employs a contingent choice (binary choice), similar to CE.

Likert scale responses into the model using three methods: 1) recoding it as a binary choice, 2) having three levels – tunnel, uncertain and bridge and 3) maintaining the five point preference scale. For each of these methods, different forms of model calibration were used depending on the type of distributions assumed for the coefficients. The authors found that while the mixed model revealed significant variation in the threshold parameter estimates, no significant variation was obtained in the relative attribute valuations. The authors also concluded that in order to obtain better precision of preferences, respondents must be allowed to state their level of preference. While this study extended the ordered logit approach by acknowledging the variation in respondent's understanding of the preference levels, the analysis of the preference data however, still followed the common method of analysis (i.e., OL), although the importance of capturing preference uncertainty was rightly stated.

While some studies can be found in the CE literature where preference levels were elicited, the common method of treating this preference uncertainty data with CE has been through recoding or elimination.

In order to examine the effect of outcome uncertainty on respondent uncertainty in the context of avoiding poor state of the Gulf of Finland in the future, Kosenius (2009) developed three questionnaire versions based on the different certainty framing of the alternatives. While in the base version (without uncertainties), all alternatives were defined as 100% certain, in case of the outcome uncertainty version, either the outcomes of the policy options or the business-as-usual (BAU) options were defined as less than 100% certain. In case of the presence of outcome uncertainty, respondents, following all choice tasks, were asked to state their level of choice certainty on the scale: certain, quite certain, quite uncertain and uncertain, corresponding to each choice task. Applying recoding (where uncertain preference data is defined as business-as-usual or best available alternative) as well as preference elimination techniques, it was found that the elimination method yielded a better model fit and hence the authors recommended this technique to handle respondent uncertainty in CE. However, elimination of uncertain preference data has a clear disadvantage in the loss of information along with reduced insight on the specific impacts of preference uncertainty.

In two studies, one to elicit preferences for conversion of natural areas (forest, wetland, heath, arable land) to motorways and another on the preferred type of national park, Lundhede *et al*. (2009) asked respondents to state their post-decisional certainty level by asking them to indicate their level of certainty on the scale: very uncertain, uncertain, neither certain nor uncertain, certain, very certain and don't know for the motorway survey and very certain, certain, uncertain, very uncertain and don't know for the national parks survey, following each choice scenario. The certainty data thus obtained was analysed using three different methods of recoding and two models with scale parameters. Modelling with scale parameters took the following forms: 1) the variation in scale parameter was linked directly to the respondent's stated level of preference uncertainty and 2) the scale parameter was a function of factors that were found to correlate with the level of preference certainty. Under the recoding methods, the authors used the following approach: 1) elimination of uncertain responses, 2) recoding the uncertain response as the status quo alternative and 3) recoding the uncertain response as the best alternative different from the one chosen.

Computing the scale parameter for each of the certainty levels with respect to the normalised category using the estimated scale parameter value and then calculating the respective error variance, the authors found that there is increased error variance for the uncertain responses especially in the case of the motorway survey. In terms of recoding processes used, the authors found that elimination and recoding did not affect attribute valuation while the elimination model was found to give a better model fit than the original data, reflecting the reduction of noise. The authors thus indicated that recoding uncertain responses had significant shortcomings due to change or elimination of some preference data. Using the scale parameter approach, the authors found a significant link between the certainty level and the error variance, indicating that the error variance increases as preference uncertainty increases. This approach can also be varied to form the error components model which will be later examined in the thesis. As the authors failed to capture any change in scale using respondent and choice set variables as proxies for the preference certainty levels, they recommended explicit modelling of the uncertainty data using the scale parameter approach instead of using proxies for the choice certainty levels.

3.3.2.3 Preference Elicitation in Marketing Literature

In a study to compare the effects of different rating scale formats on response styles, Weijters *et al.* (2010) examined the effects of different response formats on three response biases (net acquiescence response style (NARS), extreme response styles (ERS) and misresponse to reverse items (MR)). NARS focuses on the extent to which respondents are inclined to agree with the item rather than disagree, irrespective of its content causing a central tendency of the rating scale measure. The ERS on the other hand arises from a tendency to disproportionately provide extreme category responses, thus causing a wider spread in the observed data. The MR bias is caused from a tendency to respond in the same direction to two items that are contrary in meaning, thus agreeing to an item and its reversal or vice versa. The main purpose of this study was to examine the effect of different types of Likert scales (based on the labelling of the *response categories and the number of response categories offered* – ranging from four point scale to seven point scale) on response formats. The effect on these biases was examined as they affect the observed means, variance and internal consistencies of scales. While different effects on biases were observed based on the different forms of response formats, a fully labelled five point scale (or seven, when respondents were students) was found to be more preferable where direct summaries of responses were needed while a five (seven) point scale with endpoint labels were more preferable in cases where linear relations such as correlation, regression and structural equation models with variables were required.

In another study to compare the effects of binary, seven point ordered and metric response formats, Dolnicar and Grun (2007) used the different answer formats to investigate whether any variations in content-unrelated error was found based on the answer scale. By conducting several surveys where the same sub-sample was exposed to the three different answer scales consecutively, the authors found that in the mapping analysis, the answers on the metric and the ordinal scales were not comparable and could not be transformed from one to another. This implied that the tendency of answering certain levels could be dependent on the type of scale used. However, it was found that irrespective of the answer format, the main conclusions drawn remained the same, implying that the analysts are free to choose the optimal answer format based on other criteria such as the required duration to complete the questionnaire and the complexity of the task. While binary response was found to be relatively faster, no difference was found between the different answer formats in terms of simplicity, relative pleasantness or the ability to express feelings. This result contradicts results obtained from previous studies which indicate that respondents prefer multiple categories as it better enables them to express feelings. Thus, it is important to compare the effect of different elicitation formats especially in the case of different representation methods.

While previous studies have mainly focussed on capturing preference information either through a Likert scale or a post-decisional certainty measure, Albaum (1997) examined the effect of different types of Likert scales in terms of the extremity and intensity of responses. While extremity and intensity can be taken to have some common variance, they need not necessarily be correlated. Thus, this study examined whether *one stage Likert under-reports most intense positions compared to two stage Likert*. The author conducted a survey eliciting respondents' attitudes towards economic systems in United States, New Zealand and Denmark with one and two stage Likert elicitation methods. In order to compare the results obtained from the one and two stage Likert methods, the responses obtained from the two methods were scored in comparable terms. The author found that the absolute value obtained from the two methods differed based on the method used. It was found that the two stage Likert method resulted in more extreme values than the one stage method. The author thus concluded that the commonly used one stage Likert method resulted in erroneous outcome due to central tendency while the two stage method allowed respondents to state their true preferences. Though this is an important finding, the application of the study was not related to environmental or transport issues. Moreover, the analysis was not based on random utility theory. The applicability of this result will thus be tested in the current study by adopting both methods of preference elicitation.

To summarise the literature reviewed on methods of preference elicitation, the following points can be noted:

1. Polychotomous choice and post-decisional certainty measures are common techniques applied in CV. In CE, these can be extended to the use of one and two stage Likert methods
2. Recoding and elimination of uncertain data have been widely applied although it is accepted that these methods result in loss of preference information
3. Analytical methods to treat the preference data have included the use of multinomial logit model, ordered probit, ordered logit, mixed ordered logit models and the scale parameters approach

Considering these points on the methods of elicitation and analysis, the following lacunae in the literature can be observed:

1. <u>The comparison of different elicitation methods within CE</u>: while the application of one stage Likert method is relatively common to elicit respondents' preference certainty, the application of two stage Likert is rare. Moreover, no study can be found within the choice modelling literature which compares the effect of the two elicitation methods
2. <u>The method of analysis of the preference data</u>: while ordered logit and ordered probit have been commonly applied to analyse the preference uncertainty levels within CE, more examination of the error structure and the effect of uncertainty on the error form can be conducted to analyse the preference data

In light of the gaps observed in the literature and in relation to the effects of different attribute representation methods, these issues will be addressed in the thesis.

3.4 Effect on Valuation and Forecasting

This section examines the effect of preference uncertainty, strength of preference and presentation effects on valuation and forecasting. A central concern for capturing preference uncertainty in CVM as well as CE studies apart from getting further information on respondents' preference level is the effect uncertainty can have on willingness to pay (WTP) estimate. Several studies have been found across these two valuation methods where the effects of preference uncertainty as well as the different methods of preference elicitation and calibration on the WTP estimate have been outlined.

Comparing preference uncertainty adjusted WTP across the literature with a dichotomous choice (DC) WTP estimate, Akter *et al.* (2008) expected the uncertainty adjusted WTP to be lower than the DC estimate as capturing preference uncertainty is considered to remove hypothetical bias which is thought to result in a higher WTP estimate. However, examining the results from different studies it was found that this expectation was not valid across all different elicitation and calibration methods. With the numeric certainty scale (NCS) elicitation method, almost all the studies that the authors reviewed found that using the different recoding methods, lower welfare estimates were obtained except for one study which obtained a higher estimate. With the polychotomous choice (PC), variable results were obtained across different studies.

Comparing PC, NCS and the composite certainty scale (CCS – *which combines verbal expressions with numerical and graphical interpretations*), (Akter and Bennett, 2010) found that a greater number of 'yes' responses were obtained with the PC format compared to the DC method while CCS obtained a greater number of certain responses. Comparing the certainty calibrated mean WTP with that obtained from the DC WTP estimate it was found that all calibrated mean WTP were lower (except in one case) than the DC mean WTP estimate. With the PC method, the mean WTP was found to be 117% higher than the DC WTP estimate when 'maybe yes', 'probably yes' and 'definitely yes' were recoded as 'yes' while in other cases where certainty cut-off points were included, the mean WTP reduced from 3-83% compared to the DC WTP estimate. The lowest change in WTP estimate compared to the DC method was found in case of the CCS method (where a reduction of 3-32% was obtained) due to the relatively small calibration scale used compared to the other methods. Thus, this study showed that different methods of preference elicitation along with different preference calibration methods can cause significant variation in the WTP estimate.

In another study where respondents were asked a double bounded (DB) WTP question on different starting bid amounts followed by an open ended (OE) question which was subsequently followed by a five point scale (extremely unlikely, fairly unlikely, not sure, fairly likely and extremely likely) on respondents' likelihood of paying on the OE WTP question, Akter *et al.* (2009) found that the mean WTP from DB question without uncertainty was 23 Euro/flight while from the OE WTP without uncertainty was 43 Euro. Moreover, a substantial variation in the WTP amounts was obtained based on the different recoding methods used. The authors also found that the OE WTP values were significantly clustered around the five point scale categories which implied that not accounting for respondent uncertainty resulted in increased estimation errors.

Alberini *et al.* (1997) and Alberini *et al.* (2003) emphasised the effect of recoding method on the WTP estimate using the MBDC response format. The two studies revealed that based on the analytical and recoding method used, the WTP estimate varied substantially. Wang (1997) found that treating 'don't know' as 'no' or deleting it lowered the mean WTP compared to the threshold models. The effect of different recoding methods as well as different estimation methods was also observed in Broberg and Brannlund (2008) who obtained a substantially high WTP amount as all levels of uncertainty were recoded as 'yes' responses. Comparing different bid orders Alberini *et al.* (2003) found that the descending order of the bid panel yielded higher mean welfare estimates while including the uncertain responses in the model also substantially increased the welfare estimates.

Examining the effect of uncertainty level on WTP, Boman (2009) found that individuals who are uncertain have a greater propensity to accept higher bids than his/her maximum WTP. This finding has been previously stated by Hanemann *et* *al.* (1996) who found that the preference uncertainty parameter in the SP study is quite high indicating that this parameter captures much of the variation in the individual WTP and hence the mean WTP amount when accounting for preference uncertainty is lower than that obtained from the conventional model which does not incorporate uncertainty information. Li and Mattsson (1995) found that the overall mean WTP estimate was reduced by about six times compared to the model without preference uncertainty implying that not accounting for preference uncertainty can cause serious upward bias of the WTP estimate.

Comparing the MBDC format with different recoding approaches along with the payment card and dichotomous choice, Welsh and Poe (1998) found that very substantial difference in the mean WTP estimate was obtained based on the different specifications of the switching interval. When the lower end of the switching interval was chosen as the highest amount at which the respondent chose ‘definitely yes’, the MBDC method gave a lowest mean WTP of $16.70 while when the lower end of the switching interval was taken as the highest amount at which the respondent chose ‘probably yes’, a mean WTP of $39.56 was obtained and with that for ‘not sure’, a highest WTP of $92.96 was obtained. The mean WTP obtained from the payment card method was $36.64 while that obtained from the DC was $98.40. Comparing the results obtained from the different elicitation methods as well as the specification of the switching interval for the MBDC, it can be seen that the mean WTP estimate is highly sensitive to both the issues. Moreover, results from the DC yield a higher mean WTP compared to the MBDC method with the ‘not sure’ model which reiterates the previous findings that not accounting for preference uncertainty can result in an upwardly biased WTP estimate.

By asking respondents whether they are ‘definitely sure’ or ‘probably sure’ of paying a particular price which was followed by a numeric certainty scale where 0 corresponds to ‘not sure’ and 10 corresponds to ‘very sure’, Blomquist *et al.* (2009) sought to find a point on the numeric scale which provides the same WTP estimate as that obtained from the ‘definitely sure’ response for three different disease management programs. It was observed that the same mean WTP estimate as obtained from the ‘definitely sure’ response was obtained for numeric scale 9.9, 10 and 9.0 for each of the programs. While the authors aimed to find the numeric certainty level that relates to the 'definitely sure' response, the numeric level given by the respondent can be highly subjective as well as context specific and hence, this finding cannot be extrapolated to other studies.

Brouwer (2009) found that using a post-decisional NCS question where respondents were asked to indicate on a 100 points scale (with 10 percent intervals) whether they are 'not at all certain' or 'completely certain', respondents who are 100% certain of their WTP amount are willing to pay more compared to respondents who are not 100% certain. Thus the author concluded that not accounting for preference uncertainty could result in an over or underestimation of the welfare measure based on the presence (or absence) of preference uncertainty. Moreover, as the statistical efficiency of the welfare estimate was found to be lower for respondents who were uncertain, the author concluded that accounting for self-reported uncertainty in econometric models leads to smaller and less precise welfare estimates.

Loureiro and Loomis (2008) asked a follow-up certainty question where respondents were asked to indicate their level of certainty on a scale from 1 (not sure) – 10 (totally sure) on a referendum WTP question. Analysing the data by classifying it into two major segments of respondents – the 'uncertain' and the 'certain' class using the finite mixture model, the authors found that the WTP estimate conditional on the segment was affected such that the 'certain' segment had a fairly higher estimate than that obtained from the 'uncertain' segment with the mean WTP of individuals belonging to the 'certain' group being Euro 82.14 while for those in the 'uncertain' group being Euro 54.08.

Examining the results obtained from the above studies it can be seen that while not accounting for preference uncertainty (such as in the DC method) can result in upwardly biased WTP estimate, 'certain' respondents are willing to pay more than 'uncertain' respondents while the modelling method employed can have a significant effect on the WTP estimate obtained as can be seen in Welsh and Poe (1998).

Shaikh *et al.* (2007) has shown that based on the analytical model used, the preference uncertainty adjusted WTP can be higher or lower than the standard model. For the Weighted Likelihood Function Method (where initial DC question was followed by a NCS method on a 0-100 interval and responses were recoded based on the certainty score), asymmetric uncertainty model (where using a scale from 1-10 respondents were asked a certainty follow-up question and all 'yes' responses were recoded as 'no' if the respondent was not completely certain) and fuzzy model (where *the post-decisional confidence of a response was used to determine the membership values of the WTP and willingness not to pay (WNTP) fuzzy sets* while the intersection of the estimated fuzzy sets determined the membership values of *the 'comfort' level of the associated welfare estimates*) approaches, incorporating respondent uncertainty was found to lower the WTP estimate while in the case of the Random Valuation Model (an individual's value of an amenity was considered as a random variable with unspecified probability distribution and thus error was incorporated into the model) and Symmetric Uncertainty Model (based on the certainty score given by the respondent, the responses were recoded - thus 'yes' with 60% certainty was recoded as 0.6), the opposite result was obtained.

Thus, studies in CVM have revealed significant effects on the WTP estimate and hence on attribute valuation based on the elicitation and calibration methods used. While comparisons of alternative preference elicitation methods is relatively more examined with the contingent valuation literature, fewer studies can be found within CE which have compared the effect of alternative preference elicitation forms on valuation and forecasting. A major area of interest in CE however has been the comparison of different modelling techniques on model fit as well as attributes valuation.

Within choice experiments, in a comparison of the effect of different treatments of the uncertain responses by recoding and eliminating them, Koseinius (2009) found that compared to the base model, recoding the uncertain responses to business-as-usual or the best available alternative decreased the mean WTP while eliminating it increased the WTP estimate. Compared to the benchmark model where respondents' uncertainty is ignored, Lundhede *et al.* (2009) found that models with uncertain choice eliminated slightly increased WTP while models with asymmetrically recoded (uncertain choices were recoded as status-quo) method decreased the total WTP for a change from the current situation. In case of symmetric recoding (uncertain choice was recoded as the best alternative different from the one chosen) as well as for models with certainty level (incorporated through a scale parameter), the effect on WTP was insignificant. Applying the five point Likert scale where respondents were asked to make a binary choice as well as indicate their strength of preference, Swallow *et al.* (2001) compared results obtained from the binary and ordered logit models to find that the ordered logit model outperformed the binary logit model in estimating the WTP.

In another application of the five point Likert scale in CE, Whelan and Tapley (2006) adapted different analytical techniques to model the data. The authors combined the 'definitely' and 'probably' levels while eliminating the 'no preference' response to form a binary choice, converted the five point scale to a three alternative choice by including the 'no preference' responses but combining the 'definitely' and 'probably' levels as well as analyse the five point scale through ordered logit. The authors noted that though comment on overall model fit was difficult to make as the dependent variable varied across the different specifications, the scale of the model coefficients decreased as the strength of preference was incorporated into the model, implying that the statistical precision of the estimation increased as preference uncertainty levels were incorporated though relative attribute valuations across the different techniques were found to be quite constant.

While comparison of alternative preference elicitation format in choice experiments on valuation and forecasting is extremely rare, the above studies have revealed that preference uncertainty can have substantial effect on valuations and methods of elicitation as well as calibration can play a significant role in the WTP estimate. While forecasting is a significant factor in marketing and transport studies which can be affected by the method of preference elicitation, a major concern in the environmental literature has been the effect of preference uncertainty on valuation. Comparison of different elicitation formats in the contingent valuation literature reveals that the method of preference elicitation does affect valuation and the WTP estimate. In terms of forecasting, the method of recoding and the treatment of uncertain responses can play a significant role in the forecast.

In a study examining alternative catchment management strategies where an attribute was defined as ‘species lost’ or ‘species present’, Kragt and Bennett (2009) found that significant differences in the implicit prices between the two versions were obtained with higher valuation obtained when the attribute was framed as ‘lost’ though with lesser precision of the WTP estimate, compared to when it is described as ‘present’.

In order to examine how framing affects the valuation of mortality risk reduction, Rheinberger (2009) conducted a split-sample survey where one half of the respondents were asked their WTP for reduction of up to 16 fatalities per year with reference to the 7.5 million residents of Switzerland while the other half was offered an identical choice where avoided fatalities was reframed in reference to the 500 road fatalities that occur in the country. Comparing the value of statistical life (VSL) across the two framing methods, it was found that the VSL was CHF 8.44 million where risk reduction was framed in reference to road fatalities while it was CHF 7.12 million when it was framed in reference to the residential population. However, while the values obtained were statistically significant, large standard error implied that the framing effect on VSL values could be random. Analysing the effect of different framing effects on the different sample based on the interaction with socio demographic variables, the author found that the different sub-samples adapted different techniques to form perceptions of risk reduction which further affected their choice though the effect on valuation of the traffic safety programs from the different framing methods was not statistically significant.

In another study conducted by Howard and Salkeld (2009) to compare the framing effects within health context, *potential benefits and harm of screening tests were presented in both positive* (number of cancer/large polyps found) *and negative* (number of cancer missed) *terms*. Three attributes – cancers found, large polyps found and test prediction where thus framed as: a) number of cancers found and missed, b) number of large polyps found and missed and c) people correctly reassured that they do not have cancer and unnecessary colonoscopies. Comparing the parameter estimates of cancer found with cancer missed when the specificity attribute of test accuracy was expressed as 'reassured' or 'unnecessary tests', the results found that with the framing 'reassured', the framing did not significantly affect the valuation of the attribute while in the case where it was specified as 'unnecessary test', the sensitivity attribute was found to significantly affect valuations. The same results were obtained when framing effects of cancer test were analysed on the number of polyps found or missed. Moreover, the WTP for test attributes was found to be significantly affected by framing. While the WTP for one extra cancer found was $11.45, for one fewer cancer missed was $9.81. Thus, the study revealed that attribute framing significantly affected respondents' WTP.

Though some of the above studies have focused on framing effects as an extension of the gain-loss asymmetry literature found on risks (Kahneman and Tversky, 2000), it is nonetheless significant to observe empirical evidence of the theory especially as applied within CE. The other studies have also revealed that different framing methods can affect attribute perception as well as valuation. While the above studies have focussed on the effect of framing on valuation, an important aspect that has been rarely examined is the effect of framing on respondent's level of preference uncertainty. The examination of this aspect thus encompasses one of the novel aspects of this research.

As uncertainty can be defined in different ways which can have an implication on the method of analysis employed, the following section will outline the different forms of uncertainty representation along with the relevance of this classification in the research.

3.4.1 Uncertainty Representations

Besides the work conducted in utility theory in context to risk and uncertainty as outlined in the previous sections, a more general classification of uncertainty can also be found in the literature. When employing analytical methods different from probabilistic measures, this classification becomes especially pertinent as will be observed later in the section.

An interesting classification of uncertainty is given in Parsons (2001) who uses Smithson's and Smets' taxonomy (which classifies sources of 'ignorance') to characterise types and sources of uncertainty. Based on Smithson's taxonomy, three causes of uncertainty specified are vagueness, probability and ambiguity. While uncertainty due to probabilistic causes can be more generally viewed as risk, uncertainty from vagueness arises from fuzziness and non-specificity (Parsons, 2001). Based on Smets' taxonomy, the state of 'ignorance' is classified into uncertainty and imprecision. In this case, while uncertainty encompasses whether a state is subject to randomness or is believable, it is imprecision that results from fuzziness (Parsons, 2001). This understanding of uncertainty is especially pertinent to the thesis. While most choice and behavioural decision theory literature focuses on uncertainty as associated with probabilistic risk, the interpretation of uncertainty in this research arise from vagueness and ambiguity, along with the interaction between the two.

Vagueness can be defined as *something that is not well-defined* while ambiguity is caused to exist due to lack of information *that distinguishes between two alternatives* (Parsons, 2001). These definitions hold significant insight in understanding the concept of uncertainty as pertaining to this research.

In case of this research which aims to apply the location and linguistic methods of attribute representation, the concepts of vagueness and ambiguity are especially interesting. Vagueness, the quality of being poorly defined, is a rather subjective term as it is based on how well the respondent relates the attribute to the method of its representation. While instinctively, the term might arouse negative connotations, as vagueness is something that needs to be avoided, the quality and application of it however, might not be entirely negative. Two forms of vagueness can thus be identified – 1) vagueness arising from the vague representation of the attribute and 2) subjective vagueness arising due to respondent's perception that the method of representation in relation to the attribute, is vague. Thus, while vagueness in the first instance is not bad as some attributes are better represented in vague terms, it is vagueness in the second instance that can create problems in the choice task.

Ambiguity refers to lack of information to distinguish between two alternatives. In terms of choice task, this concept is dependent on the choice-set characteristics as well as the respondent's prior experience with the method of representation in relation to the attribute i.e., the respondent's understanding of each attribute level based on its representation. Thus it is evident that this research provides an excellent opportunity to understand the effect of both vagueness and ambiguity through the application of the location and linguistic representation methods. It is expected that ambiguity will imply increased effort from the respondent to successfully conduct the choice task and hence has a propensity to increase preference uncertainty. Vagueness on the other hand, can increase or decrease preference uncertainty depending on which form of vagueness is present. When attributes are represented in vague or linguistic terms, but are clearer for respondents to understand, reduced preference uncertainty is expected to be observed while in the other case, contrary result is expected.

In relation to the treatment of uncertain data, several approaches can be outlined. The most well-known treatment of uncertainty to economists is the probability measure where probability is assigned to an event while other approaches include the Possibility measure and Dempster-Shafer belief functions (Halpern, 2003; Parsons, 2001).

While Dempster-Shafer belief function gives the probability that the available information can prove the truth of the proposition (Parsons, 2001), the possibility measure is based on the concept of fuzzy logic. Possibility measure bears a confusing similarity with the probability measure; however an important distinction lies in the union of all the disjoint sets under both these measures. If we assume for sake of simplicity that two disjoint sets exists, the union of these two disjoint sets in case of the probability measure then sums to unity while with the possibility measure, the union equals the maximum value of the two disjoint sets (Halpern, 2003). This concept will elaborated in more detail in the subsequent chapter.

This section has aimed to provide a summary of different representations of uncertainty. While economists and psychologists are well acquainted with probabilistic measures, an overview of other methods, especially the possibility measure was relevant as it pertains to vagueness and ambiguity which have important implications in this research. The next section will outline the research implications in relation to the literature reviewed and provide the research hypothesis which will be tested in the subsequent chapters.

3.5 Research Implications and Research Hypothesis

The previous sections have given an overview of the preference uncertainty literature based on the psychological perspective of the violation of choice theory axioms, the causes of preference uncertainty, the method of preference elicitation, the effect of preference uncertainty on valuations as well as the different forms of uncertainty representations. From the psychological insight on choice theory, it is evident that individual's preferences are neither always known or consistent or clear. Moreover, several causes affecting choice are outlined in psychology literature. This research aims to test descriptive invariance by examining the effect of different attribute representation techniques on respondent's choice.

While accounting for respondent's preference uncertainty is not a common practice in environmental SP exercise, several examples can nonetheless be found in the literature. Though the cause of preference uncertainty in CV literature has been more due to hypothetical bias, literature on CE has identified that choice set characteristics can affect the level of preference uncertainty. The choice set characteristics mostly examined in the CE literature have been attribute difference level across the alternatives as it plays an important factor on the level of task complexity. In case of environmental valuation and especially noise, different methods of attribute representation can have significant effect on respondent's understanding of the choice set and hence any empirical research to examine these effects in light of preference level can provide pivotal valuation and experimental design recommendations.

While examination of different preference elicitation techniques has been relatively well studied in CVM, this area of research has not received significant attention in CE while the preference data have commonly been recoded or eliminated. Moreover, the causes and effect of different error structure in the preference data analysis needs further implementation within CE. Thus, the literature on preference elicitation methods and its analyses show a severe need to compare different methods of preference elicitation as well as to examine the effectiveness of models based on flexible error assumption.

The concept of vagueness as one of the representations of uncertainty can be extended to the cause of preference uncertainty. While previous research conducted in this area has focussed on choice-set and respondent characteristics on preference uncertainty, distinction on level of uncertainty needs to be made based on the role of the stochastic error. Thus, it is vital to know whether uncertainty arises from stochastic effects or whether the respondent's stated level of uncertainty is a true reflection of the respondent's willingness to commit towards the choice of an alternative. Relating this concept with that of the choice set characteristics, it is expected that when the choice set is easier for respondents to understand, there will be lesser dependence on the stochastic effects to explain the level of preference uncertainty.

Based on the research implications arising from the literature review, the following research hypotheses relating to the concept of preference uncertainty will be tested in the thesis:

Hypothesis 1b: The type of choice set is expected to have an effect on the level of respondent's preference certainty with choice sets that are easier to understand will result in lower preference uncertainty than those that are more difficult

Hypothesis 2: Stated level of preference certainty can be due to deterministic as well as random effects. Thus, respondents' stated preference certainty could either be due to their true level of commitment or arising from respondent or choice-set characteristics which affects the random error

Hypothesis 3: Different methods of preference elicitation can have varying capabilities to capture preference information

3.6 Conclusions

The purpose of this chapter was to provide an overview of the concept of preference uncertainty in SP, the underlying psychological reasons for inconsistency in choice, causes of preference uncertainty commonly found in the literature as well as the method of preference elicitation and types of uncertainty representation along with the effect of different preference elicitation and attribute representation methods on valuation. It was seen that experimental psychologists have acknowledged the inconsistency in preferences as well as have pointed out several circumstances where the assumptions underlying choice theory have been violated. Taking this into consideration, studies examining preference uncertainty were critically reviewed based on the causes of preference uncertainty they have focussed on as well as on the methods of preference elicitation and its subsequent analysis. The literature reviewed indicated a significant lack in the examination of different attribute representation methods on preference uncertainty, a comparative analysis of the effects of different methods of preference elicitation within CE, as well as the application of different modelling techniques where the preference uncertainty information is explicitly included in the model. Based on this observed lack as well as the different forms of uncertainty representation that can be found in the literature, this chapter also outlined the research hypotheses which will be tested in the subsequent chapters.

4 ANALYTICAL METHODS

4.1 Introduction

The purpose of this chapter is to provide an analytical framework for the experimental design and the subsequent data analysis. To this end, this chapter serves as a theoretical bridge between the basis of the research and the methods to test the hypotheses. The chapter will provide an introduction to the Stated Preference (SP) methods which forms an important basis of the experimental design and the survey while giving an overview of the various analytical methods applied in the thesis.

Section 4.2 will give an overview of the SP method along with the common methods of experimental design while Section 4.3 will detail the various logit models relevant to the analysis in this research. This section will also give an overview of the relevance of the specific models to capture preference uncertainty. Section 4.4 will provide theoretical methods and application of the fuzzy logic technique.

4.2 The Stated Preference Method

The Stated Preference (SP) exercise is a method to elicit respondents' preferences for a given set of attributes under a hypothetical framework (Ortuzar, 2000). This exercise can take two distinct forms – the contingent valuation method (CVM) and attribute based methods. While CVM focuses primarily on a specific attribute to be valued and the elicitation method seeks to capture the respondents' maximum willingness to pay (or minimum willingness to accept) for an environmental benefit (or loss) through open-ended questions, dichotomous choice, bidding game or a payment card (Hanley *et al*., 1997; Hoevenagel, 1994; OECD, 1994), attribute based methods can take the form of ranking, rating as well as choice experiments (Brown, 2003). While the ranking method seeks to elicit respondents' preference ordering, the rating exercise allows the respondent to state their degree of preference level on a semantic scale (Ortuzar, 2000). In case of choice experiment, valuation of associated attributes is obtained by including a monetary cost or benefit in the choice task (Brown, 2003). The main method of elicitation employed in this research is through the choice experiment (CE) and henceforth SP and CE will be used to denote each other interchangeably.

The SP technique has its root in mathematical psychology and was first adopted by market researchers in the 1970s under the guise of ‘conjoint analysis’. Many applications of it were found especially to estimate market shares of new products. Since then, this technique was further developed to discrete choice theory by following the conceptual foundations of McFadden’s economic theory (Holmes and Adamowicz, 2003). SP techniques have since been widely used in transport studies as well as in environmental valuation where the principal aim has been to evaluate non-market goods.

The SP exercise comprises of several steps such as the selection of attributes to be valued, assignment of attribute levels, experimental design, construction of choice set and model estimation (Hanley *et al*., 2001). Several methods are available for assigning attribute levels and designing the choice experiment along with various analytical models for data analysis.

Following the selection of attributes which will form the part of the choice experiment, the number and values of the attribute levels forms a crucial aspect before the experimental design is formed. Choosing the attribute levels is a complex task requiring several considerations such as the form of familiarity with the respondents, the number of necessary levels as well as the range of the extreme values (Hensher *et al*., 2005). While much onus is laid on the analyst on the choice and consideration of these, Fowkes and Wardman (1988) developed the boundary value method to decide on the values of the attribute levels. The boundary value method suggests taking into account the relative valuation at which respondents are indifferent to the offered alternatives. In order to achieve a satisfactory design by considering taste heterogeneity and uncertainty, the authors suggested drawing a range of boundary values around the mean.

Following the selection of the attribute levels, the SP experimental design is generated based on the number of attributes and their levels as well as the types of effects considered, full or fractional factorial designs can be adopted. Several methods of coding are available along with different design forms to ensure optimality (Hensher *et al*., 2005). As discrete choice theory is based on random utility theory (RUM), different logit models are normally applied to analyse the data based on the type of theoretical requirements and the error assumption made.

In terms of preference elicitation, a standard approach in SP has been to elicit respondents' discrete choice from the alternatives presented in the choice set. While this probably forms the crux of the SP experiment, deeper insight on respondents' preference certainty or preference commitment cannot be obtained. However, an important question that needs to be addressed is the application of relevant analytical methods to model the data obtained from the preference elicitation exercise.

The next section details the analytical methods applied in the thesis along with their role in preference elicitation. As the methods applied in this research are not exhaustive, a passing reference will also be made to other models that could be used for future analysis. However, these models will not be scrutinised in detail here.

4.3 Logit Methods

This section outlines the theory underlying discrete choice analysis while detailing each model applied in the data analysis. While this section is not aimed to be exhaustive of all the available models to analyse the data, some of the methods that could be used in future analysis will also be outlined here.

As a prelude to the analytical methods outlined in this section and its significance for the specific research question, it becomes imperative to allude to the methods of preference elicitation applied in the SP experiment. The choice exercise comprised of three different methods of preference elicitation which included binary choice, a one stage five point Likert scale across the binary alternatives to capture respondents' preference certainty levels and a two stage Likert question where respondents were asked to indicate the preference certainty level following the choice elicitation. Several logit models have been applied to analyse the choice data, the theory of which is given subsequently following the basic theory underpinning discrete choice analysis.

4.3.1 Random Utility Theory

The fundamental basis of choice modelling lies in random utility theory (RUT) where individuals are assumed to follow utility maximisation and utility comprises of a deterministic component along with a random error. The error term in this case represents uncertainties caused from imperfect information, unobserved attitudes and taste variations (Vythoulkas and Koutsopoulos, 2003). Considering two alternatives *i* and *j*, the choice of alternative *i* over *j* is thus observed when:

$U_i > U_j$ For $\forall i \neq j$ (Bates, 1988; Meyer and Miller, 1984)

Where U_i is the utility for alternative *i* and U_j is the utility for alternative *j*.

Each utility in this case can be decomposed such that:

$$U_n = V_n + \varepsilon_n$$

where,
U_n = stochastic utility
V_n = deterministic utility
ε_n = random error (Ben Akiva and Lerman, 1985)

As utility maximisation theory based on deterministic choice does not consider the effect of uncertainty or stochasticity, it implies that an individual will make the same choice every time the alternatives are offered or all individuals with similar tastes and preferences and similar socio-economic variables, will make the same choices. However, this is not observed empirically thus indicating the need to consider the error terms in the estimation process. As the random error for an individual cannot be known, which alternative will yield the maximum utility and will therefore be chosen also cannot be estimated with certainty. Thus, using the distribution of the error term, the probability of an individual's choice is computed. This forms the probabilistic choice model under which choice is predicted based on the probability that an individual will choose a particular alternative (Koppelman and Bhat, 2006).

Two main assumptions of the random error term are generally made for most standard forms of coefficient estimation: 1) the variance of the error terms is identically distributed and therefore, homoscedastic and 2) the error terms are independent (Bates, 1988). This has been commonly known as the i.i.d. assumption i.e., error terms are independent and identically distributed (Train, 2003). Different forms of choice models are formed based on the assumption of the error structure.

Another important assumption pertaining to probabilistic choice models is the Independence from Irrelevant Alternative (IIA) axiom which states that the probability of choosing one alternative over another is independent from other alternatives in the choice set (Louviere *et al.*, 2000; Richards and Ben Akiva, 1975). Based on whether the IIA axiom is relaxed, several forms of choice models are formed.

Most choice models belong to the Generalised Extreme Value (GEV) family of models which assume Gumbel distributed error terms (extreme value type-1 distribution) with closed form models allowing for different levels of correlation between the error terms across different alternatives.

Under GEV, allowing the function G to depend on Y_j for all j where $Y_j = e^{V_j}$ such that $G = G(Y_1, ..., Y_j)$. Let the derivative $G_i = \partial G / \partial Y_i$ then the choice probability which can be given as:

$$P_i = Y_i G_i / G$$

is consistent with utility maximisation. Any model that can be derived in this manner is a GEV model. A GEV model has the following important characteristics:

1. G always has a positive value
2. G is homogenous such that if each Y_j is raised by α, G is also raised by α
3. $G \to \infty$ as $Y_j \to \infty \ \forall j$
4. The cross partial derivative of G change signs in a particular form such that $G_i \geq 0 \forall i$, $G_{ij} \leq 0 \forall i \neq j$, $G_{ijk} \geq 0$ and so on
5. If G has the parameters x and β such that $Y_i = e^{\beta_i x_i}$ and two choice sets A and B have parameter vectors such that $x_a = \left\{ x_{a_j} \forall a_j \in A \right\}$ and $x_b = \left\{ x_{b_j} \forall b_j \in B \right\}$ then

   if $x_a = x_b$ implies that

   $G = ((Y_{a_1}, ..., Y_{a_j}), x_a, \beta) = G((Y_{b_1}, ..., Y_{b_j}), x_b, \beta)$ (Hess, 2005; Train, 2003).

The simplest form of GEV model is the multinomial logit model where the correlation between the error terms across different alternatives is zero. Thus in this case, the GEV distribution is the product of independent extreme value distributions (Train, 2003).

Looking at the theoretical foundation for logit models, the next section details the workings of each of the relevant models in context of the research question.

4.3.2 Types of Choice Model

Data from the choice experiment can be analysed using a range of logit models based on the theoretical and empirical requirements. This section details the theory underpinning those choice models which have been specifically applied to address the research question. The section will begin by the most basic choice model, the multinomial logit model and proceed successively based on the level of complexity.

4.3.2.1 Multinomial Logit

The multinomial logit model (MNL) is the most basic form of logit model which states that in any particular choice situation, the probability that an alternative *i* will be chosen from the set of alternatives is given by the following equation:

$$P(i : A_n) = e^{V_{in}} / \sum_{j \in A_n} e^{V_{jn}}$$

where,
V_{in} is the deterministic utility of alternative *i* in choice situation *n*
V_{jn} is the deterministic utility of alternative *j* in choice situation *n*
The value of the deterministic utility is assumed to depend on the values of the variables which affect choice and the functional relationship is generally assumed to be linear in parameters (Louviere *et al*., 2000).

The binary logit (BL) model is the special case of the MNL where there are only two alternatives available to the respondent. The binary logit probability therefore takes the following form:

$$P(i) = e^{\mu V_{in}} \big/ e^{\mu V_{in}} + e^{\mu V_{jn}}$$ (Ben-Akiva and Lerman, 1985)

The MNL model rests on two main assumptions –

1) The variance of the error terms is independent and identically distributed (i.i.d.) from each other and,
2) The probability of choosing an alternative is unaffected by the inclusion of another alternative in the choice set (the independence of irrelevant alternatives property - IIA)

Thus both the BL and MNL model forms do not take any form of heterogeneity or error correlation into account. In the simplest type of analysis, the binary logit method can be used to analyse the binary choice data while the data from the one and two stage Likert scales can be analysed using the MNL model.

In case of the MNL application for the Likert scale data, differences in the preference certainty level of the alternatives can be modelled by adding an alternative specific constant. An alternative specific constant (ASC) represents the mean distribution of the unobserved effects of the error components on the alternative (Louviere *et. al*, 2000) and for J alternatives, only J-1 ASCs can be defined (Train, 2003). Thus, a constant added to the option say, 'Definitely A' can be interpreted as the average effect of the unincluded factors on utility of 'Definitely A' relative to 'Probably A'.

While the value of the ASC and its associated t-ratio can be used to judge the average effect of the unobserved factors in the utility model, a MNL model does not consider the effect of the difference in the error variance structure between the relevant alternatives. Hence, though ASC can be applied to estimate the relative effect of unobserved factors, further examination of error structure is important through application of higher model forms as the MNL analysis does not provide any significant insight on the type of i.i.d. or IIA violation affecting the preference levels.

4.3.2.2 The Nested Logit

The first step to relax the IIA assumption can be carried through the nested logit (NL) model which considers the possibility of nesting alternatives with some degree of similarity between them. The rationale of the nested logit model thus lies in the fact that the commonality of characteristics between alternatives results in some parts of the random error to be correlated. Nested logit model is formed by partioning the choice set such that the alternatives that share common unobserved components are grouped together resulting in the i.i.d. assumption to be relaxed as the error component are correlated within a nest but not across the nests. Thus, NL models are applied when there are shared unobserved stochastic components associated with different choice dimensions (Ben-Akiva and Lerman, 1985). By grouping the alternatives into subsets that allow for varying error variance, the homoskedastic assumption of the MNL model is relaxed. However, though the variance of the error differs across the alternatives, the overall variance of the error of all the alternatives remains constant (Shen, 2005). Thus, homoskedasticity is maintained with the NL model.

While the nested logit structure can imply a hierarchy of choice, the derivation of the model makes no assumption of the structure of the choice process (Koppelman and Bhat, 2006). Thus, the nesting structure is reflective of the assumptions made on the correlation between alternatives and not on the sequence adopted during the decision-making process.

The assumptions of the nested logit model can be summarised as – 1) the IIA condition holds within a nest and 2) the IIA condition does not hold in general for alternatives in different nests (Train, 2003).

In case of the NL model, the joint probability of an alternative in a nest can be given as the product of the probability of choosing a nest and the conditional probability of choosing the alternative given a particular nest is chosen.

Assuming for example that a five point Likert scale (Definitely A, Probably A, Uncertain, Probably B and Definitely B) can be nested in the following way (and noting that other nesting structures are possible):

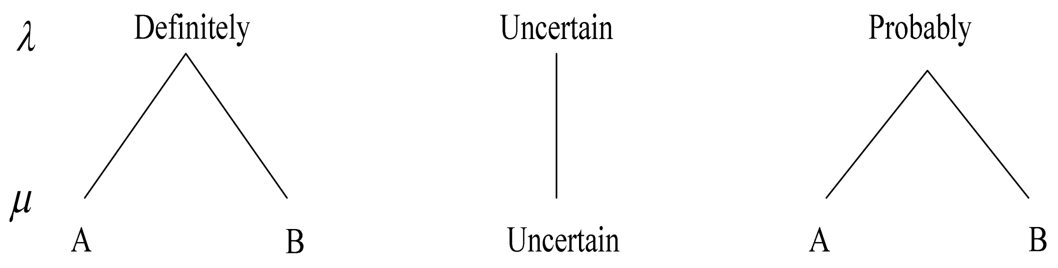


Figure 4.1Nesting Structure for the preference levels

The composite utility of selecting Definitely A can be given as:

$$U_{D,A} = U_D + U_{A|D}$$
$$U_{D,A} = V_{D,A} + \varepsilon_{D,A}$$
$$U_{D,A} = V_D + V_{A|D} + \varepsilon_D + \varepsilon_{A|D}$$

While the joint, marginal and conditional probabilities can be respectively given as:

$$P_{D,A} = P_D . P_{A|D}$$
$$P_D = \frac{e^{\lambda V_D}}{e^{\lambda V_D} + e^{\lambda U} + e^{\lambda V_P}}$$
$$P_{A|D} = \frac{e^{\mu V_{A|D}}}{e^{\mu V_{A|D}} + e^{\mu V_{B|D}}}$$

Where,

$$\lambda V_D = \lambda X_D + \phi_D \ln \sum e^{\mu V_{k|D}}$$ (Ortuzar and Willumsen, 2004)

ϕ_D is known as the logsum parameter or the expected marginal utility (EMU) or inclusive value (IV) and is equal to λ / μ_D (Ortuzar and Willumsen, 2004). The degree of independence or dissimilarity in the random error across the alternatives in a nest is given by this log-sum coefficient (also called the nesting or the structural parameter) which can differ over nests. When this parameter equals unity, the nested logit model reduces to the MNL. For the model to be consistent with utility maximising behaviour, the nesting parameter has to range from 0 to 1. When this is greater than 1, the model is consistent with utility maximising (UM) behaviour for some of the explanatory variables but not all while in the case of negative values, the model is inconsistent with UM behaviour (Train, 2003).

As data based on preference certainty level share all generic coefficients, the distinction between non-normalised nested logit (NNNL) model and utility maximising nested logit (UMNL) needs to be taken into consideration depending on the software used for analysis.

Based on the scaling of the deterministic component of the utility, the two different types of NL model can be specified as follows:

a) The Utility Maximising Nested Logit (UMNL) model is based on McFadden's Generalised Extreme Value (GEV) where marginal and conditional probabilities are multiplied by λ and μ respectively and the IV parameter is multiplied by λ
b) The Non-Normalised Nested Logit (NNNL) is the basis of the software ALOGIT and has an implicit nest specific scaling within the specification

Based on which level the specification is normalised, two forms of UMNL can be derived.

a) Normalisation on the lower level $\mu_m = \mu_n = 1$ leads the model to become NNNL (RU1)
b) Normalisation on the upper level $\lambda_m = \lambda_n = 1$ is UMNL or RU2

If a constant is added to each alternative's deterministic utility, NNNL/RU1 does not give results consistent with theory but only RU2 satisfies the condition. Consistency of RU1 with RUT can be achieved by imposing the upper level scale parameter to be held constant whereas consistency in NNNL can be achieved by holding the lower level scale parameter to be constant across the nests (Silberhorn *et al.*, 2007).

When there are no generic coefficients in the nest, both the UMNL and NNNL specifications are equivalent but the coefficient estimated with NNNL have to be rescaled with the accordingly estimated IV parameter to be equivalent to that obtained from UMNL. When generic coefficients enter the model, NNNL software cannot be used to estimate the parameters and only the RU2 normalisation of the UMNL should be carried out to be consistent with RUT.

In the UMNL specification, V_D and $V_{A|D}$ are scaled explicitly with parameters λ_m and μ_m respectively while this is automatically and implicitly done in the NNNL software. For the parameters to be consistent with RUT using NNNL, certain restrictions have to be imposed. Firstly, the coefficients in each nest have to be scaled equally and the IV parameters have to be constrained to be equal for all nests. Rescaling the parameter estimates in the restricted NNNL model with the estimated IV parameters results in the parameter estimates of RU2 (Silberhorn *et al.*, 2007).

Another possibility to guarantee consistency with RUT without imposing restriction on the IV parameters can be achieved by introducing additional dummy nests below the lowest level and the additionally estimated scale parameter has to be defined in such a way that the products of all ratios of scale parameters between levels must be identical from the root to all elemental alternatives (Koppelman and Wen, 1998).

Using the UMNL or the modified NNNL model (such that it is consistent with RUT) to analyse the preference certainty data, the NL model allows the nesting of the different alternatives based on the preference level. Thus, alternatives with similar error structure based on similar preference certainty level (such as 'Definitely A' and 'Definitely B' or 'Probably A' and 'Probably B') can be grouped together. The nest parameter estimated in this case allows examining any patterns of error correlation between nested alternatives as well as estimating the level of error variance across the different nests. While the NL model sheds some light on the correlation of the error variance, in its pure form, it does not examine the possibility of heteroskedasticity. To examine whether a further relaxation of the i.i.d. assumption is required, the mixed logit model needs to be applied in the analysis.

4.3.2.3 <u>The Mixed Logit Model</u>

The mixed logit (ML) model has been developed to account for the richness in the unobserved component by considering the potential for correlation and heteroskedasticity. Based on the cause of the flexible error assumption, two forms of the ML model can be derived – one allowing the parameter estimates to follow a non-Gumbel distribution and another allowing for heteroskedasticity of the error variance.

In the first case, the probability equation of the ML model is a weighted average of the logit probability evaluated at different values of the parameters, with the weights given by the density which is a function of the parameter. The function of the parameters can be specified to have any distribution apart from Gumbel (i.e., normal, log-normal etc.). Under this derivation of the ML model, the values of the parameter can be interpreted as representing different tastes of the individual decision makers. As the researcher cannot know each individual's parameter estimates, the unconditional choice probability is an integral of all possible values of the parameter (Train, 2003).

Considering β to be the parameter vector which can be interpreted as representing the taste of an individual decision maker, the utility of a person n from alternative j can thus be specified as:

$$U_{nj} = \beta_n^{'} x_{nj} + \varepsilon_{nj}$$

The coefficients β_n is allowed to vary across individuals and the choice probability is therefore conditional on this parameter. Though the decision maker is assumed to be aware of his own β_n and ε_n and choose based on utility maximisation, the researcher is unaware of β_n and therefore cannot form a probability function which is conditional on β. The unconditional choice probability is therefore given as an integral of the conditional probability over all possible variables of the parameter. It can therefore be given as:

$$P_{ni} = \int \left(\frac{e^{\beta^{'} x_{ni}}}{\sum_j e^{\beta^{'} x_{nj}}} \right) f(\beta) d\beta \text{ (Train, 2003)}$$

Where, $f(\beta)$ is the density function

The distribution of the parameter is specified by the researcher and the parameters then estimated. The random parameters model is most suitable to capture inter-respondent taste heterogeneity.

The second type of the ML model is built by allowing for heteroskedasticity and/or correlation in the error components across the utilities of the different alternatives. Using this Error Components Logit (ECL) model, the stochastic error portion of the utility can be decomposed into two parts – one which is correlated over alternatives depending on the specification and another which is i.i.d. Gumbel distributed. For the standard logit model, the former error component is identically zero, implying no correlation in utility over alternatives (Train, 2003).

The decomposition of the error component can be shown as follows:

$$U_{ni} = V_{ni} + (\eta_{ni} + \varepsilon_{ni})$$

Where η_{ni} is the random component with zero mean and distribution dependent on the underlying parameters and ε_{ni} is the random component with zero mean that is independent and identically distributed across alternatives for an individual *n*.

For any given value of η_{ni}, the conditional choice probability is logit as the ε_{ni} error component is assumed to be i.i.d. and is given by:

$$L_{ni}(\eta) = \exp(\beta' x_{ni} + \eta_{ni}) \Big/ \sum_j \exp(\beta' x_{nj} + \eta_{nj})$$

Since η_{ni} is not given, the unconditional probability is the integration of all values of η weighted by its density.

$$P_{ni} = \int L_{ni}(\eta) f(\eta | \Omega) d\eta$$ (Hensher and Greene, 2001)

Where Ω are the fixed parameters of the distribution.

Through the flexibility provided by the ECL model, various forms of correlation and heteroskedasticity can be examined from the preference data. This model is especially interesting due to the larger range of error effect that can be examined. An analogue of nested logit can be obtained by specifying a dummy variable for each nest that equals unity for each alternative in that nest and zero for each alternative outside the nest. The heteroskedastic random error of the ECL model can be entered as the error associated with each alternative in the nest thus inducing correlation among alternatives. By not entering for any of the alternatives in the other nests, no correlation is induced between alternatives across nests while it is done so for those within a nest. The variance captures the magnitude of correlation and it plays an analogous role to the inclusive value coefficient of the NL models (Train, 2003). However, while the NL model maintains homoskedasticity of the error variances, this is not the case with the ECL model which mimics correlation structure developed with the NL model (Munizaga and Alvarez-Daziano, 2001).

An analogue to the heteroskedastic extreme value (HEV) model can be obtained by specifying the error variance to have a different variance for each of the preference levels.

The HEV model is a relatively restrictive model as it relaxes the identically distributed error structure assumption without relaxing the independence assumption. The fundamental underpinning of this model lies in the heteroskedasticity of the error terms thus relaxing the IIA assumption. The alternative error terms are assumed to be type 1 Gumbel distributed with the variances allowed to be different across the alternatives (Bhat, 2000). This is achieved within the HEV framework by allowing different scale parameter for the error term across alternatives.

Let $z = \varepsilon_i / \lambda_i$

where, ε_i is the error and λ_i is the scale parameter.

The choice probability of the HEV model can then be given as:

$$P_i = \int_{z=-\infty}^{z=\infty} \prod_{j \in C, j \neq i} F\left[\frac{V_i - V_j + \lambda_i z}{\lambda_j}\right] f(z)dz$$ (Bhat, 2000)

By allowing for different scale parameters, the HEV model avoids the problem of the i.i.d. property. Intuitively, the random term can represent the unobserved characteristics of the alternatives and hence the scale parameter in the case of the HEV model represents the uncertainty related with expected utility of the alternative (Hensher, 1999).

4.3.2.4 <u>Ordered Logit</u>

The ordered logit (OL) model has been used to analyse the strength of respondents' preferences under the rating exercise. With this method, the respondent's preference certainty level is interpreted as an ordinal indicator of the relative size of the utility difference. Considering A and B to be two alternatives, the difference in the utilities across the alternatives can be given as:

$dU = U(X^A, Y^A) - U(X^B, Y^B) = dv - \varepsilon$ (Swallow *et al.*, 2001)

Where, X^A, X^B, Y^A and Y^B are the vector of attributes associated with the corresponding alternatives and ε is the difference in the stochastic error from the two alternatives. The ordered response implies the existence of unobserved threshold parameters θ_j that demarcate the intervals or categories within which the utility difference may fall. The ordinal indicator defines a set of dummy variables D_j that records to which category an observation belongs. Therefore,

$D_j = 1$	if $\theta_{j-1} < dU < \theta_j$
$= 0$	otherwise; where j = 1, 2… J

The ordered response implies that the respondent's strength of preference or preference certainty level will be given by the following probability function:

$$\Pr(\theta_{j-1} < dU \le \theta_j) = \Pr(dU \le \theta_j) - \Pr(dU \le \theta_{j-1}) \forall j$$

$$\Pr(D_j = 1) = F(\theta_j - dV) - F(\theta_{j-1} - dV)$$

By normalising θ_0 to $-\infty$ and θ_j to ∞, the following log likelihood can be obtained:

$$LL = \prod_{n=1}^{N} \prod_{j=1}^{J} [F(\theta_j - dV_n) - F(\theta_{j-1} - dV_n)]^{D_{jn}}$$ (Whelan and Tapley, 2006)

This log likelihood has been used for estimation to measure the strength of preference where respondents are asked to indicate how strongly they prefer the choice they have already made. Thus the respondent establishes his choice first and then rates his strength of preference (Swallow *et al.*, 2001).

For each of the model, along with the parameter coefficient, a threshold coefficient is estimated. The threshold coefficient is along the scale of the utility difference and the probability of an alternative being chosen is the probability that the utility difference is within a particular range of the associated threshold coefficient (Train, 2003). Thus in case of the five point Likert choice: Definitely A, Probably A, Uncertain, Probably B and Definitely B, the distribution of the utility difference can be given in the following manner:

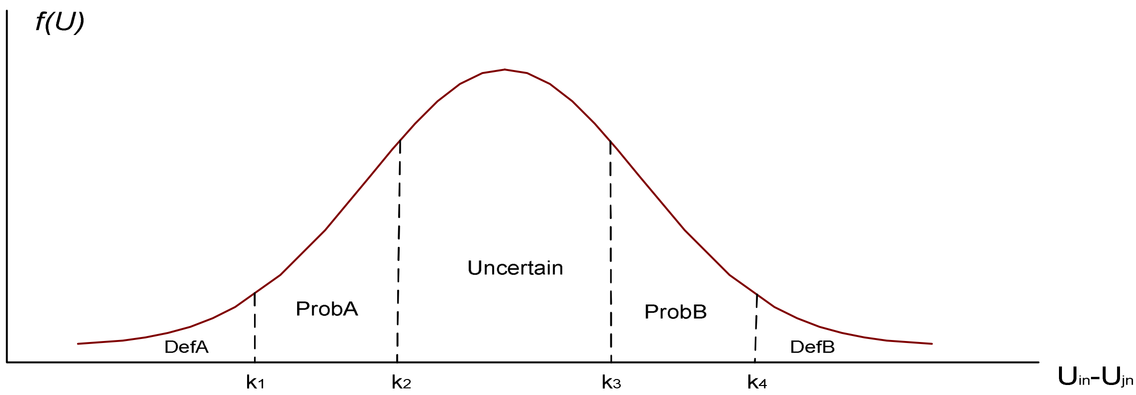


Figure 4.2 Preference Levels with Ordered Logit Analysis

The probability that 'Definitely A' is chosen in this case is given by the probability that the utility difference is less than threshold value k_1.

As the semantics of the levels of preference certainty can be different for different people, models have been developed with variable threshold coefficients using a mixed ordered logit model (Whelan and Tapley, 2006).

4.3.2.5 Other Model forms

In addition to the models outlined in the preceding sections, other model forms such as cross nested logit (CNL), ordered generalised extreme value (OGEV) and latent class models can also be used to analyse the preference uncertainty data. While this section briefly outlines the methods associated with each of these models, they are not considered in depth due to the absence of their application.

The Cross Nested logit (CNL) is an extension of the Nested Logit model allowing for ambiguous allocation of alternatives to nests. The CNL model allows nests to share some alternatives having degrees of similarity between them. Along with the nesting parameter, an additional allocation parameter is estimated in the model which reveals the degree to which an alternative belongs to a nest. An important constraint is that the sum of all allocation parameters for an alternative across all nests must be equal to unity. Thus the value of the allocation parameter lies between 0-1 (Train, 2003). Under certain specifications of the allocation parameter (which show the degree of membership of an alternative in a nest) and the nesting parameter (which indicate the degree of independence in the random error between alternatives within a nest), the CNL model reduces to NL and MNL (Hess *et al.*, 2004).

Ordered Generalised Extreme Value (OGEV) derives from the Generalised Extreme Value class of RUM models. The OGEV model differs from ordered logit in the way that the alternatives ordered in close proximity have higher correlation (Batley *et al.*, 2001). The further the outcomes are from each other, the smaller is the correlation between the associated errors. Comparing the application of the NL model on the five point Likert scale data (where alternatives with similar preference certainty level such as 'Definitely A' and 'Definitely B' are nested together), the OGEV model would give a contrary structure where 'Definitely A' and 'Definitely B' being the two extreme ends of the Likert scale, would have the lowest correlation amongst them.

While the CNL and OGEV have the ability to explicitly deal with the preference uncertainty data, the latent class (LC) model considers unobserved heterogeneity that affects individuals' choice by developing various classes. Under this model, the probability of a particular choice can be given by the conditional probability of the choice, given a particular class (Greene and Hensher, 2002). This method allows extracting any hidden characteristics that explains individual's choice of a particular preference level which can be particularly helpful when segmentation analysis is a cumbersome process.

While this and the preceding sections have detailed the techniques of stated preference method and various logit analysis, the aim of the following section is to provide a theoretical framework for the fuzzy logic technique while also providing examples of its application in choice experiments.

4.4 Fuzzy Logic

This section details the theory and application of the fuzzy logic (FL) technique in relation to its application to choice experiments. The FL theory and the fuzzy inference system will be given in Section 4.4.1while Section 4.4.2 will provide the applications of FL in CE.

4.4.1 Fuzzy Logic Theory

Though choice theory based on utility maximisation is commonly applied in understanding individual preferences, insights from other approaches to decision theory state that human decision-making process is more dependent on few normative rules and hence heuristics play an important role in the decision process. Moreover, the assumption of rationality in choice theory implies that decisions are based on a *consistent line of reasoning* which could have their basis on some normative rules (Klein and Methlie, 1990). The fuzzy logic (FL) theory is one technique which allows for heuristics in decision-making.

The FL theory is based on the fuzzy set theory. Different from the classical set theory where an element can either belong to a set or not, the fuzzy set theory allows an element to belong to multiple sets with varying levels of membership. Hence, while classical logic, based on reasoning with precise propositions, is a two-valued logic where the truth value of a statement can either be TRUE or FALSE, the truth value of a statement under fuzzy logic, which applies approximate reasoning to imprecise propositions, can assume varying values (Chen and Pham, 2001).

Let us consider a set 'OLD' and an element x = '66 years', the difference between classical and fuzzy sets can thus be given in the following manner:

Assuming the following rule pertaining to age classification:

IF $x \geq 70$ THEN OLD

The membership of x = '66 years' under the classical and fuzzy sets can be graphically represented as follows:

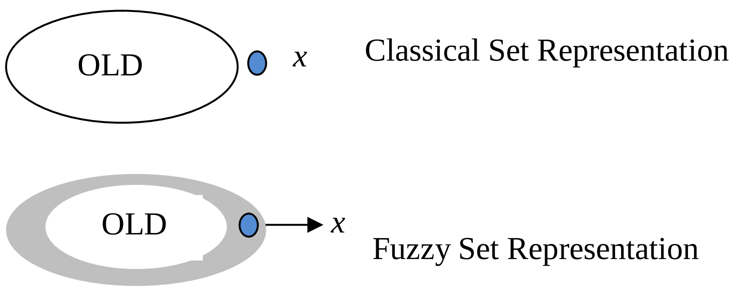


Figure 4.3 Classical and Fuzzy set representations

The above figure reveals that under the assumed rule, the element *x* does not belong to the set OLD with the classical set representation; however, in case of fuzzy set, the element belongs to the set with some level of membership.

In relation to the theory of fuzzy sets and the example given, it is evident that 'membership function' plays a crucial role in the FL theory.

Membership of an element is given by the degree to which the element belongs to a set while the function mapping the degrees of potential membership is termed as the membership function (Kasabov, 1996). Recalling the example of age, the membership function can be mathematically denoted as follows under both the set theories:

Let μ_A denote the level of membership for the element *x* in set A. Under classical set theory,

$$\mu_A(x) = 1 \text{ if } x \geq 70$$

$$\mu_A(x) = 0 \text{ if } x < 70$$

Under fuzzy set theory, the membership of the element can be given as:

$$\mu_A(x) : U \to [0,1]$$

This equation implies that set A is a fuzzy subset of the universal set U and the membership level of the element *x* in set A can take the value from 0 to 1, where 0 means that the element completely does not belong to the set while 1 implies the contrary (Kasabov, 1996).

Generating subjective membership functions for different age categories, the membership of $x = 66$ in set 'OLD' can be given as follows:

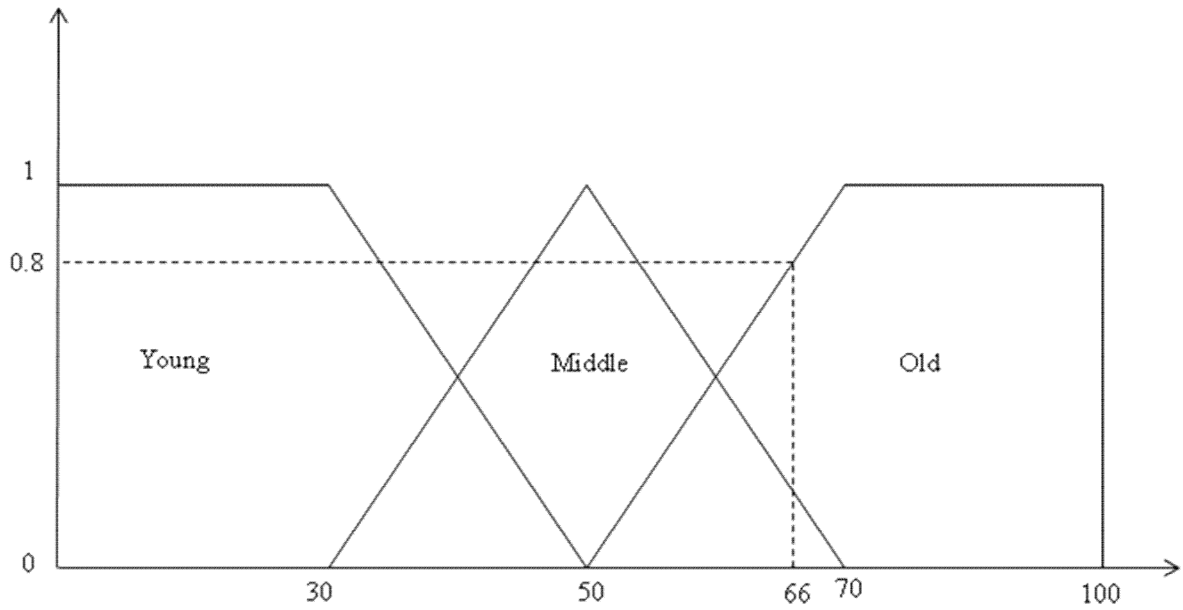


Figure 4.4 Graphical Representation of Membership functions

In this case, from the above figure, it can be seen that *x = 66* belongs to the set 'OLD' with 80% of membership. The concept of membership and membership functions thus forms a crucial component in the FL theory. Though FL is an expert decision system where many system components such as membership function and rules are specified by an expert, several methods for specifying the membership function as well as fuzzy rules can be identified. The methods of specifying these can take any of the following forms:

1. Eliciting an expert's opinion: this can either take the form of enquiring from an expert of the knowledge domain or where public surveys are concerned, eliciting required information from the respondent
2. Understanding the physical environment: the system can be designed based on the designer's understanding of the physical meaning surrounding the technique used. In this case, the designer serves as an expert of the system
3. Employing machine learning: through the use of neural networks, genetic algorithms or other methods of machine learning, rules and membership functions can be derived from the data (Kasabov, 1996)

Besides the membership functions, as necessary in the case of classical logic, rules form an integral part of the FL system. The rules in the FL system are termed as fuzzy rules as they pertain to the vagueness associated with the variables, which normally assume a linguistic form. As in classical logic, the rules follow an IF-THEN structure with the only striking characteristic that the associated variables are fuzzy representations. While the effect of operators associated with the rules base, such as AND, OR, NOT are similar to that of Boolean logic, FL rules need to be consistent (such that rules are not contradictory), complete (all possible rules are formed) and concise (there is absence of redundancy) (Chen and Pham, 2001).

Besides fuzzy rules, the FL system comprise of two other main components: the input variables and the output variable. These three components thus form the fuzzy inference system (FIS) which maps the input and output variables. To illustrate, the FIS for four inputs (view, noise, sunlight and charge), can be represented as follows:

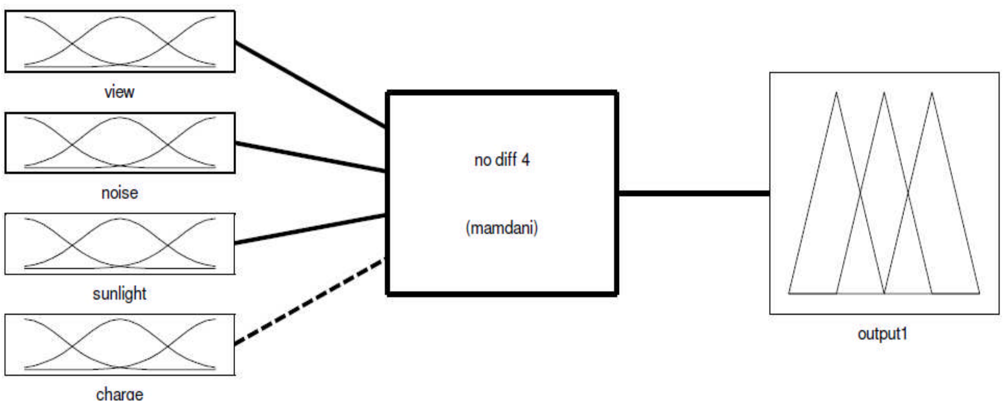


Figure 4.5 The fuzzy inference system

In this case, the four input variables as well as the output variable (choice) are linked through the fuzzy rules. In order to map the relation between the numeric input and output in FIS, several steps are conducted. As an important characteristic of FL is its ability to deal with vague and imprecise data, linguistic characterisation forms an important aspect of the FIS which is carried out through the linguistic membership function associated with each of the variables. The conversion of numeric input values to a linguistic category is termed ‘fuzzification’. The fuzzification process detects the membership function associated with each input variable and the degree to which the input variable matches the conditional element in the rule is computed. For each of the rule incorporated in the FIS, the associated values in the output membership function is computed using *minimum* (which truncates the membership function) or *product* (which scales the membership function) method. This process is called implication. Combining the effect of all rules on the output membership function forms the aggregation process while defuzzification computes a single numeric output from the aggregate membership function (Kasabov, 1996; MATLAB Handbook, R2007a).

Based on the form of the output variable and the type of fuzzy rule, two main types of fuzzy systems exist within the Fuzzy Logic Toolbox in MATLAB. These are the Zadeh-Mamdani and the Takagi-Sugeno fuzzy rules. The Zadeh-Mamdani fuzzy rule takes the following form:

IF x is A THEN y is B

Where *x* is the input variable, *y* is the output variable while A and B are fuzzy sets.

In case of the Takagi-Sugeno system, the output variable is a mathematical function and thus the rule takes the following form:

IF x is A and y is B THEN z is f (x, y)

Where *x* and *y* are the input variables, z is the output variable, while A and B are fuzzy sets (Kasabov, 1996).

It can be thus noted that in case of the Mamdani method, the output variable consists of membership functions which allows for linguistic representation in the rule base. In case of the Takagi-Sugeno method however, the output variable is a numeric computation.

The overview of the FL theory and the method of FIS have revealed that this method can be especially important in analysing linguistic data and heuristic based system. The output variable in the FIS can be specified over a range of 0-1 such that it reflects the choice associated with each rule. The specific applicability of this analytical method in relation to the research focus of this thesis stems from the examination and performance of the FL system to analyse each of the different representation data. Based on the characteristics of the FL theory and system, it can be hypothesised that data based on linguistic representation will be better analysed using this method. Moreover, the requirements for the FIS as well as the number of rules required for the system can vary based on the method of attribute input and representation method used during the survey. As a heuristic based system, this analytical method can also provide some insight on the decision-making process of individuals by examining the rules required in the system for optimum output.

While this method of analysis is commonly applied in physical systems where accurate observations can be made and more general rules developed, increasing applications are found in social surveys. The following section provides an overview of the FL application related to choice experiment as well as noise annoyance.

4.4.2 Applications of Fuzzy Logic

The application of FL in choice experiments is still relatively uncommon though increasing examples can be found. One of the early applications of FL in CE can be found in Hoogendoorn-Lanser and Hoogendoorn (2000) who applied the FL technique to analyse travel choice behaviour where the travellers' perception and appraisal of the trip attributes were considered to be vague. In this case, in-vehicle time, access, egress and waiting times (at first stop and at transfers) formed the main attributes in the choice set while fuzzy utility of the alternative formed the output variable in the FL system. The premise of this paper was that people's perception of time is subjective and hence FL theory was applied to model the impact of travel time. In order to obtain a rule base that optimises the percentage correct predicted, the authors applied Genetic Algorithm (GA) where the subset of optimal rules were selected from the total possible rules. The authors found that compared to the logit model, the FL model with GA yielded significantly better model with fewer set of decision rules required to correctly approximate travellers' choices. Thus in this case, the FL method was found to be a good descriptive tool for human behaviour.

Another application of FL in mode choice analysis was conducted by Mizutani and Akiyama (2001). In this case, the authors developed a multinomial logit model with fuzzy utility function for different mode choices and several explanatory variables. Fuzzy utility which formed the output was classified into 'positive' and 'negative' values while the effect of each explanatory variable on utility was incorporated into the system through fuzzy rules. Using trial and error as well as GA, the authors estimated the values of the linguistic variables as well as the parameters of the membership function that comprised the FIS. While this approach provided a unique perspective to model travel choice, the authors only related the effect of each input variable on the fuzzy utility. Thus, no rules were developed which incorporated the combined effect of various inputs on respondents' choice.

In yet another application of the FL method to mode choice analysis, Cantarella and Fedele (2003) developed a fuzzy utility model (FUM) with several attributes where fuzzy utility can be depicted as a choice fraction vector which is a function of the core vector. The core vector in this case reflects the difference between the deterministic utilities of the alternatives. By considering the width of the fuzzy density function, the parameters associated with the core vector were estimated by trying to reproduce the observed choices. The distance between the observed and modelled values provided an estimate on the relative effectiveness of the set of parameters with the minimum of the values leading to the final parameter estimation. While this appears as a plausible method of parameter estimation, the authors do not explicitly specify the method. Moreover, with the trial-and-error procedure, this process can be cumbersome.

Using a neuro-fuzzy method to estimate rule weights and parameters for the fuzzy membership function, Vythoulkas and Koutsopoulos (2003) applied the FL theory in modelling discrete choice from mode choice. Travel time, travel cost and access/egress time formed the main attributes in the choice experiment which were incorporated in the FIS using the difference method. The output of the FIS comprised of five membership functions reflecting the preference level for each of the alternatives (train and car). The rule base in this study was based on each alternative. By examining the effect of different rules base on the percentage correct predicted, the authors sought to emphasise the importance of optimal weights on correct prediction. It was observed that with the deterministic choice rules, the correct prediction increased from 72.3% in case of equal weights to 77% when optimal weights were used. While neuro-fuzzy method had been shown to model respondents' decision-making fairly well, the authors also acknowledged that defining the criteria for calibration, development of rules base and the calibration of parameters posed some difficulty with the implementation of this method.

Though the above examples have revealed increasing applications of FL in choice experiment, it can also be observed that the method of application as well as the output obtained from the FL system has varied with different studies. While some authors have adopted the GA and neuro-fuzzy methods to estimate parameter values as well as to calibrate the membership function and obtain optimal rule weights, little attention has been paid on the type of rules needed to develop the system. Moreover, the applicability of this method has not been examined across different types of input data. This examination along with the number and types of rules required in the FL system for different types of input variables will therefore form an important component of the FL analysis in this research.

While the previous examples of FL application have focussed exclusively on choice experiment and mode choice analysis, some research has also been conducted to apply this method to model noise annoyance.

Botteldooren *et al*. (2002) conducted an exercise to calculate noise annoyance related to several factors where the respondents were asked to provide a numeric rating from a scale of 0-10 for each linguistic noise annoyance category. Based on the exposure to noise which included the direction of the living room and bedroom window as well as the distance to the source, level of masking, sensitivity to noise along with other socio-demographic variables, rules were formed to estimate the effect on annoyance level. The authors used the *weighted percentage of the correctly predicted noise annoyance response as a measure of the model performance*. The authors found that the fuzzy rule base was better at explaining the extreme levels of annoyance than the moderate levels. Considering the same set of explanatory variables, the authors found that the fuzzy rule base system was better at predicting the noise annoyance response than the linear regression model. However, the authors noted that the theoretical and empirical basis for the expert varies depending on the cause of annoyance.

4.5 Conclusions and Implications

The main aim of this chapter has been to provide the necessary framework and background for the empirical analysis. The methods involved in the SP survey as well as the various forms of logit analysis pertaining to the specific application context were detailed in this chapter. From the examination of the available logit models for analysis, it can be concluded that different models provide insight on the preference certainty data through different key parameters. In case of the MNL model, the ASCs are expected to elicit the stochastic effects associated with each preference level while also explaining the relative preference of that level in relation to the eliminated alternative. With the NL analysis, the nest parameters capture the level of similarity between the nested alternatives, with increased need for NL analysis and greater correlation among the nested alternatives indicating greater need for flexible error assumption. In case of the ECL model, the error values obtained from the model would reveal the stochastic effects on each preference level while with the OL method, the relative probability associated with each preference level and the implicit ordering of the preference level can be examined through the threshold values.

The theory and methods of FL showed that this method is well-suited where linguistic representation and heuristic rules apply. The literature on FL applications in SP has shown that this method has been successfully applied to explain mode choice analysis while parameter estimates have also been obtained in conjunction with the GA method. It is also observed that while different types of output are obtained based on the specification of the FL system, no study has been conducted to examine the types of rules required for different types of input variables. This specific examination thus forms an interesting research area especially as it offers an opportunity to examine the effect of different types of input variable on respondents' decision making process.

5 DATA COLLECTION AND DESCRIPTION

5.1 Introduction

To test the effect of different attribute representation methods on respondent's certainty of choice, a range of SP exercises were conducted with different methods of attribute representation and choice elicitation. This chapter briefly describes the background of the survey site, details the method employed for experimental design and data collection and describes the sample socio-economic characteristics and descriptive statistics.

Section 5.2 will outline the background and rationale for conducting the survey on the particular site, along with a brief overview of the previous study carried out in the area. Section 5.3 will detail the experimental design and simulation process for both phases of the survey while Section 5.4 will provide the details of the data collection method and the survey procedure. The descriptive statistics results obtained from the data will then be given in Section 5.5 while Section 5.6 gives the chapter summary.

5.2 Survey Background and Previous Study

In order to examine the effects of attribute representation and preference elicitation techniques on choice certainty in context of traffic noise pollution, a Stated Preference (SP) survey was conducted in the residential area of Telheiras, Lisbon with view, noise, sunlight and housing service charge as attributes in the SP exercise. The survey site selected for conducting the SP exercise formed a very interesting study as the residential area is surrounded by three major traffic roads (Avenue Norton de Matos, Eixo Norte Sul and Avenue Padre Cruz) which form an important cause of traffic noise in the area.

A previous computer aided personal interview (CAPI) was carried out in the same area and with the same attributes in 1999 by Arsenio (2002) with a binary choice SP exercise. That survey aimed to evaluate environmental externalities from traffic, particularly in the context of residential noise. The same site was selected to conduct the survey for this research due to its location, the possibility to compare results obtained with the previous study and the relatively low cost of conducting a face-to-face interview compared to the U.K. The previous exposure of respondents to a similar study also implied less dependence on the need for a detailed pilot study.

The survey area along with the three main traffic roads is shown in the following figure:

Figure 5.1 Map of the survey area surrounded by three main traffic roads - Av.Gen. Norton de Matos, Eixo Norte-Sul and Av.Padre Cruz

The levels of the attributes view, noise and sunlight in the survey conducted by Arsenio (2002) was represented in the SP exercise based on the location of the apartment in the block, with the base level being the current apartment location. Thus, a singular method of attribute representation and preference elicitation was employed in the choice exercise by Arsenio (2002).

Though the attributes, general questionnaire design and survey site choice of the current study followed from the Arsenio (2002) SP study conducted in the same area, several variations in the experimental design and survey implementation were made in order to meet the research objectives. The next section details the experimental design and simulation carried out for both the phases of the survey.

5.3 Experimental Design and Simulation

This section details the experimental design and simulation carried out in the survey. To examine the effect of attribute representation and preference elicitation on attribute understanding and choice certainty, different experiments were conducted in the current study based on varied techniques of attribute representation and preference elicitation. The SP experiments varied based on the method of attribute representation used as different attribute levels were employed for each of the phase. The attributes considered for the choice experiment were: view, noise, sunlight and housing service charge. The levels of the attributes ‘view’, ‘noise’ and ‘sunlight’ were conveyed in the choice experiment in the first phase of the survey based on the relative location of the apartment in the block (location method) while in the second phase, these attributes were represented using linguistic categories. In both the phases, different experiments were conducted based on the method of choice elicitation. Three different methods of choice elicitation were employed in the current survey – the binary choice, one stage five point Likert scale and the two stage Likert question[10].

[10] The two different Likert elicitation methods were not combined to capture further precision in the uncertainty level as eliciting respondents’ preference certainty over two different uncertainty elicitation methods for each of the 16 choice scenarios would affect the level of perceived complexity and the resultant fatigue experienced by the respondent,

The binary choice elicitation method asked the respondent to choose an option from a binary set, while the one stage five point Likert method asked the respondents to indicate their level of choice certainty by choosing from scale: Definitely A, Probably A, Uncertain, Probably B, and Definitely B[11]. The two stage Likert question asked the respondents to choose from a binary option and then reveal their level of choice certainty by indicating whether they have 'absolute certainty' or 'have some doubts'[12]. The details of the preference certainty levels provided in Portuguese and its English equivalent is given in APPENDIX B.

For each of the phase, the summary of different experiments conducted can be given as follows:

thus affecting data quality. Moreover, choice data obtained would be more complex thus possibly limiting the types of analytical models applied.

[11] While these were the levels in the original English questionnaire, they were translated to Portuguese whose exact back translation to English will correspond to: Decidedly A, Probably A, I do not have the certainty, Probably B and Decidedly B using online translation service available on uk.babelfish.yahoo.com. However the original levels will henceforth be used in the rest of the thesis as indicative of the respondents' preference certainty levels as the human translation to Portuguese indicates the semantic link between the original and the translated levels, allowing for the use of the original levels in the thesis.

[12] While the original English questionnaire had the levels 'very certain' and 'somewhat certain' as indicators of the respondents' certainty level, these were translated in to Portuguese whose exact back translation to English using online translation service at uk.babelfish.yahoo.com are exemplified in Table 5.13. However, in terms of certainty, translating this back to English would correspond to 'absolutely certain' and 'not so certain' levels, which will be henceforth used in the rest of the thesis.

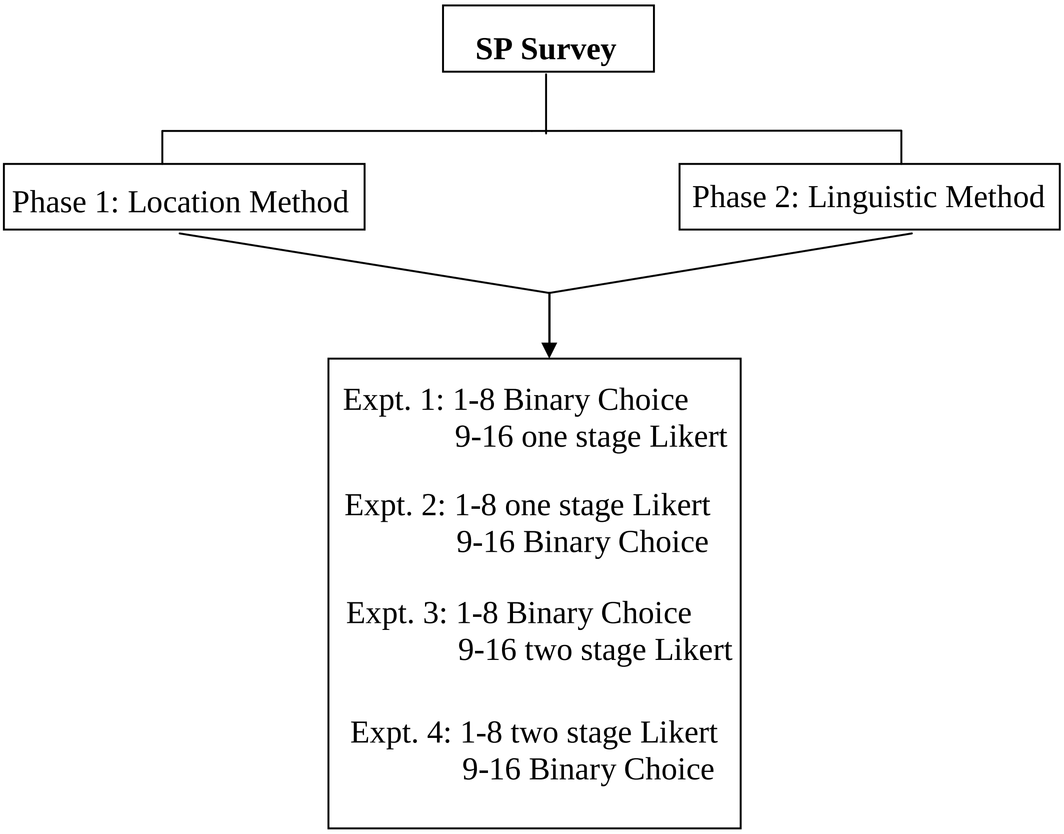


Figure 5.2 Survey Phases and Experiments

Thus, experiments 1 and 2 employed the binary and one stage Likert choice elicitation with different orderings while experiments 3 and 4 employed the binary and two stage Likert choice elicitation methods. A split sample design was used to conduct the experiments with the two methods of choice elicitation such that the levels of as many attributes as possible were split evenly between the two methods. For both the phases, a total of 16 choice scenarios were employed in the SP survey. As all the experiments involved two different methods of choice elicitation, all respondents were offered the complete set of 16 choice scenarios. A fractional factorial design was used for both the phases of the survey. Factorial design is based on the factorial count of all possible combinations of the attribute levels. As the use of full combination of levels in experimental design significantly increases the number of choice scenarios and is suitable only in the case of very small number of attributes or levels, fractional factorial design is generally employed (Louviere *et al*., 2000).

The number of choice scenarios employed in the SP survey is dependent on the number of attributes and their levels. The fractional factorial orthogonal main effects design was used for the experimental design for both the phases of the survey though some variations were then carried out in the second phase design, making it non-orthogonal.

While the Arsenio (2002) study included the 'current alternative' as the base alternative around which the SP design was based, the current survey did not employ this method of design. This is mainly due to the difference in the methods of data collection used in the two surveys. While Arsenio (2002) adapted a computer assisted personal interview (CAPI) to collect the data which provided relative ease in the use of a variable base alternative, the current survey relied on a paper based face-to-face interview for data collection (due to the limitation of resources) and hence it was not possible to have a respondent specific base alternative and a pivotal design. Moreover, as the focus of the current study is to examine the effect of different attribute representation techniques rather than examining the effect of a reference point, the non-application of the 'current alternative' was not expected to be either significant or affecting the power/applicability of the location method.

The SP experimental design is generally tested through simulation which examines the statistical property of the SP design and checks whether the t-statistics are acceptable and whether any changes in the SP design are required. The simulation process is carried out by generating choice from synthetic data and random error terms, and conducting statistical tests to verify the significance of the results and the experimental design. The choice generated in the simulation process is a function of the attribute level, the assumed attribute coefficient (informed from previous/pilot study) and the error term. For the logit model, the error term is assumed to be Gumbel distributed.

The details of the experimental design and simulation employed for each of the phases will be given in the subsequent sub-sections. Section 5.3.1 will detail the experimental design and simulation of the first phase of the survey which used the location method of attribute representation while the experimental design and simulation of the second phase (linguistic representation) will be given in Section 5.3.2.

5.3.1 The Location Method

This section will detail the experimental design and simulation process employed in the first phase of the survey which used the location method of attribute representation. This method of attribute representation was formed with attribute levels based in the context of apartment location in the block.

The survey conducted by Arsenio (2002) revealed that the average number of floors in a block in the area were six. Based on this information, the levels of the attributes using the location method were set. Thus under the location method, view, noise and sunlight were categorised in to four levels based on their location in the block. A slight variation from the method employed by Arsenio (2002) was carried out for this attribute representation method. On the basis of the noise levels obtained for the different floors in that study, the locations on the sixth floor (front and back façade of the block in relation to the main traffic road, thus 6F – apartment on the sixth floor facing the main traffic road and 6T – apartment on the sixth floor facing in the opposite direction of the main traffic road) and similarly those on the third floor (3F and 3T) were selected as levels for the choice experiment in the current study.

As pivotal design and computer aided interview were not employed in the current survey due to resource limitations, the levels for the third and the sixth floor of the block were fixed in the experimental design as these floors showed marked variation in noise levels as given in Arsenio (2002). Moreover, an average of six floors in a block in the area as stated by Arsenio (2002) implied that sufficient blocks would be available for the survey if the sixth floor was chosen as an upper floor level. While the absence of considering the respondent's current apartment location in the experimental design can be regarded as a design limitation, this method was employed considering the resource availability for the survey.

Moreover, care was also taken that the residential blocks selected for the survey contained at least six floors in order to keep the choice scenarios as realistic as possible.

A fractional factorial orthogonal design based on attribute differences was developed using the experimental plan given by Kocur *et al*. (1982) and explained in section 5.3.

The four apartment locations used in the experimental design can be summarised as follows:

Table 5.1 Apartment location description

Levels	**Description**
6F	Apartment on 6th floor on façade facing the main traffic road
6T	Apartment on 6th floor on façade opposite the main traffic road
3F	Apartment on 3rd floor on façade facing the main traffic road
3T	Apartment on 3rd floor on façade opposite the main traffic road

From the information gathered from the blocks as well as from the estate agencies in the area, it was gathered that the average charge value in the area was between 40 – 60 Euros per month. Based on this figure and in order to have the signs of the charge difference evenly distributed across the choice scenarios, both positive and negative signs of charge difference were employed in the experimental design. The levels of all the attributes in the experimental design can be given as follows:

Table 5.2 Experimental design levels for view, noise and sunlight

Attribute	Choice Options		Difference
	Option A	Option B	
View			
Level 0	6F	6F	-
Level 1	6F	6T	-
Level 2	6F	3F	-
Level 3	6F	3T	-
Noise			
Level 0	6F	6F	-
Level 1	6T	6F	-
Level 2	3F	6F	-
Level 3	6F	3T	-
Charge			
Level 0	70	50	20
Level 1	55	80	-25
Level 2	40	75	-35
Level 3	85	40	45
Sunlight			
Level 0	6F	3F	-
Level 1	6F	6T	-
Level 2	6F	6F	-
Level 3	6F	3T	-

The fractional factorial orthogonal main effects design code for four attributes and four levels, obtained from Kocur *et al*. (1982) gave 16 choice scenarios in total which were split into two different sets of choice scenarios based on the choice elicitation method used and as summarised in Section 5.3 and in Figure 5.2.

Simulation was carried out with different levels of the charge difference as well as different coefficient values (including coefficient values obtained from Arsenio, 2002). As the perception ratings for view, noise and sunlight levels were not already known, random numbers were generated for these between ranges (20-80) as well as (0-100) for the simulation process. Though different levels of charge differences were experimented during the simulation process, levels that made practical sense in terms of the absolute values as well as the charge levels prevalent in the area were chosen. Simulation was carried out with this level to test the goodness of model fit and the statistical significance of the coefficient values. Different values of standard deviation (s.d.) during the simulation process were also experimented. Table 5.3 gives the results obtained from the simulation with different levels of attribute values and attribute coefficient with the standard deviation of 1.28 and 2.56.

The variations between the different types of models run during the simulation were based on different attribute and coefficient values. Each of the models given in the table has two different input values (20-80) and (0-100). Thus model 1 has one set of input coefficient and two different sets of input values while model 2 again has the same sets of input values but different input coefficients from model 1. Effect of different standard deviation in simulation has also been given here for some input values and coefficient. Thus, models 5, 7 and 10 provide the effect of standard deviation 2.56 compared to models 4, 1 and 8 respectively which apply the standard deviation of 1.28.

Table 5.3 Simulation results for location method with different input and coefficient values

		INPUT COEFFICIENT					**ESTIMATED COEFFICIENT**				
		V	***N***	***C***	***S***	***s.d***	***V (t-ratio)***	***N***	***C***	***S***	***Adj.*** ρ^2
Model 1	Input (20-80)	.024	.031	-.015	.017	1.28	.026 (10.7)	.028 (11.6)	-.014 (-7.9)	.012 (5.4)	.153
	Input (0-100)	.024	.031	-.015	.017	1.28	.025 (13.7)	.032 (16.9)	-.016 (-7.7)	.017 (10.1)	.32
Model 2	Input (20-80)	.012	.015	-.007	-.008	1.28	.014 (6.6)	.013 (6.3)	-.007 (-4.7)	.009 (4.4)	.053
	Input (0-100)	.012	.015	-.007	-.008	1.28	.012 (8.7)	.017 (11.9)	-.008 (-4.9)	-.008 (-5.8)	.126
Model 3	Input (20-80)	.012	.031	-.015	.017	1.28	.015 (6.6)	.032 (13.1)	-.013 (-7.7)	.014 (6.4)	.142
	Input (0-100)	.012	.031	-.015	.017	1.28	.014 (8.8)	.031 (16.4)	-.014 (-7.2)	.019 (11.9)	.267
Model 4	Input (20-80)	.048	.24	-.2	-.03	1.28	.039 (7.3)	.19 (14.9)	-.16 (-15.3)	-.028 (-5.7)	.78
	Input (0-100)	.048	.24	-.2	-.03	1.28	.038 (9.5)	.18 (14.3)	-.15 (-13.9)	-.019 (5.8)	.805
Model 5	Input (20-80)	.048	.24	-.2	-.03	2.56	.021 (5.6)	.12 (17.2)	-.10 (-18.5)	-.009 (-2.7)	.63
	Input (0-100)	.048	.24	-.2	-.03	2.56	.026 (8.7)	.12 (16.3)	-.10 (-15.3)	-.014 (-5.8)	.73
Model 6	Input (20-80)	.024	.12	-.1	.03	1.28	.024 (6.6)	.12 (16.6)	-.098 (-18)	.03 (8.0)	.62
	Input (0-100)	.024	.12	-.1	.03	1.28	.024 (8.2)	.12 (16.6)	-.09 (-15.9)	.027 (9.2)	.715
Model 7	Input (20-80)	.024	.031	-.015	.017	2.56	.009 (4.0)	.017 (7.66)	-.0096 (-6)	.007 (3.4)	.054
	Input (0-100)	.024	.031	-.015	.017	2.56	.012 (8.8)	.018 (12.7)	-.004 (-2.4)	.008 (5.8)	.127
Model 8	Input (20-80)	.048	.062	-.03	.034	1.28	.049 (14.8)	.063 (16.8)	-.028 (-12)	.034 (11.3)	.382
	Input (0-100)	.048	.062	-.03	.034	1.28	.046 (16.6)	.061 (18.1)	-.028 (-10)	.034 (13.6)	.557
Model 9	Input (20-80)	.024	.031	-.015	.008	1.28	.026 (10.9)	.030 (11.9)	-.012 (-6.9)	.012 (5.3)	.163
	Input (0-100)	.024	.031	-.015	.008	1.28	.022 (13.2)	.032 (16.8)	-.017 (-8.5)	.0085 (5.7)	.276
Model 10	Input (20-80)	.048	.062	-.03	.034	2.56	.025 (10.6)	.028 (11.7)	-.014 (-7.6)	.0156 (6.8)	.158
	Input (0-100)	.048	.062	-.03	.034	2.56	.023 (13.0)	.031 (16.0)	-.015 (-7.6)	.021 (12.4)	.320

Based on the results obtained from the simulation, it can be seen that the estimated coefficient values are close to the input values and are also statistically significant in case of standard deviation of 1.28 (which is the assumption of the MNL model). In case of standard deviation of 2.56, the effect of using that value can be observed in the coefficient estimates which show an almost halving of the input coefficient values used. Using the input values, it is seen that most of the models give an acceptable fit in terms of rho-square values (an acceptable fit is considered to be between 0.1 – 0.2). As the values of the attributes view, noise and sunlight are unknown due to its dependence on respondent's perception ratings, the main aim of the simulation was to determine whether the levels of charge difference selected gave reasonable results. The results obtained indicate that this is the case for almost all the specifications experimented. Based on the simulation exercise, it can thus be concluded that the design developed is fit for purpose.

5.3.2 The Linguistic Method

The respondents in the location representation survey were asked to provide a numeric rating and the associated linguistic category for their perception of the attribute for each of the levels employed in the experimental design (see Table 5.11). Using the results obtained from the perception ratings as well as the average for the linguistic category of view, noise and sunlight from the location representation method, membership functions were developed using the Fuzzy Logic Toolbox in MATLAB. The design of the choice scenarios using the linguistic representation method closely followed from the previous location method. Thus, the decision for developing the different linguistic levels depended on each attribute's numeric rating and linguistic category as defined from the respondent's perception in the location method survey as well as on the linguistic category counts obtained from that method.

The aim of the fuzzy logic membership function was to have much variability in the categories of the attributes along with some overlap in the membership functions. In order to develop the membership function, average values for each of the attribute levels were considered.

For each of the attribute levels in the location method, the following average values were obtained across 220 respondents:

Table 5.4 Average perception rating value obtained for each attribute at each level from the Location method[13]

	6F	6T	3F	3T
View	59.5	60.1	46.8	51.6
Noise	31	50.4	29.8	48.4
Sunlight	70.7	66.4	63.5	59.8

While each respondent was asked to provide a numeric rating for the different location levels, they were also asked to indicate a linguistic level associated with each of the level. Combining the ratings as well as the linguistic categories indicated, the average ratings associated with each linguistic category was evaluated. Thus, the following information was obtained:

Table 5.5 Average perception category values obtained for each attribute for each category from the Location method

	v.bad	bad	neither	good	v.good
View	6.7	28.9	51.5	70.6	90.6
Noise	9.7	29.3	51.4	70.5	84.6
Sunlight	13.8	32.5	49.6	70.6	88.6

The above table indicates the average rating associated with each linguistic category for each attribute. Based on these values, membership functions were developed for each of the linguistic categories using the Fuzzy Logic Toolbox in

[13] While the average perception ratings for each of the apartment locations for noise do not reveal a substantial difference between the perception ratings for the apartments on the third and the sixth floor of a same façade, these levels were selected based on the floor noise levels given in Arsenio (2002). Thus in order to meet the sample size requirements, the average floors in a block in the area as well as the noise levels as given in Arsenio (2002) were considered to form the attribute levels under the location representation method.

MATLAB. The following membership functions were developed for each of the attributes:

Table 5.6 Membership function level for each category of each attribute

	Very bad	**Bad**	**Neither**	**Good**	**Very good**
View	[0 0 10 30]	[10 30 50]	[30 50 55]	[50 55 75]	[55 75 100 100]
Noise	[0 0 15 35]	[15 35 45]	[35 45 55]	[45 55 85]	[55 85 100 100]
Sunlight	[0 0 20 30]	[20 30 50]	[30 50 60]	[50 60 80]	[60 80 100 100]

The above table indicates that in case of the extreme categories ‘very bad’ and ‘very good’ a trapezoidal membership function was developed while in the case of the middle ranges, a triangular membership function was developed. It can be noted that for some of the linguistic categories (such as ‘good’), the middle values were lower than the corresponding averages obtained from the sample. As the range of the ratings for most of the linguistic categories spanned across the entire available scale, some subjective considerations had to be applied in order to form the membership functions.

Though by applying the perception ratings of the attribute levels as well as the statistics associated with each linguistic category, the membership functions were developed, this evaluation was not sufficient in order to develop the levels for the linguistic representation. In order to relate the linguistic levels with the general perception of the sample, the category counts were thus also taken into consideration. The category counts estimated the counts associated with each attribute level and each linguistic category that the respondent was asked to identify with the level. The following table provides the category counts associated with each attribute level and the linguistic category:

Table 5.7 Category counts for attributes in the location representation method

	View				Noise				Sunlight			
	6F	**6T**	**3F**	**3T**	**6F**	**6T**	**3F**	**3T**	**6F**	**6T**	**3F**	**3T**
Very bad	11	2	18	7	70	21	73	22	2	2	6	7
Bad	28	28	53	44	94	45	90	43	14	22	18	24
Neither	67	82	101	102	37	80	46	87	36	43	52	63
Good	91	92	45	60	18	68	10	64	107	114	110	105
V. good	25	18	5	9	2	7	2	5	63	41	35	22

Based on the maximum category counts for each attribute and each linguistic level, the sample's general perception (taken by the maximum number of people selecting a particular linguistic level) of the linguistic category was deciphered[14]. As the general perception of the linguistic category for some attributes resulted in too few linguistic levels (for example, most people thought that sunlight for all the attribute levels belonged to the linguistic category 'good') and while using all the five linguistic levels would make the classification of membership functions difficult in the subsequent analysis, the information obtained from the membership functions was incorporated to obtain the attribute levels for the linguistic representation with as much variation and sample representation as possible while also considering that the total number of choice scenarios in the choice experiment should be the same as that under the location method.

Thus, in order to form the attribute levels for the linguistic method, the category counts, the average perception ratings obtained for each of the attribute locations as well as the fuzzy logic membership functions were considered. For each of the attributes, the following criteria were thus taken into consideration:

[14] Category count from one respondent for noise and for '3F' and '3T' levels of sunlight for another, were excluded due to seemingly illogical categories given (such as lower numeric rating and better linguistic category or vice versa).

- View: based on the category counts obtained, most respondents considered 'view' on '3F' and '3T' to be 'neither good nor bad' while this attribute on '6F' and '6T' was considered to be 'good'. Moreover, across the different apartment locations, the linguistic levels 'neither' and 'good' in combination accounted for majority of the counts. Considering the perception ratings obtained for each of the apartment locations as well as the membership functions developed, 'view' for '6F' and '6T' was classified as 'good' while for '3F' and '3T' was classified as 'neither good nor bad'. Thus two levels for 'view' ('neither' and 'good') were developed for the linguistic representation method

- Noise: the category counts obtained for this attribute in Table 5.7 revealed that most respondents regarded 'noise' for '6F' and '3F' as 'noisy' and for '6T' and '3T' as 'neither noisy nor quiet'. However, in case of '6T' and '3T', the level 'quiet' followed closely as the second most chosen level by the respondents. Moreover, the total category counts for the levels 'noisy', 'neither' and 'quiet' across the different attribute levels constituted a significant proportion of the counts while the average perception rating obtained for the level '6T' (50.4) belonged in the fuzzy logic membership function of both 'neither' and 'quiet'. Hence based on these considerations, three linguistic levels ('noisy', 'neither' and 'quiet') were developed for this attribute

- Sunlight: while the category counts for this attribute revealed that most of the respondents considered 'sunlight' to be 'good' across the different apartment locations, the average perception ratings obtained for '6F' (70.7) belonged to the fuzzy logic membership function of 'good' as well as 'very good'. Moreover, in case of the location '3T', the average perception rating for that location also had some level of membership in the category 'neither' while a significant proportion of the category counts was also obtained for that level. Hence, 'neither', 'good' and 'very good' linguistic levels were formed for this attribute

Based on these factors, the number of associated linguistic levels thus got reduced for each of the attributes. Hence the following levels for each of the attributes were developed based on the criteria listed:

Table 5.8 Dummy levels of attributes under linguistic method

	Level 1	**Level 2**	**Level 3**
View	good	neither	
Noise	noisy	neither	quiet
Sunlight	v.good	good	neither

Fractional factorial orthogonal design was again used with the attributes levels listed in Table 5.8 and with same charge levels as that employed in the location method. However, it was observed that keeping the charge levels same as in the location method, resulted in some dominant choices and hence the sign of the charge difference was changed in these choice scenarios to eliminate the dominant choice problem. A split sample design was again used for the choice experiment and a simulation was conducted with different attribute input and coefficient values.

As the number of attribute levels is fewer with the linguistic representation method compared to the location method, a potential contestable issue can arise from the apparent simplicity in the task with the linguistic representation method. However, the effect of attribute levels on task complexity has not been found to be a significantly crucial factor compared to the number of attributes or the number alternatives within the choice-set (Caussade *et al.*, 2005)[15]. Moreover, though the

[15] Though task complexity in CE has been mostly examined by evaluating the effects of alternatives and attribute number on choice consistency and perceived choice complexity (DeShazo and Fermo, 2002; Greifeneder *et al.*, 2010), the effect of attribute levels on perceived choice complexity was examined by Caussade *et al.* (2005) along with other choice set characteristics such as the number of alternatives, the number of attributes and the range of attribute levels as well as the number of choice scenarios. Using the heteroskedastic logit model the authors found that though the number of attribute levels had an effect on the error variance, the effect observed was three times less than that observed from increasing the number of attributes. The authors thus ranked the effect on increasing error variance in the following sequence: number of attributes, number of alternatives, range of attribute levels, number of attribute levels and the number of choice scenarios.

number of attribute levels has been identified as one of the factors affecting task complexity, the level of perceived complexity is dependent on several factors and especially pertaining to this research, could be dependent on the method of attribute representation. Though the location method consisted of four levels for each of the attributes (view, noise, sunlight and housing service charge), based on the linguistic categories as well as the numeric ratings obtained from the location method, the linguistic method consisted of two levels of view, three levels of noise and sunlight and four levels for the housing service charge, keeping the number of choice scenarios constant over the two representation methods.

While the linguistic method has a lower number of attribute levels for view, noise and sunlight, these levels were formed based on the attribute perception data obtained from the location method. The perception data obtained from the location method revealed that it is not possible to form four equivalent levels in terms of linguistic representation. As the formation of the linguistic levels based on respondents' perceptions obtained from the location data was an important factor in the design of the linguistic representation method, the number of attribute levels was not forced to equate those employed in the location method.

A simulation model was generated with each attribute level category taking different range of values based on the category constructed in the experimental design. The following category values are used for simulation for model 1 – 4:

Table 5.9 Attribute category values used in simulation for the linguistic method

	VIEW		**NOISE**			**SUNLIGHT**		
	Good	**Neither**	**Noisy**	**Neither**	**Quiet**	**V.good**	**Good**	**Neither**
Range	[60, 100]	[30, 60]	[0, 40]	[30, 60]	[60,100]	[80, 100]	[55, 80]	[30, 60]

All models in Table 5.10 have the same attribute coefficient values as well as the standard deviation values as that used in the Location method simulation models, with two different sets of attribute input values. The model indicating *variable* *(var.)* input values incorporate range given in Table 5.9, whereas the other model has a random number ranging from 0 – 100 as input value in the simulation process. The following results are obtained from the simulation exercise:

Table 5.10 Simulation results for linguistic method with different input and coefficient values

		INPUT COEFFICIENT					**ESTIMATED COEFFICIENT**				
		V	***N***	***C***	***S***	***s.d***	***V (t-ratio)***	***N***	***C***	***S***	***Adj.*** ρ^2
Model 1	Input (var.)	.024	.031	-.015	.017	1.28	.024 (9.69)	.031 (17.5)	-.015 (-7.1)	.019 (8.03)	.214
	Input (0-100)	.024	.031	-.015	.017	1.28	.029 (11.2)	.034 (18.1)	-.022 (-9.2)	.0197 (8.2)	.238
Model 2	Input (var.)	.012	.015	-.007	-.008	1.28	.012 (5.79)	.015 (11.2)	-.008 (-4.2)	.0065 (3.18)	.063
	Input (0-100)	.012	.015	-.007	-.008	1.28	.011 (8.35)	.015 (11.0)	-.009 (-5.1)	-.009 (-6.5)	.107
Model 3	Input (var.)	.012	.031	-.015	.017	1.28	.016 (6.74)	.033 (18.0)	-.014 (-6.3)	.018 (7.58)	.235
	Input (0-100)	.012	.031	-.015	.017	1.28	.013 (8.47)	.028 (15.6)	-.013 (-7.2)	.017 (10.5)	.240
Model 4	Input (var.)	.048	.24	-.2	-.03	1.28	.053 (9.26)	.23 (14.39)	-.19 (-13.3)	-.028 (-5.4)	.794
	Input (0-100)	.048	.24	-.2	-.03	1.28	.047 (9.6)	.23 (12.68)	-.19 (-12.4)	-.033 (-7.9)	.836
Model 5	Input (var.)	.048	.24	-.2	-.03	2.56	.024 (6.36)	.12 (17.4)	-.10 (-14.6)	-.012 (-3.2)	.666
	Input (0-100)	.048	.24	-.2	-.03	2.56	.027 (8.97)	.12 (16.4)	-.10 (-15.9)	-.017 (-6.3)	.705
Model 6	Input (var.)	.024	.12	-.1	.03	1.28	.026 (6.95)	.12 (17.7)	-.10 (-15.4)	.03 (7.97)	.625
	Input (0-100)	.024	.12	-.1	.03	1.28	.026 (8.39)	.13 (16.18)	-.11 (-15.3)	.031 (9.53)	.732
Model 7	Input (var.)	.024	.031	-.015	.017	2.56	.011 (5.23)	.014 (10.6)	-.006 (-3.2)	.010 (5.03)	.064
	Input (0-100)	.024	.031	-.015	.017	2.56	.013 (9.54)	.015 (10.8)	-.009 (-5.5)	.007 (5.08)	.122
Model 8	Input (var.)	.048	.062	-.03	.034	1.28	.045 (13.6)	.059 (20.0)	-.028 (-9.8)	.029 (9.36)	.442
	Input (0-100)	.048	.062	-.03	.034	1.28	.050 (16.9)	.062 (18.2)	-.029 (-10)	.030 (12.78)	.557
Model 9	Input (var.)	.024	.031	-.015	.008	1.28	.024 (10.2)	.028 (16.8)	-.015 (-6.8)	.006 (2.8)	.182
	Input (0-100)	.024	.031	-.015	.008	1.28	.025 (14.1)	.028 (16.1)	-.017 (-8.8)	.009 (5.78)	.287
Model 10	Input (var.)	.048	.062	-.03	.034	2.56	.029 (11.2)	.034 (18.1)	-.022 (-9.2)	.0197 (8.2)	.238
	Input (0-100)	.048	.062	-.03	.034	2.56	.023 (13.3)	.030 (16.1)	-.015 (-7.6)	.0178 (10.9)	.297

The purpose of conducting the simulation was again to test whether the selected charge levels gave significant values. From the simulation results obtained, it can be seen in case of standard deviation of 1.28, the estimated coefficients give reasonably close estimates to the input coefficient values. In case of the standard deviation of 2.56, the effect is again observed in almost halving of the estimated coefficients. Across the different models however, it can be observed that statistically significant estimates are obtained along with a reasonable rho square value implying that the design is quite robust and the charge values selected for the experimental design and simulation can be used in the choice experiment.

This section detailed the experimental design and simulation procedure used in both the phases of the survey. A fractional factorial orthogonal design was created for both the phases based on difference design. For the second phase of the survey, a slight relaxation was made in the orthogonal design (by changing the sign of the charge difference for three of the scenarios) to avoid the dominant choice problem. Based on the results obtained from the simulation, it was seen that the charge levels selected for both the phases of the survey yielded statistically significant results in terms of the t-statistics of the estimated coefficients as well as the model fit. The designs were thus considered satisfactory.

Considering the design and simulation process in this section, the next section will detail the method of data collection and survey.

5.4 Data Collection and Description

5.4.1 Survey Objectives

The survey comprised an important aspect of the research whose objective was to capture respondent’s preference uncertainty and identify the causes of it in terms of the SP choice design. The following aims were thus identified to meet the objectives:

- Test different methods of preference elicitation techniques to capture respondent's preferences
- Check the effect of attribute representation technique on respondent's choice certainty
- Test the application of a heuristics based approach on respondent's choice and attribute representation technique

The survey procedure followed to meet these objectives is detailed in the subsequent section.

5.4.2 Survey Method

The main survey was conducted during February – April 2008 using paper based face-to-face residential interview. Prior to the pilot and the main survey, the interviewers were trained in the method to conduct the survey and briefed on health and safety matters with necessary personal safety system in place when conducting the survey. The candidate was involved in several aspects during the design and management of the survey[16].

---

[16] The candidate's role in survey and management included survey and questionnaire design, preparing the questionnaire and the associated cards used during the survey along with forms for number of completed questionnaires and interviewers' payment, employing and training interviewers and conducting health and safety inductions. The level of housing service charge generally levied in the area was also researched through the housing agents in the area as well as the service charge information provided in the various blocks of the site. Making arrangements for payments as well as travel allowances for the interviewers was also conducted by the candidate during the survey. Prior to the pilot survey, the survey site was well examined and blocks were identified where surveys would be conducted. Interviewers were also familiarised with the survey site and the different locations of the site (based on the nearest main traffic road). Initial letters were prepared informing the residents of the site on the survey with the tentative dates when it would be conducted in the blocks and these were then put into the residents' letter box with the help of few other interviewers.

On each survey day, interviewers were assigned specific area and blocks where they would conduct the interviews and were handed the relevant list of apartments that were already interviewed in the blocks of that area. The candidate accompanied an interviewer (selected randomly or based on interviewer/survey requirements) every survey day during the pilot as well as the main survey. The purpose of accompanying the interviewer was to ensure the quality of the survey as well as to help the interviewer during the survey process by lifting the questionnaires, showing the choice cards to respondents etc. The candidate was present on the survey site at all times and days when the survey was conducted and acted as

A short pilot study over a period of two days with about 13 respondents was carried out in the area to test the effect of the questionnaire design. The main survey for the first phase was conducted between mid February to early March while for the second phase was conducted from late March to mid April. The surveys were mostly undertaken between 6:30 – 9:00 pm on weekdays and between 2:00 – 8:00 pm on weekends in order to have as much variety in respondents as possible. Only one respondent per household was interviewed.

The questionnaires and the associated choice cards were colour coded based on the experiment type (questionnaire and choice cards for each experiment were printed on different colour paper based on the experiment type) as given in Figure 5.2 for easy recognition during and after the survey. The survey site was divided into three parts based on the nearest main traffic road. As one of the questions was related to the respondent's perception of noise in the apartments of the block in relation to one of the major traffic roads, the blocks closer to the periphery of the enclosed residential area, and hence nearer to one of the main traffic roads were selected for the survey. A minimum of 50 completed questionnaires were aimed for each of the experiment of each phase, thus bringing the number of respondents for each experiment at every specified site to about 17.

Blocks closer to the particular main traffic road were identified and each interviewer was given a set of blocks where they could undertake the survey. The interviewers were asked to keep a list of the block and the apartments, along with the information on which households have undertaken the survey. This was done in order to avoid repeating the survey in the same household. As only the households that participated in the survey were noted during the data collection process and

---

a site supervisor as well as a point of immediate contact for the interviewers in case of any needs/emergency. At the end of each survey day, the candidate also received feedback from the interviewers based on their experiences or comments from the respondents, discussed any important or relevant aspects with other interviewers as well as replied to any queries or doubts raised by the interviewers. The completed questionnaires of the day were then collected from the interviewers and sorted while the interviewers were asked to indicate the number of completed questionnaires on the relevant forms.

information on households that refused to the survey was not recorded[17], the rejection rate cannot be estimated. The list of all the blocks and households that took part in the first phase of the survey was examined closely to identify the blocks and households for the second phase. For the entire survey, the aim was to interview a household only once though this had to be relaxed slightly in the second phase due to lower availability of respondents. In this case, some people who participated in the first phase of the survey had to be contacted to participate in the second phase. Though some re-interviews were conducted, the number of re-interviews were very few (seven out of 204 interviews obtained in the second phase) which amounted to about 3.4% of the sample size.

Before the pilot survey of the first phase, letters were sent out to the residents of selected blocks in the site, informing them of the tentative period when the survey will be undertaken along with the objectives of the survey. During the course of the first phase of the survey, it was learnt that this practice did not affect the respondent's willingness to participate as those who did not receive the letter also seemed eager to take part in the survey, moreover the availability of respondents in their apartment was also not found to be significantly determined by the tentative survey period sent out in the letter. Hence this practice was discontinued for the second phase of the survey.

The questionnaire consisted of two parts: one, with the socio-economic, residential and noise characteristics of the concerned apartment and second, a set of choice cards for the SP exercise. The interviewer was asked to fill the questionnaire on the respondent's behalf during the interview. During the SP exercise, the choice cards were offered to the respondents, one at a time, and the choice responses were noted in the questionnaire by the interviewer. The face-to-face interview allowed the respondents to discuss or clarify any issues that emerged during the process. The

[17] The interviewers experienced that in some cases, the households' willingness to participate in the survey depended on the person who answered the door-bell. Thus, in those cases, while a member of the household had refused to participate in the survey, another household member showed willingness to participate when asked on another day. Also, when a household was unable to participate in the survey on a particular day when the interviewer was visiting the block, appointment was taken with the household member for another day and time when they would be able to participate. Thus record of households that definitely did not want to participate in the survey was not collected.

respondents were explained the meaning of the choice scenario elaborately in the beginning of the choice experiment and they were strongly encouraged to reply to all the choice experiment questions. When respondents could not reply to all the 16 choice scenarios of the choice experiment, the survey was terminated and not included in the data input process. Thus only those questionnaires which had all responses to the choice experiment section were regarded as a completed questionnaire.

While 55 completed questionnaires were obtained for experiments 1 and 2 of the first phase, 56 completed questionnaires were obtained for experiments 3 and 4, bringing the total respondents for the first phase to 222 while for the second phase, 51 completed questionnaires were obtained for each of the experiments bringing the total number of respondents to 204.

5.4.3 Questionnaire format

The initial questionnaire design followed closely to that used by Arsenio (2002). The interviewers were asked to mark the position of the block with respect to the main traffic road and also indicate the name of the traffic road. This was then followed by a brief description of the objective of the survey which took the following form:

This research is being conducted to characterise some flat attributes and the local environment in this residential area. For this purpose, we would like to ask you some questions and we thank you in advance for your cooperation.

This was followed by the address of the apartment and its position in relation to the main traffic road. Based on the questionnaire design adapted from Arsenio (2002), the household familiarity question followed where the respondents were asked to indicate the number of years they have lived in the apartment, the number of people and children living in the apartment and the reasons for choosing the location and the particular apartment as well as the information on the apartment tenure. However, during the pilot study it was observed that these questions considerably increased the amount of time spent on the questionnaire before the choice experiment was conducted and hence the interviewers were trained to ask this question later in the interview. Thus the choice experiment question took precedence over this question during the main survey of both the survey phases.

In the first phase of the survey which used the location method of attribute representation, respondents' familiarity with flat characteristics was assessed by asking them to indicate the rent, service charge, number of rooms, area and garage parking space for the locations 6F, 6T, 3F and 3T that were used as levels in the choice experiment (and are described in Section 5.3.1 of this chapter). During the pilot study it was known that respondents faced much problem in answering this question and they also required some time to respond. As the question was not considered to be extremely significant to the objectives of the survey, the interviewers were trained not to spend too much time on it if the respondent had difficulties in the familiarity of these characteristics. The purpose to include this question was to accustom the respondents to the different apartment locations that would be used in the choice experiment as the different attribute levels. Thus during the main survey, care was taken that considerable time is not spent on this particular question if the respondents were not familiar with the apartment characteristics.

Following the question on the apartment characteristics for the apartment locations used in the survey, the respondents were asked to classify attributes such as view, noise and sunlight for each of the apartment locations in terms of a linguistic category as well as a perception rating (from 0 - 100). It was assumed that as respondents might have difficulty to imagine the attribute levels of the specified apartments during the decision-making for the choice experiments, they would anchor the levels for the apartment based on the ratings they give. Thus, they were asked to give the attribute ratings prior to the choice experiment. During the course of the survey, it was also observed that some respondents cross-referenced their attribute ratings with the interviewer before the decision-making.

Thus for example, for view, respondents were asked to give a linguistic category and a rating in the following form:

Table 5.11 Attribute rating exercise for the location method

<table>
<tr><td colspan="6">Characteristic of the apartment</td><td>View</td></tr>
<tr><td colspan="7">Very bad (0) ____ Very good (100)</td></tr>
<tr><td></td><td>Very bad</td><td>Bad</td><td>Neither</td><td>Good</td><td>Very good</td><td>Ratings</td></tr>
<tr><td>6F</td><td>●</td><td>●</td><td>●</td><td>●</td><td>●</td><td></td></tr>
<tr><td>6T</td><td>●</td><td>●</td><td>●</td><td>●</td><td>●</td><td></td></tr>
<tr><td>3F</td><td>●</td><td>●</td><td>●</td><td>●</td><td>●</td><td></td></tr>
<tr><td>3T</td><td>●</td><td>●</td><td>●</td><td>●</td><td>●</td><td></td></tr>
</table>

The linguistic category and ratings followed suit for noise and sunlight. The SP choice experiment followed the question on attribute perception. The SP exercise was in a form of a payment card where the respondents were shown a card of each choice scenario. Depending on the type of the experiment, different forms and sequences of choice elicitation method were used as described in section 5.3.

An example of the choice scenario with a five point Likert scale choice elicitation method can be given as:

Table 5.12 Example of a five point Likert scale choice scenario with the location method

<table>
<tr><th colspan="2">OPTION A</th><th></th><th colspan="2">OPTION B</th></tr>
<tr><td colspan="2">View: 6F</td><td></td><td colspan="2">View: 6F</td></tr>
<tr><td colspan="2">Noise: 3F</td><td></td><td colspan="2">Noise: 6F</td></tr>
<tr><td colspan="2">Housing service charge:
€ 40</td><td></td><td colspan="2">Housing service charge:
€ 75</td></tr>
<tr><td colspan="2">Sunlight: 6F</td><td></td><td colspan="2">Sunlight: 3T</td></tr>
<tr><td>Definitely
A</td><td>Probably
A</td><td>Uncertain</td><td>Probably
B</td><td>Definitely
B</td></tr>
</table>

While an example of a choice scenario with two stage Likert scale question can be given as the following:

Table 5.13 Example of a two stage Likert scale choice scenario with location representation

OPTION A	**OPTION B**
View: 6F	View: 6F
Noise: 3F	Noise: 6F
Housing service charge: € 40	Housing service charge: € 75
Sunlight: 6F	Sunlight: 3T
A	B

Degree of certainty/confidence:

I have the absolute certainty ☐ **I have some doubts** ☐

Comparing the levels used in the two elicitation methods it can be noted that in the case of the two stage Likert method, the '*I have some doubts*' (not so certain) could possibly include aspects of 'probably' and 'uncertain' levels of the five point Likert question. As in the case of the two stage Likert the preference elicitation question followed a post-choice certainty format, a separate 'uncertain' option would not be logical in the second stage of the choice question (where choice certainty was elicited). Though the 'uncertain' alternative could have been included in the first stage of the elicitation method where respondents were offered the two alternatives, this was not conducted for two reasons – 1) the inclusion of 'uncertain' alternative at this stage could have possibly increased task complexity/fatigue in respondents for later choice scenarios as it would involve examining the alternatives at three levels of choice (A, uncertain or B) and then indicating the level of preference certainty in cases where they chose A or B and 2) the exclusion of the 'uncertain' alternative in the two stage while its inclusion in the one stage Likert method can allow for the examination of differences between the two elicitation methods and the effect of the 'uncertain' alternative on respondents' choices.

Though the two stage Likert elicitation process is different from the five point Likert method and can be argued by some as a 'debriefing' exercise, it would however not be entirely appropriate to term it as such[18]. Moreover, terming it as a two stage Likert question follows from the notation used by Albaum (1997) who compared the effects of one and two stage Likert methods, with the split question technique employed in the two stage Likert method working as an 'unfolding' process where respondents first indicates their choice and then the preference/strength level for their choice. As Albaum (1997) revealed a difference in the effect of one and two stage Likert methods based on their ability to capture the direction and intensity of preferences, a purpose of using these different elicitation methods in the current research is to examine the effects and differences of the two methods within CE and in the specific experimental context. An alternative nomenclature of the two stage Likert method could be the 'post-decisional certainty' measure as has been termed by Li and Mattsson (1995). However, deviating from the current nomenclature of the method would be of little value as the research aims to compare the effects of one and two stage Likert methods outlined by Albaum (1997).

Another possible aspect of contestation could arise from the terminology of the choice elicitation question. In the preference uncertainty literature, examples can be found where the elicitations of respondents' preferences have been termed as the strength of preference (SOP) exercise or as eliciting preference uncertainty, sometimes used interchangeably. A distinction of the two terms is thus required in order to clarify which technique has been applied in this research.

[18] In contingent valuation as well choice experiments, debriefing questions are asked in a separate section from the choice/WTP elicitation question in order to estimate respondent's attitudes or factors that influenced respondent's choice or a WTP level as well as to identify 'protest' bidders and respondents who did not understand the experiment (Alpizar and Carlsson, 2001; Navrud, 2002; Romano and Vigano, 1998). However Li and Mattsson (1995) ask respondents a follow-up question after a dichotomous WTP elicitation (yes/no) question asking respondents to indicate their level of confidence on the WTP question. Terming this follow-up question as a follow-up debriefing question could possibly result in confusion concerning the usage of terms for the preference elicitation exercise in choice experiments. While the question adapted by Li and Mattsson (1995) could be termed as a debriefing question within the CVM framework to some extent, following the standard notations used in the valuation questionnaire formats, terming the follow-up certainty question from a binary choice exercise in CE as a 'debriefing' question could be inappropriate as well as confusing.

While the strength of preference (SOP) elicits respondents' level of preference (strong vs. mild), it does not account for any uncertainty. Eliciting respondents' strength of preference, Johnston and Swallow (1999) offered respondents a trichotomous choice (accept, neutral, reject) where respondents were asked to indicate their preference strength as strongly, moderately or slightly accept (or reject). Another example of an application of SOP indicator can be found in Swallow *et al.* (2001) who offered respondents a binary choice and for either of the choice made, asked the respondent to indicate their preference strength on a five point scale ranging from 'slightly prefer' to 'strongly prefer'.

Confusion in the literature on the use of SOP for the level of preference uncertainty could possibly be from Whelan and Tapley (2006) who defined a five point Likert scale of preference certainty between bridge and tunnel of: definitely tunnel, probably tunnel, no preference, probably bridge and definitely bridge as respondents' SOP. In Whelan and Crockett (2009) the SOP method was correctly applied where respondents were asked to indicate whether they: strongly prefer A, prefer A, prefer B or strongly prefer B. Thus, different from the preference certainty level, a 'moderate' or 'slight' preference indicates the level of respondent's preference for an alternative (but not the level of certainty). In case where the level of certainty is explicitly sought either in the case of a five point Likert scale or a two stage Likert method where respondents are asked to indicate their certainty level after a choice is made, this is to be considered as a preference uncertainty measure and not a strength of preference indicator. A SOP indicator is thus to be understood to depict the level of preference strength (mostly as a matter of taste) without any uncertainty involved. The choice elicitation method applied in this research (mainly the one and two stage Likert elicitation methods) will thus henceforth be termed as methods to capture preference uncertainty rather than methods to elicit SOP.

While some distinction between the SOP indicator and the preference uncertainty measure has been outlined, it is important to recognise that these two measures could also be intricately connected. Thus, in case where an individual 'strongly' prefers alternative A over B, he could indicate that he 'definitely' prefers A. Thus, some ambiguity remains between the preference uncertainty and the strength of preference terminology which can be recognised as a limitation of the study.

The questionnaire for both phases of the survey included questions on reasons for choosing the particular apartment and the residential location, number of hours spent in the apartment during the week and on weekends, the levels and cause of noise annoyance during day and night, perception ratings for the attribute levels for view, noise and sunlight and socio-economic questions such as education, income, occupation and gender. The choice experiment section of the questionnaire varied according to the phase and the experiment involved. The phases of the SP survey and the types of different experiments are graphically represented in Figure 5.2**.**

Following from the choice experiment question, questions on the respondent's perception and causes for day and night time noise at home (from external sources), the level of noise annoyance, the main causes of noise as well as any measures undertaken to reduce the noise impact within the household were asked. These questions were followed by the respondent's household characteristics such as the number of people and children living in the household and the reasons for choosing the location and the particular apartment followed by socio-economic questions such as education, age, income, gender and occupation. A sample of the questionnaire (translated in English) can be found in APPENDIX A.

To recollect, in the location method pilot study, following from the questionnaire design employed by Arsenio (2002), question was asked to know the respondents' level of knowledge about other flat characteristics in the block. Thus the respondents were asked to indicate their knowledge of rent, housing service charge, number of rooms, area and the availability of garage for the apartments 6F, 6T, 3F and 3T (which are as described in section 5.3.1 and which formed the levels for the choice experiment in phase 1 of the survey).

For the second phase of the survey which employed the linguistic representation of attribute levels, the question on the knowledge of the respondent about other flat characteristics was excluded from the questionnaire as it was irrelevant. Thus, compared to the location method, only one additional question (eliciting respondents' familiarity with price/rent, housing tax, number of apartments, flat area and availability of garage for the different apartments used in the location method) did not precede the choice experiment question in the linguistic method as it was not required for the choice exercise. This question was included with the location method as it was needed to accustom the respondents to the attribute levels that would be used in the choice exercise and hence this question was used to define the different levels (in terms of the different apartments) to the respondents; the responses for the question were not deemed to be extremely necessary for the survey. Moreover, the interviewers were trained not to spend more than a couple of minutes on the question especially if the respondent found it difficult to answer. The important purpose of the question for the location representation survey was to accustom the respondent to the apartment locations that would be used in the survey.

The effect of exclusion of this question with the linguistic method was not expected to be extremely significant in terms of affecting the survey length and consequently, the respondent's performance as – 1) respondents with the location method questionnaire were asked not to spend too much time (more than a few minutes) on thinking about the answers if they did not know the apartment characteristics and 2) in most of the cases, respondents chose to reply by a 'don't know' though the question was still posed in order to define the levels. Thus, no substantial time variation across the two survey types was obtained due to this question. However the variation in the survey length across the two phases could imply increased fatigue in some cases prior to the choice experiment question for the first phase of the survey and thus can be regarded as a limitation of the study.

In the second phase of the survey (employing linguistic representation), the respondents were again asked to give a rating from 0 (very bad) – 100 (very good) for all the linguistic categories (very good, good, neither, bad and very bad) and for all the attributes in the block. Thus, they were asked, for example, 'if you were to give a rating from 0-100 for "very good" view in this block what rating would you give?' The rating question again preceded the choice experiment although it was considered that the linguistic representation of attributes would be relatively simpler to understand and the respondents will not have to anchor their understanding of the attribute level to the numeric ratings that they have given. This question was of particular relevance as the ratings obtained from the respondents for each of linguistic level would be used to design the membership function in the fuzzy logic analysis. Thus the rating question outlined in Table 5.11 was modified as the following for the phase 2 experiments:

Table 5.14 Attribute rating exercise for phase 2

Characteristic of apartment	View
Very bad (0) ________________________ Very good (100)	
	Ratings
Very bad	
Bad	
Neither	
Good	
Very good	

Compared to the location method where respondents are asked to rate the attribute levels in terms of verbal as well as a numeric scale, it can be noted that only numeric ratings are elicited against a word scale in the linguistic method. The basis of this variation lies in the fundamental difference between the location and the linguistic methods. As the location method used different apartment locations in the block as the basis of the attribute levels for view, noise and sunlight, respondents' perceptions on the attribute levels were sought on a numeric rating scale (0-100) along with a linguistic category that they were asked to identify. The purpose of eliciting both the linguistic category as well as the numeric rating for the location method was to obtain the link between the different attribute levels, the linguistic categories as well as the numeric ratings in order to develop the experimental design for the linguistic method.

The method of developing the levels in the choice experiment for the linguistic method using the attribute levels, linguistic category and numeric ratings data obtained from the location method is outlined in Section 5.3.2. In case of the linguistic method which represented the attribute levels in semantic terms, asking respondents to identify a linguistic level on a linguistic category would be a redundant and confusing exercise for the respondents and hence only numeric ratings information, which would be used in the modelling process, was elicited with this representation method. Thus, for the linguistic method, respondents were only asked to provide a subjective numeric rating associated for each of the attributes represented in linguistic categories.

This section detailed the survey method and the questionnaire design. The next section will now describe the initial descriptive statistics of some of the socio-economic as well as noise perception characteristics of the respondent.

5.5 Descriptive Statistics

This section provides an overview of the socio-economic characteristics observed in the collected sample, across the different experiments. As the survey was conducted in two phases with different experiments for each of the phase, the descriptive statistics of the socio-economic and choice characteristics will be given for both the phases and all the different experiments. The main socio-economic characteristics that will be focussed on are: income, education, gender and occupation.

Section 5.5.1 of the chapter will focus on phase 1 of the survey where location representation of the attributes was used in the choice experiment. The descriptive statistics of the socio-economic and choice characteristics at the aggregate level will be given in Table 5.15 of Section 5.5.1.1. The descriptive statistics for each of the related experiments in this phase will be given in the section subsequently. Section 5.5.1.2 will detail respondents' noise perceptions for the first phase of the survey. Section 5.5.2.1 will describe the socio-economic characteristics of respondents in the second phase of the survey while section 5.5.2.2 will detail the noise perception of the respondents in that phase.

While obtaining a representative sample of Lisbon was not a particular aim of the study, some comparisons of the descriptive statistics with the survey conducted by Arsenio (2002) as well as the statistics for Lisbon are provided in Section 5.5.3 along with comparative statistics across the location and linguistic representation data.

5.5.1 The Location Method

5.5.1.1 Socio-economic characteristics

This section gives the descriptive statistics of the socio-economic characteristics for the experiments conducted in the first phase of the survey. Table 5.15 will give the aggregate descriptive statistics of all the socio-economic attributes over all the experiments while Table 5.16 will outline the descriptive statistics for experiments 1 and 2 which employed the binary and five point Likert scale responses with different ordering for each of the experiments and experiments 3 and 4 which employed the binary and two stage Likert scale choice scenarios with different ordering.

The aggregate statistics of the characteristics gender, age, income, education and occupation over all the experiments carried out in the first phase of the survey can be given as the following:

Table 5.15 Socio-economic characteristics for location method - aggregate data of all experiments

RESPONDENT CHARACTERISTICS	NUMBER OF RESPONDENTS 222	PERCENTAGE
Gender		
Male	97	43.7
Female	125	56.3
Age		
18 – 25	41	18.5
26 – 40	63	28.4
41 – 55	65	29.3
56 – 75	51	22.9
> 75	2	0.9
Household Income €/month		
Less than 1K	16	7.2
1K – 2K	37	16.7
2K – 3K	42	18.9
3K – 4K	35	15.8
4K – 5K	28	12.6
More than 5K	17	7.6
No answer	47	21.2
Education		
Primary	6	2.7
Secondary	43	19.4
Graduate	134	60.4
Post-graduate	39	17.6
Occupation		
Part-time	32	14.4
Full-time	130	58.5
Unemployed	7	3.1
Retired	26	11.7
House based	3	1.3
Student	23	10.4
No answer	1	0.4

From the above table it is evident that the majority of respondents in this phase were female. The distribution across the different age groups from 18-75 is fairly even with a slightly higher proportion of 41-55 year olds and lesser proportion of above 75 years old. It is also observed that most of the respondents replied to the household income question, with the majority stating the income to be between Euro 2000 - 3000 per month. From the information gathered before the survey, it was anticipated that there would be higher number of older people with Primary school education and hence this level was included in the questionnaire. However, it was found that only 2.7% of the sample was in that category, while the majority were Graduates. While the survey did not aim for a representative sample, care was taken to incorporate as much variety as possible.

The summary of relative proportion of respondents with different socio-economic characteristics across the two sets of preference elicitation experiments can be given as follows:

Table 5.16 Percentage of respondents with different socio-economic characteristics across different preference elicitation experiments with Location method

RESPONDENT CHARACTERISTICS	BINARY + ONE STAGE LIKERT	BINARY + TWO STAGE LIKERT
Gender (in percentage)		
Male	41.8	45.5
Female	58.2	54.5
Age		
18 – 25	17.3	19.6
26 – 40	27.3	29.5
41 – 55	33.6	25
56 – 75	20.9	25
> 75	0.9	0.9
Household Income €/month		
Less than 1K	5.4	8.9
1K – 2K	15.4	17.9
2K – 3K	17.3	20.5
3K – 4K	19.1	12.5
4K – 5K	15.4	9.8
More than 5K	10.9	4.5
No answer	16.4	25.9
Education		
Primary	0.9	4.5
Secondary	18.2	20.5
Graduate	61.8	58.9
Post-graduate	19.1	16.1
Occupation		
Part-time	15.4	13.4
Full-time	60	57.1
Unemployed	2.7	3.6
Retired	10	13.4
House based	0.9	1.8
Student	10.9	9.8
No answer	0	0.9
Number of respondents	**110**	**112**

Comparing the characteristics for the category with maximum number of respondents reveal that there are slightly higher differences in age and income categories across the two columns and lesser so in the case of education and occupation. For the first two sets of the experiment (binary and one stage Likert) it seen that the maximum number of respondents lie in the age group of 41-55 while for income, the maximum number of respondents who answered the income question earn Euro 3000-4000 per month. For the other two sets of the experiment which employed binary and two stage Likert choice elicitation question, the maximum number of respondents lie in the age group of 26-40 and for the income characteristic, for those who answered the question, in the Euro 2000-3000 category. Though there are some differences in the exact percentage values across the two columns for education and occupation, the same category for both these characteristics has the maximum number of respondents across the experiments. Thus most respondents across the two different experiments are graduates and in full-time employment.

5.5.1.2 Noise characteristics

Respondent's perception of the noise level as well as their level of noise annoyance is given in this section along with the different causes of noise pollution. Table 5.17 reveals the percentage of respondents at the aggregate level of phase 1 of the survey with different levels of day and night noise as well as the disturbance perceived during these time periods.

Table 5.17 Day and night noise source and disturbance level of respondents with location method

RESPONDENT CHARACTERISTICS	NUMBER OF RESPONDENTS 222	PERCENTAGE
Day noise		
Very noisy	26	11.7
Noisy	77	34.7
Neither	81	36.5
Quiet	30	13.5
Very quiet	8	3.6
Day noise disturbance		
Very much	30	13.5
Moderate	71	31.9
Little	82	36.9
None	39	17.6
Night noise		
Very noisy	15	6.7
Noisy	61	27.5
Neither	75	33.8
Quiet	62	27.9
Very quiet	9	4
Night noise disturbance		
Very much	35	15.8
Moderate	63	28.4
Little	75	33.8
None	49	22
Measure		
Self	85	38.3
Non-self/ none	137	61.7

It is seen that for the day time noise level perceived by the respondent, there is an almost even split between the categories ‘noisy’ and ‘neither noisy nor quiet’ with about 35% for each of these. The majority also consider night time noise to be ‘neither noisy nor quiet’ while there is an almost even split between the number of respondents in the categories ‘noisy’ and ‘quiet’. For both day and night time noise annoyance, it is seen that most respondents seem to be ‘little’ annoyed by the levels. From the above table it is also evident that most people have not taken any extra measures by themselves to reduce the amount of external noise in the apartment.

Comparing the different causes of noise on the ranking given by the respondents, it is seen in Table 5.18 that most respondents ranked road traffic as the main cause of noise while the lowest rank was given to ‘neighbours’ and ‘construction’ factors of the externality.

Table 5.18 Ranking of noise causes of respondents in Location method

	NOISE CAUSE				
Rank	Road traffic	Aircraft	Neighbour	Construction	Others
1	77%	2.2%	10.4%	7.2%	3.1%
2	10.4%	21.2%	17.6%	22.1%	1.8%
3	4.9%	8.5%	11.7%	11.7%	3.1%

5.5.1.3 Preference Levels

The choice elicited from the one and two stage Likert methods for the location method can be outlined as below:

Table 5.19 Preference levels elicited across the two Likert methods

RESPONDENT CHARACTERISTICS	PERCENTAGE
One stage Likert (total observations: 880)	
Definitely A	25.6
Probably A	30.9
Uncertain	8.4
Probably B	23.7
Definitely B	11.4
Two stage Likert (total observations: 896)	
Absolutely certain A	43.2
Not so certain A	18.3
Not so certain B	12.2
Absolutely certain B	26.3

While in the case of one stage Likert method, the ‘Probably’ options have been selected more than the ‘Definitely’ options, in case of the two stage Likert method, the contrary is true. This outlines some possible effect of the two elicitation methods on the level of preference certainty.

5.5.2 The Linguistic Method

5.5.2.1 Socio-economic characteristics

This section gives the socio-economic characteristics of the respondents in the second phase of the survey which employed the linguistic attribute representation technique. The aggregate statistics from all the experiments carried out in this phase is given in Table 5.20 while an overview of the statistics across the two sets of experiments with different preference elicitation methods is given in Table 5.21.

The aggregate statistics for the socio-economic characteristics across all the experiments is as follows:

Table 5.20 Socio-economic characteristics of respondents with the linguistic method

RESPONDENT CHARACTERISTICS	NUMBER OF RESPONDENTS 204	PERCENTAGE
Gender		
Male	81	39.7
Female	123	60.3
Age		
18 – 25	49	24
26 – 40	49	24
41 – 55	64	31.4
56 – 75	37	18.1
> 75	5	2.4
Household Income/month		
Less than 1K	12	5.9
1K – 2K	33	16.2
2K – 3K	45	22
3K – 4K	20	9.8
4K – 5K	14	6.9
More than 5K	20	9.8
No answer	60	29.4
Education		
Primary	1	0.5
Secondary	40	19.6
Graduate	128	62.7
Post-graduate	35	17.1
Occupation		
Part-time	12	5.9
Full-time	119	58.3
Unemployed	7	3.4
Retired	27	13.2
House based	8	3.9
Student	31	15.2
No answer	0	0

The table reveals that female respondents again comprise a large portion in the survey. Moreover, the majority of respondents lie in the age group of 41-55. In the second phase of the survey, a slightly higher portion of the sample refused to answer the income question as compared to those in the first phase. From those who answered, majority indicated the household income to be between Euro 2000 - 3000 per month, as observed with the location method. Again similar to that obtained from the location method, majority of the respondents in the sample were graduates and full-time workers.

The following table gives the percentage of different socio-economic characteristics across the different experiments:

Table 5.21 Percentage of respondents with different socio-economic characteristics across different preference elicitation experiments with the linguistic method

RESPONDENT CHARACTERISTICS	BINARY + ONE STAGE LIKERT	BINARY + TWO STAGE LIKERT
Gender (in percentage)		
Male	46.1	33.3
Female	53.9	66.7
Age		
18 – 25	25.5	22.5
26 – 40	22.5	25.5
41 – 55	33.3	29.4
56 – 75	16.7	19.6
> 75	1.9	2.9
Household Income/month		
Less than 1K	4.9	6.9
1K – 2K	15.7	16.7
2K – 3K	23.5	20.6
3K – 4K	7.8	11.8
4K – 5K	6.9	6.9
More than 5K	9.8	9.8
No answer	31.4	27.4
Education		
Primary	0	0.9
Secondary	21.6	17.6
Graduate	60.8	64.7
Post-graduate	17.6	16.7
Occupation		
Part-time	6.9	4.9
Full-time	57.8	58.8
Unemployed	2.9	3.9
Retired	11.8	14.7
House based	6.9	0.9
Student	13.7	16.7
No answer	0	0
Number of respondents	**102**	**102**

While the number of female respondents is relatively higher across the different experiments, in case of the binary and two stage Likert sample, this proportion is found to be relatively larger. The income, education and occupation distribution across the different experiments follows a similar pattern. Thus, no substantial variation in socio-economic characteristics can be observed on average across the two Likert elicitation methods.

5.5.2.2 Noise characteristics

The perception of respondents to day and night time noise levels as well as the level of annoyance experienced from these is given in the following table:

Table 5.22 Day and night noise source and disturbance level of respondents with the linguistic method

RESPONDENT CHARACTERISTICS	NUMBER OF RESPONDENTS 204	PERCENTAGE
Day noise		
Very noisy	27	13.2
Noisy	52	25.5
Neither	82	40.2
Quiet	39	19.1
Very quiet	4	1.9
Day noise disturbance		
Very much	37	18.1
Moderate	68	33.3
Little	65	31.9
None	34	16.7
Night noise		
Very noisy	15	7.3
Noisy	44	21.6
Neither	47	23
Quiet	82	40.2
Very quiet	16	7.8
Night noise disturbance		
Very much	34	16.7
Moderate	44	21.6
Little	54	26.5
None	72	35.3
Measure		
Self	70	34.3
Non-self/ none	134	65.7

From the above table it is seen that most respondents indicate the day time noise level to be ‘neither noisy nor quiet’ while the level of disturbance experienced from this level is ‘moderate’. Most respondents consider the night time noise level to be ‘quiet’ with most showing no disturbance from the night time noise level.

The causes for noise in terms of rank are given in Table 5.23.

Table 5.23 Ranking of noise causes of respondents with the linguistic method

	NOISE CAUSE				
Rank	Road traffic	Aircraft	Neighbour	Construction	Others
1	73.5%	2.9%	7.8%	10.8%	3.9%
2	8.8%	9.8%	16.2%	15.7%	7.3%
3	1.9%	1.5%	3.9%	9.3%	4.9%

It is seen that most respondents again attribute the top cause of noise in the apartment to road traffic while the least cause in the ranking is given to construction activities. Thus, noise from road traffic can be attributed to be regarded as the major noise cause by the residents in the area as this finding has been obtained in case of both the location and the linguistic representation surveys.

5.5.2.3 Preference Levels

The following table outlines the preference levels obtained from the linguistic method across the one and two stage Likert elicitation methods. It can be observed that with the one stage Likert method, more respondents choose ‘Definitely’ options over ‘Probably’, while in the case of the two stage Likert method, a similar pattern is followed with higher number of respondents choosing ‘Absolutely certain’ alternatives over ‘Not so certain’ alternatives.

Table 5.24 Ranking of noise causes of respondents in Linguistic method

RESPONDENT CHARACTERISTICS	PERCENTAGE
One stage Likert (total observations: 816)	
Definitely A	33.7
Probably A	22.7
Uncertain	4.5
Probably B	18.0
Definitely B	21.1
Two stage Likert (total observations: 816)	
Absolutely certain A	48.0
Not so certain A	13.7
Not so certain B	8.9
Absolutely certain B	29.3

5.5.3 Comparative Statistics and Sample Representativeness

Sections 5.5.1.1 and 5.5.2.1 gave the descriptive statistics of the socio-economic characteristics of the two phases at the aggregate level. This section briefly combines the information in the previous sections in order to give a comparison of the socio-economic characteristics of the respondents across the two phases. A comparison of gender, age and income characteristics across the two phases is given in Figure 5.3 while that for education and occupation is given in Figure 5.4. Figure 5.5 gives a comparison of respondents' perception and annoyance from day and night-time noise across the two phases of the survey.

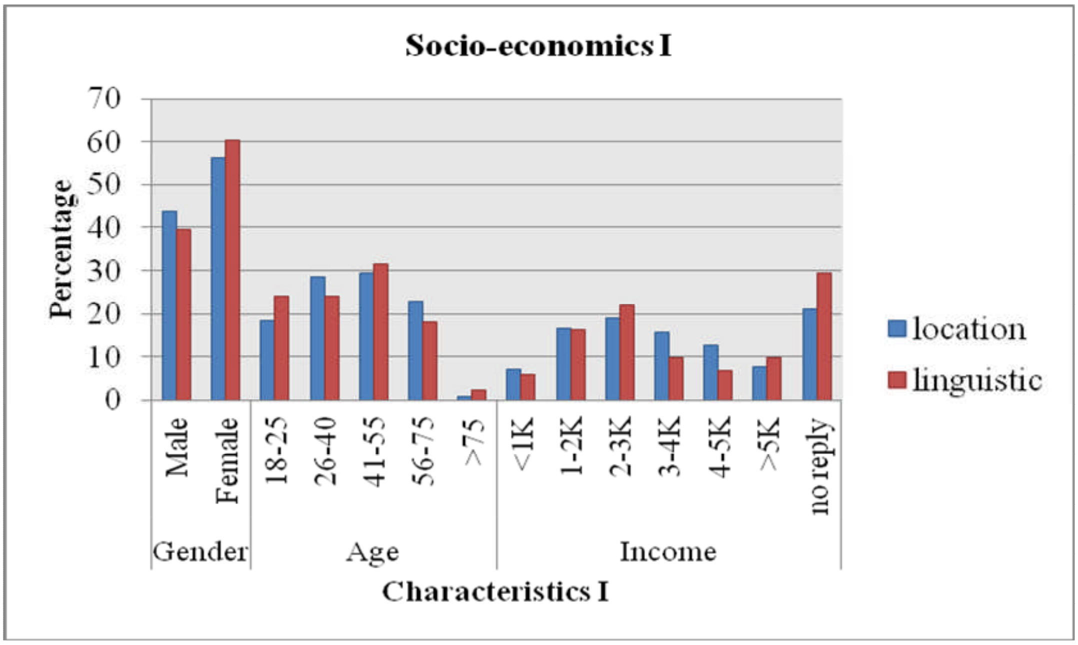


Figure 5.3 Gender, Age and Income characteristics across location and linguistic representation methods

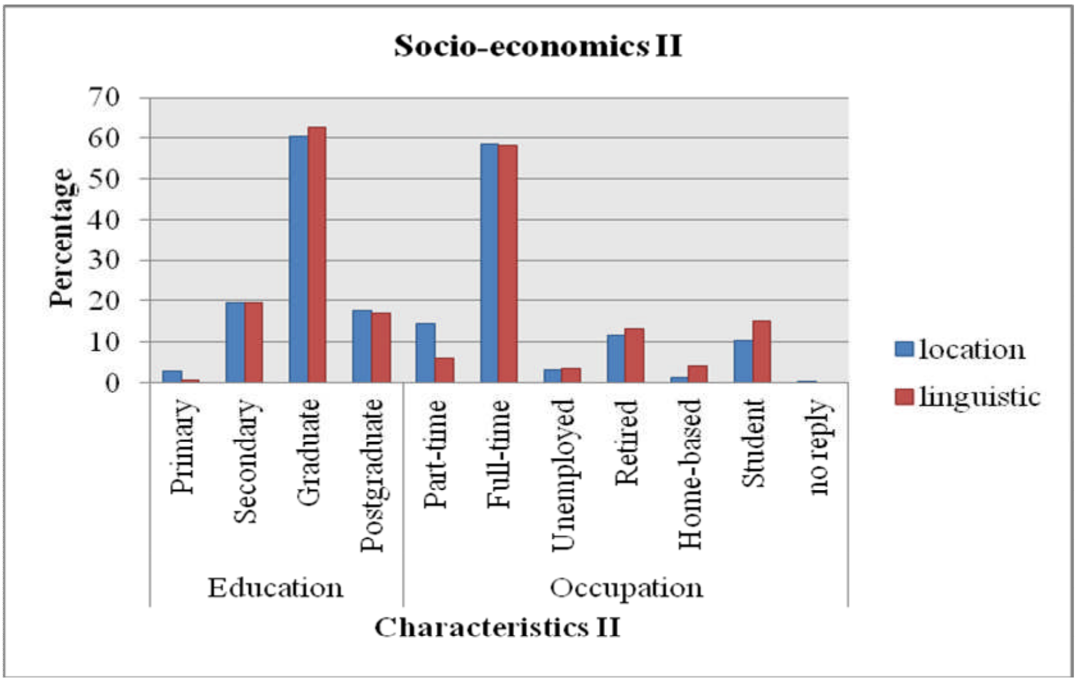


Figure 5.4 Education and Occupation characteristics across location and linguistic representation methods

For the figures given it is seen that there is slight difference in the number of respondents in terms of gender across the two phases (location and linguistic representation methods), with more female respondents in the second phase of the survey. Figure 5.3 also indicate that in both the phases, most respondents were found to be in the age group of 41 – 55 and in the income category of Euro 2000-3000 per month. In case of the linguistic representation method, it is also seen that quite a significant proportion of respondents are unwilling to answer the income question. The education and occupation characteristics across the two phases show quite similar number of respondents except that there are slightly more graduates, home-based people and students in the second phase while there are slight more part-time workers in the first phase of the survey.

The noise characteristics and perceptions across the two phases of the survey indicate that in both the phases of the survey, higher number of people considered daytime noise level to be ‘neither noisy nor quiet’. Most people in the first phase indicated little annoyance by noise levels experienced during the daytime while in the second phase, most people claimed to be moderately annoyed by noise levels. In the first phase of the survey, it can be seen that almost equal number of people considered night-time noise level to be 'noisy' and 'quiet' while majority considered it to be in the category 'neither noisy nor quiet'. For the second phase, most people drastically claimed to consider night-time noise level to be 'quiet', with no annoyance from night-time noise levels.

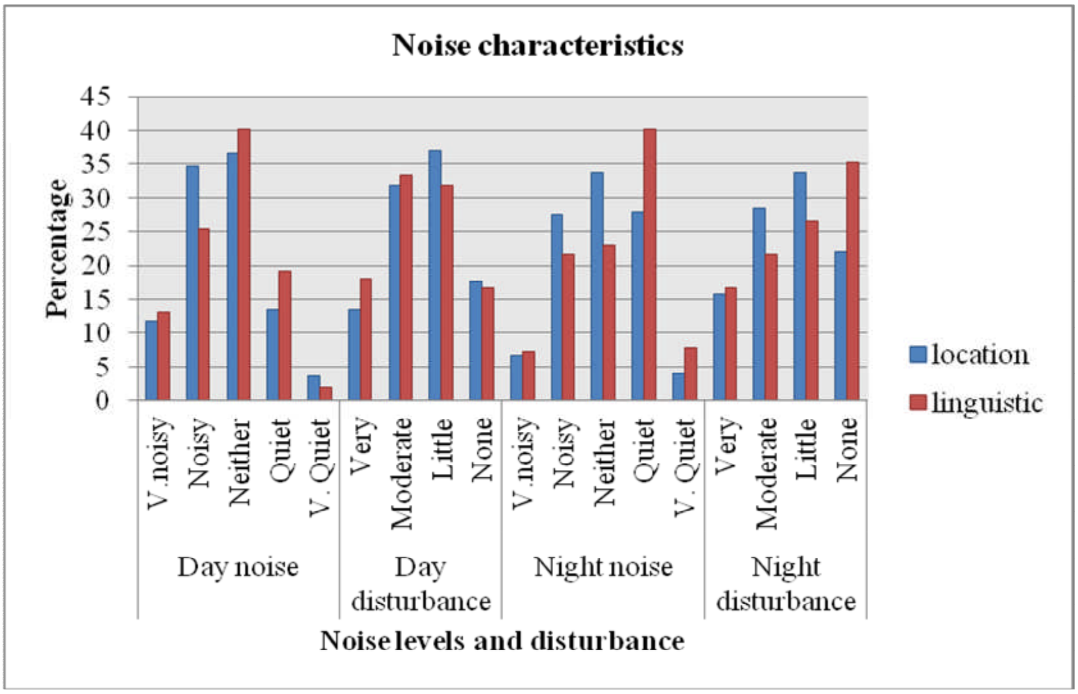


Figure 5.5 Noise characteristics and perceptions of respondents across the two phases of the survey

Comparing the preference certainty level indicated from the one and two stage Likert techniques across the two representation methods, it can be observed that in case of the location method with one stage Likert elicitation, respondents tend to chose the 'probably' options more compared to the 'definitely' alternatives while in case of the two stage Likert method, extreme preference certainty levels tend to predominate. For the linguistic representation method however, it is observed that respondents tend to select the higher preference certainty levels over the lower certainty levels for both the different elicitation methods.

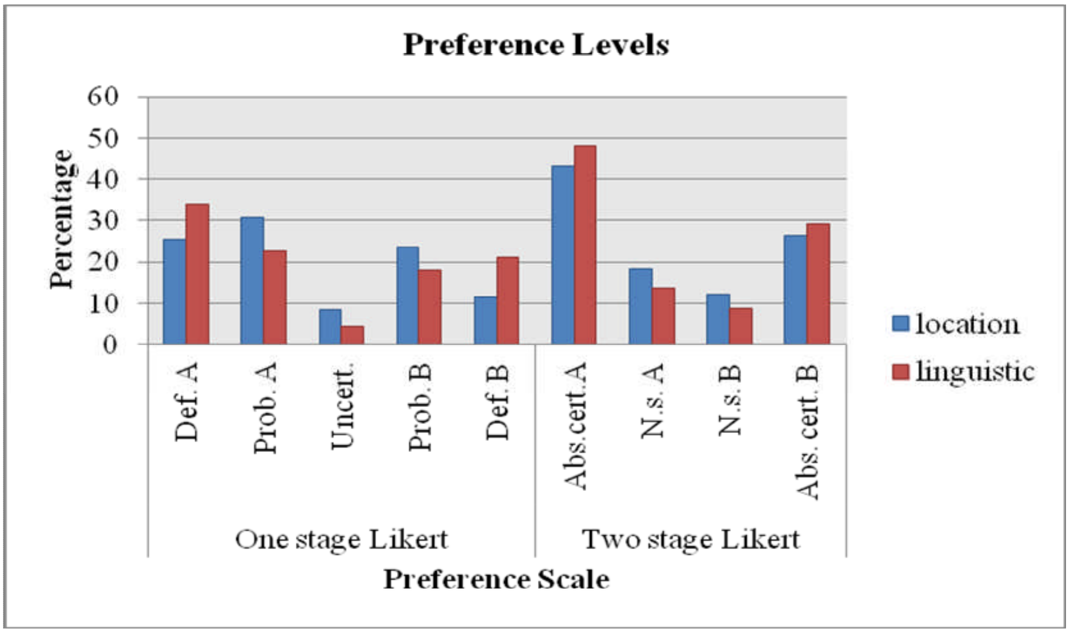


Figure 5.6 Preference levels obtained across the two elicitation and representation methods

While the main aim of the analyses conducted in this research was to examine the effect of different representation and elicitation methods on choice and the representativeness of the sample compared to Lisbon's population was not a specific focus of the study, it can be observed that the sample is well-represented across the different categories of the socio-economic characteristics and sufficient level of diversity is obtained across the sample for different socio-economic variables. Moreover, in order to examine any variation in the socio-economic characteristics observed across the current study and that conducted by Arsenio (2002), the following table provides a summary descriptive statistics of the various socio-economic characteristics.

Table 5.25 Socio-economic characteristics across different studies and experimental phases

RESPONDENT CHARACTERISTICS	ARSENIO (2002) %	LOCATION (CURRENT) %	LINGUISTIC (CURRENT) %
Gender			
Male	37.6	43.7	39.7
Female	62.4	56.3	60.3
Age			
18 – 25	9.5	18.5	24
26 – 40	36.2	28.4	24
41 – 55	36.4	29.3	31.4
56 – 75	~8.4	22.9	18.1
> 75	0	0.9	2.4
No answer	9.5	0	0
Household Income €/month			
Less than 1K	~1	7.2	5.9
1K – 2K	3.2	16.7	16.2
2K – 3K	8.7	18.9	22
3K – 4K	14.1	15.8	9.8
4K – 5K	14.8	12.6	6.9
More than 5K	38.4	7.6	9.8
No answer	19.9	21.2	29.4
Education			
Primary	1.9	2.7	0.5
Secondary	17.5	19.4	19.6
Technical	4.9	n/a	n/a
Polytechnic	7.8	n/a	n/a
Graduate	45.4	60.4	62.7
Postgrad. Master level	2.7	17.6	17.1
PhD or equivalent	0.7	n/a	n/a
No information	9.7	0	0
Occupation			
Part-time	~8.4	14.4	5.9
Full-time	68.5	58.5	58.3
Unemployed	~1	3.1	3.4
Retired	~2.5	11.7	13.2
House based	6.8	1.3	3.9
Student	3.4	10.4	15.2
No answer	9.7	0.4	0

Examining the above table, it can be seen that the socio-economic characteristics of the respondents for the both the location and the linguistic method show a similar pattern of variation as compared to that observed in Arsenio (2002). As some of the values for some socio-economic categories were not clearly elucidated in Arsenio (2002), based on the values and other information that were given, an approximate estimate was computed for the missing categories. Moreover, it can be seen that in comparison to Arsenio (2002), fewer education categories were included in the current survey. Thus, in this case, the number of graduates also comprised of respondents with technical and polytechnic degrees while the postgraduate education category in the current study also comprised of doctorate degrees. Taking these factors in to consideration, it can be observed that a similar pattern is observed across the different samples. In case of the income characteristics, it can be observed that for Arsenio (2002), a significant proportion belongs to the higher income category while in the both the phases of the current study, lower-middle income categories (Euro 1000-5000 per month) are more dominant. Compared to Arsenio (2002), it can also be observed that more respondents in the current study and especially in case with the linguistic representation sample are unwilling to divulge information on their household income.

Based on the national statistics of Portugal as on 31st December 2007, it was found that there are 48.05% male and 51.94% female in Lisbon with 16% population up to 14 years of age, 10% belonging to the age group 15-24, 57% in the age group of 25-64 and 17% older than 65 years. The unemployment rate for that period in Lisbon was found to be 8.9% while the mean monthly earning for 2007 was found to be Euro 1,245.3 (Statistics Portugal, 2007; Statistics Portugal, 2008). Comparing these statistics with that obtained from the current study, it is seen that the different age and gender groups are well represented in the current study with a slightly higher proportion of female respondents found for the linguistic representation sample. Compared to the national statistics information, it can also be seen that the unemployment rate in the area selected (as reflected in the data collected for both the phases) is much lower (about 3.5%) compared to that for the capital (8.9%).

Based on the comparative statistics across the different phases of the current study, it can be concluded that the sample obtained sufficiently represents the different categories while comparison with Arsenio (2002) reveal that a similar pattern of representation is obtained across gender, education and occupation categories.

Comparing with the statistics obtained for the city, it was found that the different age categories are also well-represented in the current study. While lower and middle income categories are more dominant in the current study compared to that found in Arsenio (2002), this could be more indicative of the mean monthly earning (for 2007) found for the capital in the national statistics.

5.6 Summary

The aim of this chapter was to detail the data collection method used as well as provide initial descriptive statistics of the socio-economic and noise characteristics of the respondents. The first section described the type of experiments carried out in the survey based on the attribute representation and choice elicitation methods used. The experimental design used for both the phases in the survey was then provided along with results from the simulation process. It was seen that as levels of view, noise, and sunlight were based on the respondent's perception ratings, the main aim of the simulation was to check the charge levels set for the experiment. For both the phases, the charge levels chosen were acceptable based on the results obtained from the simulation exercise. However for the second phase, non-orthogonal design was created in order to avoid the dominant choice problem. The methods adopted during the survey were also outlined in this chapter and the descriptive statistics for the socio-economic and noise characteristics were given along with an examination of sample representativeness.

6 LOGIT DATA ANALYSIS – I

ATTRIBUTE REPRESENTATION AND ATTRIBUTE UNDERSTANDING

6.1 Introduction

The previous chapter outlined survey methodology employed and the different choice experiments conducted based on the different methods of attribute representation and preference elicitation. This chapter gives the results obtained from the logit analysis of each of the different choice experiment data. The data from the different choice elicitation methods (binary, one stage Likert and two stage Likert) are pooled for each of the different attribute representation methods (location and linguistic representations). Thus, the pooled binary-multinomial model forms the primary model of this chapter.

The chapter will begin with a brief recollection of the preference elicitation methods used along with the research hypothesis. This along with the structure and rationale for the applied methods is given in Section 6.2 while Section 6.3 provides the model results along with the associated attribute valuations.

6.2 Research Hypotheses and Model Structure

The effect of different methods of attribute representation on attribute understanding will be examined in this chapter along with the analysis of different preference data. Towards that aim, this section provides the outline and rationale for each analytical model used in this chapter in Section 6.2.1 while the relevant research hypotheses are recollected in Section 6.1.1.

6.1.1 Research hypothesis

The research hypotheses have been outlined in Section 2.6. In light of the analytical models discussed in this chapter, the research hypotheses pertaining to this chapter can be recollected as follows:

Hypothesis 1a: The location and linguistic representation methods can have varying effects on the understanding of the different attributes. In case of noise, linguistic representation method could be clearer for respondents to understand than the location method

Hypothesis 3: Different methods of preference elicitation can have varying capabilities to capture preference information

One and two stage Likert elicitation techniques are used to capture respondent's preference level. This hypothesis tests whether either of the preference elicitation method is better than another. This hypothesis will be examined through the scale parameter for each of the elicitation methods.

6.2.1 Model Structure and Rationale

In order to test the hypotheses, several analytical models have been developed. The primary model, which accounts for panel effects, is a pooled model for the binary, one and two stage Likert elicitation methods. In this case, the binary data is analysed using a binary logit model while the one and two stage Likert data are analysed using a multinomial logit (MNL) model. The MNL model treats each preference level on the Likert scale as a separate alternative. Thus, in the case of the one stage Likert data, five preference levels are obtained and are treated as separate alternatives (Definitely A, Probably A, Uncertain, Probably B and Definitely B) while in the case of the two stage Likert data, four preference levels are obtained and are treated as distinct alternatives (absolutely (Abs.) certain A, not so (N.s.) certain A, Abs. certain B and N.s.certain B).

In the pooled BL-MNL model, the attribute coefficients are held common across the different elicitation methods while different scale parameters across the elicitation methods are estimated. Based on the estimated scale parameters, one can thus evaluate the effect of different preference elicitation techniques. By fixing each of the scale parameters to unity and comparing the model results obtained, the reference scale parameter was thus chosen for the final model. Thus, in case of the location ratings, dummy and linguistic dummy specifications, the scale parameter for ‘one stage Likert’ (reference scale parameter) is held at unity while in the case of the linguistic ratings model, the scale parameter of ‘binary’ elicitation is fixed at unity.

For each of the models developed, the level of the attributes in the utility function is incorporated into the model using the attribute ratings method and the dummy specification method. As explained in Section 5.4 of the previous chapter, respondents were asked to give a perception rating from 0 – 100 (very bad – very good) for view, noise and sunlight attributes. The average ratings obtained for each of the attribute levels across the different representation methods is given in Table 6.3.

The numeric rating obtained from the perception rating exercise is used in the attribute ratings model as the data input method. For both the attribute ratings as well as the dummy data input method, the ‘housing service charge’ is in the units of Euro. In case of the dummy specification method, the dummy categorical level of each of the attributes (except charge) is incorporated into the model. The number of dummy levels for each of the attributes varied based on the method of attribute representation. For the location method, the number of dummy levels for view, noise and sunlight were fixed to four, based on the number of apartment locations. Thus, the following four dummy levels are observed for the location method:

Table 6.1 Attribute dummy levels with location method

	Level 1	**Level 2**	**Level 3**	**Level 4**
View, Noise, Sunlight	6F	6T	3F	3T

In the case of the linguistic representation method, the number of levels for each of the attributes varied and in this case, the following levels were incorporated into the model:

Table 6.2 Attribute dummy levels with linguistic method

	Level 1	**Level 2**	**Level 3**
View	good	neither	
Noise	noisy	neither	quiet
Sunlight	v.good	good	neither

The following average ratings were obtained for each of the attribute levels across the different representation methods:

Table 6.3 Average ratings for attribute levels across different representation methods

Location Attributes	**Location Ratings**	**Linguistic Attributes**	**Linguistic Ratings**
View		**View**	
6F	59.57	Good	64.29
6T	59.68	Neither	44.93
3F	46.60		
3T	51.53		
Noise		**Noise**	
6F	30.70	Noisy	25.96
6T	50.58	Neither	42.47
3F	29.39	Quiet	60.74
3T	48.80		
Sunlight		**Sunlight**	
6F	70.49	Very good	86.17
6T	66.24	Good	69.56
3F	63.77	Neither	50.00
3T	59.66		

For each of the model structure created, results from both the ratings and dummy input methods are reported.

For the binary choice, the general utility expression for a linear in parameters model with attribute ratings can be given as:

$$U_A = \alpha VA + \beta NA + \gamma SA + \eta CA + ASC_A$$
$$U_B = \alpha VB + \beta NB + \gamma SB + \eta CB$$

where,

VA, NA, SA and CA are view, noise, sunlight and housing service charge for option A;

VB, NB, SB and CB are view, noise, sunlight and housing service charge for option B and,
ASC is the alternative specific constant

The utility function using the attribute dummy level is constructed by omitting one level as the 'reference case' for each attribute. In case of the location method thus, the utility function using the dummy level specification can be given as follows where the fourth level of the 'view', 'noise' and 'sunlight' attributes is fixed as the reference level:

$$U_A = \alpha_1 VA_1 + \alpha_2 VA_2 + \alpha_3 VA_3 + \beta_1 NA_1 + \beta_2 NA_2 + \beta_3 NA_3 + \gamma_1 SA_1 + \gamma_2 SA_2 + \gamma_3 SA_3 + \eta CA + ASC_A$$
$$U_B = \alpha_1 VB_1 + \alpha_2 VB_2 + \alpha_3 VB_3 + \beta_1 NB_1 + \beta_2 NB_2 + \beta_3 NB_3 + \gamma_1 SB_1 + \gamma_2 SB_2 + \gamma_3 SB_3 + \eta CB$$

As the number of attribute dummy levels in the case of the linguistic representation is lesser than that in the location method, it should be noted that the number of levels in the corresponding utility functions will also be less compared to the location method. Thus, in this case, there is one dummy level for view while two dummy levels for noise and sunlight.

The one stage, five point Likert choice elicitation was carried out by offering the respondents the following range of preference levels to indicate their choice of alternative – Definitely A, Probably A, Uncertain, Probably B and Definitely B.

The general utility expression for the MNL model using attribute ratings can be given as:

$$U_{DA} = \alpha VA + \beta NA + \gamma SA + \eta CA + ASC_{DA}$$
$$U_{PA} = \alpha VA + \beta NA + \gamma SA + \eta CA + ASC_{PA}$$
$$U_U = ASC_U$$
$$U_{PB} = \alpha VB + \beta NB + \gamma SB + \eta CB + ASC_{PB}$$
$$U_{DB} = \alpha VB + \beta NB + \gamma SB + \eta CB$$

Where,

DA = definitely choose A

PA = probably choose A

U = uncertain

PB = probably choose B

DB = definitely choose B

As only difference in the utilities of the alternatives matter for choice (Train, 2003), the utility for the 'uncertain' alternative could be stated as a difference between the utilities of alternative A and alternative B. However, as in the case of the linguistic representation method, the orthogonal design was compensated in order to avoid the dominant choice problem, the difference specification cannot be employed due to non-orthogonality conditions. Moreover, the constant associated with the 'uncertain' alternative is set to capture any effects associated with the choice of the alternative, without incorporating the utility difference design in modelling, thus this specification was adapted for both the location and the linguistic methods.

The utility expression for the dummy variables approach can be given as the following (it is to be noted here that the following equations are for the utility expression for the location method and the utility expressions for the linguistic representation method will vary based on the linguistic dummy levels provided in Table 6.2, where (n-1) levels will be incorporated in the equation for each of the attributes):

$$U_{DA} = \alpha_1 VA_1 + \alpha_2 VA_2 + \alpha_3 VA_3 + \beta_1 NA_1 + \beta_2 NA_2 + \beta_3 NA_3 + \gamma_1 SA_1 + \gamma_2 SA_2 + \gamma_3 SA_3 + \eta CA + ASC_{DA}$$
$$U_{PA} = \alpha_1 VA_1 + \alpha_2 VA_2 + \alpha_3 VA_3 + \beta_1 NA_1 + \beta_2 NA_2 + \beta_3 NA_3 + \gamma_1 SA_1 + \gamma_2 SA_2 + \gamma_3 SA_3 + \eta CA + ASC_{PA}$$
$$U_U = ASC_U$$
$$U_{PB} = \alpha_1 VB_1 + \alpha_2 VB_2 + \alpha_3 VB_3 + \beta_1 NB_1 + \beta_2 NB_2 + \beta_3 NB_3 + \gamma_1 SB_1 + \gamma_2 SB_2 + \gamma_3 SB_3 + \eta CB + ASC_{PB}$$
$$U_{DB} = \alpha_1 VB_1 + \alpha_2 VB_2 + \alpha_3 VB_3 + \beta_1 NB_1 + \beta_2 NB_2 + \beta_3 NB_3 + \gamma_1 SB_1 + \gamma_2 SB_2 + \gamma_3 SB_3 + \eta CB$$

The two stage Likert choice elicitation was carried out by asking respondents to choose between options A and B and then indicate their level of choice certainty (by indicating whether they have 'absolute certainty' or 'have some doubts' in their choice). Each level of the preference certainty was again treated as a separate alternative. A full set (n-1) of ASCs were used in the model building with the following utility functions for the ratings model:

$$U_{AcA} = \alpha VA + \beta NA + \gamma SA + \eta CA + ASC_{AcA}$$
$$U_{nsA} = \alpha VA + \beta NA + \gamma SA + \eta CA + ASC_{nsA}$$
$$U_{nsB} = \alpha VB + \beta NB + \gamma SB + \eta CB + ASC_{nsB}$$
$$U_{AcB} = \alpha VB + \beta NB + \gamma SB + \eta CB$$

The ASCs capture any other effects affecting the alternative and in this case are used to examine additional effects related to each preference level. It is hypothesised that in case where less randomness is associated with choice, more plausible ASC values will be obtained across the different preference levels. Thus the ASC values are used to examine additional effects associated with the choice of the alternative as well as the level of randomness involved in choice across the different preference elicitation and attribute representation methods.

For the sake of simplicity in the above utility models, 'Abs. certain' A is denoted by AcA and 'N.s.certain' A is denoted by nsA. The same procedure is followed for Option B.

The utility functions for the dummy model followed closely to that developed for the MNL model with one stage Likert choice data, with the exception on the number of utility functions developed.

6.3 Model Results

6.3.1 Pooled Binary-Multinomial Logit Model

Based on the utility specifications provided in Section 6.2.1, a pooled parsimonious BL-MNL model was developed where the attribute coefficients were held constant across the different preference elicitation methods but a scale parameter was estimated for two of the three different elicitation methods, with the scale parameter for one of them held constant (at unity). This approach allowed investigating whether different preference elicitation methods captured different preference information and had varying levels of variance associated with them while allowing for a more parsimonious specification. The choice of the elicitation method for the fixed scale parameter depended on the overall model results. While the scale parameter can be fixed arbitrarily, it was found that other scale parameters did not have any estimation problem when the lowest scale parameter was fixed at unity. Thus, several models were estimated with alternative specifications of the scale parameter in order to find the best specification structure.

In order to account for panel effects, the error component specification was used in BIOGEME 2.0 (Bierlaire, 2003; Bierlaire 2008; Yanez *et al.*, 2010) where an error component was added to (n-1) alternatives.

A panel data set comprises of a series of repeated observations from the same unit (individuals, households or firms) over a number of periods. Though the availability of repeated observations on the same unit allows for more complicated and realistic models, the repeated nature of the data implies that it is no longer appropriate to assume that the different observations are independent. In comparison to time-series or cross-sectional data, an important advantage with panel data lies in the possibility to estimate certain parameters without making restrictive assumptions. However, given that the repeated observations are obtained from the same individuals, it is unrealistic to assume that the error terms over different time periods would be uncorrelated (Verbeek, 2008).

In case of the discrete choice model where repeated observations from an individual are obtained (panel data), a correlation of the disturbances (serial correlation) or heterogeneity due to the variations in the unobserved effects across individuals is observed (Abdel-Aty *et al.*, 1997). While this panel data offers certain advantages over cross-sectional observations from the same individual as it allows for more accurate measurements (Yanez *et al.*, 2010), the repeated measurement data introduces correlation of the unobserved terms which needs to be explicitly treated.

For a standard linear regression of the form: $y_{it} = \beta_0 + x'_{it}\,\beta + \varepsilon_{it}$ where x_{it} is a K-dimensional vector of explanatory variables and which excludes the intercept term and *i* is an index of all individuals (*i* = 1, . . ., N) and *t* is an index of all time periods (*t* = 1, . . ., N), the model implies that the intercept β_0 and the coefficient β are identical for all individuals and time periods while the error term varies over individuals and time periods and captures all the factors that affects y_{it}. A typical panel data model then assumes: $\varepsilon_{it} = \alpha_i + u_{it}$

where, u_{it} is assumed to be homoskedastic and not correlated over time while the component α_i is time invariant and homoskedastic over individuals. This model is referred as the error components or **random effects** model. In a **fixed effects** model, this problem is addressed by including an individual specific constant term in the model which is estimated along with other regressors of the model. In this case, the model is given as:

$$y_{it} = \alpha_i + x'_{it}\,\beta + u_{it}$$

where, α_i (*i* = 1, . . ., N) are fixed unknown constants estimated in the model and u_{it} is the error term assumed to be i.i.d. over individuals and time periods. The overall intercept term is omitted and is replaced by the individual specific constants, which are referred as the fixed (individual) effects. These effects capture all unobservable time-invariant differences across individuals. Most panel data models are estimated either using fixed or random effects model.

Factors that affect the dependent variable but which have not been included as the explanatory variable can be appropriately summarised by a random error term. In this case, it leads to the assumption that α_i are random factors, i.i.d. over individuals. The error component thus comprises of two parts: an individual specific component which does not vary over time and a remainder component which is assumed to be uncorrelated over time. Thus, all correlation of error terms over time is attributed to α_i (Verbeek, 2008).

In application of panel analysis methods for SP data, it can be seen that the implementation of the fixed effects method can be complicated and unrealistic as it requires an estimation of *n* individual-specific constants. This leads to the incidental parameter problem where the number of estimated parameters increases with the sample size (Verbeek, 2008). In order to avoid this problem, the random effects model is more frequently applied in case of the SP data to account for the serial correlation and heterogeneity observed.

The random effects model can either take the form of the random parameters logit (RPL) model or the error components logit (ECL) model. In case of the RPL model, the coefficient vector β_n is the coefficient associated with individual *n*, representing that person's tastes which vary across different individuals and the density of this distribution is given by θ, which represent *the population parameters that describe the distribution of the individual parameters* (Revelt and Train, 1998). Under the panel analysis framework, this implies that the correlation among observations obtained from an individual causes correlation among taste parameters which can be captured using the RPL method and where the variance of the taste parameters reflects inter-respondent heterogeneity caused due to the panel effect.

Under the ECL framework, an error component is introduced in the model which accounts for correlation obtained across observations from an individual (Abdel-Aty *et al.*, 1997; Mabit *et al.*, 2008; Yanez *et al.*, 2010). For each of the utility functions, if the stochastic utility is taken to be the sum of the deterministic utility and an associated random error which can be further decomposed into the form $\varepsilon_{it} = \alpha_i + u_{it}$, where α_i captures the correlation over individuals (i.e., panel effect), and u_{it} is i.i.d. Gumbel distributed error term, then *n* alternative specific error terms for the panel effect can be specified. However the estimation of *n* error terms can cause identifiability issues and hence this procedure requires estimating *(n-1)* error variances (Yanez *et al.*, 2010) where one procedure to identify the reference alternative is by estimating all error components and holding the alternative with the lowest value of error variance along with the associated low t-statistics, as the reference alternative (Walker *et al.*, 2007). If a common error variance is used to capture panel effects across *(n-1)* alternatives, a correlation among these alternatives is induced and hence alternative specific *(n-1)* error variances are estimated in this study to capture the panel effects for each of the alternatives. While the RPL approach can also be applied to account for panel effects, this approach was not undertaken as it would imply estimating the variances for each of the coefficients, over each of the alternatives, thus substantially increasing the number of estimated parameters in comparison to the ECL method.

As given in Section 4.3.2.3 of Chapter 4, the estimation of the ECL or RPL model requires drawing parameters from a density. Several types of draws are available as has been outlined in the Section and the Modified Latin Hypercube Sampling (MLHS) procedure as developed by Hess *et al.* (2006) has been used to generate the draws.

Using the error component specification to account for panel effects in BIOGEME (Bierlaire, 2003; Bierlaire 2008; Yanez *et al.*, 2010), an error component was added to *(n-1)* alternatives. In order to select the base alternative for the panel specification, a model with all error components and 500[19] MLHS draws was estimated for all the different model specifications. The alternative with the least error component variance and a low associated statistical significance was selected as the base alternative. In case of location ratings 'Alternative A' was selected as the base alternative while in case of the location dummy model, the base alternative for the panel specification was 'Alternative B' of the binary elicitation method. For the linguistic ratings and linguistic dummy models, the base alternative was 'Abs. Certain A' of the two stage Likert elicitation method.

Compared with a model with different parameter estimates for each of the elicitation methods but with a common charge estimate across the elicitation methods (while other parameters are allowed to freely vary), it was found that without considering the panel effects, no significant improvement was observed in the model fit compared to the common parameters model, thus paving way for the

[19] Both MLHS as well as Halton draws were experimented with 500 and 1000 draws. It was found that MLHS draws were largely efficient with stable estimation obtained from 500 draws. The use of 1000 draws increased the estimation time without any substantial improvement in the model fit.

implementation of a more parsimonious model where each of the attribute coefficients are held to be same across the different elicitation methods while the scale parameters for the different elicitation methods (binary and two stage Likert, in relation to one stage Likert) are estimated. The following table thus provides the results of the pooled BL-MNL parsimonious model considering panel effects. It is to be noted that while the t-statistics for each of the parameters, ASC estimates and error variances are with respect to zero, those for the scale parameters are with respect to one. The panel error component obtained for each of the alternatives across the different model specifications, by considering the panel effects is provided in Table 6.5 and discussed subsequently thereafter.

The following results are obtained from the pooled BL-MNL panel models:

Table 6.4 Binary-Likert pooled MNL model for Location and Linguistic data with specific availability conditions and panel effects, with each variable level common across the elicitation methods

Location attributes	Location ratings	Location dummy	Linguistic attributes	Linguistic ratings	Linguistic dummy
ASC			**ASC**		
Binary			**Binary**		
A	0.406 (5.28)	0.820 (4.80)	A	0.319 (4.94)	0.184 (1.71)
B	Fixed	Fixed	B	Fixed	Fixed
One Stage			**One Stage**		
Definitely A	1.06 (3.27)	1.45 (3.78)	Definitely A	0.319 (2.15)	0.160 (0.95)
Probably A	1.59 (5.64)	2.03 (5.80)	Probably A	-0.12 (-0.79)	-0.28 (-1.56)
Uncertain	1.01 (2.41)	-2.25 (-4.83)	Uncertain	-0.49 (-0.91)	-4.22 (-8.55)
Probably B	1.25 (4.33)	1.37 (4.12)	Probably B	-.075 (-0.55)	-0.07 (-0.52)
Definitely B	Fixed	Fixed	Definitely B	Fixed	Fixed
Two stage			**Two stage**		
Abs. cert. A	0.407 (2.32)	0.865 (3.47)	Abs. cert. A	0.241 (2.53)	0.139 (1.14)
N.s.cert. A	-0.41 (-2.22)	-0.03 (-0.15)	N.s.cert. A	-1.18 (-5.40)	-1.26 (-5.36)
N.s.cert. B	-0.58 (-3.03)	-0.64 (-3.03)	N.s.cert. B	-1.53 (-5.30)	-1.51 (-5.29)
Abs. cert. B	Fixed	Fixed	Abs. cert. B	Fixed	Fixed
View	.0148 (6.29)		**View**	.0172 (4.84)	
6F		0.314 (2.62)	Good		0.596 (5.76)
6T		0.384 (3.26)	Neither		
3F		-0.17 (-1.49)			
3T					
Noise	.0209 (8.52)		**Noise**	.037 (13.46)	
6F		-1.04 (-7.25)	Noisy		-1.97 (-11.1)
6T		-0.68 (-3.49)	Neither		-0.64 (-5.04)
3F		-1.26 (-5.59)	Quiet		
3T					
Sunlight	.0190 (6.48)		**Sunlight**	.0155 (7.57)	
6F		-.197 (-1.71)	Very good		0.824 (7.16)
6T		.379 (2.46)	Good		0.585 (3.62)
3F		-.248 (-2.43)	Neither		

3T					
Charge	-.024 (-9.81)	-.026 (-9.59)	**Charge**	-.02 (-10.84)	-.027 (-9.98)
ρ^2 w.r.t. 0 **adjusted ρ^2**	0.225 0.219	0.213 0.205	**ρ^2 w.r.t. 0** **adj. ρ^2**	0.168 0.161	0.196 0.188
Scale parameter (w.r.t. 1) **Binary** **1 Stage** **2 Stage**	 1.34 (2.22) 1.00 (fixed) 1.25 (1.55)	 1.20 (1.39) 1.00 (fixed) 1.15 (0.95)	**Scale parameter (w.r.t. 1)** **Binary** **1 Stg.** **2 Stg.**	 1.00 (fixed) 1.01 (0.09) 1.10 (0.76)	 1.03 (0.30) 1.00 (fixed) 1.12 (0.90)
FLL	-3012.899	-3061.215	**FLL**	-2974.987	-2875.937
no. of obs.	3552	3552	**no. of obs.**	3264	3264
no. of indiv.	222	222	**no. of indiv.**	204	204
MLHS draws	500	500	**MLHS draws**	500	500

Under the location ratings method, the respondents were asked to give a rating for the attributes ‘view’, ‘noise’ and ‘sunlight’ from 0 (very bad) – 100 (very good) while the location dummy model incorporated the various dummy levels of the apartment into the model. In case of the location representation method, comparing the different model specifications (ratings versus dummy), it is observed that the ratings model provides a better model fit in terms of the final log-likelihood (-3012.899 compared to -3061.215 obtained from the location dummy model) and adjusted ρ^2 values (0.219 versus 0.205). Thus under this attribute representation method it can be concluded that the ratings approach is a more suitable data input method. As the respondents were asked to provide a numeric rating for each of the apartment locations for the attributes ‘view’, ‘noise’ and ‘sunlight’, this finding implies that the ratings provided by the respondents are a better indicator of their choice under this attribute representation method than the dummy location levels.

Prior to the choice exercise with the linguistic representation method too, the respondents were asked to provide a numeric rating (from 0 (very bad) – 100 (very good)) for each of the attribute levels in reference to their block which was used in the ratings specification. Thus, in case of 'view' for example, the respondents were asked to give a numeric rating for 'good' view in terms of their block (i.e., what they perceive as 'good' view for their specific block) on a scale from 0 (very bad) – 100 (very good) where the extremes of the scale represented the worst (0) and the best (100) view that were possible/available from their block. Under this representation method thus it is found that the dummy model specification provides a better statistical fit compared to the ratings model in terms of the final log-likelihood value (-2875.937 versus -2974.987) and the adjusted ρ^2 values (0.188 versus 0.161).

A better model fit for the dummy specification for this representation shows that in this case the respondents' choices are better explained by the dummy linguistic levels. This result was expected as the linguistic representation of the attribute levels was expected to be easier for the respondents to understand. Moreover, a lower statistical fit of the ratings model under this representation method indicates higher variation in ratings provided for the linguistic levels across all the respondents. Thus in this case, though 'good' view is considered better than 'neither good nor bad' view, the numeric ratings obtained for each of these levels could have a higher variation under this representation method, though the absolute differences between the two could be same across different individuals.

For each of the attribute representation and data input methods, the different preference elicitation methods were pooled while the attribute coefficients were fixed for each of the attributes across the different elicitation methods. The difference in the two data input methods (the ratings and the dummy methods) based on the model fit, thus reflects which data input method is a suitable technique for each of the attribute representation methods. In case of the location representation method, improved model fit is obtained when respondents' numeric rating is incorporated in the model while in case of the linguistic representation method, the model fit is found to be better when the data is incorporated as dummy linguistic variables. Thus, the results indicate that for the location representation method, the ratings method is more appropriate as a modelling technique while for the linguistic method, the dummy method is more appropriate.

As the scale parameter in BIOGEME is the inverse of the theoretical scale parameter, which in turn is inversely related to variance, a higher scale parameter value in the table implies a higher associated variance. Comparing the scale parameter across the different models, it is found that in the case of the location ratings method, a value significantly different from 1.0 is found for the binary method which indicates that this elicitation method is significantly different from the one stage Likert method. The scale parameter value for the two stage Likert method does not show a high statistical significance indicating that this elicitation technique does not capture substantially different preference information than the one stage Likert method. However, compared to the binary and two stage Likert elicitation methods, the one stage Likert method is found to capture the respondents' preferences most precisely under the location ratings method. In case of the location dummy model, no significant scale parameter values are obtained for the different elicitation methods. Thus, in this case the different preference elicitation methods do not capture substantially different preference information while in the case of the location ratings model, different preference information is obtained through the binary and Likert elicitation methods.

In case of the linguistic ratings as well as the linguistic dummy specifications, no significant difference is obtained across the different elicitation methods implying that this representation method does not affect preference elicitation and hence each of the different preference elicitation methods is able to capture the respondent's true preferences equally well under this representation method. Comparing the scale parameter values obtained across the different models, it can be concluded that where the continuous granular independent variables are properly modelled as in the case of the location ratings method, the model allows granularity to be captured in preferences when presented as a one-stage Likert choice. In other cases, the preference certainty effect is not captured through the scale parameters.

As the location ratings incorporated the ratings from 0 (very bad) to 100 (very good), a positive sign is obtained for noise coefficient under this modelling approach. In case of the dummy specification however, the parameter estimates are obtained for each dummy level which is independent of the numeric ratings. Thus under this modelling framework negative coefficient estimates are obtained for the dummy levels of 'noise' (as both 'noisy' and 'neither noisy nor quiet' are perceived as negative attributes).

Examining the coefficient estimates obtained along with their associated t-statistics across the location and linguistic ratings method, it can be observed that in line with the research hypothesis 1a, 'noise' coefficient has a higher value and greater statistical significance under the linguistic representation method than that compared to the location method while the 'view' coefficient has a more precise estimate under the location representation method. This result was expected as it was considered that the linguistic representation of 'noise' would be easier for respondents to understand while in the case of 'view', the location method would provide a clearer understanding of the 'view' attribute in comparison to the linguistic method which could result in greater subjectivity as to what constitutes a 'good' view. The 'sunlight' attribute has a higher coefficient value under the location method while the t-statistics for that attribute is slightly lower than that obtained from the linguistic ratings model, implying that this attribute is better represented linguistically.

For the location dummy specification and the 'view' attribute, it can be observed that respondents perceive 'view' in '6F' and '6T' to be better than '3T'. This was expected as height would affect the perception of good view. The value for '3F' is just about insignificant and close to zero which implies that the perception of this attribute does not vary much across the two apartments on the third floor.

In relation to '3T', 'noise' has a significantly high negative value at both '6F' and '3F' which implies that respondents perceive noise on the façade facing the main traffic road to be worse than the opposite façade and this perception applies irrespective of the height of the apartment in relation to the block. The value for 'noise' for '6T' is negative and also statistically significant which interestingly implies that respondents perceive 'noise' on the 6th floor (for the façade not facing the main traffic road) to be worse than that on the 3rd floor of the same façade .

This finding is in contrast to what is obtained for the façade facing the main traffic road where '3F' is perceived to be slightly worse than '6F'. Thus the results indicate that in terms of the 'noise' attribute, the perception of noisier apartment is more dependent on the façade , while in terms of height of the apartment (in relation to its location in the block), the perception varies based on which façade the apartment is located in.

The coefficient estimates for the 'sunlight' attribute indicate that respondents perceive sunlight for '6T' to be better than '3T' while for both '6F' and '3F' this is perceived to be worse. Hence for this attribute it is observed that the side of the façade does have an effect on the relative perceptions.

For linguistic dummy model, it is observed that reasonable signs and values of the coefficient estimates are obtained across all the attribute levels. Comparing the coefficient estimates obtained across the location and linguistic dummy specification, it can be observed that the dummy levels for the linguistic method can represent combined levels from the location method. In case of 'noise' for example, 'noisy' noise can be taken as a combination of levels '3F' and '6F' from the location dummy method while 'good' view can be taken as combined levels '6F' and '6T'. A high t-statistics value for all the coefficient estimates under the linguistic dummy specification implies that these levels are well-understood by the respondents.

Though it can be observed that in case of the location dummy specification, some attribute coefficients do not give a significant value, these levels are not excluded from the estimation procedure as they provide some insight on the effect of different representation and modelling methods on the relative parameter estimates.

The statistical significance of each parameter estimate reveal the level of precision associated with the estimation of that coefficient, and can thus reflect the level of understanding associated with that attribute under each of the representation methods. In case where only 'noise' would be represented with alternative methods (location and linguistic) while holding the representation of 'view' and 'sunlight' to the location method, pure effect of varying 'noise' presentation could have been identified. However, alternative methods of representation were employed for all the three attributes in order to detect the best means of representation for each of these attributes.

The alternative specific constant (ASC) signify the average effect of all unobserved factors that affect choice but are not included in the utility function (Koppelman and Bhat, 2006). In case of the one stage Likert elicitation method, the ASC values are set in relation to the base level ‘Definitely B’ whose ASC is set to zero, while in case of the two stage Likert, this is set in relation to ‘Abs. certain B’. Thus for each level of the Likert scale, the ASC value indicates the preference towards that alternative in relation to ‘Definitely/Abs. certain B’. As the individual specific effects are assumed to be captured through the panel specification, the ASC values are considered to reflect the effect of different choice set characteristics as plausible values of ASC (reflected through their sign and value in relation to the base alternative) indicates that the respondents understand the choice set well and their choices are based more on rational choice and are less random. Thus where the choice set is clearer for respondents to understand, it is expected that the ASCs will have more plausible values while in other cases the contrary is expected.

In case of the one stage Likert elicitation method for the location ratings model it is seen that all values of ASC show a positive sign in relation to the base alternative, while the ‘uncertain’ alternative shows a negative sign for the location dummy specification, the ASC values indicate that there is more randomness associated with choice in case of the location representation and one stage Likert elicitation method. While based on the scale parameter, one stage Likert method is found to capture respondents’ preferences more precisely, the ASCs reveal that the choice of different preference alternatives could be more due to stochastic effect in case of this representation and elicitation method.

With the two stage Likert method for the location ratings model, it can be observed that a plausible ASC values are mostly obtained as ‘N.s.certain B’ is considered worse than ‘Abs. certain B’, while ‘Abs. certain A’ is considered to be better than ‘Abs. certain B’. The sign for ‘N.s.certain A’ in relation to ‘Abs. certain B’ under a rational framework was expected to be positive while the value was expected to be lower than ‘Abs. certain A’. However, in this case, the sign of the ‘N.s.certain A’ alternative is contrary to that expected. For the location dummy model, a similar pattern of ASC values are obtained as that from the location ratings specification except that the sign for the ‘Uncertain’ ASC with the one stage Likert elicitation method is correctly negative.

In case of the linguistic dummy specification, it can be observed that most of the ASC values are insignificant for the one stage Likert elicitation method. However, the ASC value for ‘Uncertain’ alternative under this specification is correctly negative in relation to ‘Definitely B’. With the two stage Likert elicitation method, it can be observed that for all the model specifications, the signs and values of the ASC coefficient consistently reveal that ‘N.s.certain B’ is considered to be worse than ‘Abs. certain B’ while ‘Abs. certain A’ is considered to be better than ‘Abs. certain B’. In case of ‘N.s.certain A’ however, this is perceived to be slightly worse than ‘Abs. certain B’. Thus, the ordering effect is seen to be more maintained under the two stage Likert elicitation method.

Comparing the ASC values across all the model specifications, it is seen that there is a higher level of randomness in case of the location ratings method and the one stage Likert elicitation method, while in the case of the linguistic representation method, this is observed to be less so. However, in case of the linguistic representation method, a lesser dependence on the ASC values is observed for the one stage Likert elicitation, as many of the ASC estimates with this representation method are found to be insignificant. In case of the two stage Likert elicitation method, the ordering of the levels is maintained. Thus, it can be concluded that in case of the linguistic representation, there is lower randomness in choice than found with the location representation and hence there is lesser dependence on stochastic effects to explain choice. In order to test the effect of different attribute representation and elicitation methods on implicit ordering of the Likert scale levels, an ordered logit model will also be subsequently applied while the level of stochasticity associated with each of the representation methods will be further explored using an error components logit model.

Thus to summarise the effect of different response scales and representation methods on scale and coefficients, the following points are pertinent. By holding the coefficients for each of the attributes to be the same across the different representation methods, setting the data availability conditions in the model as required and pooling across the different preference elicitation methods across each of the representation and data input methods, *(n-1)* scale coefficients were estimated to examine the differences across each of the preference elicitation methods. Based on the model structure thus defined, it was found that the scale parameters capture the difference in the preference elicitation methods whereas the difference across the attribute representation methods is given by the variation in the coefficient estimates. The variation in the uncertainty level on the other hand was obtained by the *(n-1)* ASC values estimated for each of the preference elicitation methods.

As the model accounts for panel effects where an alternative specific panel error component (panel sigma) was added to *(n-1)* alternatives, this panel error component can be examined to assess which alternative and/or preference certainty level has a higher panel error component value. The panel error component value captures the level of correlation among observations obtained from an individual and thus also captures heterogeneity/variance across individuals. In case of alternative specific panel sigma specification, this value depicts the level of individual effects on the choice for that alternative. A high and statistically significant panel error component value thus indicates a higher correlation among observations from an individual, implying that the choice for that alternative is more dependent on individual characteristics. For each of the different model specifications, the following alternative specific panel error components were observed:

Table 6.5 Panel variance across different model specifications

Panel Sigma Var. (t-value)	**Location ratings**	**Location dummy**	**Linguistic ratings**	**Linguistic dummy**
Binary A - σ_1	Base	-.833 (-6.85)	.0463 (0.16)	.110 (0.32)
Binary B - σ_2	-.590 (-6.25)	Base	.0271 (0.09)	.085 (0.21)
1 stg. DA - σ_3	-1.55 (-7.25)	-1.61 (-7.60)	.464 (2.74)	.538 (3.48)
1 stg. PA - σ_4	.613 (3.19)	.560 (2.57)	.584 (3.16)	.572 (3.22)
1 stg. Un. - σ_5	-1.46 (-5.74)	-1.37 (-4.91)	1.86 (4.52)	1.53 (4.34)
1 stg. PB - σ_6	.691 (2.87)	-1.02 (-5.41)	.050 (0.15)	.056 (0.16)
1 stg. DB - σ_7	1.65 (5.45)	2.03 (5.74)	.577 (3.17)	.615 (3.53)
2 stg.AcA - σ_8	-.886 (-4.19)	-1.12 (-5.07)	Base	Base
2 stg. nsA - σ_9	-.720 (-3.03)	-.648 (-2.81)	.922 (4.69)	.911 (4.67)
2 stg.AcB - σ_{10}	.938 (4.03)	1.26 (5.08)	-.0322 (-0.13)	-.0438 (-0.13)
2 stg. nsB - σ_{11}	-.521 (-1.93)	0.487 (1.49)	1.11 (4.34)	1.08 (4.31)

Examining the panel error component for each of the elicitation methods across the different representation and data input methods and ignoring its sign as it is of no significance in this case, it can be observed that in the case of the one stage Likert elicitation methods ($\sigma_3 - \sigma_7$) for location ratings as well as location dummy specifications, a higher panel component and statistical significance is obtained for 'Definitely A' and 'Definitely B' alternatives. Hence in case of these preference levels, a higher correlation among intra-respondent observations is obtained resulting in a higher variance across respondents. This implies that these levels of choice certainty have a close relation with respondents' characteristics. In case of the 'Uncertain' alternative as well, a high component and associated t-statistics is obtained implying that some individuals are more uncertain than others. The 'Probably A' and 'Probably B' alternatives do not show a high level of variance (in comparison to the other preference levels in the Likert scale) which implies that the choice of the 'Probably' preference level is less dependent on individual characteristics. Thus, based on the panel error component values obtained it can be concluded that the preference level of ‘Definitely’ and ‘Uncertain’ are more associated with individual characteristics in case of the location representation method while this is less true for the ‘Probably’ preference level.

In case of the two stage Likert method for the location ratings specification, it is observed that the panel values and significance of ‘Abs. certain’ options is higher than that of the ‘Not so certain’ options again indicating that respondent characteristics affect ‘Abs. certain’ levels of preference certainty. Similar results are obtained in case of the two stage Likert location dummy model.

In case of the one stage Likert elicitation across the linguistic ratings and dummy specification it is observed that compared with the ‘Uncertain’ alternative, all other preference certainty levels for the one stage Likert elicitation method have lower values and statistically significance for the panel error components. This result reveals that respondent characteristics affects the choice of the ‘Uncertain’ alternative the most, followed by the ‘Definitely’ alternatives. However, compared to the location ratings method, it can be observed that except for the ‘Uncertain’ alternative, the panel error components in case of the linguistic representation method have a lower value and statistical significance compared to the location method. This result could imply that the choice set is easier for respondents to understand under the linguistic representation method and hence respondent characteristics have a lesser effect on choice.

In case of the two stage Likert elicitation method for the linguistic representation, a lower and statistically insignificant panel component is observed across the ‘Abs. certain’ alternatives while a higher value is observed for the ‘Not so certain’ alternatives. This finding is contrary to that obtained from the location representation method, however it does reveal the effect of correlation among respondents’ observations on the level of preference certainty. The presence of panel effects across the different attribute representation and preference elicitation methods thus implies that this effect needs to be considered for model estimation as significant correlations are found between the repeated observations.

With the BL-MNL model it is seen that we obtain three important parameters besides the attribute coefficients which are the ASC, the scale parameter and the panel error component. While the ASC values reveal whether the ordering of the preferences follow a logical pattern and which representation, elicitation as well as data input method has a higher randomness in choice, the panel error component reveals the amount of correlation in the choice of a particular alternative/preference level with respondents' characteristics while also revealing the amount of heterogeneity found across individuals for each of those alternatives. The scale parameter value on the other hand, examines whether there are substantial differences between the different preference elicitation methods for each of the representation and data input methods and which preference elicitation method has a higher associated variance. Thus, for each of the representation and data input methods, the most suitable preference elicitation method can be selected based on the scale parameter value obtained.

To summarise the findings obtained from the BL-MNL model, it has been observed that comparing across the two representation method, 'view' is more precisely estimated (has a higher statistical significance) in case of the location representation method while 'noise' and 'sunlight' have a higher statistical significance with linguistic representation. Thus, based on these findings, the location method is a better representation technique in case of the 'view' attribute whereas 'noise' and 'sunlight' are better represented linguistically. Comparing the model fit obtained from the different data input methods, it can be concluded that ratings data input is recommended for the location representation method while the dummy data input is better for the linguistic representation. As the one stage Likert elicitation method is found to capture preferences more precisely for the location representation method, while in case of the linguistic method, no significant difference is observed across the different preference elicitation methods, the one stage Likert method is the recommended preference elicitation technique. Thus, in terms of data collection, it is recommended that a mixed attribute representation technique should be adopted with the one stage Likert as the preferred elicitation method.

6.3.2 Segmentation and Interaction Analyses

To examine whether any socio-economic characteristics affects the choice of a particular preference certainty level, segmentation analyses were conducted with different socio-economic characteristics such as age, education, income and occupation, in the utility function for different levels of preference certainty levels in the pooled BL-MNL model. However, across the different attribute representation methods as well as the preference elicitation methods, it was found that most of the different socio-economic characteristics did not have a statistically significant estimate. Thus, across the different representation and elicitation methods, most of the respondents' socio-economic characteristics were not found to significantly affect their preference certainty levels.

In order to examine whether respondents' noise perception and annoyance levels during the day and night affect their noise valuation, an interaction model was estimated with the attribute ratings data input method. As the noise perception and annoyance levels were found to be significantly correlated to each other, individual models for noise perception and noise annoyance interaction was estimated. In case of the location ratings model, noise interaction model with the day and night time noise annoyance levels was found to have a better statistical fit (in terms of the final log-likelihood value) while in case of the linguistic ratings model, the noise interaction model with day and night time noise perception levels was found to perform better statistically. Moreover, significant interaction with noise annoyance was found in case of the location method while significant interactions with noise perception levels was found for the linguistic method. Thus, the following table provides results of the interaction of the 'noise' variable with day and night time noise annoyance level in case of the location ratings model and with the day and night time noise perception level for the linguistic ratings model.

Table 6.6 BL-MNL pooled panel model with noise interaction

Attributes	**Location ratings**	**Linguistic ratings**
ASC		
Binary		
A	0.418 (5.30)	0.297 (4.07)
B	Fixed	Fixed
One Stage		
Definitely A	1.07 (3.27)	0.336 (2.39)
Probably A	1.62 (5.74)	-0.114 (-0.71)
Uncertain	1.08 (2.66)	-0.615 (-1.15)
Probably B	1.27 (4.45)	-0.0749 (-0.55)
Definitely B	Fixed	Fixed
Two stage		
Very certain A	0.463 (2.54)	0.236 (2.53)
Not so cert. A	-0.379 (-2.10)	-1.12 (-4.92)
Not so cert. B	-0.563 (-2.83)	-1.46 (-4.90)
Very certain B	Fixed	Fixed
View	.0151 (6.37)	.0169 (4.59)
Noise	.0156 (3.93)	.00975 (1.33)
Day Noise Perception		
V.noisy		.0242 (2.37)
Noisy		.0132 (1.51)
Neither		.0218 (2.64)
Quiet		.0417 (4.77)
V. quiet		Base
Night Noise Perception		
V.noisy		-.000801 (-0.09)
Noisy		.00736 (0.98)
Neither		.00778 (1.12)
Quiet		.00252 (0.41)
V. quiet		Base
Day Noise Annoyance		
V.much	.0205 (2.48)	
Moderately	.0056 (0.98)	

A little Not at all **Night Noise Annoyance** V.much Moderately A little Not at all	.0039 (0.75) Base -.00708 (-1.06) .00885 (1.63) -.00064 (-0.13) Base	
Sunlight	.0196 (6.59)	.0151 (6.60)
Charge	-.0245 (-10.06)	-.0195 (-8.75)
ρ^2 w.r.t. 0 **adjusted ρ^2**	0.228 0.220	0.174 0.165
Scale parameter (w.r.t. 1) **Binary** **1 Stage** **2 Stage**	 1.32 (2.15) 1.00 (fixed) 1.21 (1.36)	 1.08 (0.58) 1.00 (fixed) 1.15 (0.90)
FLL	-3004.333	-2952.879
no. of obs.	3552	3264
no. of indiv.	222	204
MLHS draws	500	500

Examining the results obtained from the day and night time noise annoyance level with the 'noise' variable in case of the location ratings model, it can be observed that compared to the 'not at all' day time noise annoyance, respondents who regard the day time noise annoyance as 'very much' values noise higher. In case of the other levels of day time noise annoyance ('moderate' and 'a little'), statistically insignificant results were obtained implying that these levels do not have a significant effect on the overall 'noise' variable estimate. In case of the night time noise annoyance for the location ratings model, statistically insignificant coefficient estimates were obtained for all the night time noise annoyance levels in relation to the 'not at all' noise annoyance category. Thus in case of the location representation method, only one significant noise interaction effect (the effect of 'very much' day noise annoyance) was found.

For the linguistic ratings model, ‘noise’ interaction with day and night time noise perception level revealed that in relation with the ‘very quiet’ day time noise perception level, respondents who perceive the day time noise as ‘quiet’ value the ‘noise’ variable higher than other day time noise perception categories. Based on the coefficient estimate and statistical significance obtained for the other day time noise perception levels, it can also be noted that respondents who perceive the day time noise level as ‘very noisy’ and ‘neither noisy nor quiet’ also value ‘noise’ highly in comparison with the base category. In case of the night noise perception, no significant effects were observed across the different perception levels in relation to the ‘very quiet’ noise perception category. Thus with the linguistic representation method, it can be concluded that different day time noise perception levels significantly affect the ‘noise’ coefficient estimate while no effect of night time noise perception categories is found on the ‘noise’ variable.

Across the location and linguistic representation methods, it can thus be concluded that while ‘very much’ day time noise annoyance level significantly affects the ‘noise’ coefficient estimate for the location method, in case of the linguistic representation method, significant values were obtained for different levels of day time noise perception categories thus distributing the valuation of ‘noise’ across these different levels.

6.4 Attribute Valuation

As the utility functions take a linear form, the valuation (willingness to pay - WTP) for each of the attribute can be computed by considering the coefficient estimate of the attribute and the coefficient estimate obtained for ‘charge’. In case of ‘view’ for example, the WTP estimate can be thus given as:

$$E(WTP)_{View} = -\frac{\alpha}{\eta}$$

where, α is the coefficient estimate for view and η is the coefficient estimate for charge.

The estimated variance around the WTP estimate was calculated using the delta method (Langford, 1994). Thus, in case of the 'view' attribute, the estimated variance around the WTP estimate can be given as:

$$Var(WTP)_{View} = \frac{\alpha^2}{\eta^2}\left[\frac{\text{var}(\alpha)}{\alpha^2} + \frac{\text{var}(\eta)}{\eta^2} - \frac{2\,\text{cov}(\alpha,\eta)}{\alpha\eta}\right]$$

While the t-statistics with respect to zero is calculated as:

$$t_{View} = \left(E(WTP)_{View} - 0\right) \Big/ \sqrt{\text{var}(WTP)}_{View}$$

Based on the coefficient estimates obtained for each of the attributes across the different representation and data input methods, the following valuations for each of the attributes along with their associated t-statistics can be given:

Table 6.7 Willingness to pay (in Euro and 2008 prices) for each attribute across the different representation and data input methods

Location attributes	**Location ratings**	**Location dummy**	**Linguistic attributes**	**Linguistic ratings**	**Linguistic dummy**
View	0.62 (5.29)		**View**	0.86 (4.43)	
6F		12.22 (2.52)	**Good**		22.49 (4.99)
6T		14.94 (3.08)	**Neither**		
3F		-6.46 (-1.47)			
3T					
Noise	0.88 (6.43)		**Noise**	1.87 (8.47)	
6F		-40.5 (-5.80)	**Noisy**		-74.3 (-7.44)
6T		-26.6 (-3.28)	**Neither**		-24.1 (-4.49)
3F		-49.0 (-4.83)	**Quiet**		
3T					
Sunlight	0.80 (5.40)		**Sunlight**	0.77 (6.23)	
6F		-7.66 (-1.69)	**Very good**		31.09 (5.82)
6T		14.75 (2.38)	**Good**		22.07 (3.39)
3F		-9.65 (-2.36)	**Neither**		
3T					

Following on from the results obtained in Table 6.4, the above table reveals that in the case of the linguistic ratings model, both ‘view’ and ‘noise’ are valued higher than the location ratings method while ‘sunlight’ has a higher WTP with the location representation method. Based on the ratings method of data input, it can be observed that under both the location and linguistic representations, ‘noise’ has a higher WTP (0.88 Euro per unit change for the location method and 1.87 Euro per unit change for the linguistic method) than ‘view’ and ‘sunlight’.

The ordering of the WTP values for each of the dummy levels under the location and linguistic dummy models reflect the ordering obtained with the attribute coefficient in Table 6.4. Thus, in the case of location dummy model, ‘6T’ is valued higher in case of ‘view’, ‘3F’ shows a higher level of disutility in case of ‘noise’ and ‘6T’ gives a higher WTP for ‘sunlight’. Based on the values obtained, it can be concluded that the apartment on the sixth floor, on the opposite façade is more preferable in terms of ‘view’ and ‘sunlight’ while the ‘noise’ disutility is also lower than the apartment on the façade facing the main traffic road. For the linguistic dummy specification, ‘noisy’ noise shows a higher disutility, while ‘very good’ sunlight shows a higher WTP. The WTP values in this case reveal that highest valuation is obtained for ‘noisy’ noise followed by ‘very good’ sunlight and then ‘good’ view.

In order to obtain comparable estimates across the ratings and the dummy models for each of the attribute representation methods, respondents’ mean ratings for each of the attribute levels was computed from the data[20]. Using the mean ratings obtained for each of the levels, the value for each level in relation to the base level was computed. Thus, in case of ‘view’ for example under the location dummy specification and taking ‘3T’ as the base level and ‘6F’ as the preferred level (whose relative valuation is aimed to be estimated), the difference in ratings between the base level and the preferred level can be computed as:

$$6F_{diff} = 6F_{rat.} - 3T_{rat.}$$

[20] Recall that respondents were asked to provide a rating from 0-100 for each of the attribute levels across the two representation methods.

Where, $6F_{rat}$ is the mean rating obtained for the level '6F' and $3T_{\mathrm{rat}}$ is the mean rating obtained for the level '3T'.

Multiplying the difference obtained with the coefficient estimate for 'view' under the location ratings specification, the value for '6F' ratings in relation to '3T' was obtained. Thus, the estimate for the level can be computed as:

$$Est._{6F} = 6F_{diff} * V_{Est.} \qquad \text{(in relation to 3T)}$$

Where, $V_{Est.}$ is the WTP estimate obtained for 'view' under the location ratings specification.

This valuation can be compared across with the valuation obtained from the location dummy specification where the attribute levels are incorporated into the model through the dummy levels.

In order to illustrate the method employed, let us consider the case of 'view' under the location representation method. The following table shows that the average ratings (μ_{rat}) obtained for the level '6F' was 59.57 while the average ratings obtained for '3T' was 51.53. Taking the difference of the two such that,

$$6F_{diff} = 6F_{rat.} - 3T_{rat.}$$

we obtain the value for $6F_{diff}$ which is 8.04. This value is then multiplied with the WTP estimate obtained for 'view' (0.62) to obtain the value for '6F' in relation to '3T' which in this case is 4.98 Euro. The following table reveals the values obtained for each of the attribute levels:

Table 6.8 Willingness to pay (in Euro and 2008 prices) for each attribute level using the average ratings method

Location Attributes		**Loc. Rat. values**	**Loc. Dum. values**	**Linguistic Attributes**		**Ling. Rat. values**	**Ling. Dum. values**
View		0.62 (5.3)		**View**		0.86 (4.4)	
	$\mu_{rat.}$				$\mu_{rat.}$		
6F	59.57	4.98	12.2 (2.52)	**Good**	64.29	16.65	22.49 (4.9)
6T	59.68	5.05	14.9 (3.08)	**Neither**	44.93	Base	
3F	46.60	-3.06	-6.5 (-1.47)				
3T	51.53	Base					
Noise		0.88 (6.4)		**Noise**		1.87 (8.3)	
	$\mu_{rat.}$				$\mu_{rat.}$		
6F	30.70	-15.93	-40.5 (-5.8)	**Noisy**	25.96	-65.04	-74.3 (-7.4)
6T	50.58	1.57	-26.6 (-3.3)	**Neither**	42.47	-34.16	-24.1 (-4.5)
3F	29.39	-17.08	-49.0 (-4.8)	**Quiet**	60.74	Base	
3T	48.80	Base					
Sunlight		0.80 (5.4)		**Sunlight**		0.77 (6.4)	
	$\mu_{rat.}$				$\mu_{rat.}$		
6F	70.49	8.66	-7.7 (-1.7)	**V.good**	86.17	27.85	31.09 (5.8)
6T	66.24	5.26	14.7 (2.4)	**Good**	69.56	15.06	22.07 (3.4)
3F	63.77	3.29	-9.6 (-2.4)	**Neither**	50.00	Base	
3T	59.66	Base					

Using the average ratings method to compute the relative valuation for each attribute level, it is seen that in the case of both the location as well as the linguistic representation methods, there is some variation in the values obtained through the average ratings and the dummy approach. In case of the location representation method, the difference between the values obtained from the ratings and the dummy methods is more pronounced than that observed in the case of the linguistic representation. Under the location representation method, it is found that in the case of 'view' and 'noise', a higher variation is observed for the level '6T' while in the case of 'sunlight' a higher variation is observed for the level '6F'. This finding reveals that the level ‘6T’ is valued higher for both ‘view’ and ‘noise’ with the dummy input method than is reflected in the respondents’ attribute ratings obtained for that level. However based on the model fit obtained in Table 6.4, the attribute ratings method is a more suitable data input method for the location representation method. Table 6.8 thus indicates that with the dummy input method, the attribute values can be over-estimated under the location representation method. Moreover, the implied perception of the levels can also significantly vary based on the data input method used as can be observed in the case of the level ‘6T’ for ‘noise’ where the ratings model indicates that the traffic noise level for the apartment on the sixth floor on the opposite façade does not cause a significant amount of disutility but values obtained from the dummy model show the contrary.

In case of the linguistic representation, about similar values between attribute ratings and dummy approach are found across ‘view’, ‘noise’ and ‘sunlight’. While the results obtained in Table 6.4 indicate that the dummy method provides a better model fit for the linguistic representation method, it is observed that a relatively lower discrepancy is present in respondents’ perceptions of the attribute levels based on the values obtained with the ratings and the dummy models. Thus, it can be concluded that for the linguistic representation method, the implied perception derived through the ratings and the dummy models are less varied than that obtained from the location method. However, except for ‘neither’ noise, it can be observed that in this case the attribute values can be under-estimated using the average ratings for calculation of WTP for each dummy level. This finding thus indicates that the effect of data input method has a relative significance based on the method of attribute representation adopted.

While discussion on the results obtained in Table 6.4 revealed the effects of different data input methods on the different attribute representation methods in terms of the relative model fit as well as the effect of varying representation methods on attribute valuation, this exercise has revealed that the implied subjective perception of the attribute categories and their associated ratings varies based on the attributes, their levels as well as the method of attribute representation. Hence, the finding emphasises the importance of using the most indicative data input method, which is closely related to the method of attribute representation used, within the sample collected.

6.5 Comparison with Previous Study

To recollect, the current SP survey was conducted in the area where a previous noise valuation study was conducted by Arsenio (2002) using the location representation method. Physical noise measurement was conducted inside and outside the apartments and models were built based on both the noise measurement as well as the location ratings given by the respondents. The respondents were asked to provide a numeric rating on a scale from 0 to 100 for the attributes view, noise and sunlight while the elicitation method employed was a binary choice. The base model in this study took the following form of utility functions using the attribute ratings given by the respondents:

$$U(A) = \alpha VA + \beta NA + \gamma SA + \delta CA$$
$$U(B) = \alpha VB + \beta NB + \gamma SB + \delta CB$$

As this study used only the location representation method and binary choice elicitation method, comparison with the current study can only be made for the binary choice location ratings model.

The following coefficient estimates were obtained for the base model from the previous study as well as the current binary logit model with location ratings specification. It is to be noted that the results reported for the binary logit model for the comparative purpose is without the inclusion of the ASC and hence has a different specification than that applied in 6.3.1. As the results for the 2002 study accounts for repeated observations, it is compared with an equivalent binary logit model with panel specification.

Table 6.9 Comparison of BL results from current and 2002 studies

Attributes	**2002 BL study**	**Current BL study**
View	0.02437 (9.39)	0.0222 (6.52)
Noise	0.03107 (8.40)	0.0326 (9.85)
Sunlight	0.01782 (6.24)	0.0302 (6.97)
Charge	-.00007932 (-2.96)[21]	-0.0312 (-14.57)
ρ^2 w.r.t. 0 **ρ^2 w.r.t. constant**	.149 .088	.245 .241
FLL	-2915.257	-928.897
no. of obs.	4944	1776
Sigma – Alt. B	NA	-.957 (-9.18)
MLHS draws	NA	500
no. of indiv.	NA	222

Comparing the results obtained from the two studies, it can be seen that the current study gives a better fit in terms of the ρ^2 value. Considering the smaller sample size but relatively better t-statistics in the current study, it can be concluded that the results obtained are at least as good as that obtained by Arsenio (2002). In case of the current study, while 'noise' has a higher overall coefficient estimate, the coefficient estimate for the 'charge' attribute is almost double in value compared to the previous study. The reasons for this variation in the charge coefficient as well as its impact on valuation will be discussed subsequently.

As the costs for the 2002 study are expressed in 1999 prices/household/month, the inflation factor was computed in order to compare with the values from the current study where the costs are in 2008 prices. Taking 1997 as the base year, the consumer price index (CPI) by *special aggregates* for 1999 was 105.2 while that for 2008 was computed to be 137 (Statistics Portugal). The following method was thus applied to compute the inflation factor of 1.30

[21] Considering 1 Euro = 200.482 Escudos as given in Arsenio (2002), this value is Euro -0.0159 based on 1999 conversion rate.

Inflation factor (1999) = 2008 CPI value / 1999 CPI value (Defra)

Based on the coefficient estimates obtained across the two studies, the conversion rate to Euro in the 2002 study as well as the inflation factor, the following WTP values (in Euro) can be computed in 2008 prices:

Table 6.10 WTP values for 2002 and current study

Attributes	**2002 Study (in Euro for 2008 prices)**	**Current Study (in Euro for 2008 prices)**
View	1.992	0.711
Noise	2.540	1.045
Sunlight	1.457	0.967

Comparing the values obtained it can be seen that the difference in the attribute WTP across the two studies is lowest for 'sunlight' (difference of Euro 0.49) and highest for 'noise' (with a difference of Euro 1.495). While a high difference is obtained in case of 'noise', it is to be noted that in the current study, the coefficient estimate for the attribute has a higher statistical significance compared to the previous study. Table 6.9 reveals that the 'charge' coefficient in the current study has almost double the value than that in Arsenio (2002) with much a higher significance, which has a bearing in the valuation of the other attributes in the current study. However the statistical significance of the charge coefficient in the current study also implies that this attribute is more precisely estimated than in Arsenio (2002).

Based on the sample's socioeconomic characteristics, the variation in the charge coefficient across the two studies can be largely attributed to the income categories represented across the two samples. While the current study has a larger proportion of individuals (51.4%) belonging to the income categories of Euro 1000-4000 per month, a larger proportion of respondents in the Arsenio (2002) belonged to the higher income category. Table 5.25 reveals the number of respondents in each of the income categories across the previous and the current studies. Based on that table while about 38.4% of respondents belonged to the income category of more than Euro 5000 per month in the previous study, it is important to note that this proportion is estimated based on 1999 prices. Converting the income categories to the 2008 price level by accounting for inflation, it was estimated that about 53.2% of the respondents in the previous study belonged to the income category of more than Euro 5000 per month. This compares with a modest 7.6% of respondents found in that income category for the location method in the current study. Moreover based on the 2008 values and considering the midpoint of each of the income categories[22] as well as the percentage of respondents choosing that category, while the average monthly household income for Arsenio (2002) was computed to be about Euro 4947, this value for the current study under the location representation method was found to be Euro 2449. This variation in the household income level across the two studies can largely explain the variation in the charge coefficient obtained across the two studies. While a marked variation in the number of respondents across the different income categories of the previous and the current studies is obtained, an increased proportion of respondents with higher education levels is obtained in the current study while a large portion of the sample is also in full-time employment. However other factors such as education and employment status are not found to increase the valuations over time.

While this section has only compared the results for the location representation method, results from the linguistic representation method in the current study has revealed that linguistic representation is a more suitable technique for the ‘noise’ attribute. This research thus presents an interesting extension to the work conducted by Arsenio (2002) by examining the effect of different attribute representation and preference elicitation methods.

[22] The mid-point for the income category > Euro 5000 was taken to be Euro 7500 for computation purposes.

6.6 Summary and Conclusions

This chapter aimed to examine whether different methods of attribute representation have different effects on respondents' understanding of the attributes and whether different methods of preference elicitation can have varying capabilities to capture preference information. Besides the examination of the two main hypotheses listed in the chapter, analyses were also conducted to examine the effect of including the panel error component in the model and what it implies for the different certainty levels as well as which is the most preferred apartment location based on the coefficient estimates obtained from the location dummy model.

To address the first hypothesis which examined the effect of different attribute representation methods on respondents' understanding, surveys were conducted with two different attribute representation methods (the location and the linguistic method) and the attribute coefficients as well as their significance were examined to estimate which representation method was more suitable for view, noise and sunlight attributes. As for each of the representation methods, two different data input methods (the attribute ratings and the dummy method) were possible, the suitability of each of these data input methods for the different representation methods were also examined.

Based on the coefficient estimates and their significance levels obtained, it was found that 'noise' has a higher value and a greater significance under the linguistic representation method while 'view' has a more precise estimate for the location method. Thus, the results revealed that 'noise' is easier to understand under linguistic representation while the location method provides a clearer understanding of the 'view' attribute. Thus, the results obtained supported the hypothesis 1a as the location and linguistic representation methods were found to be more suitable for 'view' and 'noise' respectively. Moreover, all the three attributes 'view', 'noise' and 'sunlight' were represented with location and linguistic methods in order to detect the most preferable representation method for each of these attributes, though it could be contested that this technique would not be able to capture pure effects of alternative ‘noise’ representations.

The results also revealed that based on the attribute representation method used, different data input methods were suitable. Thus, in case of the location representation method, the ratings method was a more suitable data input method while in the case of the linguistic representation method, a better FLL was obtained with the dummy method. This finding emphasised the importance of using the right data input technique based on the attribute representation method in order to obtain an optimal model fit in terms of FLL.

Another hypothesis examined in this chapter was the effect of different preference elicitation methods and their varying abilities to capture preference information. By pooling the binary, one stage and two stage Likert models for each of the representation methods and examining the scale parameters, it was observed that in case of the location representation method, the one stage Likert method is the most suitable elicitation method while in case of the linguistic representation no significant difference was found across the different elicitation methods. This finding thus revealed that the linguistic representation method does not affect the method of preference elicitation. Thus the results showed that this hypothesis is supported in case of the location representation method while in case of the linguistic representation method, no substantial difference in preference elicitation abilities was obtained across the different methods.

The panel error component also revealed that different preference certainty levels showed a high correlation level with respondents’ characteristics based on the method of attribute representation and preference elicitation chosen. In case of the location representation method, higher levels of preference certainty as well as the ‘uncertain’ alternative in case of the one stage Likert method showed a high correlation with respondents’ characteristics while in case of the linguistic representation method, lower levels of preference certainty had a high correlation with respondents’ characteristics. Thus, while in case of the location representation method certain individuals were more prone to choose high certainty levels and/or the ‘uncertain’ alternative, in case of the linguistic representation method, this relation to respondents' characteristics was obtained for lower levels of preference certainty. Variation in panel effects was thus observed based on the method of attribute representation used.

In addition to the results obtained for the hypotheses tests, it was also found that in case of both the location as well as the linguistic ratings model, 'noise' had a higher coefficient estimate as well as greater statistical significance, implying that this attribute is valued higher than 'view' and 'sunlight' under both forms of attribute representation. Moreover, with the location dummy model it was found that varying effects of apartment locations (in terms of height and façade) was obtained on the valuation of the different attributes. Thus, the relative height of the apartment was found to affect respondents' perceptions of a good 'view' while 'noise' was perceived to be worse on the façade facing the main traffic road.

In case of the location dummy model, 'view' and 'sunlight' were found to be valued higher for apartment on the sixth floor on the façade not facing the main traffic road ('6T') compared to the other levels while 'noise' was found to have the lowest disutility for this apartment, implying that this apartment location is most preferable objectively in case of the attribute 'view', 'noise' and 'sunlight'.

Conducting segmentation analysis to examine respondents' socio-economic characteristics on preference certainty level, no significant effect was obtained across the different characteristics, preference elicitation and attribute representation methods. In case of the noise interaction analysis, it was found that while some effect of 'very much' day noise annoyance level was found on the 'noise' variable in case of the location ratings method, no other significant effects were obtained for other perception and annoyance categories. In case of the linguistic representation method, relatively higher effects of these were observed on the 'noise' variable albeit with some variability in the type of effect found. However, across the two representation methods it was found that higher levels of day noise annoyance resulted in increased sensitivity for the 'noise' variable with greater value given to the level of quiet.

As part of the current survey was closely followed from that conducted by Arsenio (2002), comparisons were made with the valuations obtained from the two studies which revealed that considering the inflation factor, a higher difference in the WTP estimate across the two studies was obtained in case of ‘noise’ while least difference was obtained for ‘sunlight’, with values obtained from the 2002 study to be higher than the current study. However, despite the difference in the WTP values obtained, it was found that the current study has a significantly better model fit with a smaller sample size, implying that the data gathered in the current study provides a more reliable estimate than that found in the previous study.

Thus, based on the results obtained from the pooled binary-MNL model, it can be concluded that the location representation method is better for ‘view’ while the linguistic representation method is better for ‘noise’ and ‘sunlight’. Hence, in terms of data collection, a mixed representation method could be utilised. As the ratings data input method was found to be more suitable for the location representation method while the dummy method was more favourable for the linguistic representation, these considerations need to be taken into account during the modelling process. Where mixed representation methods are used, the attributes could be incorporated into the model through their numeric ratings where it is represented with the location method while in case of the linguistic representation they could be incorporated through the dummy levels. In case of the preference elicitation method, the one stage Likert elicitation is recommended as it allows for the granularity in location representation data to be captured adequately in the different preference levels while in case of the linguistic representation data, it is found to be no worse than the binary and two stage Likert elicitation methods. This chapter has thus shown that the method of attribute representation affects attribute understanding and valuation. Moreover, the effect of different elicitation methods varies based on the method of attribute representation technique chosen.

7 LOGIT DATA ANALYSIS – II

PREFERENCE UNCERTAINTY AND ERROR STRUCTURE

7.1 Introduction

The previous chapter examined the effects of attribute representation method on attribute understanding and valuation. The preference uncertainty information was also incorporated in the previous chapter through the use of the MNL model. This chapter will further examine the preference uncertainty data through ordered logit (OL) as well through complex error assumptions. The chapter will begin by recollecting the hypothesis in light of the model applied followed by a brief structure and rationale underlying each model in Section 7.2. This will then be followed by the results obtained from the nested logit (NL) and the error components logit (ECL) model in Section 7.4 while the results from the OL model will be given in Section 7.5. Section 7.6 will examine the effect of the NL, ECL and OL models on attribute valuation.

7.2 Research Hypotheses

While the previous chapter mainly dealt with the first and third hypotheses which focussed on the effect of attribute representation method on attribute understanding through the relative significance of the coefficient estimates as well as the ability of different preference elicitation techniques to capture preference information through the scale parameter, the main hypotheses examined in this chapter are the following:

Hypothesis 1b: The type of choice set is expected to have an effect on the level of the respondent's preference uncertainty

This hypothesis is related to the effect that different representation techniques have on respondent's understanding of the choice set and hence their level of preference certainty. It is expected that the linguistic representation method would generally be more comprehensible to the respondents, thus causing more certain choices.

In relation to the model developed in this chapter, the effect of different choice sets on preference uncertainty can be examined by the need to make more complex error assumptions. This is tested through the Likelihood Ratio (LR) test between the higher model forms and the base MNL model. It is expected that the LR test will have a higher and significant estimate in the case of the location method than the linguistic method as the latter method is expected to be less complex for choice set comprehension.

Hypothesis 2: Stated level of preference certainty can be due to deterministic as well as random effects.

This hypothesis focuses on the different causes of the stated preference levels. As respondents are given an opportunity to state their level of preference certainty, it has to be acknowledged that the choice of a particular preference level could be due to several reasons.

Respondents can have different levels of preference certainty which is dependent on deterministic and stochastic factors. Under deterministic cause of preference uncertainty, there is lesser need for complex model structures while in the case of higher stochastic factors affecting the level of certainty, there is a higher need for relaxing the error assumption.

Thus the preference uncertainty associated with a choice set can be due to two factors. First, the level of preference certainty indicated by the respondent can reflect the true level that the respondent is willing to commit to a particular alternative based on the characteristics of the choice set. Another cause of differing level of preference certainty is the respondent's choice uncertainty. This uncertainty can be attributable to various factors such as respondent's inherent indecisiveness, fatigue or boredom effects, lack of understanding of the choice task, other respondent characteristics and choice set characteristics; thus, in this case, the choice set characteristics can result in increased randomness.

In case where the cause of the stated preference levels is more deterministic, there would be lesser dependence on complex error assumption to explain choice. Where the stated preference level is due to actual uncertainty and randomness, there would be higher dependence on complex model forms. This hypothesis is tested through the LR test between MNL and higher model forms.

7.3 Model Structure and Rationale

In order to analyse the preference uncertainty data, two main approaches were undertaken. While the first involved relaxation of error assumption to analyse the main factor (deterministic versus stochastic) for preference uncertainty as well as the structure of error variance based on the level of preference certainty, the other approach involved the application of a commonly used analytical technique for Likert type data, namely the Ordered Logit (OL) model. The rationale underlying each model form can be given as follows:

- Relaxation of the error assumption, which provides insight on whether the chosen choice certainty level is due to the respondent's certain level of commitment or due to other stochastic factors and whether specific choice certainty levels have more/less stochastic factors influencing them than others can be achieved by accounting for correlation and/or heteroskedasticity. Two different models, the NL and the ECL, can thus be applied based on whether correlation or heteroskedasticity is examined. By comparing the model fit obtained from flexible error assumption with that obtained from the MNL model using the likelihood-ratio (LR) test statistic, the relative improvement in specifying a flexible error structure can be estimated. The rationale for each of these models is as follows:

    - The NL model allows for a correlation pattern based on the level of preference certainty. Thus, two alternatives such as 'Definitely A' and 'Definitely B' can be nested together based on their preference certainty level. The rationale for applying the NL model is to examine whether any correlation in the error variance exists based on the similarity of the preference certainty level

- In case of the ECL model, the standard deviation associated with the error values for each of the preference levels reveals the extent of the error variance associated with that preference level. Thus this model can be applied to examine which alternatives and preference certainty levels have a higher error component associated with it

- The purpose of applying the OL method is to implement a commonly used technique for analysing preference uncertainty data. As the one stage Likert data has a natural ordering in its structure, the OL method is commonly applied for its analysis. The application of the OL method here seeks to examine whether this analytical technique is suited to both one and two stage Likert data. The threshold parameters from this model will be examined across the two elicitation methods to inspect whether the ordering of the preferences is maintained while the comparison of the log-likelihood values across the OL and its MNL counterpart will reveal the suitability of the analytical technique for the one and two stage Likert preference data in comparison to the MNL model

7.4 Flexible Error Structures

The MNL model assumes that the error variances in the model are independent and identically Gumbel distributed. However, in the specific context of varying choice certainty levels, there is a possibility that the error variances across similar certainty levels (option Definitely A and option Definitely B, for instance) could have some correlation among them or the error variances of each of the choice certainty levels could be different from each other. Thus, in order to relax this i.i.d. assumption, each of these properties (viz. independent and identical distribution) can be relaxed and any improvement on the log-likelihood (LL) value from the MNL model can be examined.

While the ‘independent’ assumption of the i.i.d. property can be relaxed through the NL model where patterns of correlation are introduced following the nest structure, ECL model allows for added flexibility in error assumption by allowing for both correlation and heteroskedastic error variances within the model structure. However, an important point of distinction between the NL and the ECL correlation patterns lie in the property of the error variances. While in the case of the NL model, the correlated error variances are by definition homoskedastic, this property is not maintained in case of the ECL correlated error variances (Munizaga and Alvarez-Daziano, 2001).

This section incorporates both NL and ECL analyses in order to examine whether the different preference certainty levels have correlated or heteroskedastic error variances. Section 7.4.1 focuses on results obtained from the NL model while Section 7.4.2 provides the ECL output and its discussion.

7.4.1 The Nested Logit (NL) model

The NL model is the first step to accommodate the violations of the IIA assumptions. The possibility of each alternative to contain information in the stochastic portion of the utility and its effect on different choice outcomes can be captured by categorising the alternatives into nests with some level of correlation (Hensher *et al.*, 2005). The nest or scale parameter of the NL model indicates the degree of independence or dissimilarity of alternatives across the nests. Thus, the nest parameter estimated from the NL model can also be used to judge the patterns of correlation that exists between alternatives which are grouped together based on similar error structures. When considering preference uncertainty, the possibility of a specific error structure and error correlation between particular preference certainty levels can be examined through the nest parameter. The theoretic nest parameter lies between 0-1. As the nest parameter increases towards unity, there is little correlation in the error variance among the nested alternatives, resulting in the model to collapse to a MNL. When the nest parameter is equal to null, the model becomes a degenerate case. Thus, for the NL model, the nest parameter needs to be between the specified range, with the value being significantly different from unity (Louviere *et al*., 2000).

In order to test whether any patterns of correlation exists between error variance of the Likert scale levels based on the degree of the respondent's certainty, nest were built on the basis of the certainty levels, while also accounting for panel effects to segregate correlation from intra-respondent observations. Thus, in case of the one stage Likert elicitation method, the alternatives 'Definitely A' and 'Definitely B' were grouped into a nest 'Definitely' while 'Probably A' and 'Probably B' were nested into 'Probably'. In case of the two stage Likert elicitation method, 'Absolutely certain A' and 'Absolutely certain B' were nested as 'Absolutely certain' while 'Not so certain A' and 'Not so certain B' were nested into 'Not so certain'.

As conducted for the MNL model, the data was pooled across the different elicitation methods and *(n-1)* scale parameters associated with the elicitation methods were estimated along with the nest parameters obtained for *(n-1)* different preference certainty levels. The model also accounted for panel effects. In order to set a reference scale as well as the nest parameters, different models were estimated with 500 MLHS draws and a full set of panel error components. The aim of this exercise was to find a model that estimated as many of the remaining scale and nest parameters as possible, without any identifiability problems. Selecting the most preferred model from this exercise for each of the attribute representation and data input methods, the alternative with the lowest panel error component was set as the reference alternative to estimate the model with *(n-1)* panel error components.

Conducting this exercise it was found that in case of the location ratings and dummy models, when the nest coefficient for ‘uncertain’ was fixed at unity, the nest coefficient for the ‘probably’ alternatives was also estimated to be unity. While fixing the nest coefficient for the ‘probably’ alternatives to unity resulted in the estimation of the ‘uncertain’ alternative which was not significantly different from unity, the model had identifiability problems arising from the inclusion of the ‘uncertain’ alternative in the estimation process[23]. Thus, in case of the one stage Likert, location representation models, the ‘uncertain’ nest parameter was fixed at unity[24] while in case of the one stage Likert, linguistic representation models, the ‘probably’ nest parameter was fixed at unity as this specification allowed for the estimation of the other two nest parameters without any identifiability problems. The models thus estimated are reported in Table 7.1. The models were estimated in BIOGEME 2.0 where the nest parameter is the inverse of the theoretic scale parameter. Hence in this case, higher values of the nest parameter indicate increased correlation between the nested alternatives.

The following results were thus obtained from the NL model:

---

[23] This was found to be the case in two out of three location ratings models that varied based on the reference scale parameter. In the one model where identifiability problem was not present, the resulting model fit (in terms of FLL) was poorer compared to the model where the ‘uncertain’ nest parameter was fixed at unity. For the location dummy model, all three models thus specified showed identifiability problems.

[24] As the nest parameter for the ‘Definitely’ alternatives was estimated under both the model specifications, though not significantly different from unity but with a larger significance value than either ‘Probably’ or ‘Uncertain’ nests, this nest parameter was not set as the reference parameter.

Table 7.1 NL results for pooled data across different representation and data input methods

Location attributes	**Location ratings**	**Location dummy**	**Linguistic attributes**	**Linguistic ratings**	**Linguistic dummy**
ASC			**ASC**		
Binary			**Binary**		
A	.314 (5.12)	.666 (4.68)	A	.099 (2.55)	.083 (1.74)
B	Fixed	Fixed	B	Fixed	Fixed
One Stage			**One Stage**		
Definitely A	1.09 (2.51)	1.30 (3.18)	Definitely A	.056 (1.57)	.02 (0.37)
Probably A	1.64 (3.21)	1.77 (3.75)	Probably A	-.552 (-5.98)	-.54 (-5.59)
Uncertain	.963 (1.71)	-2.03 (-3.75)	Uncertain	-.126 (-.59)	-1.48 (-4.2)
Probably B	1.32 (2.54)	1.21 (2.65)	Probably B	-.615 (-6.46)	-.52 (-5.79)
Definitely B	Fixed	Fixed	Definitely B	Fixed	Fixed
Two stage			**Two stage**		
Abs. cert. A	.229 (3.27)	.651 (3.97)	Abs. cert. A	.055 (1.79)	.045 (0.96)
Not so cert. A	-.128 (-.92)	.258 (1.56)	N.s.cert. A	-1.85 (-10.0)	-1.81 (-9.7)
Not so cert. B	-.496 (-4.64)	-.459 (-3.68)	N.s.cert. B	-2.36 (-9.43)	-2.28 (-9.1)
Abs. cert. B	Fixed	Fixed	Abs. cert. B	Fixed	Fixed
View	.0129 (6.74)		**View**	.005 (2.58)	
6F		.276 (2.73)	Good		.242 (3.69)
6T		.321 (3.25)	Neither		
3F		-.159 (-1.69)			
3T					
Noise	.0172 (8.80)		**Noise**	.0120 (3.14)	
6F		-.871 (-7.68)	Noisy		-.79 (-4.47)
6T		-.586 (-3.57)	Neither		-.29 (-3.60)
3F		-1.07 (-5.88)	Quiet		
3T					
Sunlight	.0159 (6.77)		**Sunlight**	.0051 (3.03)	
6F		-.129 (-1.33)	Very good		.341 (4.01)
6T		.329 (2.56)	Good		.228 (2.81)
3F		-.178 (-2.06)	Neither		
3T					

Charge	-.019 (-9.89)	-.022 (-10.2)	**Charge**	-.006 (-3.1)	-.011 (-4.4)
ρ^2 w.r.t. 0 **adjusted ρ^2**	0.235 0.228	.218 .210	**ρ^2 w.r.t. 0** **adj. ρ^2**	0.175 0.168	.203 .195
Scale parameter (w.r.t. 1) **Binary** **1 Stage** **2 Stage**	 1.65 (3.48) 1.00 (fixed) 7.62 (2.10)	 1.44 (2.66) 1.00 (fixed) 4.15 (3.35)	**Scale parameter (w.r.t. 1)** **Binary** **1 Stg.** **2 Stg.**	 3.13 (2.09) 3.74 (2.42) 1.00 (fixed)	 2.52 (2.60) 3.48 (2.96) 1.00 (fixed)
Nest Coeff. (w.r.t. 1) **1 stage** Definitely Uncertain Probably **2 stage** Abs. certain N. s. certain	 1.02 (0.07) 1.00 (fixed) 1.00 (0.00) 1.82 (3.56) 1.00 (fixed)	 1.08 (0.29) 1.00 (fixed) 1.00 (0.00) 1.49 (2.42) 1.00 (fixed)	**Nest Coeff. (w.r.t. 1)** **1 stage** Definitely Uncertain Probably **2 stage** Abs. certain N. s. certain	 4.54 (2.30) 1.01 (0.00) 1.00 (fixed) 4.26 (2.32) 1.00 (fixed)	 3.29 (2.87) 1.01 (0.05) 1.00 (fixed) 3.23 (2.89) 1.00 (fixed)
FLL	-2975.003	-3040.376	**FLL**	-2948.325	-2849.312
no. of obs.	3552	3552	**no. of obs.**	3264	3264
no. of indiv.	222	222	**no. of indiv**	204	204
MLHS draws	500	500	**MLHS draws**	500	500

The final log-likelihood values from the NL model results reveal that in line with the findings obtained from the MNL model, the ratings specification is better for the location representation method while the dummy specification is better for the linguistic representation method. The coefficient estimates obtained across the different model specifications show that 'noise' is valued higher in both location and linguistic ratings method, in comparison with other attributes within each of the two models.

Across the location and the linguistic ratings models it can be seen that the relative significance of the attributes across the two representation methods is different than that obtained from the BL-MNL model given in Chapter 6. In case of the linguistic representation method it is observed that the t-statistics obtained for all the attributes is lower than that obtained from the BL-MNL model. Moreover, the scale parameter values obtained in this model show a higher and more statistically significant values than that obtained in the BL-MNL model. Thus, in case of the linguistic representation method it can be implied that the variation in the t-statistics of the coefficient estimates is accounted by the change in the scale parameter values and its significance. While in the case of the location representation method, a relatively lower variation is found in the case of the coefficient estimates as well as the scale parameter values. The relative coefficient estimates and its significance for the linguistic representation method is thus lower than that found for the location representation method compared to that obtained from the BL-MNL model.

In terms of the nest parameters value, for the location representation method, it is observed that statistically insignificant nest coefficient values are obtained for the 'Definitely' and the 'Probably' alternatives of the one stage Likert elicitation method while a significant nest coefficient value is obtained for the 'Absolutely certain' alternatives of the two stage Likert elicitation method across both the ratings and dummy data input methods. For the linguistic representation method however, it can be seen that for the one stage Likert elicitation method, a high and statistically significant nest coefficient value is obtained for the 'Definitely' nest while the 'Absolutely certain' nest of the two stage Likert method also yields a high and significant nest coefficient value. These findings reveal that in case of the location representation and two stage Likert elicitation method, the error variance between the alternatives 'Absolutely certain A' and 'Absolutely certain B' are correlated. Thus the choice sets during the choices of these alternatives are perceived in a particular similar light. In case of the linguistic representation method on the other hand, this finding is revealed for both the 'Definitely' nest of the one stage Likert elicitation as well as the 'Absolutely certain' nest of the two stage Likert elicitation method. Hence in case of the linguistic representation method, it can be concluded that the correlation in error variance across alternatives with high levels of certainty is evident irrespective of the preference elicitation technique used. These results reveal the structure of error associated with each level of preference certainty with greater similarity in the way alternatives are perceived when respondents are absolutely certain and presence of higher randomness when complete certainty is not present, which is especially evident in case of the two stage Likert elicitation method and for the one and two stage Likert methods under the linguistic representation.

Without accounting for the panel effect in the NL model it was observed that a high and significant correlation is obtained for the 'Definitely' alternatives of the one stage Likert elicitation for the location representation method. This implies that the correlation obtained for this certainty level in absence of panel specification is due to the correlation in intra-respondent observations. By accounting for panel effects, the nest parameters obtained does not include correlation from intra-respondent observations and hence the nest parameter value obtained with panel specification is lower than that obtained when panel effects are not considered in the model. In terms of the NL model application therefore, this finding emphasises the importance of considering panel effects in modelling as the absence to account for it could result in an overestimation of the nest coefficients.

In order to examine whether any improvements are gained by relaxing the IIA assumption and thus whether the NL model provides a better fit than the MNL model, the likelihood ratio (LR) test was conducted. The LR test can be given as:

$$-2(LL_{base\,\mathrm{mod}el} - LL_{estimated\,\mathrm{mod}el})$$

Where,

$LL_{base\ model}$ = final log-likelihood from the MNL model

$LL_{estimated\ model}$ = final log-likelihood from the NL model

The value obtained from the above equation is compared with χ^2 estimate where the degrees of freedom equal the number of new parameters in the estimated model.

The estimated model is deemed to be better than the base model when the LR value exceeds the critical χ^2 estimate (Hensher *et al*., 2005).

The following results were obtained from the LR test for the NL model:

Table 7.2 LR test for NL model versus MNL model

	Location		**Linguistic**	
	Ratings	**Dummy**	**Ratings**	**Dummy**
LR value	75.79	41.68	53.32	53.25
d.f.	3	3	3	3
Significance at $\alpha = 0.05$	NL better	NL better	NL better	NL better

Examining the results obtained from the LR test, it is seen that the NL model is better than the MNL model for all the different attribute representation as well as the data input methods. Based on the LR value, it is seen that there is greater improvement in model fit for the location ratings specification by accounting for correlation in the preference certainty levels. Thus, the finding reveals the importance to account for correlation between error variances for similar levels of preference certainty while also accounting for panel effects in the model.

In order to further relax the IIA assumption, the error components logit (ECL) model was applied. The following section provides the structure and results of the ECL model.

7.4.2 The Error Components Logit (ECL) model

The main aim of the ECL model application was to examine the presence of heteroskedastic error variances. To that end, the binary, one stage Likert and two stage Likert data were again pooled and panel effect was considered in the model formation. To incorporate heteroskedastic error variances, an alternative specific error variance $\sim N(0,\sigma^2)$ was added to each alternative and the model was estimated with 500 MLHS draws. Based on the error components obtained during this estimation, the base alternative was determined for the final model estimation. Thus, the alternative with the least error component in the initial estimation was fixed as the base alternative (with no error component) during the final estimation process. It has to be noted that using this methodology, the base alternative in the final estimation varied for the different model specifications. Thus, in the case of the location ratings as well as the linguistic ratings and the linguistic dummy models, the base alternative was 'Absolutely certain A' while in the case of the location dummy model this was 'Definitely A'. Thus, the ECL model structure took the following form for the location ratings specification:

$$
\begin{aligned}
U_A &= \alpha VA + \beta NA + \gamma SA + \eta CA + ASC_A + \xi_A \\
U_B &= \alpha VB + \beta NB + \gamma SB + \eta CB + \xi_B \\
U_{DA} &= \alpha VA + \beta NA + \gamma SA + \eta CA + ASC_{DA} \\
U_{PA} &= \alpha VA + \beta NA + \gamma SA + \eta CA + ASC_{PA} + \xi_{PA} \\
U_U &= ASC_U + \xi_U \\
U_{PB} &= \alpha VB + \beta NB + \gamma SB + \eta CB + ASC_{PB} + \xi_{PB} \\
U_{DB} &= \alpha VB + \beta NB + \gamma SB + \eta CB + \xi_{DB} \\
U_{AcA} &= \alpha VA + \beta NA + \gamma SA + \eta CA + ASC_{AcA} + \xi_{AcA} \\
U_{NsA} &= \alpha VA + \beta NA + \gamma SA + \eta CA + ASC_{NsA} + \xi_{NsA} \\
U_{NsB} &= \alpha VB + \beta NB + \gamma SB + \eta CB + ASC_{NsB} + \xi_{NsB} \\
U_{AcB} &= \alpha VB + \beta NB + \gamma SB + \eta CB + \xi_{AcB}
\end{aligned}
$$

Where, alternative specific error component $\xi \sim N(0,\sigma^2)$ and the notation for the utility functions and the individual attributes follow from the notation introduced in Section 6.2.1.

The following results were obtained from the ECL model with heteroskedastic error variances:

Table 7.3 ECL model with heteroskedastic error variances across different model specifications

Location attributes	Location ratings	Location dummy	Linguistic attributes	Linguistic ratings	Linguistic dummy
ASC			**ASC**		
Binary			**Binary**		
A	.435 (5.09)	.956 (4.85)	A	.326 (4.60)	.212 (1.93)
B	Fixed	Fixed	B	Fixed	Fixed
One Stage			**One Stage**		
Definitely A	1.62 (3.40)	1.57 (3.83)	Definitely A	.345 (1.70)	.093 (0.56)
Probably A	2.33 (5.25)	1.98 (5.20)	Probably A	.004 (0.02)	-1.12 (-2.5)
Uncertain	1.24 (1.82)	-2.38 (-5.00)	Uncertain	-.751 (-1.07)	-4.8 (-6.55)
Probably B	1.87 (4.20)	1.19 (3.36)	Probably B	-.472 (-1.72)	-.16 (-0.91)
Definitely B	Fixed	Fixed	Definitely B	Fixed	Fixed
Two stage			**Two stage**		
Abs. cert. A	.423 (2.54)	.580 (2.02)	Abs. cert. A	.232 (2.47)	.155 (1.27)
Not so cert. A	-.558 (-2.52)	-.184 (-0.68)	N.s.cert. A	-1.12 (-4.85)	-1.1 (-4.96)
Not so cert. B	-1.32 (-3.06)	-.875 (-3.55)	N.s.cert. B	-2.70 (-3.78)	-3.2 (-4.29)
Abs. cert. B	Fixed	Fixed	Abs. cert. B	Fixed	Fixed
View	.0183 (6.48)		**View**	.0173 (4.53)	
6F		.353 (2.57)	Good		.637 (6.57)
6T		.433 (3.22)	Neither		
3F		-.202 (-1.58)			
3T					
Noise	.0235 (8.27)		**Noise**	.0397 (11.5)	
6F		-1.18 (-6.99)	Noisy		-2.1 (-13.8)
6T		-.778 (-3.48)	Neither		-.70 (-5.64)
3F		-1.43 (-5.48)	Quiet		
3T					
Sunlight	.0229 (6.44)		**Sunlight**	.0156 (6.90)	
6F		-.211 (-1.62)	Very good		.868 (7.89)
6T		.451 (2.57)	Good		.626 (3.82)
3F		-.294 (-2.52)	Neither		
3T					

Charge	-.027 (-9.40)	-.0297 (-8.7)	**Charge**	-.0206 (-9.5)	-.028 (-12)
ρ^2 w.r.t. 0 **adjusted ρ^2**	0.229 0.220	0.213 0.203	**ρ^2 w.r.t. 0** **adj. ρ^2**	0.170 0.161	0.199 0.189
Scale parameter (w.r.t. 1) **Binary** **1 Stage** **2 Stage**	1.19 (1.32) 1.00 (fixed) 1.22 (1.20)	1.05 (0.38) 1.00 (fixed) 1.22 (1.16)	**Scale parameter (w.r.t. 1)** **Binary** **1 Stg.** **2 Stg.**	1.00 (fixed) 1.16 (0.97) 1.16 (1.05)	1.00 (fixed) 1.08 (0.67) 1.18 (1.37)
ε Comp. **Binary** A B **One Stage** Definitely A Probably A Uncertain Probably B Definitely B **Two stage** Abs. cert. A Not so cert. A Not so cert. B Abs. cert. B	.105 (0.50) .238 (1.50) .451 (1.51) .766 (2.50) -1.64 (-2.91) -.683 (-2.13) -1.16 (-3.85) Fixed .0688 (0.24) 1.36 (3.07) -.308 (-0.74)	.129 (0.62) -.029 (-0.18) Fixed .953 (2.11) -.193 (-0.47) -.829 (-2.69) -.207 (-0.63) -1.19 (-3.75) .696 (2.21) .911 (3.08) .0015 (0.01)	**ε Comp.** **Binary** A B **One Stage** Definitely A Probably A Uncertain Probably B Definitely B **Two stage** Abs. cert. A N. s. cert. A N. s. cert. B Abs. cert. B	.268 (1.34) -.490 (-2.01) .646 (1.64) -.354 (-0.92) .854 (2.26) 1.52 (3.90) .577 (2.21) Fixed -.243 (-0.57) 2.01 (3.21) .115 (0.45)	-.073 (-0.4) .295 (1.45) -.196 (-0.8) 2.03 (3.48) -.854 (-2.2) .673 (1.76) .328 (1.15) Fixed -.198 (-0.5) 2.37 (4.28) .005 (0.02)
FLL	-2998.785	-3059.694	**FLL**	-2967.125	-2864.493
no. of obs.	3552	3552	**no. of obs.**	3264	3264
no. of indiv.	222	222	**no. of indiv**	204	204
MLHS draws	500	500	**MLHS draws**	500	500

Examining the results obtained, it can be observed that the relative attribute coefficients obtained as well as their statistical significance follows the pattern obtained from the MNL model. Thus, in the case of both the location and the linguistic ratings specifications, 'noise' has a higher value and statistical significance compared to 'view' and 'sunlight'. However, compared to the results obtained from the MNL model, it can now be observed that the scale parameters obtained have a very low value and are not significantly different from unity across all the model specifications. This result is obtained due to the introduction of heteroskedastic error variances. Thus the level of commonality within the nested alternatives is low resulting in lower scale parameters across the nests.

While the introduction of heteroskedastic error variance for the binary alternatives does not have a direct empirical underpinning, this structure was adapted in order to have a comparable ECL model of the base MNL model. However, focussing on the error components obtained for the one and two stage Likert alternatives, it is found that in case of the one stage Likert method for the location ratings model, a high and significant error component is obtained for the 'Probably', 'Uncertain' as well as the 'Definitely B' alternatives, with the highest error component value for the 'Uncertain' alternative. These results indicate that in the choice of these preference levels, there is higher randomness associated with the choice. While increased randomness in choice is expected at lower levels of preference certainty, the presence of this for the level 'Definitely B' however is atypical.

In case of the two stage Likert elicitation for the location ratings model, it is observed that 'Not so certain B' has the highest error component value with high statistical significance. For the location dummy one stage Likert model, it is found that the 'Probably' certainty levels have a high and significant error component while the 'Definitely' and 'Uncertain' levels do not have a significant value. The two stage Likert model also reveal that the 'Not so certain' certainty levels have a high randomness associated with them. However, anomalously it is also found that 'Absolutely certain A' has a high and significant error component. Thus, in case of the location representation method, it is found that high certainty levels in some cases also have high randomness associated with their choice which is contrary to expectation.

With the linguistic representation method, it is found that in case of the one stage Likert elicitation method, a high and significant error component is obtained for the level 'Probably B' for the ratings specification while for the dummy specification a high and significant error component is obtained for the 'Probably A' alternative, along with the 'Uncertain' alternative for both the data input methods. For the two stage Likert elicitation method, it is observed that the 'Not so certain B' alternative has a high and significant error component across both the ratings and the dummy specifications. Thus, in case of the linguistic specification it can be observed that the preference levels with lower certainty have a relatively higher randomness associated with their choice.

The presence of a significant error component indicates higher randomness associated with that preference level. While the link between lower levels of preference certainty and randomness seems intuitively more plausible, randomness in choice however need not be restricted to these certainty levels alone. Thus, high and significant error components for higher levels of preference certainty indicate that there is considerable amount of randomness associated with the choice of those alternatives. Besides obtaining a high and significant error component for the lower preference certainty levels, in case of the location representation method, it is observed that 'Definitely B' and 'Absolutely certain A' also have a high and significant error component for the ratings and dummy models respectively indicating that there is higher level of randomness associated in the choice of these alternatives. In case of the linguistic representation method, it is seen that across the different data input as well as different preference elicitation methods, higher randomness is associated with lower levels of preference certainty. Thus, across the two attribute representation methods, it can be concluded that while lower levels of preference certainty are associated with randomness, some aberrancy is also observed in case of the location representation method.

In order to examine whether the ECL model offers any advantages over the MNL model, the likelihood ratio test was conducted. The following results are obtained from the LR test for the ECL model:

Table 7.4 LR test for ECL model versus MNL model

	Location		Linguistic	
	Ratings	**Dummy**	**Ratings**	**Dummy**
LR value	28.23	3.04	15.72	22.89
d.f.	10	10	10	10
Significance at $\alpha = 0.05$	ECL better	ECL not better	ECL not better	ECL better

Examining the above results it is seen that the ECL model offers significant advantages over the MNL model for the location ratings and the linguistic dummy models. As the ratings method is more suitable data input technique for the location representation method while the dummy method is more suitable for the linguistic representation, the LR test results imply that it is important to allow for randomness in choice through the ECL specification. The heteroskedastic error specification sheds more light on the different error structure and the amount of randomness for each of the preference certainty levels. Similar to the findings obtained from the NL model, the LR values from the ECL model reveal that by accounting for heteroskedastic error variances, significant improvement in the model fit can be obtained for the location ratings specification.

Examining the results obtained from the NL as well as the ECL models it can be concluded that in terms of LR test, both these models offer some advantages over the MNL model. While the NL model reveals that there is higher correlation between alternatives with high preference certainty levels, the ECL results indicate that the preference certainty levels such as ‘Probably’, ‘Uncertain’ and ‘Not so certain’ have higher randomness associated with their choice in most of the cases. Moreover, as panel data can result in correlation within intra-respondent observations as well as affect inter-respondent variability, it is important to consider these effects while accounting for both correlation and heteroskedasticity through the NL and ECL models.

7.5 The Ordered Logit Model

The ordered logit (OL) model has been traditionally and commonly used to analyse ordered strength of preference and preference uncertainty responses, where preferences are explained in terms of utility differences and preference intervals, demarcated by threshold values (Johnston and Swallow, 1999; Johnston, 2003; Whelan and Tapley, 2006). The primary focus of the OL analytical technique is to predict the proportion of individuals choosing each level on the attitude scale model (Gob *et al*., 2007). The application of this technique on Likert scale data is based on the assumption that individuals choose a continuous strength of preference response based on a single, underlying, fixed preference function and there is no preference heterogeneity or asymmetry due to respondent's socio-economic characteristics or tastes (Johnston, 2003).

For the one stage and two stage Likert data, an ordered logit model was used to analyse the implicit ordering among the preference levels across the two elicitation methods based on the estimated threshold values. As the ordered logit model needs to be specified for each elicitation method (one and two stage Likert methods), explicit pooling of data based on the different elicitation methods is not possible with this analytical technique. Moreover, as BIOGEME does not allow panel specification for ordered logit model, the models were estimated without considering panel effects.

In case of the one stage Likert data, the following ordering of preferences was incorporated into the model: Definitely A, Probably A, Uncertain, Probably B and Definitely B. The threshold values obtained from the OL model thus depicted the relative ordering of the categories.

The following results were obtained from the OL analysis for the one stage Likert data:

Table 7.5 OL results for one stage Likert data

Location attributes	**Location ratings**	**Location dummy**	**Linguistic attributes**	**Linguistic ratings**	**Linguistic dummy**
View	.017 (5.90)		**View**	.019 (3.32)	
6F		.173 (0.95)	Good		.83 (5.66)
6T		.292 (1.63)	Neither		
3F		-.285 (-1.64)			
3T					
Noise	.017 (5.79)		**Noise**	.032 (10.25)	
6F		-.775 (-4.39)	Noisy		-1.66 (-10.0)
6T		-.433 (-1.41)	Neither		-.460 (-2.28)
3F		-1.27 (-4.13)	Quiet		
3T					
Sunlight	.026 (6.86)		**Sunlight**	.016 (5.17)	
6F		-.043 (-0.25)	Very good		.813 (5.13)
6T		.237 (0.99)	Good		.592 (2.29)
3F		-.080 (-0.50)	Neither		
3T					
Charge	-.019 (-9.70)	-.018 (-9.34)	**Charge**	-.0165 (-6.97)	-.023 (-8.50)
ρ^2 w.r.t. 0	0.134	0.112	**ρ^2 w.r.t. 0**	0.132	0.153
adjusted ρ^2	0.128	0.102	**adj. ρ^2**	0.126	0.145
Threshold values			**Threshold values**		
τ_1	-1.42 (-16.0)	-1.04 (-4.96)	τ_1	-.949 (-9.41)	-1.03 (-6.08)
τ_2	.153 (2.04)	.450 (2.15)	τ_2	.108 (1.12)	.098 (0.59)
τ_3	.575 (7.46)	.848 (4.05)	τ_3	.322 (3.33)	.325 (1.95)
τ_4	2.21 (19.59)	2.42 (10.86)	τ_4	1.31 (12.19)	1.35 (7.82)
FLL	-1226.968	-1257.459	**FLL**	-1139.356	-1112.576
no. of obs.	880	880	**no. of obs.**	816	816

In order to maintain the implicit ordering of the preference levels, ordered logit requires that the threshold values for the middle parameters should not be statistically different from zero (Swallow *et al.*, 2001). Comparing the threshold values obtained across the different representation and data input methods, it can be observed that the correct sign of the values are obtained across the different representation and data input methods.

In case of the linguistic dummy specification, the threshold values for the middle level ($\boldsymbol{\tau}_2$ and $\boldsymbol{\tau}_3$) are not statistically different from zero. This finding implies that in case of this specification, the implicit ordering of the preference levels is maintained under the OL method. In case of the location ratings model on the other hand, it is found that all the threshold values give a statistically significant value albeit with a value close to zero for $\boldsymbol{\tau}_2$, implying that the implicit ordering while present to some extent, is not as maintained in this case as observed for the linguistic dummy method. In case of the location dummy specification, the middle threshold parameters have a high value as well as a high statistical significance implying that the implicit ordering is not maintained in this case. For the linguistic ratings model it is seen that one of the middle threshold parameters is insignificant, while another has a lower value compared to those obtained from the location representation method. Thus in this case, the implicit ordering of preferences is seen to be relatively maintained.

For the two stage Likert data, the threshold parameters signify the following ordering of preferences: Absolutely certain A, Not so certain A, Not so certain B and Absolutely certain B. The results from the OL analysis of the two stage Likert data can be given as follows:

Table 7.6 OL results for two stage Likert data

Location attributes	**Location ratings**	**Location dummy**	**Linguistic attributes**	**Linguistic ratings**	**Linguistic dummy**
View	0.021 (6.91)		**View**	.0121 (2.02)	
6F		-.132 (-0.69)	Good		.322 (2.06)
6T		.168 (0.90)	Neither		
3F		-.277 (-1.54)			
3T					
Noise	0.018 (5.71)		**Noise**	.036 (10.85)	
6F		-.658 (-3.67)	Noisy		-1.88 (-10.5)
6T		-.178 (-0.56)	Neither		-.664 (-3.10)
3F		-.649 (-2.05)	Quiet		
3T					
Sunlight	0.024 (5.98)		**Sunlight**	.0172 (5.19)	
6F		-.014 (-0.08)	Very good		.992 (5.79)
6T		.448 (1.81)	Good		.828 (2.94)
3F		.0457 (0.28)	Neither		
3T					
Charge	-.0239 (-11.5)	-.0228 (-11.2)	**Charge**	-.0198 (-7.4)	-.0237 (-7.9)
ρ^2 w.r.t. 0	0.176	0.143	**ρ^2 w.r.t. 0**	0.204	0.227
adjusted ρ^2	0.170	0.133	**adj. ρ^2**	0.198	0.219
Threshold values			**Threshold values**		
τ_1	-.669 (-8.07)	-.127 (-0.58)	τ_1	-.439 (-4.31)	-.575 (-3.26)
τ_2	.264 (3.25)	.745 (3.40)	τ_2	.216 (2.13)	.119 (0.68)
τ_3	.978 (11.20)	1.40 (6.33)	τ_3	.689 (6.63)	.622 (3.51)
FLL	-1023.606	-1064.272	**FLL**	-900.235	-874.316
no. of obs.	896	896	**no. of obs.**	816	816

The results reveal that correct sign of the threshold values are obtained from the model. Based on the value of the threshold parameters and its associated significance level, it can be observed that the implicit ordering is maintained in case of the linguistic dummy specification. In case of the location ratings specification, the middle threshold parameter has a value close to zero (0.264) with high statistical significance which implies that though the implicit ordering is obtained in this case, it is less maintained than is found in the linguistic dummy specification. In case of the location dummy specification it is observed that the implicit ordering is not maintained as the value for the middle threshold parameter is high and significantly different from zero, while one of the extreme parameters reveal a value not significantly different from zero. For the linguistic ratings model, it is seen that the implicit ordering is relatively maintained. The results thus indicate that the ordered logit analysis is best applied in the case of the linguistic dummy specification as well as the location ratings model. The results for the threshold parameters reveal that the two stage Likert method is also well analysed using the OL analysis.

Results from the one stage Likert model reveals that in line with the results obtained from the MNL model, the location ratings model gives a better model fit in terms of FLL value and the adjusted ρ^2 than the location dummy specification (FLL -1226.968 versus FLL -1257.459). In case of the linguistic representation method, a better model fit is obtained in case of the dummy specification than that obtained from the ratings specification (FLL -1112.576 versus FLL -1139.356). In case of the location ratings specification, it is observed that ‘view’ and ‘sunlight’ attributes provide a higher statistical significance than the ‘noise’ attribute, though the coefficient estimates obtained for ‘view’ and ‘noise’ are the same. In case of the linguistic dummy specification, a high value is obtained for ‘noisy’ noise with high level of significance with other attributes also providing statistically significant estimates. With the linguistic ratings specification it is seen that ‘noise’ has a higher coefficient estimate and higher significance than ‘view’ and ‘sunlight’ as obtained from the MNL model.

Results from the two stage Likert model also provide similar insight in terms of the relative suitability of data input methods based on the different attribute representation methods (i.e., location ratings model is better than location dummy while linguistic dummy model is better than linguistic ratings specification). Similar findings are also obtained for the relative valuations of the different attributes based on the different representation and data input methods. It can thus be noted that the OL analysis for both the one and two stage Likert data reiterates the findings obtained from the MNL model in terms of the relative effects of different representation and data input methods on model fit and coefficient estimates.

Comparing across both one and two stage Likert data it can be concluded that implicit ordering of preferences is best maintained for the linguistic dummy specification as well as the location ratings specification, to some extent. Thus, when the data input method is chosen correctly based on the attribute representation method, the OL model can be well-applied irrespective of the preference elicitation (one or two stage Likert) method used. However, current limitations of the OL model specification in the software used, implies that pooling or incorporation of panel effects cannot be conducted in this case, giving the BL-MNL model a comparative advantage over the OL model.

7.6 Effect on Valuation

Based on the coefficient estimates obtained from the nested, error components and ordered logit models, the WTP value for 'view', 'noise' and 'sunlight' attributes were computed using the valuation method outlined in Section 6.4. This was done in order to examine the effect of these models on attribute valuation.

In case of the NL model, the following attribute valuations were obtained:

Table 7.7 Willingness to pay (in Euro and 2008 prices) for each attribute across the different representation and data input methods from the NL model

Location attributes	**Location ratings**	**Location dummy**	**Linguistic attributes**	**Linguistic ratings**	**Linguistic dummy**
View	.675 (8.33)		**View**	.803 (5.03)	
6F		12.60 (2.81)	**Good**		21.80 (7.65)
6T		14.66 (3.43)	**Neither**		
3F		-7.26 (-1.70)			
3T					
Noise	.900 (11.39)		**Noise**	1.91 (16.81)	
6F		-39.8 (-9.38)	**Noisy**		-71.6 (-15.2)
6T		-26.8 (-3.67)	**Neither**		-25.8 (-7.02)
3F		-48.9 (-6.42)	**Quiet**		
3T					
Sunlight	.832 (8.13)		**Sunlight**	.819 (7.81)	
6F		-5.89 (-1.34)	**Very good**		30.72 (7.99)
6T		15.02 (2.65)	**Good**		20.54 (3.55)
3F		-8.13 (-2.12)	**Neither**		
3T					

Compared to the WTP estimates obtained from the BL-MNL model as shown in Table 6.7, it can be observed that for all the representation and data input methods, the statistical significance of the WTP estimate is higher. In case of the location ratings model, a slightly higher WTP estimate is obtained for each of the different attributes. In case of the linguistic dummy model, while a higher statistical significance is obtained across the different attributes and attribute levels, the effect on the WTP estimate is variable with a slightly higher value for ‘neither noisy nor quiet’ noise level and a slightly lower estimate for the other attribute levels.

With the ECL model, the following attribute valuations were obtained across the different representation and data input methods:

Table 7.8 Willingness to pay (in Euro and 2008 prices) for each attribute across the different representation and data input methods from the ECL model

Location attributes	**Location ratings**	**Location dummy**	**Linguistic attributes**	**Linguistic ratings**	**Linguistic dummy**
View	.678 (8.13)		**View**	.839 (5.17)	
6F		11.88 (2.68)	**Good**		22.5 (7.75)
6T		14.58 (3.42)	**Neither**		
3F		-6.80 (-1.61)			
3T					
Noise	.870 (11.01)		**Noise**	1.93 (14.82)	
6F		-39.73 (-9.3)	**Noisy**		-73.8 (-15.3)
6T		-26.19 (-3.6)	**Neither**		-24.8 (-6.6)
3F		-48.15 (-6.4)	**Quiet**		
3T					
Sunlight	.848 (8.11)		**Sunlight**	.757 (7.28)	
6F		-7.10 (-1.63)	**Very good**		30.7 (7.87)
6T		15.18 (2.69)	**Good**		22.1 (3.75)
3F		-9.89 (-2.62)	**Neither**		
3T					

While compared to the BL-MNL model the WTP estimates show a higher statistical significance under the ECL model, compared to the NL model, the WTP estimates obtained from this model have a slightly lower statistical significance for most of the attributes and attribute levels across the different data input methods. The WTP estimates obtained from the ECL model is closer to that obtained from the BL-MNL model with slightly higher values in case of ‘view’ and ‘sunlight’ attributes for the location ratings model and slightly lower values for ‘noisy’ noise and ‘very good’ sunlight under the linguistic dummy model.

The attribute valuation estimated from the one and two stage Likert OL model is provided in Table 7.9.

Table 7.9 Valuation (in Euro and 2008 prices) across the different representation and data input methods from the one and two stage OL models

Location attributes	Loc. rat. – 1 stg. Likert	Loc. rat. – 2 stg. Likert	Loc. dum. – 1 stg. Likert	Loc. dum. – 2 stg. Likert	Linguistic attributes	Ling.rat. – 1 stg. Likert	Ling.rat. – 2 stg. Likert	Ling. dum. – 1 stg. Likert	Ling. dum. – 2 stg. Likert
View	.90 (5.23)	.879 (6.12)			**View**	1.17 (3.49)	.611 (2.16)		
6F			9.45 (0.95)	-5.79 (-0.68)	**Good**			36.04 (6.24)	13.59 (2.25)
6T			15.96 (1.6)	7.37 (0.90)	**Neither**				
3F			-15.57 (-1.6)	-12.15 (-1.5)					
3T									
Noise	.88 (5.09)	.761 (5.35)			**Noise**	1.93 (7.85)	1.82 (8.89)		
6F			-42.3 (-4.05)	-28.86 (-3.6)	**Noisy**			-72.17 (-7.6)	-79.32 (-7.5)
6T			-23.7 (-1.39)	-7.81 (-.56)	**Neither**			-20.0 (-2.47)	-28.02 (-3.5)
3F			-69.4 (-3.81)	-28.5 (-2.02)	**Quiet**				
3T									
Sunlight	1.36 (5.85)	1.00 (5.62)			**Sunlight**	.94 (4.54)	.869 (4.67)		
6F			-2.37 (-0.25)	-.627 (-.078)	**Very good**			35.35 (4.72)	41.86 (5.05)
6T			12.9 (0.99)	19.65 (1.82)	**Good**			25.74 (2.23)	34.94 (2.79)
3F			-4.38 (-0.50)	2.00 (0.28)	**Neither**				
3T									

Comparing the WTP estimates obtained from the one and two stage Likert OL model, it can be observed that in case of the location and linguistic ratings model, a slightly lower WTP estimate is obtained for each of the different attributes in case of the two stage Likert OL model. For the linguistic dummy specification, variable effect of attribute valuation is observed across the two Likert elicitation methods with higher valuation for 'good' view with the one stage Likert method and a relatively lower valuation for the attribute levels for 'noise' and 'sunlight'. The statistical significance of the WTP estimates is also seen to vary across the different model specifications and attribute levels.

While the OL model specified here does not allow for pooled data across the different preference elicitation methods, compared to the BL-MNL model it can be observed that in case of the location and linguistic ratings specifications, the WTP estimate obtained for the 'noise' attribute from the one stage as well as the two stage OL models is closer to that obtained from the BL-MNL model. The WTP estimates for the other attributes under the ratings specification has a larger variation from that obtained with the BL-MNL model. While most of the attribute levels under the location dummy specification do not give a statistically significant WTP estimate, under the linguistic dummy specification it can again be observed that least difference in WTP values between OL and BL-MNL models is obtained for the 'noise' attribute levels.

This section outlined the attribute valuations obtained from the NL, ECL and OL models. It has been observed that the valuations obtained from the NL model are higher than those obtained from the BL-MNL model for most of the attributes in case of the location ratings model and lower than the BL-MNL model in case of the linguistic dummy model. Valuations obtained in case of the ECL model were found to be closer to those from the BL-MNL model. As the LR test reveals that the NL model provides substantial improvements over the base BL-MNL model across all the different model specifications while the ECL model yields significant improvement over the base model in case of the location ratings and linguistic dummy specifications, the valuations obtained from the NL and ECL models needs to be considered when accounting for flexible error structure.

7.7 Summary and Conclusions

One of the hypotheses examined in this chapter focussed on whether the different types of choice set have an effect on respondent's preference uncertainty. It was hypothesised that the different types of choice sets defined by the different methods of attribute representation could have different effect on respondents' preference uncertainty and as the linguistic representation method would be easier for respondents to understand, this representation method would result in more certain choices and hence there would be lesser need for more complex error assumption. This hypothesis was tested by applying the nested and error components logit model and estimating the LR value between the more complex models and the base BL-MNL models. It was expected that as linguistic representation method would be clearer for respondents to understand, there would be a lesser need to make more complex error assumptions and hence the LR value for this representation method would be lower than that obtained from the location representation method.

Results from the LR test for the NL model revealed that this model showed substantial improvement over the BL-MNL model for all the different representation as well as the data input methods. Thus, it was found that accounting for error correlation is required for all the representation and data input techniques. However comparing the LR value obtained from the location ratings and the linguistic dummy specifications (which are the best models based on the data input type for each of the respective representation methods), it was seen that the LR value obtained from the linguistic dummy model is lower than that obtained from the location ratings model, implying that there is greater improvement in the model fit by applying the NL model for the location ratings specification.

The ECL model further relaxed the error assumption by allowing for heteroskedastic error variances. The LR value for the ECL model showed that the ECL model offered improvements over the BL-MNL model only for the location ratings and the linguistic dummy specifications. Moreover, comparing the LR value obtained from the location ratings and the linguistic dummy models it was found that greater improvement in model fit was obtained by applying the ECL model for the location ratings specification than that from the linguistic dummy model. Thus, the ECL model application also revealed that there is a greater need for relaxing the error assumption in case of the location ratings model than the linguistic dummy model. Thus, examining the results obtained from both the NL and the ECL models, it can be noted that the hypothesis is supported and there is relatively lesser improvement in the model fit by applying more complex error models in case of the linguistic dummy specification. This finding thus implies that the linguistic representation method is relatively easier for respondents to understand.

The second hypothesis examined in this chapter dealt with the different causes of preference uncertainty. It was hypothesised that the stated level of preference certainty could be due to deterministic as well as random effects. Thus the preference certainty levels could reflect the true preferences of the respondents as well as the randomness in choice. In order to dissociate the causes of preference uncertainty to true preferences of the respondents or that arising from randomness, it was hypothesised that in case of the former, there would be lesser need for complex error assumption while in case of the latter there would be higher need for relaxing the error assumption. Examining the LR value across both the NL and ECL models, it was seen that in case of the linguistic dummy specification, there is lesser need for error relaxation implying that in this case there is lesser randomness associated with respondents' choices while in case of the location ratings method, there is higher randomness in respondents' preferences. By applying the NL and ECL models it was thus seen that the linguistic method is easier for respondents to understand and the preference certainty levels chosen by respondents with this method is more due to their true preferences, with lesser effect of stochastic factors.

In addition to the error relaxation, the OL method was applied as a commonly used analytical model on Likert type data and the results from this model revealed that the implicit orderings of the preferences are maintained when appropriate data input method is used for each of the attribute representation methods. The results also revealed that the OL method which is commonly used to analyse one stage Likert data can also be appropriate for the two stage Likert elicitation method.

Results from the valuation exercise from the NL and ECL models showed that the WTP estimates obtained from these models varied from those obtained from the BL-MNL model. In case of the NL model, this variation was found to be greater with higher valuations for most of the attributes compared to the base model. Comparison of the one and two stage Likert OL model's WTP estimate revealed a lower valuation for all the attributes in case of the two stage Likert model while comparison with the WTP values obtained from the BL-MNL model showed least difference in the valuations for the 'noise' attribute across the two model specifications. The valuation exercise thus revealed the effect of different error assumptions, modelling techniques as well as the attribute representation and preference elicitation methods on the WTP estimate.

The following table summarises the results obtained from each of the models while also recommending the suitability of each model based on the model characteristics as well as the research findings:

Table 7.10 Model results summary and application recommendation

Model	Model Parameter	Conclusions	Model Recommendation
Nested Logit	Nest Parameter	Higher correlation between higher certainty preference levels	This can be a suitable model to apply when examining the correlation between different preference levels is needed. Results from the NL model give substantial improvement over the MNL model, implying this factor should be considered.
Error Components Logit	Error Components	Higher error variance mostly for lower certainty preference levels	For correct data input method, ECL offers substantial improvement over MNL. Can be a suitable model to apply when estimating the difference in error variance for different levels of preference certainty is needed.
Ordered Logit	Threshold Parameter	Implicit ordering of preferences maintained for models with appropriate data input method for each of the representation methods	Is a commonly used method for analysing Likert type data. Can be used for models with correct data input method based on the representation technique; used to examine whether the implicit ordering of preferences is maintained.

8 FUZZY LOGIC ANALYSIS

8.1 Introduction

The previous chapter examined the effect of attribute representation and choice elicitation on respondent's choice certainty using various logit model forms. However, as noted by Swait and Adamowicz (2001), individuals might adopt different decision strategies as a result of choice task complexity and hence examining whether respondents' choices can be adequately modelled using logical rules can be pertinent. With the application of heuristic rules, the model allows the researcher to examine which variables and criteria have most affected respondents' decision-making, without strict assumptions of linearity and compensatory decision-making in choice. This chapter applies a heuristic based approach to examine the effect of different representation techniques on the fuzzy system requirements while also comparing the model performance against results obtained from the standard logit models.

The method and theory of fuzzy logic analysis and fuzzy inference system was provided in Chapter 4. This chapter will focus on the results obtained from fuzzy logic analysis of different representation techniques. The chapter will begin with the fuzzy inference system (FIS) for the location representation input data, followed by the results from the linguistic representation data. For both these methods, the fuzzy inference system developed will be explained. Section 8.2 will focus on the results obtained from the location method while section 8.3 will detail the results obtained from the linguistic method of attribute representation. The analysis for both the representation techniques is conducted with and without the inclusion of stochastic error in the fuzzy inference system. The effect of error inclusion will also thus be examined in each of the sections. Discussion and chapter conclusions will be given in Section 8.4. All rules developed for the main models reported in this chapter are given in APPENDIX C.

8.2 Fuzzy Inference System with Location Representation Data

This section will detail the results obtained from the fuzzy inference system using attribute location representation data as an input in the FIS. Different models were developed based on the type of inputs and outputs as well as the inclusion of error. The models were developed with binary choice data as the FIS output.

Section 8.2.1.1 details the results obtained without error incorporated in the FIS while section 8.2.1.2 will focus on the results obtained with Gumbel and Normal distributed error in the FIS.

As explained in Chapter 4, the fuzzy inference system (FIS) mapping technique consists of three main components: the input, fuzzy rules and the output. For both the attribute representation methods used in the choice experiment exercise, ‘view’, ‘noise’, ‘sunlight’ and ‘housing service charge’ formed the main inputs for the FIS. Based on the type of model developed, an additional input in the system was the stochastic error.

Keeping the certainty of rule effect (i.e., the FIS correct prediction ability) constant across different models, the number of rules required in a FIS is dependent on the number of inputs, the number of input membership function as well as the number of output membership function. For the location representation method, the respondents’ perception of the attributes (especially, ‘view’ and ‘sunlight’) was expected to vary for a particular location, depending on individual tastes and preferences. Considering the choice modelling assumption that only attribute differences are effectual in respondents’ decision-making, the difference between the attribute ratings were chosen as a more appropriate input in the FIS. Moreover, in cases of high variation in the effect of input on choice, the number of rules required in the FIS is directly proportional to the number of inputs. Hence, using attribute differences in this case implied a substantial reduction in the number of rules required in the system.

To recollect, both the attribute representation techniques comprised of two different experiments. The first experiment had binary and one stage Likert choice scenarios while the second experiment consisted of binary and two stage Likert choice elicitation method. For each of the representation techniques, the binary choice data from both the experiments were combined to form a joint binary choice model in the FIS.

As FIS without error input would be a completely deterministic model, stochastic error was also included in the system. Thus, the following three models were developed based on the type of error incorporation - without error inclusion, with Gumbel distributed error and with Normal distributed error. The following section details the results obtained from the three models, with the binary choice data as the FIS output.

8.2.1 Binary Choice Data

The attribute input values and choice data from the binary choice experiments were used to develop the FIS. The attribute ratings obtained from the respondents were incorporated into the FIS using the difference values. Hence for each of the attributes, ‘view’, ‘noise’ and ‘sunlight’, the difference of the individual’s perception ratings between option A and option B were computed and incorporated into the FIS, as the input value, along with ‘charge’ level difference for each of the choice scenario. This was done in order to reduce the number of rules required for the FIS. Three models were created – a) without error input, b) with Gumbel distributed error and c) with Normal distributed error. Section 8.2.1.1 focuses on the results obtained from the fuzzy logic analysis without error inclusion in the FIS while section 8.2.1.2 focuses on the results obtained from the stochastic error inclusion. For each of these models however, the characteristics of the membership functions for the other input variables (‘view’, ‘noise’, ‘sunlight’ and ‘housing service charge’) remain the same.

Based on the difference between levels of the attributes across the two options (Option A – Option B) for binary choice data, the maximum and minimum difference values were noted and used as a range for the input variable in the FIS. For 'view', the range of [-100 80] was incorporated into the FIS while the range [-80 80] was used for 'noise' and 'sunlight'. For the 'housing service charge', four levels were used in the SP experimental design whose range based on their differences was [-35 45]. This range was used for the 'charge' difference input in the FIS.

Considering the range specified, membership functions for each of the attributes were developed based on the level of differences (attribute value in option A – attribute value in option B). Thus for each of the attributes, three membership functions were developed – high negative difference, medium difference and high positive difference. Following several experimentation with Gaussian, trapezoidal and triangular membership function shapes, the trapezoidal shape was selected for the extreme membership functions (high positive and high negative) in case of 'view', 'noise' and 'sunlight' while a triangular shape was selected for the middle level. Considering that for these attributes, difference levels below and above a certain value act as a threshold, the trapezoidal membership function provides a plausible explanation. In case of 'charge' difference, all the three levels had a triangular shape. The membership function values for each of the level and each attribute can be given as following:

Table 8.1 Membership function values for attributes with location representation method

	View	Noise	Sunlight	Charge
High Negative	[-100 -100 -40 0]	[-80 -80 -40 0]	[-80 -80 -40 0]	[-35 -35 0]
Medium	[-40 0 40]	[-40 0 40]	[-40 0 40]	[-35 0 45]
High Positive	[0 40 80 80]	[0 40 80 80]	[0 40 80 80]	[0 45 45]

The graphical illustration of the input level range and membership function values given in the above table with 'view' as an example is as follows:

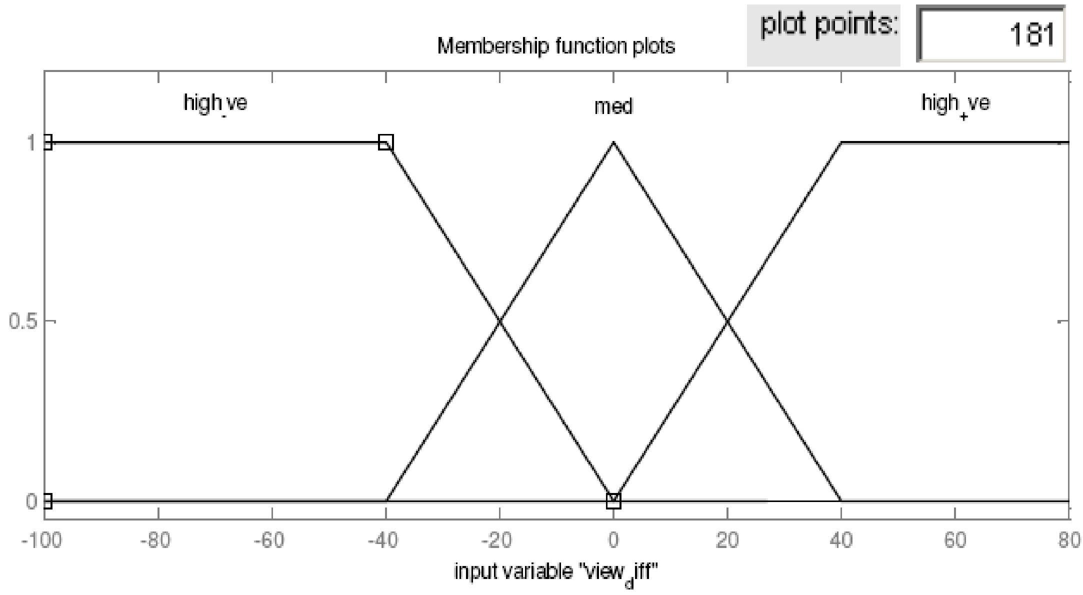


Figure 8.1 Membership function shapes for ‘view’, ‘noise’ and ‘sunlight’

Based on whether error is included in the FIS, different criteria and number of rules were developed. In both cases however (with and without error inclusion), only two choice possibilities were allowed as choice uncertainty was not incorporated into the system. This reflected the binary choice elicitation conducted in the survey. Thus, the output for the FIS consisted of two levels – choice for option A and choice for option B. The range for the output membership function was from [0 1] where higher membership function value indicated greater preference for option B. After experimenting with various shapes of the output membership functions, a smooth Gaussian shape was selected. Graphically, the output variable took the following range and membership function shapes:

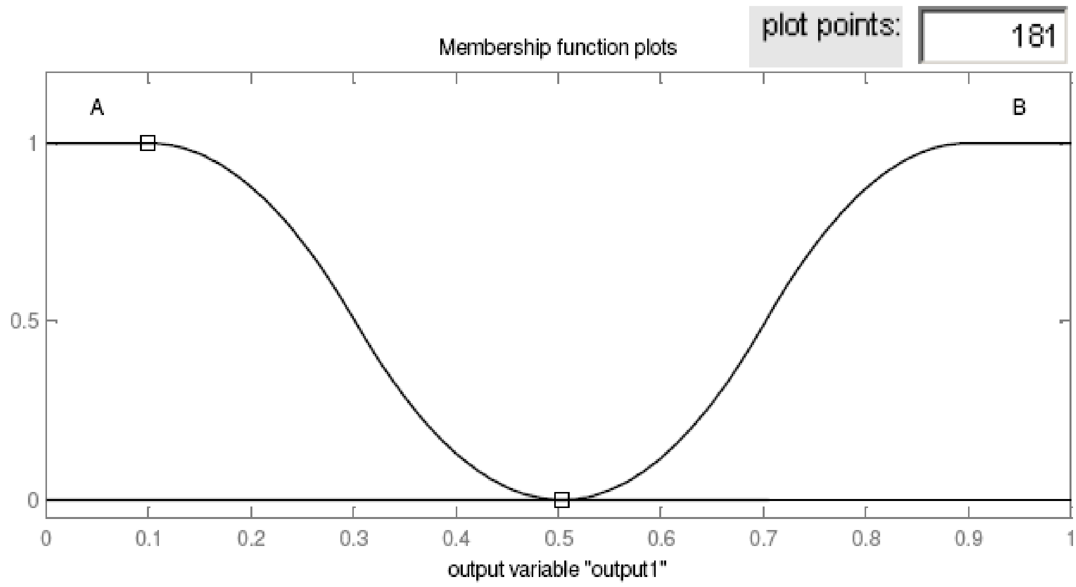


Figure 8.2 Output membership functions

The following sections details the FIS, fuzzy rules and results obtained from the various models. Section 8.2.1.1 focuses on the results obtained without error inclusion in the FIS while Section 8.2.1.2 details the methods and results obtained with error inclusion.

8.2.1.1 Without Error

With only ‘view’, ‘noise’, ‘sunlight’ and ‘housing service charge’ differences as inputs, 20 rules were developed based on the levels and membership function shapes of the input and output, as well as the subjective most likely influence of the input levels on the output. The rules created followed an if-then structure and were formed using the following criteria:

- When the difference level for two or more attributes across the two options is ‘high negative’, irrespective of the ‘charge’ level, the choice is option B
- When the difference level for two or more attributes across the two options is ‘high positive’, irrespective of the ‘charge’ level, the choice is option A
- When the difference level for attributes across the two options is ‘medium’, the ‘charge’ level is deemed to be a crucial factor and hence the choice is

based on the sign of the ‘charge’ difference, where a negative sign implies that option A is chosen and a positive sign implies that option B is chosen

- When three out of four attributes has ‘medium’ level, choice is dependent on the level of the fourth attribute

As this model did not take uncertainty into account, no rules were developed for cases when all the attributes, including ‘charge’, have a medium value.

Moreover, it is to be noted that this model does not take individual tastes and preferences into account. As FIS is an expert system, the model form is dependent on human expertise. The source of expertise can either be from external sources, such as the respondent or from the modeller. As information on the attribute weights were not obtained from the respondents and as each respondent’s tastes and preferences were unknown, the effect of individual attribute on choice could not be estimated and incorporated into the FIS. Hence, rules were only formed for cases where one option was clearly better than another on more than one attribute.

The defuzzified output obtained from the FIS was in terms of the choice for option B. The values were obtained in the range of [0 1] where a higher value indicated a greater preference for option B.

In order to compare the results obtained from the binary logit model, two different statistical criteria were used: the percentage correct predicted and the root mean square error (RMSE). To compute the percentage correct predicted, comparison between the FIS result and the real choice obtained from the data was conducted by transforming the FIS result into a discrete choice. This was carried out by employing the following rule:

If the computed output value is less than 0.5 then choice is option A, if the computed output value is greater than 0.5 then choice is option B, else it is ‘uncertain’

To compute the percentage correct predicted for the BL data, a BL model was estimated without considering panel effects (in order to form an equivalent model to that developed using the fuzzy logic analysis). Based on the coefficient estimates obtained for each of the attributes with this model, the deterministic utility for each of the alternatives was computed (without incorporating the alternative specific constant), which was used to calculate the deterministic forecast for each alternative. The deterministic forecast compares the relative values of the deterministic utilities to predict which alternative will be chosen. Hence, where the deterministic utility of option A was higher than that of option B, the choice using this method was predicted to be option A while in the contrary case, the choice was predicted to be option B. Comparing the choice output thus computed with the real choice gathered from the data, the percentage correct predicted was calculated.

In case of the RMSE calculation for the FL analysis, the difference between the FIS computed output and the real choice obtained from the data required the following transformation:

As the real choice from the data was coded as ‘2’ for option B and ‘1’ for option A while the FIS output was obtained within the range [0 1] where values closer to ‘1’ indicated higher preference for option B while values closer to ‘0’ indicated higher preference for option A, the FIS output was converted to an equivalent scale as that used for the real choice data by adding unity to each computed output. The RMSE was then calculated using the following equation:

$$RMSE = \sqrt{\sum_{i=1}^{n} (a_i - c_i)^2 \Big/ n}$$

Where,
a_i is the real choice observed
c_i is the transformed FIS computed output
n is the total number of observations

To compare the RMSE from the FIS computed output with that obtained from the BL model, the deterministic utility for each of the alternatives was again computed based on the coefficient estimates obtained from the BL analysis. Based on the values of the deterministic utilities, the probabilistic forecasting (Fowkes and Preston, 1991) using the following equations was derived:

$$P(A) = e^{V_A} \big/ (e^{V_A} + e^{V_B})$$

$$P(B) = e^{V_B} \big/ (e^{V_A} + e^{V_B})$$ (Preston and Wardman, 1988)

Where,

P(A) is the probability of option A being chosen

P(B) is the probability of option B being chosen

V_A is the deterministic utility for option A

V_B is the deterministic utility for option B

As the values from the probabilistic forecasting ranged from [0 1][25], while the real choice from the data was coded as [1 2][26], the RMSE was computed by transforming the probabilistic choice forecast computed for option B by adding unity. This was then incorporated into the equation for RMSE and deducted from the real choice obtained from the data[27].

Using the percentage correct predicted, it was found that while the BL model provided 71.85% correct prediction, the FL model without error inclusion gave 71.90% correct prediction. Thus, the FL model gave a slightly higher correct prediction compared to the BL model. However, as the estimation of percentage correct predicted for the FL model required the transformation of the FIS output choice to a discrete choice; the rule employed for the transformation resulted in some 'uncertain' choices as the FIS choice output values in these cases were exactly equal to 0.5. These 'uncertain' choices comprised of 1.86% of the total observations which affected the percentage correct predicted for the FL model.

---

[25] Where values closer to 0 indicated lower choice probability while values closer to 1 indicated higher choice probability for that alternative.

[26] Where '1' denotes choice of option A while '2' choice of option B.

[27] An alternative method of calculating the RMSE would be to transform the real choice obtained from the data into 0 (when the alternative is not chosen) and 1 (when the alternative is chosen.). Using any one of the two alternatives, the probabilistic choice forecast obtained for that alternative can then be deducted from the thus transformed real choice in the RMSE calculation.

Comparing the RMSE across the BL and the FL models, it was seen that while the RMSE for the FL model was 0.436, that obtained from the BL model was 0.427. As the RMSE value closer to zero is considered to be better, it is seen that based on this criterion, the BL model performs slightly better than the FL model. This could possibly be due to the exactly 0.5 choice output value obtained from the FL model which could increase the difference between the real and the computed value in the RMSE computation at least in some of the cases compared to the equivalent BL model.

Based on the percentage correct predicted as well as the RMSE criteria, the results thus reveal that the FIS developed without error inclusion yields as good prediction as the BL model while the RMSE is slightly higher. However, in order to allow for differences in respondents' tastes and preferences, a FIS was also developed with error difference as an additional input. The following section details the results obtained with error inclusion in the FIS.

8.2.1.2 With Error

Stochastic error is an important component in random utility models as it represents those respondent's factors that are unobserved by the researcher. Thus any assumption of the error components has important conceptual underpinnings. In order to allow for the variation in respondent's tastes or other unobserved factors affecting choice, error was incorporated into the FIS. A fuzzy logic model was again built with error difference (option A – option B) as an additional input. Two different distributions of error variance were experimented. In order to facilitate comparison with the results obtained from the logit model, Gumbel distributed errors were generated. While normal distributed errors were also tested in the FIS in order to examine which error distribution better represented the data and gave closer results to the real choices made by the respondents. Both the error distributions were incorporated into the FIS as error differences. Hence for both the choice alternatives, error values were generated with the specific distribution and an error difference across the two options was computed. In case of the Gumbel distributed error, it has to be noted that the difference between the error values had a logistic distribution.

Based on the specification used for the logit model in terms of standard deviation and distribution, stochastic error values for both the options were simulated and the difference computed in order to be incorporated in the FIS. Thus, Gumbel distributed error values with standard deviation of 1.28 was used to simulate the error values for both the options in order to compare FIS results with that obtained from the logit model.

The rules for the FIS model with both the different forms of error distribution were based on the following criteria:

1. For all rules with certain choice outcome, no error was incorporated into the fuzzy rule. Thus, the 20 rules outlined in the model without error inclusion were included in this model without any modification
2. For rules where there were uncertain outcomes based on the difference level of the attributes across the two options, the effect of error was included such that when the error difference (option A – option B) is high –ve, option B will be selected, while when it is high +ve, option A will be chosen

Thus this model considered the error effect on choice when the attribute difference value could not give a certain choice outcome based on the rules criteria employed in Section 8.2.1.1. Based on these criteria listed for incorporating error effect, 42 rules were developed.

By computing the maximum and minimum values for the error difference across the two options, the range for the error input was formed. For the Gumbel distributed error values, a range of [-10 10] was used in the FIS and three membership functions were formed which took the following shape:

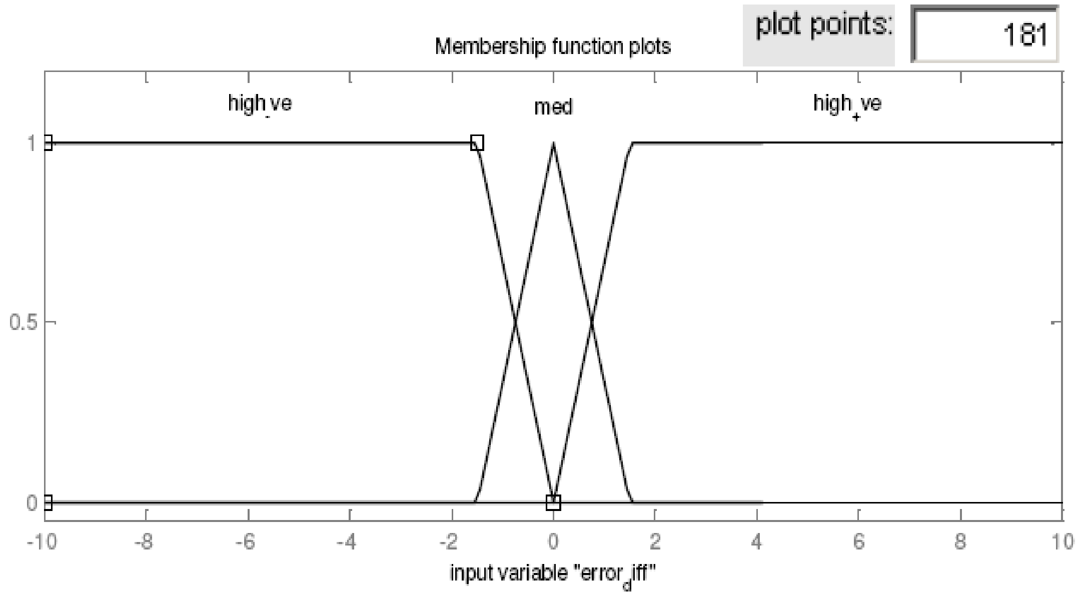


Figure 8.3 Membership functions for Gumbel distributed error input

As the difference between the Gumbel distributed errors has a logistic distribution, the membership function for 'medium' error difference has a very narrow range. The output membership function for the FIS followed from that used in the model 8.2.1.1. As uncertain outcomes were not used in the rules, only two membership functions were developed as is seen in Figure 8.2. Again using the same rule for choice transformation as outlined in Section 8.2.1.1, and comparing the results with the choice made by the respondents, it was observed that the FIS developed gave 71.73% correct prediction, which is slightly lower than the result obtained without considering the error effect.

As the Gumbel distributed error with standard deviation 1.28 does not affect the choice outcome in logit simulation for sufficiently large sample, the effect of error in logit choice simulation was not incorporated for comparison and hence the stochastic utility was not used to compute the percentage correct predicted for the equivalent logit model with Gumbel distributed errors.

In order to examine the effect of a different error distribution on fuzzy logic choice outcome, Normal distributed error for both the options were simulated and the difference between the two was included as an input in the FIS. The range and the shape of the membership function for the Normal distributed error difference varied from that for the Gumbel distributed error terms. The range of the error difference varied due to the values of the simulated stochastic error while the shape of the membership function was altered in order to follow the specification of the Normal distribution. The range and shape of the membership function is given in Figure 8.4. Using the same set of rules as that with the Gumbel distributed error terms; a FIS was run which yielded 72.07% correct prediction, which is slightly better than the model without error inclusion as well as with Gumbel distributed error.

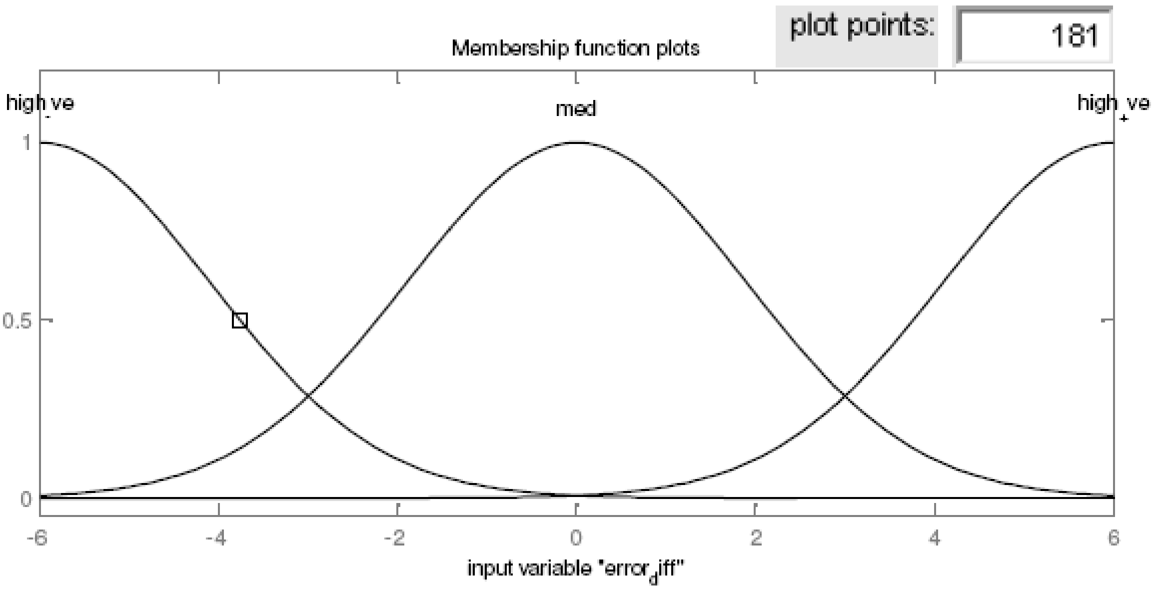


Figure 8.4 Membership function shape and range for normal distributed error

To summarise, the following results were obtained from the different models:

Table 8.2 Summary of results for binary choice, location representation method in percentage correct predicted as well as RMSE[28]

	Logit (% correct predicted)	FL (% correct predicted)	Binary Logit (RMSE)	Fuzzy Logic (RMSE)
Without Error	71.85%	71.90%	0.427	0.436
Gumbel Error	---	71.73%	---	0.445
Normal Error	---	72.07%	---	0.436

From the above table it can be seen that the FL model with Normal distributed error as an additional input gave the highest correct prediction. It is also evident that the results obtained from the three different models with the FL method do not show large difference in their respective rates of correct prediction. Thus, the FIS developed without error inclusion provides satisfactory rate of correct prediction compared to the models with error inclusion especially as the number of rules developed for the former FIS model is substantially less than those for the latter two models.

While this section focussed on the analysis of the binary choice data using the location method, the next section will outline the results obtained from the linguistic representation method.

[28] With methods employed for the respective models for each of the statistical measure as detailed in Section 8.2.1.1

8.3 The Linguistic Method

The linguistic method of attribute representation used linguistic categories of the attributes. Using this representation technique as well, respondents were asked to provide a perception rating from 0-100 for all the linguistic categories of the attributes. Using the relevant perception ratings for each of the choice scenarios, different models were built based on the method of attribute input in the FIS. As developed for the location method, a FIS was built based on the level of difference between the attribute perception ratings. As an important characteristic of the FL method is its ability to handle linguistic or imprecise data and use heuristic rules to judge choice outcomes, a FIS was also built based on the absolute linguistic categories of the attributes as used in the choice scenario. The results from both the methods are given in this section.

8.3.1 Binary Choice Data

The methods developed to build this model followed closely to that developed for the location method. However, some variation was carried out based on the range and number of the input membership categories. As the attribute levels using this representation technique were different than that with the location representation method, the input levels in the FIS changed accordingly. As there were two levels for ‘view’ while three for ‘noise’ and ‘sunlight’ in the SP choice experiment, the FIS developed had the following levels and ranges for the input:

Table 8.3 Attribute input levels for linguistic difference design

	View	Noise	Sunlight	Charge
Low	[-30 -10 0 0]			
High –ve		[-60 -60 -20 0]	[-80 -80 -20 0]	[-45 -45 0]
Medium	[-60 -60 -30 0]	[-20 0 20]	[-20 0 20]	[-45 0 45]
High +ve		[0 20 80 80]	[0 20 60 60]	[0 45 45]

From the above table it is seen that two levels were formed for ‘view’: low and medium difference which have a trapezoidal shape while for ‘noise’, ‘sunlight’ and ‘housing service charge’, three levels: high –ve, medium and high +ve were formed. For ‘noise’ and ‘sunlight’ the levels high –ve and high +ve have a trapezoidal shape while the level ‘medium’ has a triangular shape. All levels for ‘charge’ have triangular membership function shape.

8.3.1.1 Without Error

Based on the input ranges and membership functions outlined in Table 8.3, 21 rules were developed. The choice output again had two levels as ‘uncertain’ outcome was not incorporated into fuzzy rules. The membership functions of the output levels were as given in Figure 8.2. From Table 8.3 it can be seen that depending on the levels for view, option B is either much better or slightly better than option A. Taking this into consideration, rules were developed such that when the level for ‘view’ is ‘low’ and ‘sunlight’ is ‘medium’, the level for ‘noise’ (in case when it is not ‘medium’) formed a crucial factor in the choice of the output. When ‘view’ was ‘medium’, and ‘noise’ ‘high –ve’, option B would be selected. In the case where ‘view’ was ‘medium’ and ‘noise’ was not ‘high –ve’, the levels of ‘sunlight’ and ‘charge’ were the deciding factors in the choice of the outcome.

The percentage correct predicted and the RMSE were again used to compare the results obtained from the FIS as well as the BL model. The methods applied to compute these statistical measures followed closely as that outlined in Section 8.2.1.1. Thus, in order to convert the FIS choice output to a discrete choice for the percentage correct prediction, the following rule was used while the computation of the RMSE comprised of adding unity to the FIS choice output and deducting this from the real choice obtained from the data in the RMSE computation equation:

If the output value is less than 0.5 then choice is option A, if the output value is greater than 0.5 then choice is option B, else it is ‘uncertain’

Following the method outlined in Section 8.2.1.1 to compute the percentage correct predicted and the RMSE from the linguistic BL model, these statistical measures were then compared to that obtained from the FL model. Comparing the percentage correct predicted across the two models it was found that the FL model provided 69.73% correct prediction while the equivalent linguistic ratings BL model gave 67.16% correct prediction. With this method of attribute representation it is therefore seen that using the percentage correct prediction criterion, the FL model performs slightly better than the BL model.

Comparing the results obtained from the RMSE computation, it was found that in case of the FL model, a RMSE of 0.473 was obtained while in case of the BL model, a RMSE of 0.515 was obtained. Thus, under this statistical measure too, the FL model is again found to perform better than the equivalent BL model.

As the FL method is well-suited for imprecision and vagueness, the effect of linguistic variables in the FIS design was experimented by defining the FIS input variables based on absolute linguistic categories used in the SP choice scenarios. In this case, the attributes ‘view’, ‘noise’ and ‘sunlight’ were incorporated in the FIS for both the options (A and B) while ‘charge’ was incorporated through the difference method (option A – option B). Thus, in this case, a total of seven inputs were considered in the FIS. Based on the levels adopted in the SP experiment, two levels of ‘view’ and three for ‘noise’, ‘sunlight’ and ‘charge’ formed the input membership functions.

While the pi shaped spline based curve was used for the absolute linguistic categories for the ‘view’, ‘noise’ and ‘sunlight’ attributes (with the range [0 100]), the ‘charge’ difference membership functions took a Gaussian combination[29] shape (with the range [-45 45]). With this specification of FIS, several rules were experimented considering the choice scenario used in the survey. It was found that with the following three rules, about 64.03% correct prediction was obtained without error inclusion:

---

[29] While the symmetric Gaussian function is based on the values that determine its mean and the spread, the combination Gaussian function uses two of these values each to determine the shape of the left and the right most curves (MATLAB HANDBOOK, 2007).

1. If NOISEA is noisy and SUNLIGHTB is very good then choice is B
2. If VIEWA is good then choice is A
3. If NOISEB is quiet and SUNLIGHTB is very good then choice is B

The RMSE with these set of rules was found to be 0.479 which is slightly worse than that obtained from the FL ratings difference model but better than that obtained from the BL model.

By adding another five rules to the preceding three rules, it was observed that the percentage correct predicted improved to 65.56% while the RMSE slightly lowered to 0.471.

Based on this analysis, it can be observed that the FL model deals very well with linguistic categories of attribute input. Moreover, the ability of only three rules to correctly predict about 64% of the real choice obtained has some implications for the representation method as well as the respondents' decision making process. As the FL system is better able to deal with linguistic forms of data, the high correct prediction could be due to the relative ease for the modeller to deduce the consequence of each rule. However, besides the ease of developing FIS rules with this type of attribute input, the high prediction rate obtained using these rules also sheds some light on the decision-making criteria largely adapted by the respondents. Based on this type of analysis therefore, it is clearly seen that combinations of attribute levels for 'noise' and 'sunlight' play an important role in respondents' decision-making while the level of 'view' for alternative A on its own is a significant contributor to the decision-making process. Using the FL method in this case, the significant effect of a few rules with important combination of attributes and their levels have been thus observed.

Besides the strong effect of certain rules as predictors of respondents' choice, the model results also indicate that the choice process could be a relatively easy exercise for the respondents with this type of attribute representation. Moreover, a lower amount of randomness in respondents' choice when using this form of representation is also indicated. Thus, it is observed that the consequence of the attribute level combinations is more deterministic in this case. This inference will be tested in the next section when error difference forms an additional input in the FIS model and the choice output obtained is compared to the equivalent FIS model without error inclusion.

This section was aimed at providing results from the FIS model developed for the linguistic representation data when the FIS input comprised of attribute ratings difference as well as the absolute linguistic categories of ‘view’, ‘noise’ and ‘sunlight’. In both these cases, stochastic error did not form an additional input in the FIS. Comparing the results obtained from the former model (FL with attribute ratings difference as input) with the equivalent BL (linguistic ratings) model in terms of percentage correct predicted and RMSE, it was seen that for both these statistical measures, the FL model performed slightly better than the BL model. Examining the model developed with the absolute linguistic categories as input, it was seen that few heuristic rules were able to correctly predict a large amount of respondents’ choices. Thus it can be concluded that for the linguistic representation data, the FL analysis offers a valuable insight and complements the logit analysis.

8.3.1.2 With Error

In order to incorporate the effect of stochastic error on respondent’s choice, a FIS was built with Gumbel and Normal distributed error differences across the two alternatives as an additional input.

The input variable’s membership functions for ‘view’, ‘noise’, ‘sunlight’ and ‘charge’ were as given in Table 8.3. For the Gumbel distributed error, the additional error input had three levels with the range [-9 10] and the following shapes of the membership function:

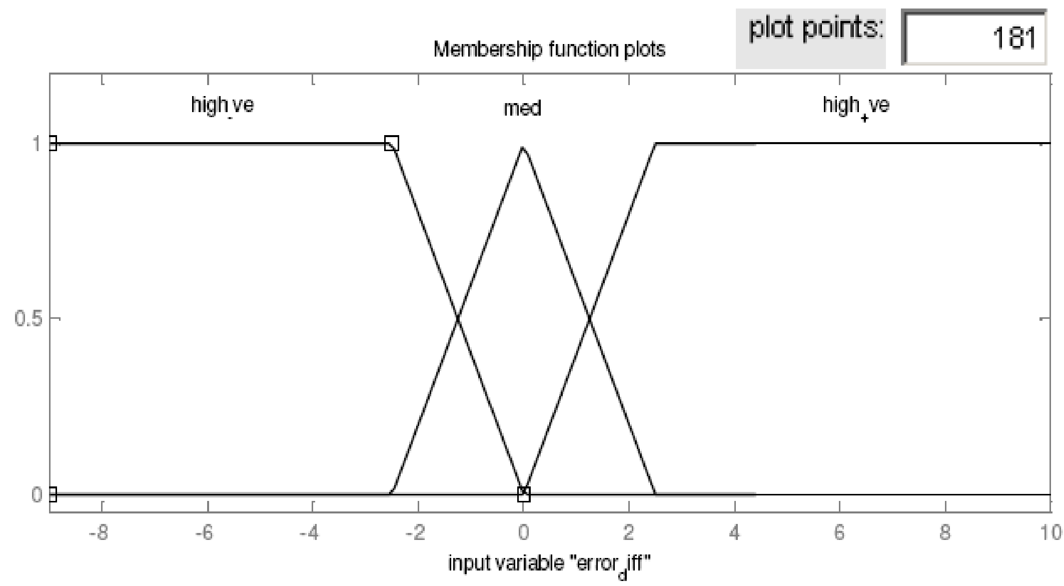


Figure 8.5 Gumbel error input for linguistic representation

In case of the Normal distributed error, the same range and membership functions were used as given in Figure 8.4 for the location method. The error values included in the FIS in this case were same as those used for the location method. Though the choice of the error values does not imply same error assumption for different people across the two representation methods, using the same error values under a particular standard deviation assumption should not theoretically have a major effect on the aggregate choices and hence the same values were used. The rules were formed such that when two attributes (from 'view', 'noise' and 'sunlight') have values favourable towards a particular alternative, that alternative would be selected. However, when two of the attributes gave conflicting values in terms of preference towards an alternative, the value of 'charge' difference was used as a deciding factor in the choice while when the 'charge' difference was 'medium' and the other attributes suggested conflicting preferences, the values of the error difference were used in the fuzzy rule. Thus, a total of 26 rules were created for the FIS. With the Gumbel distributed error assumption, the FIS model gave 69.73% correct prediction. Using the same Normal distributed error values as that in the location method, the model gave 69.79% correct prediction which is slightly higher than that obtained using the Gumbel distributed errors.

To summarise, the following results were obtained from the different models:

Table 8.4 Summary of results for binary choice, linguistic representation method for percentage correct predicted and RMSE[30]

	BL ratings (% correct predicted)	FL – diff. model (% correct pred.)	FL – abs. 3 rules (% correct pred.)	BL (RMSE)	FL – diff. model (% correct pred.)	FL – abs. 3 rules (% correct pred.)
Without Error	67.16%	69.73%	64.03%	0.515	0.473	.479
Gumbel Error	---	69.73%	---	---	0.471	---
Normal Error	---	69.79%	---	---	0.471	---

Compared to the results obtained from the logit ratings model, it can be seen that the FL model gives a slightly higher improvement in the percentage correct predicted as well as lower RMSE. However, interestingly, not much difference is found between the FL models without error and with Gumbel distributed error in terms of percentage correct prediction, though the model with Normal distributed error performs slightly better than the other two. Comparing the RMSE value obtained across the different FL models, it is observed that the RMSE value for model without error inclusion is very slightly worse than that obtained for models with error inclusion. These findings imply that with the FL model, there is again lesser dependence on random error to predict choice using this form of attribute representation.

Examining the results obtained from the logit and FL models, using both the location and linguistic methods, it can be seen that higher improvement of FL over logit ratings model is found in the case of the linguistic method rather than the location method, while the difference between the FL models (with and without error inclusion) for each attribute representation is slightly higher in the case of the

[30] With methods employed for the respective models for each of the statistical measure as detailed in Section 8.2.1.1.

location method using the percentage correct predicted measure. Thus, it can be concluded that the linguistic method is easier for respondents to understand, with choice decisions based on more heuristic rules and lesser level of randomness in choice.

When the percentage correct predicted and RMSE are compared across the two attribute representation methods, it is seen that the FL linguistic model gives a slightly lower prediction than the FL location method. This could possibly be due to the variation observed in respondents' subjective ratings for the linguistic method. Thus, considerations on the variation in respondents' ratings need to be taken into account in related future application of FL.

Based on the comparisons made across the results obtained for the linguistic and location FL models, two important inferences can thus be drawn:

a) In case of the location method, the error affecting choice prediction could be more due to the variation in respondents' preferences, which could be dependent on the choice task
b) The error affecting choice prediction with the linguistic method could be more due to the respondents' variation in attribute ratings

Thus, in case of both the location as well as the linguistic representation methods, the FL method has been observed to predict respondents' choices sufficiently well, with additional insight on the heuristics that significantly affect respondents' decision-making under the linguistic representation method.

As FL has the capability to handle multiple linguistic outputs, a FIS was developed for the five point Likert scale data. The next section outlines the result obtained for one stage Likert choice elicitation technique.

8.3.2 Five Point Likert Data

Using the five point Likert choice elicitation data, a FIS was developed to check the suitability of the FL technique on this type of choice data. Using the data from the location representation method and five point Likert choice elicitation, a FIS was developed based on the differences in the attribute levels across the two alternatives, without error inclusion. The membership functions of the attribute inputs followed from the shape adopted in the binary choice location representation model. The output membership function however had five levels according to the levels used in the choice experiment. Thus, the choice outcomes were: Definitely A, Probably A, Uncertain, Probably B and Definitely B. The membership function shape of the output can be given as follows:

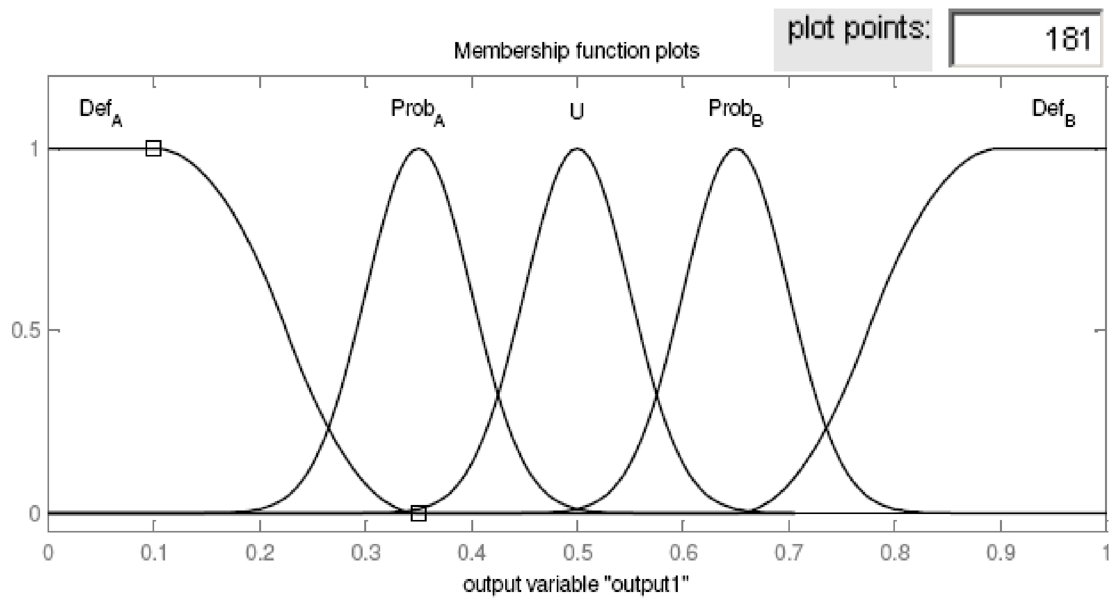


Figure 8.6 Output membership function shapes for five point Likert data

50 rules were formed based on the attribute levels and the output obtained from the FIS was converted to choice outcome using the following rule:

If output value < 0.3 then choice is Definitely A, if output value is between 0.3 – 0.45 then choice is Probably A, if output value is between 0.45 - 0.55 then choice is Uncertain, if output value is between 0.55 – 0.7 then choice is Probably B and if output value is > 0.7 then choice is Definitely B

Using the FIS developed and this conversion rule to obtain definite choice outcomes, it was found that the FIS for the location method obtained 33.4% correct prediction. The low level of correct prediction is due to higher number of output levels that need to be correctly matched. Conducting the FIS Likert analysis for the linguistic method with 24 rules, gave 28.06% correct prediction for the model where attribute differences were considered as inputs and 28.19% correct prediction where absolute levels design was adopted for ‘view’, ‘noise’ and ‘sunlight’ with ‘charge’ difference, using 15 rules. In case of the linguistic method, vast variation in the attribute ratings given by the respondents could result in further deterioration in the prediction levels using the Likert scale elicitation. Hence it can be concluded that though FL has the capability to handle linguistic output categories, this method in its current application form is not suitable in case of the Likert elicitation method. In order to successfully apply this analytical method to Likert data, further work needs to be done in terms of FIS development and application of genetic algorithm techniques which could form interesting future research work in this area. Moreover, in order to aid the membership function formation for the Likert data, future surveys applying this technique on Likert data can elicit respondents’ numeric rating along the Likert scale. Further work can also be done to develop methods for calibrating the preference certainty levels and developing methods to link the input and output through Sugeno type FIS.

8.4 Discussion and Conclusions

Fuzzy logic techniques are suitable to analyse data with imprecision, vagueness and uncertainty. This chapter employed the FL method to develop a heuristic based approach and to evaluate its pertinence to analyse choice data. FL analysis was conducted across different representation and elicitation techniques, with findings giving significant insight for choice modelling in the context of the experiments and the application.

For the location representation method using the attribute level differences as input, the FL model was seen to perform nearly as well as the BL model in terms of percentage correct predicted as well as the RMSE. In case of the linguistic representation method, where the attribute level differences were incorporated in the FIS as inputs, the FL model showed higher improvement over the BL model as compared to the location representation data. Thus with both the location and linguistic representation methods, with attribute level difference as the input in FIS, the FL model was seen to perform sufficiently well, implying that this analytical method can be adequately applied when choice prediction level and the associated market share are the required output from the exercise. In case where coefficient estimates are required using the FL method, help from genetic algorithm and/or fuzzy c-means clustering could be required to form complex hybrid systems.

In terms of the decision making process, different requirements were discerned based on the type of model developed. For both the location and the linguistic representation methods with difference in attribute ratings as an input in the FIS, considering the effect of all attributes on the choice outcome was found to be more appropriate for model development while in case of the linguistic model with absolute levels of the attributes across the two options as input, it was found that certain attributes and their combinations were more significant in respondents' decision making.

As an important characteristic of the FL method is its compatibility with linguistic variables, it was observed that in case of the linguistic representation where the attribute levels were incorporated in the FIS based on their absolute linguistic categories, the FL method provided sufficiently high rate of correct prediction based on very few rules. Though the FL method as applied in this thesis does not provide coefficient estimates that can be compared with the output obtained from the logit analyses, the application of the FL method however sheds a different light on how respondents conduct the decision-making process. In case of the linguistic FL model with absolute categories, it was seen that only three rules with certain combinations of attribute levels correctly predicted about 64% of choice. In this case therefore, the FL model revealed that certain combinations of attributes such as ‘noise’ and ‘sunlight’ significantly affected respondents’ decision-making while the level of ‘view’ for option A on its own had a major effect on respondents’ choices. While the logit model provides coefficient estimates which sheds light on the relative effect of each attribute on choice, the detail of which combinations of attributes and levels are more significant in the decision-making have been easily captured by the FL model. Moreover, the percentage correct predicted obtained from this model is more striking when it is considered that only three rules were present in its FIS rule structure. As the role of heuristics in choice modelling is gaining increasing attention (Leong and Hensher, 2011), findings such as these can be helpful to enhance the logit model where alternative model forms of estimation such as those incorporating non-linearity and heuristics could be experimented.

Based on the results obtained from the FL model with attribute differences as input as well as the absolute linguistic category model and examining the rules incorporated in their FIS, it can be concluded that the FL method could be used to capture any potential non-compensatory decision-making or heuristic rules employed by the respondents and with the insight gained from its structure and result, help to develop a more realistic choice model for estimation.

9 CONCLUSIONS

9.1 Thesis Summary

This chapter gives an overview of the thesis and the conclusions obtained from the research. The research objectives are given in Section 9.2 followed by the main research findings in Section 9.3. Section 9.4 will outline the research implications while the limitations of the research are outlined in Section 9.5. Future research work arising from this study will be given in Section 9.6.

From the various external effects of transportation, noise is one of the key factors adversely affecting people and several methods have been applied to estimate economic value of noise. Increasingly, SP methods have been applied to evaluate transport noise. While this method offers an innovative approach for noise valuation compared to indirect valuation techniques, a crucial problem lies in the optimal representation of the attribute. Another factor affecting the SP method is the key assumption of respondent's known and consistent preferences, which have been refuted by various experimental economics and psychology studies. The aim of this research has thus been to encompass and link the aspects of descriptive invariance violation and uncertain, inconsistent preferences. By conducting a SP survey in Portugal, one of the countries with highest traffic noise level in Europe, this research incorporated the different methods of attribute representation along with varying methods to elicit respondents' preference uncertainty within the framework of noise valuation. The specific research objectives met in order to follow the research aim are outlined in the next section along with the results obtained.

9.2 Research Objectives

Traffic noise in SP studies has been represented using several methods such as physical measurement, percentage changes as well as in some cases, through exposure of sound levels to the respondents. The reference to relative noise levels based on different locations has also been applied in choice experiments. While noise representation as physical measure or percentage change is difficult for respondents to understand, methods such as noise level exposure and reference to location offers an interesting alternative. However, few examples are found in the literature which compares the effect of different representation methods on the respondents. Moreover, though the effect of choice set on task complexity has been relatively well-studied, little attention has been paid to the effect of choice-set characteristics on the level of respondents' certainty. The specific aim of this research was thus the following:

Research Aim: to examine the effects of different attribute representation and preference elicitation methods on respondents' choice set understanding and level of preference uncertainty

This research aim comprised of several levels of evaluation as the effect of different representation techniques could affect the significance of the parameter estimates, the level of respondent uncertainty as well as the decision making process employed by the respondent. Thus, in order to realise the aim, the following objectives were outlined:

Research Objective 1: *to examine whether different methods of attribute representation affects respondents' understanding of that attribute as well as its valuation*

The purpose of this objective was to identify whether the method of attribute representation affects the level of attribute understanding. As the statistical significance of the coefficient estimate was taken to reflect the level of attribute understanding, *ceteris paribus*, a statistically higher and significant coefficient estimate implied a better level of comprehension.

In order to meet the objective, attributes were represented in location and linguistic terms in the choice experiment. Based on the attribute ratings obtained from the respondents, results from the pooled BL-MNL model indicated that the 'noise' parameter estimate had a higher and more significant value under the linguistic representation while 'view' had a higher significant value under the location representation method.

In terms of attribute valuation for 'noise', a difference of about Euro 0.98 was obtained across the linguistic and location ratings method, with higher valuation obtained under the linguistic method.

Thus, the coefficient estimate for 'view' was found to have a higher significance with the location method while that for 'noise' and 'sunlight' had a higher significance value under the linguistic method; the valuation obtained for 'view' and 'noise' was higher under the linguistic method while no substantial difference was obtained for valuation for the 'sunlight' attribute across the two representation methods. These findings thus revealed the effect of different attribute representation methods on the level of attribute understanding (as reflected by the statistical significance of the coefficient estimates) as well as on the relative attribute valuation. In case of 'noise' the findings indicate that this attribute is better represented by the linguistic method with higher valuations also obtained from this representation method. Based on the findings obtained from the statistical significance, it can be concluded that 'view' is better represented using the location method while 'noise' and 'sunlight' are better represented using the linguistic method.

<u>Research Objective 2</u>: *to assess which method of attribute input is more suitable for each attribute representation method*

In order to examine whether certain modelling techniques are more suitable for a specific representation method, data from the location and linguistic representation methods were incorporated in the model using subjective ratings as well as dummy categories. Across all the different logit models developed, it was consistently found that better model fit was obtained with the ratings method in case of the location method while incorporating the data as dummy categories provided a better fit for the linguistic method. Thus the results clearly revealed that based on the different forms of representation, different methods of data modelling needs to be applied. Thus, in case of the location method, the ratings data input method is more appropriate in application while in case of the linguistic representation method, the dummy data input method should be applied.

Research Objective 3: *to examine whether different methods of attribute representation affects respondents' preference certainty level*

Ceteris paribus, one of the stochastic reasons for respondents' different levels of preference uncertainty was taken to arise from choice-set characteristics viz., the method of attribute representation. In order to test this hypothesis, respondents were offered Likert scale questions in addition to the binary choice question to indicate their level of preference certainty. The pooled binary and Likert choices were analysed using the NL and ECL models to examine the need for more complex error assumptions.

Results from the NL model showed substantial improvement over the base BL-MNL model for all the different specifications while the ECL model showed improvement in model fit for the location ratings and the linguistic dummy specifications. Comparing the LR value obtained from the NL and ECL models it was found that a greater improvement in model fit is obtained for the location representation method, indicating that this method of attribute representation affects the stochastic factor associated with respondents' preference certainty level. Thus, the location method is more complex for respondents to understand and applying this method of representation would imply examining whether complex error assumptions need to be made through the application of the NL and the ECL models.

In case of the linguistic dummy specification, a relatively lesser improvement in model fit was obtained by applying the NL and ECL models, implying that the linguistic representation method is relatively easier for respondents to understand, with lesser dependence on the stochastic component to explain preference certainty levels.

Research Objective 4: *to test the ability of different preference elicitation methods in capturing choice information*

While eliciting respondents' preference uncertainty formed a crucial aspect of the research, it was also important to test whether different methods of preference elicitation can affect choice formation. Hence, three different methods of preference elicitation (the binary, one stage and two stage Likert methods) were adopted in the survey. The choice data obtained from both these elicitation methods were pooled and scale parameters were estimated to indicate the relative effect of the different preference methods in terms of their associated variance. Based on the scale parameter obtained from the BL-MNL model, it was found that in case of the location ratings method, the one stage Likert method captured respondents' preferences most precisely while in case of the location dummy as well as the linguistic ratings and dummy specifications, no significant difference in preference elicitation methods was obtained. In context of mixed attribute representation method as well as selecting the single most preferred elicitation method, this finding implies that the one stage Likert method is the most suitable preference elicitation technique to apply irrespective of the attribute representation and the data input method used as it is capable of capturing the respondents' preference most precisely in case of the location ratings method, without any significant difference with the other preference elicitation techniques in case of the linguistic representation method.

Research Objective 5: *to examine the performance of heuristic based method to predict choice outcome across various methods of attribute representation*

In order to examine the role of heuristics in respondents' decision-making across the two representation methods, a heuristic based approach was experimented through the application of the fuzzy logic (FL) technique. Analysing the fuzzy inference system requirements as well as the rules developed, it was observed that in case of the linguistic representation method where absolute categories of the attribute levels were incorporated in the design, fewer rules were required to obtain a reasonably high level of correct prediction. Models based on error inclusion revealed that in case of the location method, error effect could be more due to respondents' variation in tastes while in case of the linguistic method this could be more due to the variation in respondents' subjective ratings for each linguistic level. Results from the FL analysis revealed that while this method can provide as good prediction as the equivalent logit model in terms of percentage correct predicted and RMSE, in case of the linguistic representation method, certain combinations of attributes and levels had a significant effect on respondents' choices where the use of heuristics in decision-making was strikingly seen to be in effect, when absolute categorical levels of attributes were incorporated as input. The FL analysis thus provided insight on the decision-making process largely employed by the respondents in case of the linguistic representation method, which can have implications for enhancing future logit models by examining the presence of any potential heuristics or non-linearities in the choice model.

9.3 Main Research Findings

The main research findings obtained can be summarised as follows:

- Based on the results obtained from the BL-MNL model, it was seen that the coefficient estimate for 'noise' had a higher statistical significance with the linguistic representation method while that for 'view' had a higher statistical significance with the location method, implying that 'noise' is easier to understand under the linguistic representation while the location method provides a clearer understanding of the 'view' attribute

- Results from the different modelling techniques revealed that while the ratings method was more suitable for the location representation method, the dummy variables modelling technique was more suitable for the linguistic representation method

- Computation of attribute valuation revealed that ‘noise’ is the most valued attribute under both the location as well as the linguistic ratings method

- Based on the ASC values, a higher level of randomness in choice was found in case of the location ratings, one stage Likert method while a lower dependence on ASC was found in case of the linguistic representation method, implying a lesser dependence on stochastic effects to explain choice in case of this representation method

- The panel error component from the BL-MNL model revealed that correlation of preference certainty levels with respondents’ characteristics was present for all the different attribute representation and data input methods. However, the preference certainty levels which had higher correlation with respondents’ characteristics varied based on the model specification (in terms of representation and preference elicitation method used)

- Interaction analysis of noise variable with day and night time noise perception and annoyance levels revealed that while effect of ‘very much’ day time noise annoyance level was found for both location and linguistic representation methods, the level and type of effect of noise perception and night time noise annoyance levels varied across the different representation methods as well as the different perception and annoyance categories

- Compared to the findings obtained by Arsenio (2002), it was found that a higher difference in WTP estimate is obtained for ‘noise’ while least difference is obtained for ‘sunlight’. Moreover, the current study also showed a significantly better model fit than that obtained by Arsenio (2002). While Arsenio (2002) demonstrated that the location ratings method gives a better model fit than model with physical measure of noise, this study has found that the linguistic representation of noise is clearer for respondents to understand compared to the location method; thus furthering the research finding obtained by Arsenio (2002)

- Results from the NL analysis revealed that a high correlation among error variances was obtained for the high preference certainty levels across all the preference elicitation methods with the linguistic representation while in case of the location representation, this finding was obtained only in the case of the two stage Likert method. Correlation in error variance for the one stage Likert elicitation with this representation method was mainly due to correlation in the intra-respondent observations which when captured by the panel effects, resulted in lower nest coefficient

- Results from the ordered logit model revealed that this method can be used to analyse both the one and two stage Likert elicitation data provided suitable data input method is used for each of the attribute representation methods

- Results from the LR test for both the NL as well as the ECL models showed a larger improvement in model fit compared to the BL-MNL model in case of the location representation method. The two analytical models thus indicated the need for greater flexibility in error assumption with the location method than the linguistic method, revealing the effect of different attribute representation techniques on respondents' understanding and randomness

- While the NL model revealed higher correlation among higher preference certainty levels, the ECL model showed higher error variance for lower levels of preference certainty. These results thus reveal the structure of error associated with each level of preference certainty with greater similarity in the way alternatives are perceived when respondents are more certain and higher randomness when complete certainty is not present

- Fuzzy logic analysis gave a further insight on the method of decision-making associated with each attribute representation method. While both the location and linguistic methods with attribute ratings difference as input, required consideration of all the attributes in the FL rules structure, in case of the linguistic representation where absolute attribute levels were used as inputs, the FL analysis gave a relatively high value of correct prediction from very few rules, highlighting the role of heuristics that could be employed with the linguistic method.

9.4 Research Implications

Based on the research results obtained, the following implications can be drawn on the design and analysis of surveys:

1. 'Noise' and 'sunlight' are better represented linguistically while 'view' is better represented with the location method. This implies that using a mixed representation mode can probably provide relatively better understanding to the respondents and thus improve the model fit

2. In terms of modelling the attribute values, attribute ratings from the location representation method are more suited to be treated as continuous variables whereas the application of dummy variables is more appropriate in the case of the linguistic representation method. This implies that when applying a mixed representation method, 'noise' and 'sunlight' are better modelled as dummy variables whereas 'view' is better modelled using the ratings method

3. For most of the model specifications (in terms of attribute representation and data input method), no substantial difference was obtained for the different preference elicitation method in their ability to capture preference information based on their scale parameter value. However, as in case of the location ratings method, the one stage Likert method was found to be

most suitable. This method can thus be applied as the preferred method of preference elicitation

4. Results from the panel error component revealed that failing to account for panel effects could affect model conclusions as well as the nest coefficients in case of the NL model

5. In terms of model analyses, results have revealed that ordered logit analysis is suitable for both one and two stage Likert method if the most suitable data input method is used for each of the representation methods. As different insights were obtained from the NL and ECL models, the application of both these methods become imperative to gain a complete understanding of the error effects

9.5 Research Limitations

As in any research, several research limitations have been recognised and can be outlined as follows:

- While the study employed alternative means of attribute representation for the ‘view’, ‘noise’ and ‘sunlight’ attributes in order to examine which method of representation is most suitable for each of the attributes, this technique implied that pure effect of alternative ‘noise’ representation techniques could not be isolated and this can thus be regarded as a limitation of the study

- The one and two stage Likert methods were employed in order to examine the effect of alternative methods of preference uncertainty elicitation. While the scale and language used in the two stage Likert method for lower preference certainty level closely combined the ‘uncertain’ and ‘probably’ levels of the one stage Likert method, the exclusion of the ‘uncertain’ alternative in the two stage Likert and some variation in the scale and language of the preference certainty levels implied that these two elicitation methods cannot be regarded as exactly same linguistically

- While some distinction between the preference uncertainty and strength of preference indicator has been outlined in the thesis, some ambiguity remains between the two measures as a strength of preference level could be indicative of the level of preference certainty. As the respondents in the study were not provided with a distinction of the two measures during the survey, this ambiguity can thus be recognised as a limitation of the study

- As the face-to-face paper based interview employed in the current study due to resource limitations did not incorporate the respondent's current apartment location in the choice experiment, this aspect limits the comparison with the previous study which employed a pivotal design (with a computer assisted personal interview) based on the respondent's current apartment

- The exclusion of the question to evaluate respondents' awareness of other apartment characteristics in the second phase of the survey implied that the questionnaire for the linguistic representation method was shorter than that for the location representation method. While the question was excluded in the second phase as it was irrelevant to the survey and respondents in the first phase were encouraged not to spend too much time on the question if they were unaware of the apartment characteristics, the difference in the questionnaire length could affect respondents' fatigue level prior to the choice experiment

9.6 Conclusions and Future Research

This research provided significant insight on the effect of attribute representation and preference elicitation method on respondents' choice. It was found that while different representation methods yield varying model fit, the optimum representation method is highly attribute-specific. In case where the stated level of preference certainty is the true preference that the respondent holds, a lower dependence on complex error assumption was observed while in the case of uncertainty arising from randomness, higher model forms gave a better model fit. The different causes of preference uncertainty were also seen to be dependent on the method of attribute representation. Thus, in case of the location method, greater preference uncertainty was observed due to stochastic reasons while in case of the linguistic representation method, the preference uncertainty obtained was a reflection of respondents' true preferences. Findings from the FL analysis revealed that heuristics could have played an important role in respondents' decision-making process in the case of the linguistic representation method.

While result from the logit analysis provided insight on factors affecting valuation and choice certainty, whereas the FL analysis shed light on the possibility of employing heuristics in the choice process, an interesting application of the FL technique was not successfully conducted. Though it was initially considered that the FL method will be able to analyse preference uncertainty data, the analysis conducted in the research showed a very low percentage correct predicted implying that the analytical method employed is not entirely suitable for this form of preference data, and more complex hybrid structure needs to be developed in this case. Another shortcoming of the research arose from the simulation process for the location representation method. The experimental design was based on the location ratings which were generated during the simulation process. However, empirically as these are highly subjective, the simulation conducted was not entirely reflective of respondents' true rating. Though this did not result in any significant loss in model fit during the estimation process, some respondent rating from a pilot survey would have been beneficial for the experimental design.

Besides the insight provided by this research, a myriad of future research possibilities have been generated, which can be outlined as follows:

- The effect of different attribute representation methods on the same respondent offers an interesting possibility to eliminate any respondent effects arising from separate surveys for each representation method. However, as design of the linguistic method was dependent on the ratings and category information obtained from the location method, a combination of both the representation methods in the same survey was not possible in this research. In order to conduct this form of study, a pilot survey to elicit information on attribute rating and categories for both the representation methods can be carried out

- By obtaining information for location and linguistic ratings in the pilot survey, a mixed representation method (linguistic for noise while location for view and sunlight) can be experimented to estimate the relative improvement in the model fit

- To capture the pure effects of alternative representation methods on the ‘noise’ attribute, the representation method of only that attribute can be varied across the location and linguistic methods

- In order to test whether one and two stage Likert elicitation methods have varying effects on the same respondent, a split sample design with the two elicitation methods for each of the representation methods can be conducted. In this case thus, the two representation methods should be conducted in different phases

- To relate the scale and language of the one and two stage Likert methods more closely, the ‘uncertain’ alternative can be included in the two stage Likert method in the first phase of the choice elicitation process. Thus, three options can be offered in the first instance of the two stage Likert question – Alternative A, Uncertain and Alternative B. While in the second phase of the question, if the respondent has not chosen the ‘uncertain’ alternative, he can be asked to indicate whether he ‘definitely’ or ‘probably’ chooses an alternative

- The variation between preference uncertainty and strength of preference (SOP) measures can be explained to the respondent along with the emphasis that the study is eliciting their preference uncertainty prior to the choice exercise, in order to eliminate any ambiguity between the preference uncertainty and SOP measures

- The question eliciting respondents’ familiarity with apartment characteristics for different apartment locations employed in the first phase of the survey can be replaced by simple explanation of the different apartment locations in order to maintain the same level of survey length across the location and linguistic representation methods. However this method could compromise with the estimation of respondents’ familiarity of apartment characteristics

- In order to estimate specific respondent characteristics (if any) associated with each preference level, the Latent Class Analysis could be applied

- To gain a better insight on respondents’ decision making process as well as to gain further information in aid of FL rules formation, the respondents’ importance for any specific attribute (if any) during each choice scenario can be elicited. This information can be used in FL analysis to give relative weights to the attribute. Moreover, the output can also be defined as a possibility measure and the possibility-probability transformation can be conducted

- As results from the FL linguistic model with absolute attribute levels as inputs showed important role of heuristics in the decision-making process employed by the respondents, this information can be used to enhance the logit modelling by allowing for heuristics to be incorporated in the choice model during the estimation process. Moreover hybrid system can be developed in combination with Genetic Algorithm and fuzzy c-means clustering to derive coefficient estimates from the FL method

10 REFERENCES

Abdel-Aty M.A., Kitamura R. and Jovanis P.P. (1997), 'Using Stated Preference Data for Studying the Effect of Advanced Traffic Information on Drivers' Route Choice', *Transportation Research Part C: Emerging Technologies* **5** (1): 39-50

Acquisti A. and Grossklags J. (2005), 'Uncertainty, Ambiguity and Privacy', *4th Annual Workshop on Economics and Information Security*, Harvard University

Akter S., Bennett J. and Akhter S. (2008), 'Preference uncertainty in contingent valuation', *Ecological Economics* **67** (3): 345-351

Akter S. and Bennett J. (2009), 'A cognitive psychological approach of analyzing preference uncertainty in contingent valuation', *Australian Agricultural and Resource Economics Society*, 53rd Annual Conference

Akter S., Brouwer R., Brander L. and van Beukering P. (2009), 'Respondent uncertainty in a contingent market for carbon offsets', *Ecological Economics* **68** (6): 1858-1863

Akter S. and Bennett J. (2010), 'Testing construct validity of verbal versus numerical measures of preference uncertainty in contingent valuation', *4th World Congress in Environmental and Resource Economics*, Montreal

Albaum G. (1997), 'The Likert Scale Revisited: an alternate version', *Journal of the Market Research Society* **39** (2): 331-341

Alberini A., Boyle K. And Welsh M. (2003), 'Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty', *Journal of Environmental Economics and Management* **45** (1): 40-62

Alberini A., Boyle K. and Welsh M. (1997), 'Using Multiple-Bounded Questions to Incorporate Preference Uncertainty in Non-market Valuation', *W-133, Benefit and Cost Transfers in Natural Resource Planning*, 9th Interim Report, University of Nevada

Alpizar F. and Carlsson F. (2001), 'Policy Implications and Analysis of the Determinants of Travel Mode Choice: An Application of Choice Experiments to Metropolitan Costa Rica', *Working Papers in Economics no. 56*, Department of Economics, Göteborg University, Sweden

Arsenio E. (2002), 'The Valuation of Environmental Externalities: A Stated Preference Case Study on Traffic Noise in Lisbon', *PhD Thesis*, Institute for Transport Studies, University of Leeds

Arsenio E., Bristow A.L. and Wardman M. (2000), 'An Innovative Stated Preference Computer Survey Model for Valuing Noise Impacts from Road Traffic', European Transport Conference

Arsenio E., Bristow A.L. and Wardman M. (2002), 'Marginal Values of Traffic Noise Externalities from Stated Preference Methods', *European Transport Conference*, Cambridge

Babelfish website: uk.babelfish.yahoo.com

Barreiro J., Sanchez M. and Viladrich-Grau M. (2005), 'How much are people willing to pay for silence? A Contingent Valuation Study', *Applied Economics* **37** (11): 1233-1246

Bates J. (1988), 'Econometric Issues in Stated Preference Analysis', *Journal of Transport Economics and Policy* **22** (1): 59:69

Bates J.J. (1984), ‘Values of Time from Stated Preference Data’, *Proceedings from Seminar H, PTRC Annual Meeting*

Bates J.J. and Roberts M. (1983), ‘Recent Experiences with Models Fitted to Stated Preference Data’, *Proceedings from Seminar M, PTRC Annual Meeting*

Batley R., Fowkes A., Whelan G. and Daly A. (2001), ‘Models for Choice of Departure Time’, *European Transport Conference,* Cambridge

Batley R. and Ibanez N. (2009), ‘Randomness in preferences, outcomes and tastes: an application to journey time risk’, *International Choice Modelling Conference*, Harrogate, UK

Ben Akiva M. and Lerman S.R. (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge

Bhat C. (2000), ‘Flexible Model Structures for Discrete Choice Analysis’ in (eds.) Hensher D.A. and Button K.J., *Handbook of Transport Modelling,* Pergamon

Bickel P., Schmid S., Tervonen J., Hamekoski K., Otterstrom T., Anton P., Enei R., Leone G., van Donselaar P. and Carmigchlet H. (2003), *Environmental Marginal Cost Case Studies*, UNITE, IER, Stuttgart

Bierlaire, M. (2003), ‘BIOGEME: A free package for the estimation of discrete choice models’, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland

Bierlaire, M. (2008)*, An introduction to BIOGEME Version 1.6*, biogeme.epfl.ch

Bjorner T.B. (2004), ‘Combining socio-acoustic and contingent valuation surveys to value noise reduction’, *Transportation Research Part D: Transport and Environment* **9** (5): 341-356

Blomquist G.C., Blumenschein K. and Johannesson M. (2009), 'Eliciting Willingness to Pay without Bias using Follow-up Certainty Statements: Comparisons between Probably/Definitely and a 10-point Certainty Scale', *Environmental and Resource Economics* **43** (4): 473-502

Boman M. (2009), 'To pay or not to pay for biodiversity in forests – What scale determines response to willingness to pay questions with uncertain response options?', *Journal of Forest Economics* **15** (1-2): 79-91

Botteldooren D., Verkeyn A. and Lercher P. (2002), 'Noise Annoyance Modelling using Fuzzy Rule Based Systems', *Noise Health* **4** (15): 27-44

Bristow A.L. and Wardman M. (2006a), 'What Influences the Value of Aircraft Noise?', *European Transport Conference*, Strasbourg

Bristow A.L. and Wardman M. (2006b), 'Valuation of Aircraft Noise by Time of Day: A Comparison of Two Approaches', *Transport Reviews* **26** (4): 417-433

Broberg T. and Brannlund R. (2008), 'An alternative interpretation of multiple bounded WTP data – Certainty dependent payment card intervals', *Resource and Energy Economics* **30** (4): 555-567

Brouwer R. (2009), 'Stated Preference Uncertainty: Signal or Noise?' *Working Paper*, IVM, VU University, Amsterdam

Brouwer R., Dekker T., Rolfe J. and Windle J. (2010), 'Choice Certainty and Consistency in Repeated Choice Experiments', *Environmental and Resource Economics* **46** (1): 93-109

Brown T.C. (2003), 'Introduction to Stated Preference Methods', in (eds.) Champ P.A., Boyle and Brown T.C., *A Primer on Nonmarket Valuation*, Kluwer Academic Publishers, The Netherlands

Brown T.C., Kingsley D., Peterson G.L., Flores N.E., Clarke A., Birjulin A. (2006), 'Reliability of individual valuations of public and private goods: Choice consistency, response time and preference refinement', *Journal of Public Economics* **92** (7): 1595-1606

Camerer C. and Weber M. (1992), 'Recent developments in modeling preferences: Uncertainty and ambiguity', *Journal of Risk and Uncertainty* **5** (4): 325-370

Campbell D., Hutchinson G. and Scarpa R. (2007), 'Using Choice Experiments to Explore the Spatial Distribution of Willingness to Pay for Rural Landscape Improvements', *Working Paper in Economics*, University of Waikato

Cantarella G.E. and Fedele V. (2003), 'Fuzzy vs. Random Utility Models: An Application to Mode Choice Behaviour Analysis', *European Transport Conference*, Strasbourg

Cantillo V., Heydecker B. and Ortuzar J.D. (2006), 'A discrete choice model incorporating thresholds for perception in attribute values', *Transportation Research Part B: Methodological* **40** (9): 807-825, Elsevier

Carlsson F., Lampi E. and Martinsson P. (2004), 'Measuring marginal values of noise disturbance from air traffic: Does the time of the day matter?', *Working Papers in Economics no.125*, Department of Economics, Gothenburg University

Caussade S., Ortuzar J.D., Rizzi L.I., Hensher D.A. (2005), 'Assessing the influence of design dimensions on stated choice experiment estimates', *Transportation Research Part B* **39** (7): 621-640

Chen G. and Pham T.T. (2001), *Fuzzy Sets, Fuzzy Logic and Fuzzy Control Systems*, CRC Press, USA

de Palma A. (1998), 'Individual and Collective Decision Making: Application to Travel Choice', in (eds.) Gärling T., Laitila T. and Westin K., *Theoretical Foundations of Travel Choice Modeling*, Elsevier Science, UK

Defra: www.defra.gov.uk/environment/policy/natural-environ/using/valuation/steps/step5/value.htm

DeShazo J.R. and Fermo G. (2002), 'Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency', *Journal of Environmental Economics and Management* **44** (1): 123-143

Dolnicar S. and Grun B. (2007), 'How constrained a response: A comparison of binary, ordinal and metric answer formats', *Journal of Retailing and Consumer Services* **14** (2): 108-122

EC (2002), *Position Paper on Dose Response Relationships between Transportation Noise and Annoyance*

EEA (1999) *Environment in the European Union at the Turn of the Century*, Copenhagen, Office for Official Publications of the European Communities, Luxembourg and Elsevier Science Ltd., Oxford

Fields J.M., De Jong R.G., Gjestland T., Flindell I.H., Job R.F.S., Kurra S., Lercher P., Vallet M., Yano T., Guski R., Felsher-Suhr U. And Schumer R. (2001), 'Standardized General-Purpose Noise Reaction Questions for Community Noise Surveys: Research and A Recommendation, *Journal of Sound and Vibration* **242** (2): 641-679

Fiore S.M., Harrison G.W., Hughes C.E. and Rutström E.E. (2009), 'Virtual experiments and environmental policy', *Journal of Environmental Economics and Management* **57** (1): 65-86

Fishburn P.C. (1988), 'Normative Theories of Decision Making Under Risk and Under Uncertainty', in (eds.) Bell D.E., Raiffa H. and Tversky A., *Decision Making: Descriptive, Normative, and Prescriptive Interactions*, Cambridge University Press, UK

Fosgerau M. and Bjorner T.B. (2006), 'Joint models for noise annoyance and willingness to pay for road noise reduction', *Transportation Research Part B: Methodological* **40** (2): 164-178

Fowkes A.S. and Preston J. (1991), 'Novel approaches to forecasting the demand for new local rail services', *Transportation Research Part A: General* **25 (4)**: 209-218

Fowkes A.S. and Wardman M. (1988), 'The Design of Stated Preference Travel Choice Experiments with Particular Regard to Inter-personal Taste Variation', *Journal of Transport Economics and Policy* **22** (1): 27-44

Fraser I. and Balcolme K. (2010), 'Estimating WTP with Uncertainty Choice Contingent Valuation', *4th World Congress on Environmental and Resource Economics*, Montreal

Fujii K., Atagi J. and Ando Y. (2002), 'Temporal and Spatial Factors of Traffic Noise and Its Annoyance', *Journal of Temporal Design in Architecture and the Environment* **2** (1): 33-41

Galilea P. and Ortuzar J.D. (2005), 'Valuing noise level reductions in a residential location context', *Transportation Research Part D: Transport and Environment* **10** (4): 305-322

Garrod G.D., Scarpa R. And Willis K.G. (2002), 'Estimating the Benefits of Traffic Calming on Through Routes', *Journal of Transport Economics and Policy* **36** (2): 211-231

Gob R., McCollin C. and Ramalhoto M.F. (2007), 'Ordinal Methodology in the Analysis of Likert Scales', *Quality and Quantity* **41** (5): 601-626

Greene W.H. and Hensher D.A. (2002), 'A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit', *ITS Working Paper*, University of Sydney

Greifeneder R., Scheibehenne B. and Kleber N. (2010), ‘Less may be more when choosing is difficult: Choice complexity and too much choice’, *Acta Psychologica* **133** (1): 45-50

Halpern J.Y. (2003), *Reasoning about Uncertainty*, MIT Press

Hanemann W.M., Kriström B. and Li C-Z. (1996), ‘Nonmarket Valuation under Preference Uncertainty: Econometric Models and Estimation’, *Working Paper No. 794*, Department of Agricultural and Resource Economics, University of California, Berkeley

Hanley N. and Kriström B. (2002), ‘What’s It Worth? Exploring Value Uncertainty using Interval Question in Contingent Valuation’, *Working Paper 2002-10*, University of Glasgow

Hanley N., Mourato S. and Wright R.E. (2001), ‘Choice Modelling Approaches: A Superior Alternative for Environmental Valuation?’ *Journal of Economic Surveys* **15** (3): 435-448

Hanley N., Shogren J. and White B. (1997), *Environmental Economics in Theory and Practice*, MacMillan Press, UK

Hensher D.A. (1999), ‘HEV Choice Models as a Search Engine for the Specification of Nested Logit Tree Structures’, *Marketing Letter* **10** (4): 339-349

Hensher D.A. and Greene W.H. (2001), ‘The Mixed Logit Model: The State of Practice and Warnings for the Unwary’, *Working Paper*, University of Sydney

Hensher D.A., Rose J.M. and Greene W.H. (2005), *Applied Choice Analysis: A Primer*, Cambridge University Press

Hess S. (2005), 'Advanced Discrete Choice Models with Applications to Transport Demand', *PhD Thesis*, Imperial College London

Hess S., Bierlaire M. and Polak J.W. (2004), 'Development and Application of a Mixed Cross-Nested Logit model', *European Transport Conference*, Strasbourg

Hess S., Rose J.M. and Hensher D.A. (2008), 'Asymmetric preference formation in willingness to pay estimates in discrete choice models', *Transportation Research Part E: Logistics and Transportation Review* **44** (5): 847-863

Hess S., Train K.E. and Polak J.W. (2006), 'On the use of a Modified Latin Hypercube Sampling (MLHS) method in the estimation of a Mixed Logit Model for vehicle choice', *Transportation Research Part B: Methodological* **40** (2): 147-163

Hoevenagel R. (1994), 'Comparison of Economic Valuation Methods' in (ed.) Pethig R., *Valuing the Environmental: Methodological and Measurement Issues*, Kluwer Academic Publishers, The Netherlands

Holmes T.P. and Adamowicz W.L. (2003), 'Attribute-Based Methods', in (eds.) Champ P.A., Boyle and Brown T.C., *A Primer on Nonmarket Valuation*, Kluwer Academic Publishers, The Netherlands

Hoogendoorn-Lanser, S. & Hoogendoorn, S. P. (2000), 'A Fuzzy Genetic Approach to Travel Choice Behavior in Public Transport Networks', in *CD-ROM Preprints Transportation Research Board 79th Annual Meeting*, The National Academies, Washington D.C.

Howard K. and Salkeld G. (2009), 'Does Attribute Framing in Discrete Choice Experiments Influence Willingess to Pay? Results from a Discrete Choice Experiement in Screening for Colorectal Cancer', *Value in Health* **12** (2): 354-363

Jia J., Luce M.F. and Fischer G.W. (2004), 'Consumer Preference Uncertainty: Measures of Attribute Conflict and Extremity', Wharton-SMU Research Center

Johnston R.J. (2003), ‘Forecasting Support for Rural Land Use Policies: The Role of Preference Asymmetries’, *American Agricultural Economics Association Annual Meetings*, Montreal

Johnston R.J. and Swallow S.K. (1999), ‘Asymmetries in Ordered Strength of Preference Models: Implications of Focus Shift for Discrete-Choice Preference Estimation’, *Land Economics* **75** (2): 295-310

Kahneman D. and Tversky A. (2000), ‘Choices, Values and Frames’, in (eds.) Kahneman D. and Tversky A., *Choices, Values and Frames*, Cambridge University Press

Kasabov N.K. (1996), *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*, MIT Press

Kingsley D.C. (2008), ‘The Role of Preference Uncertainty in the Willingness to Pay – Willingness to Accept Disparity: An Experimental Test’, *April Mimeo*, Westfield State College

Klein M. and Methlie L.B. (1990), *Expert Systems: A Decision Support Approach with Applications in Management and Finance*, Addison-Wesley Publishing

Kliendorfer P.R., Kunreuther H.C. and Shoemaker P.J.H. (1993), *Decision Sciences: an integrative perspective*, Cambridge University Press

Kocur G., Hyman W. And Aunet B. (1982), ‘Wisconsin Work Mode-Choice Models Based on Functional Measurement and Disaggregate Behavioural Data’, *Transportation Research Record* **895**: 24-32

Koppelman F.S. and Bhat C. (2006), *A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models*, US Department for Transportation

Koppelman F.S. and Wen C.H. (1998), ‘Alternative Nested Logit Models: Structure, Properties and Estimation’, *Transportation Research B* **32** (5): 289-298

Kosenius A. (2009), ‘Causes of Response Uncertainty and its Implications for WTP Estimation in Choice Experiment’, *Discussion Paper no. 29*, University of Helsinki

Kragt M.E. and Bennett J.W. (2009), ‘What’s appropriate? Investigating the effects of attribute framing and changing cost levels in choice experiments’, *17th Annual Conference of the European Association of Environmental and Resource Economists*, Amsterdam

Langford I.H. (1994), ‘Using a Generalised Linear Mixed Model to Analyse Dichotomous Choice Contingent Valuation Data’, in (eds.) Langford I.H., Turner R.K. and Bateman I., *Environmental Decision-Making and Risk Management: Selected Essays*, Edward Elgar Publishing, U.K.

Lauraeus-Niinivaara T., Saarinen T. and Oorni A. (2007), ‘Knowledge and Choice Uncertainty Affect Consumer Search and Buying Behaviour’, *Proceedings of the 40th Hawaii International Conference on Systems Science*, Hawaii

Leiser D., Azar O.H. and Hadar L. (2008), ‘Psychological construal of economic behaviour’, *Journal of Economic Psychology* **29** (5): 762-776

Leong W. and Hensher D.A. (2011), ‘Embedding Multiple Heuristics into Choice Models: An Explanatory Analysis’, *International Choice Modelling Conference*, Leeds, UK

Li C.Z. and Mattsson L. (1995), ‘Discrete choice under preference uncertainty: an improved structural model for contingent valuation’, *Journal of Environmental Economics and Management* **28** (2): 256-269

Li H.N., Chau C.K., Tse M.S. and Tand S.K. (2009), ‘Valuing road noise for residents in Hong Kong’, *Transportation Research Part D: Transport and Environment* **14** (4): 264-271

Loomis J. and Ekstrand E. (1997), ‘Economic Benefits of Critical Habitat for the Mexican Spotted Owl: A Scope Test Using a Multiple-Bounded Contingent Valuation Survey’, *Journal of Agricultural and Resource Economics* **22** (2): 356-366

Loureiro M.L. and Loomis J.B. (2008), ‘Dealing with Preference Uncertainty in Contingent Valuation: A Mixture Model Approach’, *envecon: Applied Environmental Economics Conference*, London

Louviere J.L., Hensher D.A. and Swait J.D., (2000), *Stated Choice Methods: Analysis and Application*, Cambridge University Press, UK

Louviere J.J. and Kocur G. (1983), ‘The Magnitude of Individual-Level Variations in Demand Coefficients: A Xenia, Ohio Case Example’, *Transportation Research Part A: General* **17** (5): 363-373

Lundhede T.H., Olsen S.B., Jacobsen J.B. and Thorsen B.J. (2009), ‘Handling respondent uncertainty in Choice Experiments: Evaluating recoding approaches against explicit modelling of uncertainty’, *Journal of Choice Modelling* **2** (2): 118-147

Mabit S., Hess S. and Caussade S. (2008), ‘Correlation and willingness-to-pay indicators in transport demand modelling’, *Transportation Research Board 87th Annual Meeting*, Washington D.C.

March J.G. (1988), ‘Bounded Rationality, Ambiguity and the Engineering of Choice’, in (eds.) Bell D.E., Raiffa H. and Tversky A., *Decision Making: Descriptive, Normative, and Prescriptive Interactions*, Cambridge University Press, UK

Martin M.A., Tarrero A., Gonzalez J. and Machimbarrena M. (2006), ‘Exposure-effect relationships between road traffic noise annoyance and noise cost valuation in Valldolid, Spain’, *Applied Acoustics* **67** (10): 945-958

MATLAB Handbook (2007), The Fuzzy Logic Toolbox

MATLAB Handbook (R2007a), The Fuzzy Logic Toolbox, Tutorial on Fuzzy Inference Systems

Meyer M.D. and Miller E.J. (1984), *Urban Transportation Planning: A Decision-Oriented Approach*, McGraw-Hill, USA

Miedema H.M.E. and Oudshoorn C.G.M. (2001), 'Annoyance from Transportation Noise: Relationships with Exposure Metrics DNL and DENL and Their Confidence Intervals', *Environmental Health Perspectives* **109** (4): 409-416

Mitchell R.C. and Carson R.T. (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Resources for the Future, Washington D.C.

Mizutani K. and Akiyama T. (2001), 'Construction of Modal Choice Model with a Descriptive Utility Function using Fuzzy Reasoning', *IFSA World Congress and 20th NAFIPS International Conference* **2**: 852-856

Morrison M.D., Blamey R.K., Bennett J.W. and Louviere J.J. (1996), 'A Comparison of Stated Preference Techniques for Estimating Environmental Values', *Choice Modelling Research Report*, Australia

Munizaga M.A. and Alvarez-Daziano R. (2001), 'Mixed Logit Vs. Nested Logit and Probit Models', *5th Tri-annual Invitational Choice Symposium Workshop: Hybrid Choice Models, Formulation and Practical Issues*, Asilomar, June 2001

MVA Consultancy (1987), 'The Value of Travel Time Savings', *Project Journals*

Navrud S. (2002), 'The State of the Art on Economic Valuation of Noise', *Final Report to European Commission DG Environment*

Nijland H.A. and van Wee G.P. (2005), 'Traffic Noise in Europe: A Comparison of Calculation Methods, Noise Indices and Noise Standards for Road and Railroad Traffic in Europe', *Transport Reviews* **25** (5): 591-612

Nunes P.A.L.D. and Travisi C.M. (2007), 'Rail Noise-Abatement Programmes: A Stated Choice Experiment to Evaluate the Impacts on Welfare', *Transport Reviews* **27** (5): 589-604

OECD Document (1994), *Project and Policy Appraisal: Integrating Economics and Environment*, OECD, Paris

Olsen S.B., Lundhede T.H., Jacobsen J.B. and Thorsen B.J. (2011), 'Tough and Easy Choices: Testing the Influence of Utility Difference on Stated Certainty-in-Choice in Choice Experiments', *Environmental and Resource Economics* **49** (4): 491-510

Oppewal H., Morrison M., Wang P. and Waller D. (2009), 'Preference stability: Modelling how consumer preferences shift after receiving new product information', *International Choice Modelling Conference*, Harrogate

Ortuzar J.D. (2000), 'Fundamentals of Stated Preference', in (eds.) Ortuzar J.D., *Stated Preference Modelling Techniques*, PTRC

Ortuzar J.D. and Willumsen L.G. (2004), *Modelling Transport*, Wiley Publications, UK

Papi J. and Halleman B. (2004), 'Road Traffic Noise: The Road User's Perspective', *European Union Road Federation*, Position Paper

Parsons S. (2001), *Qualitative Methods for Reasoning under Uncertainty*, MIT Press

Passchier-Vermer W. and Passchier W.F. (2000), ‘Noise Exposure and Public Health’, *Environmental Health Perspectives* **18** (1): 123-131

Pratt C. (2002), ‘Estimation and Valuation of Environmental and Social Externalities for the Transport Sector’, *25th Australasian Transport Research Forum Incorporating the BTRE Transport Policy Colloquium*, Canberra

Preston J. and Wardman M. (1988), ‘Demand Forecasting for New Local Rail Services: A Case Study of a New Service between Leicester and Burton-on-Trent’ *Working Paper 260*, Institute for Transport Studies, University of Leeds

Restle F. (1961), *Psychology of Judgment and Choice: A Theoretical Essay*, John Wiley and Sons, USA

Revelt D. and Train K. (1998), ‘Mixed Logit with Repeated Choices: Households’ Choices of Appliance Efficiency Level’, *The Review of Economics and Statistics* **80** (4): 647-657

Rheinberger C. (2009), ‘On the Effects of Risk Framing in Choice Experiments’, *17th Annual Conference of the European Association of Environmental and Resource Economists*, Amsterdam

Richards M.G. and Ben-Akiva M. (1975), *A Disaggregate Travel Demand Model*, Saxon House, England

Romano D. and Vigano L. (1998), ‘A Review of CVM Environmental Applications in Italy’, in (eds.) Bishop R.C. and Romano D., *Environmental Resource Valuation: Applications of the Contingent Valuation Method in Italy*, Kluwer Academic Publishers, Dordrecht, The Netherlands

Saelensminde K., (1999), ‘Stated choice valuation of urban traffic air pollution and noise’, *Transportation Research* Part D **4** (1): 13-27

Salisbury L.M. and Feinberg F.M. (2008), ‘Future Preference Uncertainty and Diversification: The Role of Temporal Stochastic Inflation’, *Journal of Consumer Research* **35** (2): 349-359

Savage S.J. and Waldman D.M. (2008), ‘Learning and Fatigue During Choice Experiments: A Comparison of Online and Mail Survey Modes’, *Journal of Applied Econometrics* **23** (3): 351-371

Shaikh S.L., Sun L. and van Kooten G.C. (2007), ‘Treating Respondent Uncertainty in Contingent Valuation: A Comparison of Empirical Treatments’, *Ecological Economics* **57** (3): 507-519

Shen J. (2005), ‘A Review of Stated Choice Method’, *Discussion Papers in Economics and Business*, Graduate School of Economics and Osaka School of International Public Policy, Japan

Silberhorn N., Boztug Y. and Hildebrandt L. (2007), ‘Estimation with the Nested Logit Model: Specifications and Software Particularities’, *Working Paper*, Humboldt University, Germany

Slovic P. (2000), ‘The Construction of Preference’, in (eds.) Kahneman D. and Tversky A., *Choices, Values and Frames*, Cambridge University Press

Soguel N. (1996), ‘Contingent Valuation of Traffic Noise Reduction Benefits’, *Swiss Journal of Economics and Statistics* **132** (1): 109-123

Sommerhoff J., Recuero M. and Suarez E. (2006), ‘Relationship between loudness perception and noise indices in Valdivia, Chile’, *Applied Acoustics* **67** (9): 892-900

Staskeviciute G. and Kaklauskas A. (2007), ‘Impact of Air Pollution and Noise on Property Prices: Theoretical Aspects of Economic Valuation’, *Jaunuju Mokslininju Conference*, Lithuania

Statistics Portugal (2007), *The People*

Statistics Portugal (2008), *The People*

Statistics Portugal: www.ine.pt

Stone D.N. and Schkade D.A. (1991), 'Numeric and Linguistic Information Representation in Multiattribute Choice', *Organizational Behaviour and Human Decision Processes* **49** (1): 42-59

Stone D.N. and Schkade D.A. (1994), 'Effects of Attribute Scales on Process and Performance in Multiattribute Choice', *Organizational Behaviour and Human Decision Processes* **59** (2): 261-287

Stone D.N. and Sckade D.A. (1991), 'Numeric and linguistic information representation in multiattribute choice', *Organizational Behavior and Human Decision Processes* **49** (1): 42-59

Sun L. and van Kooten G.C. (2009), 'Comparing Fuzzy and Probabilistic Approaches to Preference Uncertainty in Non-Market Valuation', *Environmental and Resource Economics* **42** (4): 471-489

Swait J. and Adamowicz W. (2001), 'Choice Environment, Market Complexity and Consumer Behavior: A Theoretical and Empirical Approach for Incorporating Decision Complexity into Models of Consumer Choice', *Organizational Behavior and Human Decision Processes* **86** (2): 141-167

Swallow S.K., Opaluch J.J. and Weaver T.F. (2001), 'Strength-of-Preference Indicators and an Ordered-Response Model for Ordinarily Dichotomous, Discrete Choice Data', *Journal of Environmental Economics and Management* **41** (1): 70-93

Thanos S.M., Wardman M. and Bristow A.L. (2006), 'A Stated Choice Experiment Valuing Aircraft Noise in the Context of Airport Relocation', *European Transport Conference*, Strasbourg

Tinch R. (1996), *Valuation of Environmental Externalities Full Report*, Department of Transport, HSMO Publishing

Train K. (2003), *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, MA

Tversky A. and Kahneman D. (2000), 'Rational Choice and the Framing of Decisions', (eds.) Kahneman D. and Tversky A., *Choices, Values and Frames,* Cambridge University Press, UK

Urbany J.E., Dickson P.R. and Wilkie W.L. (1989), 'Buyer Uncertainty and Information Search', *The Journal of Consumer Research* **16** (2): 208-215

van Kooten G.C., Krcmar E. and Bulte E.H. (2001), 'Preference Uncertainty in Non-Market Valuation: A Fuzzy Approach', *American Journal of Agricultural Economics* **83** (3): 487-500

Verbeek M. (2008), *A Guide to Modern Econometrics*, John Wiley and Sons Ltd., England

von Neumann J. and Morgenstern O. (1944), *Theory of Games and Economic Behaviour*, Princeton University Press, USA

Vythoulkas P.C. and Koutsopoulos H.N. (2003), "Modeling discrete choice behaviour using concepts from fuzzy set theory, approximate reasoning and neural networks", *Transportation Research Part C: Emerging Technologies* **11** (1): 51-73

Walker J.L., Ben-Akiva M. and Bolduc D. (2007), 'Identification of Parameters in Normal Error Component Logit-Mixture (NECLM) Models', *Journal of Applied Econometrics* **22** (6): 1095-1125

Wang H. (1997), 'Treatment of "Don't Know" Responses in Contingent Valuation Surveys: A Random Valuation Model', *Journal of Environmental Economics and Management* **32** (2): 219-232

Wardman M. (1987), ‘An evaluation of the use of Stated Preference and Transfer Price Data in forecasting the demand for travel’, *PhD thesis*, University of Leeds

Wardman M. (1988), ‘A Comparison of Revealed Preference and Stated Preference Models of Travel Behaviour’, *Journal of Transport Economics and Policy* **22** (1): 71-91

Wardman M. and Bristow A. (2004), ‘Traffic Related Noise and Air Quality Valuations: Evidence from Stated Preference Residential Choice Models’, *Transportation Research D: Transport and Environment* **9** (1): 1-27

Wardman M. and Bristow A. (2008), ‘Valuations of aircraft noise: experiments in stated preference’, *Environmental and Resource Economics* **39** (4): 459-480

Watkins L.H. (1981), *Environmental Impact of Road and Traffic*, Applied Science Publishers Ltd., UK

Weijters B., Cabooter E. and Schillewaert N. (2010), ‘The effect of rating scale format on response styles: The number of response categories and response category labels’, *International Journal of Research in Marketing* **27** (3): 236-247

Welsh M.P. and Poe G.L. (1998), ‘Elicitation Effects in Contingent Valuation: Comparisons to a Multiple Bounded Discrete Choice Approach’, *Journal of Environmental Economics and Management* **36** (2): 170-185

Whelan G. and Crockett J. (2009), ‘An Investigation of the Willingness to Pay to Reduce Rail Overcrowding’, *International Choice Modelling Conference*, Harrogate, UK

Whelan G. and Tapley N. (2006), ‘Development and Applications of the Mixed Ordered Response Logit Model’, *European Transport Conference*, Strasbourg

Yanez M.F., Cherchi E., Heydecker B.G. and Ortuzar J.D. (2010), 'On the Treatment of Repeated Observations in Panel Data: Efficiency of Mixed Logit Parameter Estimates', *Networks and Spatial Economics* **11** (3): 393-418

Yoon S. and Simonson I. (2008), 'Choice Set Configuration as a Determinant of Preference Attribution and Strength', *Journal of Consumer Research* **35** (2): 324-336

11 APPENDICES

11.1 APPENDIX A

QUESTIONNAIRE

In order to provide an idea of the type of questions employed in the survey, an example of the questionnaire as well as the choice scenario is provided. The following questionnaire and choice scenarios relate to the Location representation method with binary and one stage Likert elicitation methods. Though the general format of the questionnaire remains similar for other phases and experiments of the survey, questions based on attribute perception rating and form of elicitation employed can differ as outlined in Section 5.4.3

1.1

Institute for Transport Studies

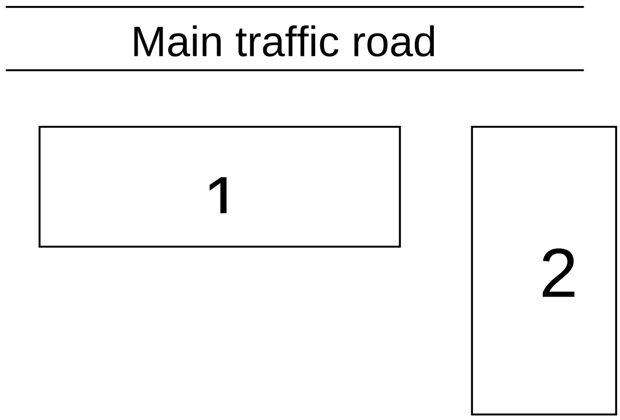
UNIVERSITY OF LEEDS

Q1. INFORMATION TO BE CODED BY THE INTERVIEWER FOR EACH RESPONDENT

Orientation of the block relative to main traffic road

Main façade is parallel to the main traffic road

Main façade is perpendicular to the main traffic road



Please refer to the name of the main road (tick as appropriate):

M1. Av Norton de Matos (2 Circular)

M2. Eixo Norte-Sul

M3. Av. Padre Cruz

Q2. OBJECTIVE OF THE RESEARCH (STATE TO THE RESPONDENTS):

This research is being conducted to characterize some apartment attributes and the local environment in this residential area. For this purpose, we would like to ask you some questions and we thank you in advance for your co-operation.

<u>Q3. RECORD THE APARTMENT INFORMATION:</u>

Street:

____

Block number:

____

Floor number:

____

Number of bedrooms/flat type:

____

Position of the bedroom in relation to the main traffic road (back ‘1’, front ‘2’ or side ‘3’):

____

Position of the living room in relation to the main traffic road (back ‘1’, front ‘2’ or side ‘3’):

____

<u>Q4. RESPONDENT’S FAMILIARITY WITH :</u>

Please refer the characteristics that you were aware at the time of your purchase, indicating values known by you:

	A1 (6th floor, front facing the main traffic road – 6F)	**A2** (6th floor, front not facing the main traffic road – 6T)	**A3** (3rd floor, front facing the main traffic road – 3F)	**A4** (3rd floor, front not facing the main traffic road – 3T)
Price/Rent				
Housing service charge				
No. of rooms				
Área (m2)				
Parking space in garage				

Q5. WE WOULD LIKE TO ASSESS YOUR FAMILIARITY WITH SOME CHARACTERISTICS OF THE FLATS. PLEASE CLASSIFY THE FLAT TYPES BASED ON YOUR PERCEPTIONS AND PLEASE ALSO GIVE A RATING FROM 0 TO 100 FOR NOISE, VIEW AND SUNLIGHT:

Characteristics of the apartment					View **(Card - 3)**	
Very bad (0) ______________________________ Very good (100)						
	Very bad	Bad	Neither	Good	Very good	Rating
6F	●	●	●	●	●	
6T	●	●	●	●	●	
3F	●	●	●	●	●	
3T	●	●	●	●	●	

Characteristics of the apartment					Noise **(Card – 3b)**	
Very noisy (0) ______________________________ Very quiet (100)						
	Very noisy	Noisy	Neither	Quiet	Very quiet	Rating
6F	●	●	●	●	●	
6T	●	●	●	●	●	
3F	●	●	●	●	●	
3T	●	●	●	●	●	

Characteristics of the apartment					Sunlight **(Card - 3)**	
Very bad (0) ______________________________ Very good (100)						
	Very bad	Bad	Neither	Good	Very good	Rating
6F	●	●	●	●	●	
6T	●	●	●	●	●	
3F	●	●	●	●	●	
3T	●	●	●	●	●	

Q6. STATED PREFERENCE CHOICE SCENARIOS – USING PAYMENT CARD METHOD – PLEASE SHOW CARDS SUCCESSIVELY TO THE RESPONDENTS

Definitions:
6F – 6th floor, front facing the main traffic road
6T – 6th floor, front not facing the main traffic road
3F – 3rd floor, front facing the main traffic road
3T – 3rd floor, front not facing the main traffic road

For the following choice scenarios, please select the preferred alternative (Interviewer: Please note the selected options for choice scenarios 1-8 in the following table):

	OPTION A	OPTION B
Scenario 1		
Scenario 2		
Scenario 3		
Scenario 4		
Scenario 5		
Scenario 6		
Scenario 7		
Scenario 8		

For the following choice scenarios, please give the strength of your preference[31] between the two options (Interviewer: Please note the selected options for the choice scenarios 9-16 in the following table):

	OPTION A		Uncertain	OPTION B	
	Definitely A	Probably A		Probably B	Definitely B
Scen. 9					
Scen. 10					
Scen. 11					
Scen. 12					
Scen. 13					
Scen. 14					
Scen. 15					
Scen. 16					

PLEASE NOTE WHETHER THE RESPONDENT CONSIDERED ALL THE ATTRIBUTES WHEN MAKING THE CHOICE OR HAS ANY ATTRIBUTE PLAYED A MORE IMPORTANT PART:

[31] In Portuguese this was translated as 'degree of preference'.

Q7. PERCEPTION OF NOISE AND LEVEL OF ANNOYANCE:

Please only consider the level of outside noise, ignoring the noise indoors

1. How would you describe the general day time (7h – 22h) noise level inside your home: **Card - 4**

1. Very noisy
2. Noisy
3. Neither noisy nor quiet
4. Quiet
5. Very quiet

2. How much does noise annoy you and your household during this time: **Card - 5**

1. Very Much
2. Moderately
3. A little
4. Not at all

3. How would you describe the general night-time (22h – 7h) noise levels inside your home: **Card - 4**

1. Very noisy
2. Noisy
3. Neither noisy nor quiet
4. Quiet
5. Very quiet

4. How much does noise annoy you and your household during this time: **Card - 5**

- Very much
- Moderately
- A little
- Not at all

5. <u>In general, which of the following are the three most important causes of noise in your home (rate from 1-3 in order of importance)</u> **<u>Card - 6</u>**

1. Road traffic
2. Aircraft noise
3. Neighbours
4. Construction work
5. Other (specify)

Q8. MEASURES TAKEN BY THE HOUSEHOLD TO REDUCE NOISE IMPACT

Have you taken any measures to reduce the impact of noise in the home?

Yes

No

If yes, please specify the type of measure taken:			
Type of measure	Cost	Year of installation	Purpose of installation
Double glazing			
Secondary glazing			
Double ceiling			
Shutters			
Other			

Q9. HOUSEHOLD FAMILIARITY QUESTIONS:

1. How long have you lived here (year, months)?

   ____

2. Please indicate the number of people in the household:

   ____

3. Number and age of the children living in the flat:

   ____

4. What are the main reasons for you and your household to move to this location (rate 1-3 according to the level of importance): **(Card - 1)**

1. Proximity to workplace
2. Price of the apartment
3. Surrounding (calm/noisy)
4. Public transport
5. No industries nearby
6. Accessible by car
7. Proximity of school
8. Quality of neighbourhood
9. Quality of housing
10. Other (specify) ______________________

5. What are the three main reasons to choose this flat in the block (rate 1-3 based on the level of importance): **(Card - 2)**

    1. View
    2. Price of the apartment
    3. Number of rooms
    4. Noise level
    5. Type of construction
    6. Sunlight/Orientation
    7. Enclosed parking
    8. Housing service charge
    9. Availability
    10. Safety
    11. Other (specify)

Q10. TYPE OF FLAT TENURE

1. What is your current flat tenure type (tick as appropriate):

    Owned (if yes, Q.3)

    Rented (if yes, Q.4)

2. How much housing service charge do you pay per month?

____

3. If flat is owned, what was the buying price and which year was it bought in:

---

4. If flat is rented, what is the current rent?

---

5. How many hours do you normally stay at home during daytime (7h – 22h) on the weekdays?

---

6. How many hours do you normally stay at home during daytime (7h – 22h) on the weekends?

Q11. SOCIO-ECONOMIC CHARACTERISTICS

1. What is your highest education level: **Card - 7**

1. Primary School	☐	3. Graduate/Technical School	☐
2. Secondary School	☐	4. Post-graduate	☐

2. What is your age group: **Card - 8**

1. 18 – 25	☐	4. 56 – 75	☐
2. 26 – 40	☐	5. 75+	☐
3. 41 – 55	☐		

3. Gender:

1. Male	☐	2. Female	☐

4. Net household income (category): **Card - 10**

1. < 1000	☐	4. 3001 – 4000	☐
2. 1001 - 2000	☐	5. 4001 - 5000	☐
3. 2001 - 3000	☐	6. > 5000	☐

5. <u>Employment Status: **Card - 11**</u>

1. Part-time employment
2. Full-time employment
3. Unemployed
4. Retired
5. House work
6. Student

11.2 APPENDIX B

This appendix provides an example of the one and two stage Likert choice scenarios in Portuguese as well as English.

The following is an example of the one stage Likert choice scenario in Portuguese. While the preference certainty levels in the original English choice scenario were: Definitely A, Probably A, Uncertain, Probably B and Definitely B, the following levels were translated into Portuguese by a native human translator:

<table>
<tr><th colspan="2">OPÇÃO A</th><th colspan="3">OPÇÃO B</th></tr>
<tr><td colspan="2">Vista: 6F</td><td colspan="3">Vista: 6F</td></tr>
<tr><td colspan="2">Ruído: 3F</td><td colspan="3">Ruído: 6F</td></tr>
<tr><td colspan="2">Taxa de condomínio: € 40</td><td colspan="3">Taxa de condomínio: € 75</td></tr>
<tr><td colspan="2">Insolação solar: 6F</td><td colspan="3">Insolação solar: 3T</td></tr>
<tr><td>Decididamente A</td><td>Provavelmente A</td><td>Não tenho a certeza</td><td>Provavelmente B</td><td>Decididamente B</td></tr>
</table>

Using online translation facility available on the Babelfish website (uk.babelfish.yahoo.com), the levels can be back translated to English as:

- Decididamente: Decidedly – this closely corresponds to the level 'Definitely'
- Provavelmente: Probably – this is translated exactly as in the original English questionnaire
- Não tenho a certeza: I do not have the certainty – closely corresponds to 'Uncertain'

As the native Portuguese translator was aware of the original English levels, these Portuguese translations were made based on the best equivalent words and hence the orginal English Likert scale was adopted in the thesis. Moreover, the translated levels corresponded well with the original levels. The corresponding choice scenario in English can thus be given as:

<table>
<tr><th colspan="2">OPTION A</th><th colspan="3">OPTION B</th></tr>
<tr><td colspan="2">View: 6F</td><td colspan="3">View: 6F</td></tr>
<tr><td colspan="2">Noise: 3F</td><td colspan="3">Noise: 6F</td></tr>
<tr><td colspan="2">Housing service charge: € 40</td><td colspan="3">Housing service charge: € 75</td></tr>
<tr><td colspan="2">Sunlight: 6F</td><td colspan="3">Sunlight: 3T</td></tr>
<tr><td>Definitely A</td><td>Probably A</td><td>Uncertain</td><td>Probably B</td><td>Definitely B</td></tr>
</table>

For the two stage Likert choice scenario, the choice scenario in Portuguese can be given as follows:

OPÇÃO A	**OPÇÃO B**
Vista: 6F	Vista: 6F
Ruído: 3F	Ruído: 6F
Taxa de condomínio: € 40	Taxa de condomínio: € 75
Insolação solar: 6F	Insolação solar: 3T
A	B

Grau de certeza/confiança:

Tenho a certeza absoluta ☐ **Tenho algumas dúvidas** ☐

While the original English questionnaire had the post-decisional certainty level as: very certain and somewhat certain, these were translated to Portuguese in the levels given above by a native human translator. Back translating these levels into English using the online translation facility available at uk.babelfish.yahoo.com, the following levels were obtained:

OPTION A	**OPTION B**
View: 6F	View: 6F
Noise: 3F	Noise: 6F
Housing service charge: € 40	Housing service charge: € 75
Sunlight: 6F	Sunlight: 3T
A	B

Degree of certainty/confidence:

I have the absolute certainty ☐ **I have some doubts** ☐

While *I have the absolute certainty* corresponds to 'very certain' level, this level has been defined as 'absolutely certain' in the rest of the thesis while the *I have some doubts* level is shortened to the certainty level 'not so certain' as it corresponds well to both the original English level as well as the back translated English level

11.3 APPENDIX C

Choice scenarios employed for location and linguistic representation methods

Choice Scenarios: Location representation, binary choice and one/two stage Likert elicitation technique

	OPTION A				OPTION B			
	View	Noise	Charge	Sun.	View	Noise	Charge	Sun.
Scen. 1	6F	6F	70	6F	6F	6F	50	3F
Scen. 2	6F	6T	55	6F	6F	6F	80	6F
Scen. 3	6F	3F	85	6F	6T	6F	40	6F
Scen. 4	6F	6F	40	6F	6T	3T	75	3F
Scen. 5	6F	3F	70	6F	3F	6F	50	6T
Scen. 6	6F	6T	85	6F	3F	6F	40	3F
Scen. 7	6F	6F	85	6F	3T	6F	40	3T
Scen. 8	6F	6F	70	6F	3T	3T	50	6F
Scen. 9	6F	3F	40	6F	6F	6F	75	3T
Scen. 10	6F	6F	85	6F	6F	3T	40	3F
Scen. 11	6F	6F	55	6F	6T	6F	80	3F
Scen. 12	6F	6T	70	6F	6T	6F	50	3T
Scen. 13	6F	6F	40	6F	3F	6F	75	6F
Scen. 14	6F	6F	55	6F	3F	3T	80	3T
Scen. 15	6F	3F	55	6F	3T	6F	80	3F
Scen. 16	6F	6T	40	6F	3T	6F	75	6T

Choice Scenarios: Linguistic representation, binary choice and one/two stage Likert elicitation technique

	OPTION A				OPTION B			
	View	Noise	Charge	Sun.	View	Noise	Charge	Sun.
Scen. 1	good	noisy	70	v.good	good	noisy	50	good
Scen. 2	good	quiet	70	v.good	good	noisy	50	v.good
Scen. 3	good	neither	55	v.good	good	quiet	80	v.good
Scen. 4	good	quiet	75	v.good	good	noisy	40	v.good
Scen. 5	neither	quiet	55	v.good	good	noisy	80	good
Scen. 6	neither	noisy	40	neither	good	noisy	75	v.good
Scen. 7	neither	neither	40	v.good	good	quiet	85	v.good
Scen. 8	neither	quiet	85	v.good	good	noisy	40	good
Scen. 9	good	quiet	55	neither	good	noisy	80	v.good
Scen. 10	good	neither	40	v.good	good	quiet	75	good
Scen. 11	good	noisy	85	v.good	good	noisy	40	v.good
Scen. 12	good	quiet	85	neither	good	noisy	40	v.good
Scen. 13	neither	neither	50	neither	good	quiet	70	v.good
Scen. 14	neither	quiet	70	v.good	good	noisy	50	v.good
Scen. 15	neither	noisy	55	v.good	good	noisy	80	v.good
Scen. 16	neither	quiet	40	v.good	good	noisy	75	v.good

11.4 APPENDIX D

This appendix provides the fuzzy rules developed for the main FL models given in Chapter 8.

1. Location rating model with attribute difference as input, without error

Rule no.	View diff (A-B)	Noise diff.	Sun. diff.	Charge diff.	Choice
1	High -ve	High -ve			B
2	High -ve		High -ve		B
3	High -ve			High +ve	B
4		High -ve	High -ve		B
5		High -ve		High +ve	B
6			High -ve	High +ve	B
7	High +ve	High +ve			A
8	High +ve		High +ve		A
9	High +ve			High -ve	A
10		High +ve	High +ve		A
11		High +ve		High -ve	A
12			High +ve	High -ve	A
13	Med	Med	Med	High -ve	A
14	Med	Med	Med	High +ve	B
15	High -ve	Med	Med	Med	B
16	High +ve	Med	Med	Med	A
17	Med	High -ve	Med	Med	B
18	Med	High +ve	Med	Med	A
19	Med	Med	High -ve	Med	B
20	Med	Med	High +ve	Med	A

2. Location rating model with attribute difference as input with Gumbel/Normal error

Rule no.	**View diff. (A-B)**	**Noise diff.**	**Sun. diff.**	**Charge diff.**	**Error diff.**	**Choice**
1	High -ve	High +ve	Med	Med	High -ve	B
2	High -ve	High +ve	Med	Med	High +ve	A
3	High -ve	Med	Med	High -ve	High -ve	B
4	High -ve	Med	Med	High -ve	High +ve	A
5	Med	High -ve	Med	High -ve	High -ve	B
6	Med	High -ve	Med	High -ve	High +ve	A
7	Med	High -ve	High +ve	Med	High -ve	B
8	Med	High -ve	High +ve	Med	High +ve	A
9	Med	Med	High -ve	High -ve	High -ve	B
10	Med	Med	High -ve	High -ve	High +ve	A
11	Med	Med	High -ve	Med	High -ve	B
12	Med	Med	High -ve	Med	High +ve	A
13	Med	Med	Med	Med	High -ve	B
14	Med	Med	Med	Med	High +ve	A
15	Med	Med	High +ve	High +ve	High -ve	B
16	Med	Med	High +ve	High +ve	High +ve	A
17	Med	High +ve	High -ve	Med	High -ve	B
18	Med	High +ve	High -ve	Med	High +ve	A
19	Med	High +ve	Med	High +ve	High -ve	B
20	Med	High +ve	Med	High +ve	High +ve	A
21	High +ve	High -ve	Med	Med	High -ve	B
22	High +ve	High -ve	Med	Med	High +ve	A
23	High -ve	High -ve				B
24	High -ve		High -ve			B
25	High -ve			High +ve		B
26		High -ve	High -ve			B

27		High -ve		High +ve		B
28			High -ve	High +ve		B
29	High +ve	High +ve				A
30	High +ve		High +ve			A
31	High +ve			High -ve		A
32		High +ve	High +ve			A
33		High +ve		High -ve		A
34			High +ve	High -ve		A
35	Med	Med	Med	High -ve		A
36	Med	Med	Med	High +ve		B
37	High -ve	Med	Med	Med		B
38	High +ve	Med	Med	Med		A
39	Med	High -ve	Med	Med		B
40	Med	High +ve	Med	Med		A
41	Med	Med	High -ve	Med		B
42	Med	Med	High +ve	Med		A

3. Linguistic rating model with attribute difference as input without error

Rule no.	**View diff (A-B)**	**Noise diff.**	**Sun. diff.**	**Charge diff.**	**Choice**
1	Low	High -ve	High -ve		B
2	Low	High -ve	Med		B
3	Low	High -ve	High +ve	High -ve	A
4	Low	High -ve	High +ve	High +ve	B
5	Low	Med	High -ve		B
6	Low	Med	Med	High -ve	A
7	Low	Med	Med	High +ve	B
8	Low	Med	High +ve	High-ve	A
9	Low	Med	High +ve	High +ve	B
10	Low	High +ve	High -ve	High -ve	A
11	Low	High +ve	High -ve	High +ve	B
12	Low	High +ve	Med		A
13	Low	High +ve	High +ve		A
14	Med	High -ve			B
15	Med	Med	High -ve		B
16	Med	Med	Med	High -ve	A
17	Med	Med	Med	High +ve	B
18	Med	High +ve	High -ve		B
19	Med	High +ve	Med	High -ve	A
20	Med	High +ve	Med	High +ve	B
21	Med	High +ve	High +ve		A

4. Linguistic ratings model with attribute difference as input with Gumbel/Normal error

Rule no.	**View diff (A-B)**	**Noise diff.**	**Sun. diff.**	**Charge diff.**	**Error diff.**	**Choice**
1	Low	High -ve	Med			B
2	Low	High -ve	High +ve	High -ve		A
3	Low	High -ve	High +ve	High +ve		B
4	Low	Med	High -ve			B
5	Low	Med	Med	High -ve		A
6	Low	Med	Med	High +ve		B
7	Low	Med	High +ve	High -ve		A
8	Low	Med	High +ve	High +ve		B
9	Low	High +ve	High -ve	High -ve		A
10	Low	High +ve	High -ve	High +ve		B
11	Low	High +ve	Med			A
12	Low	High +ve	High +ve			A
13	Med	High -ve				B
14	Med	Med	High -ve			B
15	Med	Med	Med	High -ve		A
16	Med	Med	Med	High +ve		B
17	Med	High +ve	High -ve			B
18	Med	High +ve	Med	High -ve		A
19	Med	High +ve	Med	High +ve		B
20	Med	High +ve	High +ve			A
21	Low	Med	Med	Med	High -ve	B
22	Low	Med	Med	Med	High +ve	A
23	Med	High -ve	High +ve			B
24	Med	High +ve	High -ve			B
25	Med	High +ve	High +ve			A
26		High -ve	High -ve			B

5. Linguistic model with absolute categories as input, without error – 8 rules

Rule	**VA**	**VB**	**NA**	**NB**	**SA**	**SB**	**C diff.**	**Choice**
1			noisy			v.good		B
2	good							A
3	neither			quiet				B
4			noisy		v.good			A
5			noisy		neither			B
6			noisy				high+ve	B
7	good				v.good			A
8				quiet		v.good		B