
Floresta Relíquia © FMB foto José Moura

Paula Marisa Nunes Simões

THE RECREATIONAL USE VALUE OF A NATIONAL FOREST

Dissertação de doutoramento na área científica de Economia, orientada pelo Senhor Professor Doutor Luís Cruz e pelo Senhor Professor Doutor Eduardo Barata e apresentada à Faculdade de Economia da Universidade de Coimbra.

Agosto de 2012

· U C ·

UNIVERSIDADE DE COIMBRA

FEUC FACULDADE DE ECONOMIA
UNIVERSIDADE DE COIMBRA

Paula Marisa Nunes Simões

THE RECREATIONAL USE VALUE OF A NATIONAL FOREST

Tese de Doutoramento em Economia apresentada à Faculdade de Economia da Universidade de Coimbra para obtenção do grau de Doutor.

Orientadores:

Prof. Doutor Luís Cruz

Prof. Doutor Eduardo Barata

Coimbra, 2012

FEUC FACULTY OF ECONOMICS
UNIVERSITY OF COIMBRA

Paula Marisa Nunes Simões

THE RECREATIONAL USE VALUE OF A NATIONAL FOREST

A dissertation submitted for the degree of Doctor of Philosophy (Ph.D.) in Economics to the Faculty of Economics of the University of Coimbra.

Supervisers:

Prof. Luís Cruz

Prof. Eduardo Barata

Coimbra, 2012

ACKNOWLEDGEMENTS

Regarding the completion of this dissertation, I benefited greatly from the contributions of a number of people and institutions and I especially wish to thank them.

First and foremost, I would like to take this opportunity to express my gratitude to my supervisors, Professor Eduardo Barata and Professor Luís Cruz for their ready guidance, pertinent remarks and valuable suggestions. Their continuous encouragement, enthusiastic support and availability were essential to conducting and finalizing this dissertation.

I want to thank all the Bussaco National Forest visitors who kindly spent some of their time answering the questionnaire. Their cooperation was invaluable.

My thanks to three institutions which were essential in this research:

The Fundação *Mata do Bussaco*, in particular its president, António Jorge Franco, and Liliana Duarte, whose help with administering the surveys was invaluable.

The Fundação para a Ciência e Tecnologia (FCT), whose financial support through the PROTEC made it possible to obtain release from classes for some semesters in order to conduct the research.

The Polytechnic Institute of Leiria (IPL) whose classes' management allowed me to work on this research on a full time basis for three semesters.

I must express my thanks to Professor Óscar Lourenço for his help and constructive comments regarding the econometric analysis. A special word goes to my friend Vítor Hugo Ferreira, who generously lent me important reference materials.

Finally, I wish to thank my family and my friends for their support.

ABSTRACT

National forests and woodlands are some of the environmental public resources that provide a diversity of goods and services to society. Supporting, regulating, provisioning and cultural services are all known to contribute to human well-being. As these services are not traded in regular markets because of the public or semi-public characteristics of the resources involved, their values are largely unknown. However, a deeper knowledge of the related benefits' value is expected to help to enhance management practices. The research described in this dissertation concentrates on the analysis of the benefits related to recreational activities enjoyed in national forests and in understanding the demand for these environmental services. The research was motivated by the perception that these values are largely unknown, particularly in Portugal. Bussaco National Forest was chosen as the case study area, but the conclusions are likely to be adapted and extended to other national forests.

Two non-market valuation techniques, the travel cost method and the contingent behaviour method, are used to estimate the recreational use benefits. The travel cost method, which belongs to the group of revealed preferences techniques, is used to analyse the actual behaviour and enables us to estimate recreational use values in current conditions. The individual version of the method is identified as the most accurate in the present context as we analyse the recreational demand of a forest visited by people living at different distances from it. From the management perspective, it is also important to address how people would behave if new hypothetical conditions were to be observed. It is particularly important to predict the effects on demand resulting from changes in forest access costs and from the deterioration of current conservation conditions due to a forest fire. The contingent behaviour method, which belongs to the group of stated preferences techniques, is applied jointly with the travel cost method to assess the effects of these changes.

Count data models corrected for endogenous stratification and ordered models are used in the analysis of the actual visit behaviour. Travel cost, substitute cost, income per capita, visit motivations, on-site time and visit distribution during the year were identified as the main explanatory variables of demand. Price and income elasticity of demand computed using count data models are low. This result is supported by the ordered models, as results show that the change in income/price must be quite significant to modify demand levels. Considering only the current users, the forest recreational use value estimated for the past three years is about €106 700. A count data model and a pseudo-panel specification is used to combine contingent and observed travel behaviour. The analysis reveals that visitors are sensitive to price and quality changes and that in the forest fire scenario the intended number of trips would be seriously reduced, thus imposing an important welfare loss. There are evidences of hypothetical bias in answers to future behaviour if current conditions do not change and signals of strategic bias when changes in management options are in view. There are no signals of these biases when the quality changes are exogenous.

Keywords: Travel cost method, contingent behaviour method, forest recreation demand, count data, joint estimation.

JEL classification: C83, Q26, Q51, Q57.

RESUMO

As florestas nacionais estão entre os recursos ambientais públicos que fornecem à sociedade uma grande diversidade de bens e serviços. Reconhecidamente os serviços de suporte, regulação, provisão e culturais contribuem para o bem-estar do ser humano. No entanto, como estes serviços não são transacionados em mercados convencionais dadas as características de bens públicos ou semipúblicos dos recursos envolvidos na sua provisão, parte dos seus valores permanecem desconhecidos. Contudo, espera-se que um conhecimento mais profundo do valor dos benefícios proporcionados possa contribuir para melhorar as práticas de gestão. A investigação conduzida nesta dissertação centra-se na análise dos benefícios proporcionados pelas atividades recreativas desenvolvidas nas florestas nacionais e na compreensão da procura por estes serviços ambientais. A motivação para esta investigação resulta da perceção de que estes valores permanecem largamente desconhecidos, particularmente em Portugal. A Mata Nacional do Bussaco foi escolhida para estudo de caso, mas as conclusões são suscetíveis de serem adaptadas e alargadas a outras florestas nacionais.

Na estimação do valor recreativo são usadas duas técnicas de avaliação externas ao mercado, o método dos custos de viagem e o método do comportamento contingente. O método dos custos de viagem, que pertence ao grupo das técnicas de preferências reveladas, é usado na análise do comportamento atual e permite estimar o valor de uso recreativo da floresta nas atuais condições. A versão individual do método é identificada como a mais adequada no presente contexto, uma vez que analisamos a procura recreativa de uma floresta visitada por pessoas que residem a diferentes distâncias. Do ponto de vista da gestão da floresta é também importante estimar como se comportarão os indivíduos perante condições hipotéticas. É particularmente importante prever os efeitos sobre a procura resultantes de alterações nos custos de acesso à floresta ou da deterioração nas atuais condições de conservação devido a um incêndio florestal. O método do comportamento contingente, que pertence ao grupo das técnicas de preferências declaradas, é aplicado conjuntamente com o método dos custos de viagem no diagnóstico dos efeitos destas mudanças.

Na análise do comportamento observado são usados modelos com dados de contagem, corrigidos de estratificação endógena, e modelos ordenados. As principais variáveis explicativas da procura identificadas através da análise econométrica são: os custos de viagem, o custo de acesso a locais substitutos, o rendimento *per capita*, os motivos da visita, a duração da visita e a sua distribuição ao longo do ano. A elasticidade-preço e a elasticidade-rendimento da procura estimadas com base nos modelos de contagem revelam uma procura rígida. Este resultado é corroborado pelos modelos ordenados que mostram que a alteração no rendimento/preço terá que ser muito expressiva para modificar os níveis de procura. Considerando apenas os utilizadores atuais, o valor de uso recreativo da floresta estimado para o conjunto dos três anos anteriores é de cerca de €106 700. Na combinação do comportamento contingente com o comportamento observado é usado um modelo de contagem e a especificação corresponde a um *pseudo*-painel. A análise revela que os visitantes são sensíveis a alterações no preço e na qualidade e que num cenário de ocorrência de um incêndio florestal, o número de viagens previstas seria seriamente reduzido, daí resultando uma perda significativa de bem-estar. Se não se alterarem as condições atuais, há evidência de enviesamento hipotético nas respostas relativas ao comportamento futuro e há sinais de enviesamento estratégico quando estão em causa alterações da responsabilidade dos decisores. Quando as alteraçãos de qualidade são exógenas não se identificam sinais destes enviesamentos.

Palavras-chave: método dos custos de viagem, método do comportamento contingente, procura recreativa da floresta, dados de contagem, estimação conjunta.

Classificação JEL: C83, Q26, Q51, Q57.

CONTENTS

Acknowledgements......i
Abstract......iii
Resumo......v
Contents......vii
List of Tables......xi
List of Figures......xiii
List of Abbreviations......xv
CHAPTER 1 – Introduction......1
  1.1. Motivation and purpose......3
  1.2. Dissertation outline......5
    1.2.1. Part I: Conceptual framework......6
    1.2.2. Part II: Empirical analysis......7
PART I: CONCEPTUAL FRAMEWORK......11
CHAPTER 2 – Essential concepts on non-market environmental valuation......13
  2.1. Introduction......13
  2.2. Total economic value......14
  2.3. Economic values and non-market valuation methods......19
  2.4. Willingness to pay and willingness to accept......22
    2.4.1. Individual preferences......22
    2.4.2. Welfare measures for a price change......24
    2.4.3. Welfare measures under quantity constraints......30
    2.4.4. Comparison of welfare measures......38
  2.5. Concluding remarks......42
CHAPTER 3 – Non-market environmental valuation: an appraisal focused on Portugal......45
  3.1. Introduction......45
  3.2. Valuation methods......46
    3.2.1. Contingent valuation......46
    3.2.2. Choice modelling......51
    3.2.3. Contingent behaviour......53
    3.2.4. Travel cost method......54
    3.2.5. Closing notes......56
  3.3. Empirical research focusing on Portuguese resources: a review......57
  3.4. Concluding remarks......63

CHAPTER 4 – The travel cost method: a review 65
4.1. Introduction 65
4.2. Theoretical background 66
4.3. The TCM modelling framework 68
4.3.1. Single site models 69
4.3.2. Multiple site models 74
4.4. Main issues in the TCM 88
4.4.1. Time 89
4.4.2. Substitute sites 93
4.4.3. Trip 94
4.4.4. Other issues 96
4.5. Concluding remarks 96
PART II: EMPIRICAL ANALYSIS 99
CHAPTER 5 – Non-market recreational value of a national forest: survey design and results 101
5.1. Introduction 101
5.2. Case study area: Bussaco National Forest 102
5.3. The survey 103
5.3.1. Questionnaire administration 103
5.3.2. Data 106
5.4. Concluding remarks 122
CHAPTER 6 – Application of the travel cost method using count data models accounting for endogenous stratification 125
6.1. Introduction 125
6.2. Modelling aspects 126
6.2.1. Poisson regression model 127
6.2.2. Negative Binomial model 129
6.2.3. Endogenous stratification 131
6.3. The variables 134
6.4. Estimations and results 141
6.5. Welfare estimation 144
6.6. Concluding remarks 147
CHAPTER 7 – From count data to ordered categories, data reinterpretation 151
7.1. Introduction 151
7.2. Ordered models 152
7.3. Estimation and results 154
7.4. Concluding remarks 160

CHAPTER 8 – Combining observed and contingent travel behaviour 163
8.1. Introduction 163
8.2. Combining observed and contingent travel data: a review 166
8.3. Contingent behaviour data 172
8.4. Methods 175
8.4.1. Econometric models 175
8.4.2. Parameter estimates 178
8.4.3. Welfare change in the forest fire scenario 185
8.5. Concluding remarks 187
CHAPTER 9 – Conclusion 189
Appendices 197
Appendix A: Questionnaire 199
Appendix B: NB distribution density function 203
Appendix C: Chapter 6 stata commands and outputs 207
Appendix D: Chapter 7 stata commands and outputs 213
Appendix E: RE-NB distribution function 219
Appendix F: Chapter 8 stata outputs 223
Appendix G: The LR test and goodness-of-fit statistics 235
References 237

LIST OF TABLES

Table 2.1: Compensating and equivalent variation..30
Table 2.2: Compensating and equivalent surplus...35
Table 3.1: Biases likely to affect CVM...48
Table 3.2: CVM elicitation formats 49
Table 3.3: Variants of CM 52
Table 3.4: Versions of the TCM..55
Table 3.5: Studies applying non-market valuation methods to Portuguese resources......58
Table 5.1: Household monthly disposable income 108
Table 5.2: Perceptions concerning environment 110
Table 5.3: Opinions about Bussaco 112
Table 5.4: Visits' purpose..115
Table 5.5: Visitation levels 121
Table 6.1: Potential explanatory variables..135
Table 6.2: Descriptive statistics of potential explanatory variables 140
Table 6.3: Estimation results using count data models..141
Table 6.4: Marshallian consumer surplus...145
Table 6.5: Hicksian welfare measures..146
Table 6.6: Aggregated recreational benefits...146
Table 7.1: Categories definition and distribution...155
Table 7.2: Estimation results using ordered models...156
Table 7.3: AMPE in ordered logit model...159
Table 7.4: Descriptive statistics for predicted probabilities and observed frequencies....160
Table 8.1: TCM-CB/CVM literature review 170
Table 8.2: Stated trips in different scenarios 174
Table 8.3: Descriptive statistics...178
Table 8.4: Demand models..180
Table 8.5: Stated and predicted number of trips...183
Table 8.6: Consumer surplus...186

LIST OF FIGURES

Figure 1.1: Structure of the dissertation and relation between chapters 10
Figure 2.1: Total Economic Value 16
Figure 2.2: Hicksian welfare measures from a price change 26
Figure 2.3: Demand curves and welfare measures 28
Figure 2.4: CS and ES for a quantity constrained good..32
Figure 2.5: CS and ES for a quantity-constrained good with null price 34
Figure 3.1: Non-market valuation techniques...45
Figure 5.1: Respondents' origin 109
Figure 5.2: Environmental perceptions...110
Figure 5.3: Visit counts 115
Figure 5.4: Distance categories...116

LIST OF ABBREVIATIONS

AIC	Akaike information criterion
AME	Average marginal effect
AMPE	Average marginal probability effect
BIC	Bayesian information criterion
CA	Conjoint analysis
CB	Contingent behaviour
CE	Choice experiment
CM	Choice modelling
CS	Compensating surplus
CV	Compensating variation
CVM	Contingent valuation method
DBDC	Double bounded dichotomous choice
DC	Dichotomous choice
ES	Equivalent surplus
EV	Equivalent variation
HTCM	Hedonic travel cost method
IB	Iterative bidding
ICB	Intended contingent behaviour
iid	Independent and identically distributed
ITCM	Individual travel cost model
km	Kilometre
LR	Likelihood ratio
MBDC	Multiple bounded dichotomous choice
MCS	Marshallian consumer surplus
ML	Maximum likelohood
MMS	Minimum monthly salary
MPE	Marginal probability effect
NB	Negative Binomial
NOAA	National Oceanic and Atmospheric Administration
OE	Open ended
OHBDC	One-and-one-half-bound dichotomous choice
OLS	Ordinary least squares
PC	Payment card
PGNP	Peneda-Gerês National Park
POIS	Poisson
RCB	Reassessed contingent behaviour
RE	Random effects
RP	Revealed preferences

RRDM	Regional recreation demand model
RUM	Random utility model
SP	Stated preference
s.t.	Subject to
SU	Seemingly unrelated
TCM	Travel cost method
TEEB	The Economics of Ecosystems and Biodiversity
TEV	Total Economic Value
WTA	Willingness to accept
WTP	Willingness to pay
ZTCM	Zonal travel cost model

CHAPTER 1 – INTRODUCTION

1.1. MOTIVATION AND PURPOSE

Environmental resources provide a wide diversity of goods and services that are valuable for society. Outdoor recreation opportunities are among these services and several types of activity are involved. They range from the more passive ones such as sitting and relaxing, enjoying the landscape or watching the fauna (e.g., whales or birds), to more active ones such as walking, skiing, mountain biking or climbing.

In recent years, demand for outdoor recreation in developed countries has been rising and projections point towards this increase continuing. Hence, growing demand pressure is expected for natural areas that offer outdoor leisure opportunities. A wide range of natural spaces are used for outdoor recreation purposes and national surveys conducted elsewhere suggest that forests and woodlands are among the most popular ones (Zandersen, 2005; Bell *et al.*, 2007: 35).

In Portugal there are national forests[1] which are public property and have free access for recreation. Historically, these forests were promoted by the public authorities with a dual aim: to protect soil and the coastal areas from erosion and to produce timber and related products, particularly resin. These are still objectives today, but national forests are now additionally used for recreational purposes.

Indeed, the Portuguese law recognizes the recreational goals of public forests[2] by establishing the government objectives "to promote and ensure access to the social utilization of the forest, promoting the harmonization of its multiple functions and safeguarding its landscape, recreational, scientific and cultural aspects*".* Furthermore, the national strategy relative to forests, published in 2006, highlights the need to identify the value that society attaches to woodland areas in order to maximize their value. This document also mentions that both direct use values and indirect use values should be

---

[1] For more details on the list of Portuguese national forests see www.afn.min-agricultura.pt/portal/gestao-florestal/regflo/stmatas ;

[2] For example, in Article 4(b) of the Basic Forest Law (*Lei de Bases da Floresta)* published on 17 August 2006.

included in such valuation. However, in Portugal economic values associated with outdoor recreation are largely unknown.

In general the benefits of outdoor recreation are not reflected in market prices because the areas involved are of free access, in spite of being private, quasi-public or public goods. For this reason, non-market valuation[3] techniques are required to estimate the welfare benefits accruing from their recreational use. Typically, these techniques are divided in two broad groups, the stated preference (SP) and the revealed preference (RP) methods.

The analysis reported in this dissertation makes use of both techniques to evaluate the recreational use value of a Portuguese national forest. More specifically, the travel cost method (TCM), an RP method, and the contingent behaviour (CB) method, an SP method, are used in the evaluation exercise. The Bussaco National Forest is used as the case study in the analysis, but conclusions are expected to provide significant insights relevant to other, similar, spaces.

A number of empirical studies have explored analogous questions, although in different environmental and social contexts (Cameron, 1992; Beal, 1995; Englin and Shonkwiler, 1995; Liston-Heyes and Heyes, 1999; Alberini *et al.*, 2007). Regarding Portugal, knowledge on forest recreation is poor. First of all, the level of demand is unknown. There is also a lack of information about visitors' socio-demographic characteristics, their preferences with respect to forest features and the motivation and characteristics of their visits. Reactions of visitors to changes in current conditions and/or in access prices are also unknown. To partially fill this gap, demand is estimated and welfare figures derived. Accordingly, the ultimate purpose of the empirical environmental valuation study is to

---

[3] The Concise Oxford Dictionary (1990: 1356) defines valuation as "***1 a*** *an estimation (especially by a professional valuer) of a thing's worth;* ***b*** *the worth estimated.* ***2*** *the price set on a thing*". The word 'evaluation' has been used synonymously in the literature. The same dictionary (p. 404) considers 'evaluation' equivalent to 'evaluate', which is defined as "***1*** *assess, appraise.* ***2 a*** *find or state the number or amount of*. **b** find a numerical expression for". Accordingly, we consider that any of these words would be used correctly in this text. In terms of the Portuguese language, following the Dicionário de Inglês-Português (1994: 277, 1400), we use the word '*avaliação'* whenever translation is required.

produce structured information on demand, and consequently on the value attached by users to forest recreation, in order to support forest management. As affirmed by Horne *et al.* (2005: 190) “since the goal of forest management at outdoor recreation sites is to fulfil the recreational needs of different visitors, forest managers need information on the preferences of forest recreationists”.

Two relevant contributions are expected. On the practical level, this study is intended to contribute to forest management in Portugal through an illustrative empirical study that shows decision makers the need (and also how) to expand the analysis they typically/traditionally consider, by broadening its scope to include the resource of ‘recreational value’. Accordingly, forest managers will be able to use the information on factors that drive forest recreation demand to increase the attractiveness of the forests/parks they manage, and if they use the estimates provided they will be in a better position to infer the recreational value of the forest and predict changes in visitors’ behaviour due to changes in policy variables. On the theoretical level, besides the inputs to the assessment of general trends in forest recreation, this study contributes to the ongoing debate about the issues underlying the application of TCM, CB and TCM-CB.

1.2. DISSERTATION OUTLINE

In addition to this introductory first chapter and the Conclusion (Chapter 9), this dissertation has seven other chapters which are organized in two parts. The conceptual framework underlying non-market valuation is presented in chapters 2 to 4, which comprise Part I. Part II is concerned with the empirical analysis. It consists of chapters 5 to 8, which present the empirical application and discuss the main results. Furthermore, some of the theoretical and methodological issues raised in Part I are empirically addressed in Part II.

1.2.1. PART I: CONCEPTUAL FRAMEWORK

The main aim of the first part of this dissertation is to provide the theoretical concepts and methodological framework which make up the background of non-market valuation analysis. This is achieved through three chapters, evolving from a wider to a narrower perspective.

To begin with, Chapter 2 presents some essential concepts implicit in the non-market valuation literature. Particular attention is paid to the notion of total economic value (TEV) and its components, largely because there is some disagreement about the concept's connotation, even though it is widely used. Accordingly, the concept and composition of TEV is discussed in Section 2.2 with the aim of clarifying our interpretation. As a result, a specific structure is proposed, which is implicit in the analysis performed in the rest of the chapters of this dissertation. The possibility of monetization of the environmental values is further discussed and our position on the subject established. Section 2.3 provides an overview of the relation between the non-market valuation methods and components of TEV. Finally, Section 2.4 describes the analytical framework based on neoclassical principles used in the estimation of the willingness to pay (WTP) or the willingness to accept (WTA), when applying non-market valuation methods.

Chapter 3 presents an overview of four main non-market valuation methods employed in environmental valuation: the contingent valuation method (CVM); choice modelling (CM); the CB method, and the TCM. Their main theoretical aspects are reviewed in Section 3.2. Section 3.3 provides a comprehensive survey of contributions using these methods in the valuation of Portuguese environmental resources. The survey is structured around three main research questions: "What has been done in the domain of non-market environmental valuation in Portugal?"; "What common features can be observed across different studies?" and "What do we know about the validity/reliability of the monetary values estimated?".

A deeper literature review of the TCM is left until Chapter 4 as its role in this dissertation requires a more meticulous analysis. The main objective was to examine the possible different versions of the model, helping us to select the most accurate one(s) to use in this research. Equally important, this chapter sets out the fundamental theoretical basis for the questionnaire design. It also offers some guidance on the routes to be followed in the econometric analysis. Accordingly, Section 4.2 presents the TCM theoretical background. Section 4.3 provides a comprehensive overview of the main TCM versions. Section 4.4 discusses the main empirical issues and the practical solutions adopted.

1.2.2. PART II: EMPIRICAL ANALYSIS

The second part of this dissertation is dedicated to the empirical analysis, which is performed using a case study. The four chapters which make up this part are interrelated.

Chapter 5, in particular, details important data inputs that are used in the three following chapters, which focus on the econometric analysis. Indeed, Chapter 5 presents the main features concerning the data used in the empirical analysis. First, Section 5.2 gives a brief description of the area chosen as case study – the Bussaco National Forest. The main reasons for choosing this forest for the case study are also presented here. The survey is then analysed in detail in Section 5.3. The design of the questionnaire used to collect information is described, the motivations behind each question are explained, the solutions adopted to deal with the different issues of the TCM analysis are discussed and the descriptive statistics are presented. Besides the socio-demographic characterization of the sample, this chapter offers answers to some key questions in the TCM framework, such as: “What is the average visit frequency?”, “What are the main motivations of the visit?”, “Are visits single or multiple destination?” and “What are the substitutes for this forest?”.

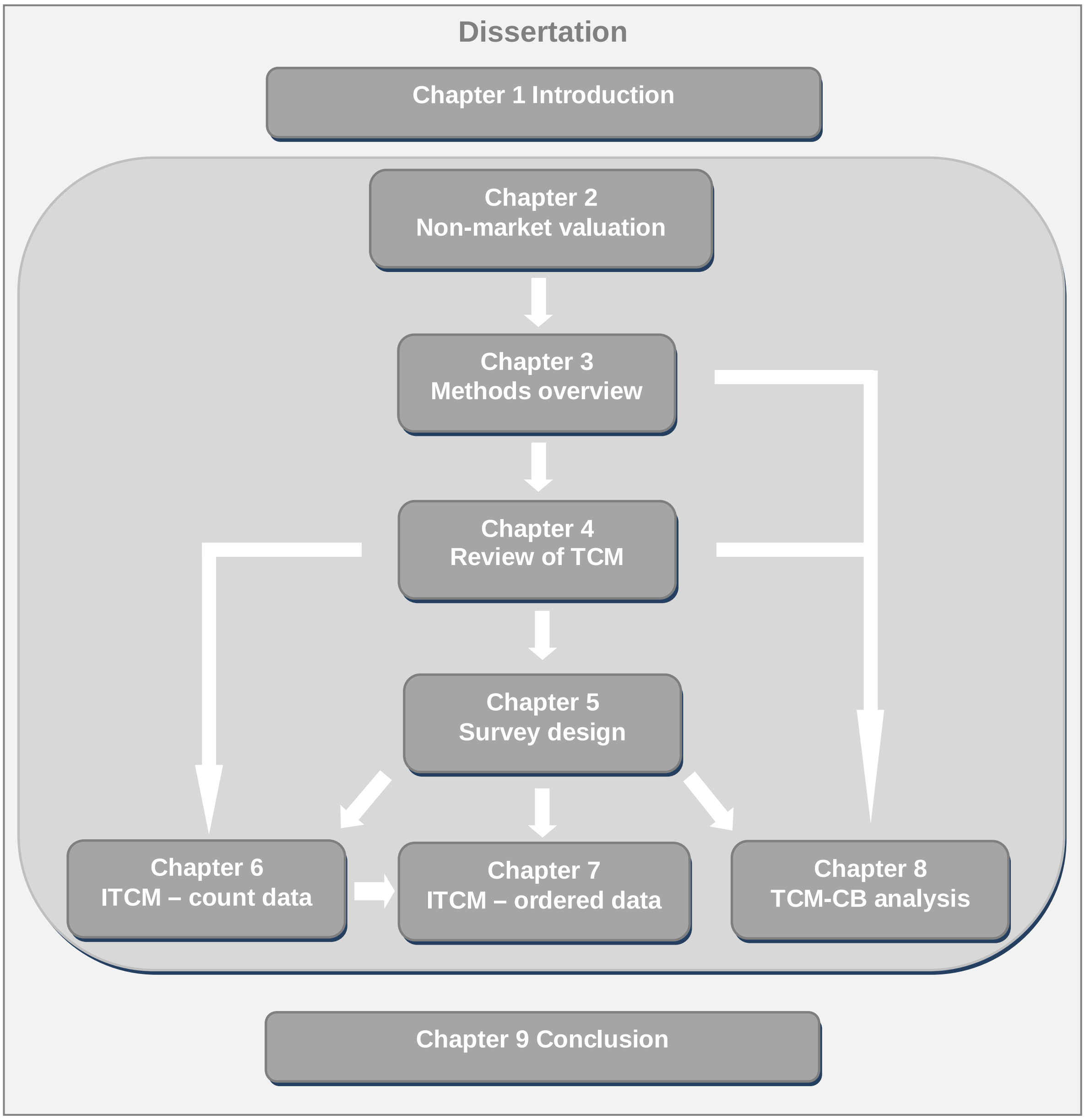
In Chapter 6, the demand for recreation in the Bussaco National Forest is derived and analysed using the individual version of the TCM. The use of this method is meant to derive the demand curve. The main questions we should be able to answer are: “What variables explain demand?” and “How much is the consumer surplus derived from the visits?” When answering the first question, the role of the substitutes and the relation between demand and income is identified. Three sets of potential explanatory variables are considered: socio-demographic variables; variables characterizing visitor preferences; and variables characterizing the visits. This chapter is organized as follows. In Section 6.2, the (econometric) modelling aspects are described. First, the main characteristics of the standard Poisson and Negative Binomial (NB) models are presented. Then, the corrections required for endogenous stratification and zero truncation are described. Section 6.3 explains and justifies the design of the variables not obtained directly from Chapter 5. In Section 6.4 the estimation results are presented and discussed. Finally, in Section 6.5 welfare measures are derived making use of the results obtained in the preceding section.

Chapter 7 provides a reinterpretation of the data using a different econometric approach – the ordered category models. This alternative has often been suggested but seldom explored. However, the classification of visits into discrete categories may better express the differences between visitors than the count if categories better characterize demand than the number of visits *per se*. Hence, this chapter looks for answers to two main interrelated research questions. One asks whether it is meaningful to divide visitors into sub-groups in accordance with their visit levels. The other asks whether the set of significant explanatory variables stays the same as when count data is considered. This chapter receives inputs from Chapter 6 because the set of explanatory variables considered in the econometric models is basically the one used before. Thus, this chapter is organized along the same lines as the preceding one. Accordingly, Section 7.2 sets out the econometric model and Section 7.3 presents and discusses the estimation’s results.

In Chapter 8, revealed and stated preferences are jointly analysed using the TCM-CB framework and applying a pseudo-panel structure in the econometric analysis. This analysis aims to provide information about the intended behaviour of current visitors in three hypothetical situations: if current conditions endure, if travel cost changes and if conservation conditions change due to a forest fire. This information is particularly relevant to forest managers as it reports the effects that these changes are expected to have on demand. For this, the chapter is organized as follows. In Section 8.2 the background literature on combination of observed and contingent travel data is presented. Section 8.3 presents contingent behaviour data on Bussaco forest visits. In Section 8.4 the econometric approach is set out, the results of the econometric models are reported and the estimations of changes in welfare resulting from hypothetical scenarios are computed.

Finally, in Chapter 9 the most important conclusions are summarized, with a special emphasis to those derived from the findings of the empirical analysis. This chapter concludes by specifying some avenues for possible future research. Figure 1.1 presents a schematic overview of the links and interdependencies between the chapters in this dissertation.

Figure 1.1: Structure of the dissertation and relation between chapters



PART I: CONCEPTUAL FRAMEWORK

CHAPTER 2 – ESSENTIAL CONCEPTS ON NON-MARKET ENVIRONMENTAL VALUATION

2.1. INTRODUCTION

People live in a world of scarcity where resources tend to be insufficient to produce the goods and services that would be necessary to satisfy all the human needs. This is one of the basic premises of economic analysis. Among the scarce resources are the natural ones on which individuals depend to satisfy a set of needs and wants. Their range is extensive, going from the most basic such as breathing pure air, to much more complex ones such as recreation and short breaks.

Moreover, as societies become richer and more urbanized, demand for environmental services tends to increase, as illustrated by the U-shaped environmental Kuznets curve (Culas, 2007: 430). This means that as the time goes by, demand for environmental services tends to increase, putting great pressure on natural resources. Hence the efficient allocation of the resources involved is increasingly urgent. However, the values of environmental services are not usually directly revealed in market transactions because many of them are non-tradable. Accordingly, non-market valuation techniques must be used to assess their economic value and promote efficiency.

One of the difficulties in non-market valuation derives from the fact that the concept of value is neither unique nor trouble-free. The analysis of environmental goods/services involves a wide diversity of aspects, which adds complexity to the concept. In view of that, in Section 2.2 we discuss the concept of TEV of natural resources and suggest a specific structuring of their different dimensions. The relation between the different dimensions of TEV and non-market valuation techniques is also examined. Then we discuss the possibility and accuracy of using monetary units to evaluate environmental resources. Section 2.3 discusses the relation between the components of value and the valuation methods. In Section 2.4 we present the analytical framework used to derive the theoretical welfare measures. Section 2.5 concludes the chapter.

2.2. TOTAL ECONOMIC VALUE

In order to clarify the concept of TEV we selected from the literature the three definitions below, because of their complementary focus. The TEV of a natural resource can be defined as the sum of all its marketable and non-marketable values (Torras, 2000). The Earthscan book of “The Economics of Ecosystems and Biodiversity” (TEEB) defines the TEV of ecosystems and biodiversity as “the sum of the values of all service flows that natural capital generates now and in the future – appropriately discounted” (Pascual and Muradian, 2010: 188). The definition of ecological values used by Norton and Noonan (2007: 666) can also be used in defining TEV. Their words are: “the whole range of values that humans derive from ecological systems, including services, provision of material resources, aesthetic values attributed to pristine and/or healthy systems, recreation, spiritual and bequest values”.

To sum up, the TEV of natural resources includes marketable and non-marketable values, their present and future values and goods provided can be either material or non-material. As observed by Plottu and Plottu (2007: 55), the concept of TEV follows from a definition and an interpretation of the environment value stemming from a neo-classical field of reflection. The TEV has been disaggregated into two main parts, use and non-use values.

Use value arises from actual, planned or possible use and consists of two branches, actual use value and option value. Actual use value reflects the utility that people derive from direct or indirect use of the resource. Direct use value concerns the active use of the resource, while indirect use value is associated with benefits that people experience indirectly or as a consequence of the primary function of the resource (Torras, 2000: 286). Taking forests and woodland as an example, direct use values are derived from the physical use such as production of timber and non-timber products and forest recreation (which is a non-consumptive use value). Indirect use values arise when society benefits from the ecological functions of forests/woods, such as watershed protection, water and air purification, erosion control, climate regulation and carbon storage (also known as regulating services). Option value is the value that people place on the potential benefits related to every use that can be realized in the future, even if they are not actual users and/or do not eventually use the resource in the future.

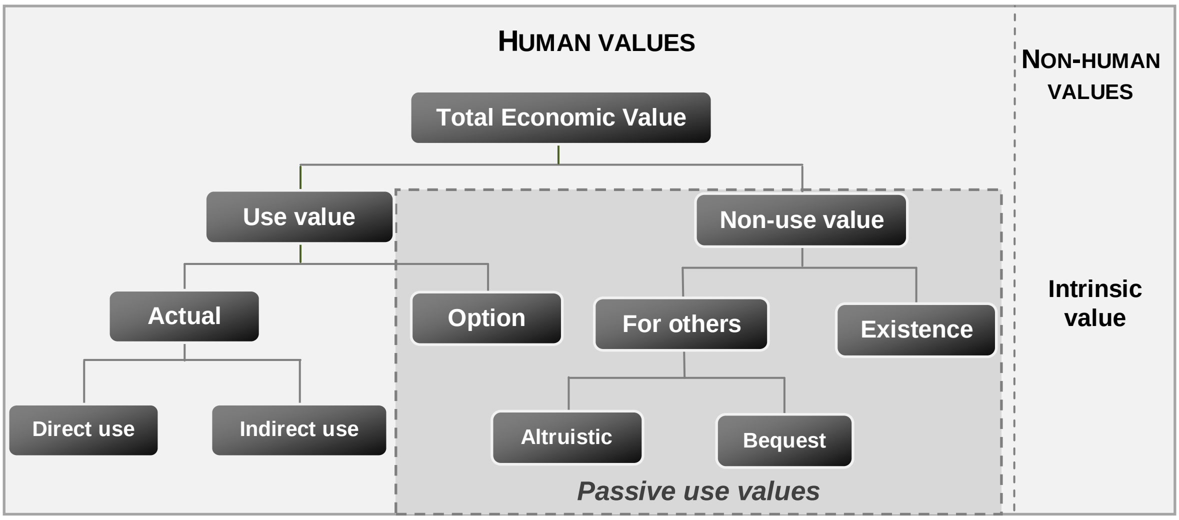
The concept of option value was first introduced by Weisbrod (1964) using visits to the Sequoia National Forest as an example. He claims that actual use values (measured by the entrance fees) should not be the only benefits accounted in a cost-benefit analysis. In his view, the amount people are willing to pay for the option to consume the commodity in the future – the option value – should also be considered in the decision. The main conditions governing people's willingness to pay for this option are: the infrequency and uncertainty of purchase and that production cannot or can only with difficulty be reinitiated once it has been closed down and inputs devoted to other uses.

Since the publication of Weisbrod's work, the concept of option value has been a major subject for discussion. It is agreed that it comes mainly from uncertainty in relation to the future availability of resources whose continued existence is in question (Bishop, 1982). Less consensual is the source of uncertainty. In line with Ojea and Loureiro (2010), we consider that option value stems from uncertainty on the supply-side. Therefore, as showed by Bishop (1982), if income and preferences are certain, the option value is positive.

Non-use values refers to the value of safeguarding some good even though there is no present or future planned use by the individual to whom the benefits accrue. This value follows from the sake of the good or from securing the opportunity for others to derive benefit, either from the use or non-use. The most common categorization separates the non-use component into existence and bequest values. But the Pearce *et al.* (2006: 86) taxonomy includes a third category, the altruistic value. We prefer this classification.

The altruistic value is the benefit people receive from knowing that the good is available to others in the current generation. The bequest/legacy value is the benefit accruing to people from the assurance that the resource will be preserved and available in the future. The concept of existence value was originally proposed by Krutilla (1967). It is the benefit derived from preserving the good in a context where the individual has no actual or planned use for himself/herself or for anyone else at the present or in the future. This is closely related to the concept of intrinsic value and sometimes not distinguished from it (e.g., Boyce *et al.*, 1992; Plottu and Plottu, 2007), but there is a fundamental difference. Existence value depends on individual preferences, while intrinsic values are not anthropocentric; they are independent of human needs and tastes. Figure 2.1 shows how TEV is structured into separate motivation-based values.

Figure 2.1: Total Economic Value



Adapted from Pearce *et al.* (2006: 87) and Bateman and Langford (1997: 573)

There is lack of consensus on how the option value should be categorized. Authors have variously classified option value as a non-use value (Walsh *et al.*, 1984; Kaoru, 1993), as a use value (Pearce *et al.*, 2006: 87) or as an autonomous component (Tietenberg, 2003: 37). The first categorization is justified on the grounds that it does not immediately lead to market transactions. The second, and more usual one, derives from the notion that benefits accrue from the possibility of use. In our view, it is primarily a passive use value, but also a use value. We borrow the concept of passive use value from Adamowicz *et al.* (1998: 64), who define it as the economic value derived from a change in environmental attribute(s) that is not reflected in any observable behaviour. It is analogous, too, to the concept of preservation benefits presented by Walsh *et al*. (1984: 14). Accordingly, besides the option value, passive use includes all the non-use values.

This aggregation is considered for three related reasons. First, option value is associated with a possible future use by the people evaluating the good; as a result it must be classified as a use value. Second, option value is, however, disconnected from any present direct or indirect use, so it is not an actual use value. Third, this motivation relates to users and non-users, as do non-use values, and its inclusion in empirical estimations is only possible when an SP technique is applied. Therefore, this is a hybrid concept which, while related to a possible future use, in practice has greater affinities with non-use values.

Plottu and Plottu (2007: 52) argue that option, use and non-use values are fundamentally different. They believe that actual use values (either direct or indirect) can have a monetary expression since it is only a question of resources allocation. Monetization of passive use values, however, is more complicated because they stem from different levels of choice. Option values belong to a higher level of decision because they will determine the availability of future options. Existence values belong to an even higher level, which will determine future sets of options, so they have an asset dimension. In fact, those authors explain that difficulties in the monetary evaluation of option and non-use values arise in part because the value ascribed by people reflects a collective concern, a preference as a member of a community and not a personal preference. In a similar line of argument, other authors (e.g. Faber *et al.*, 2002; Mill *et al.*, 2007) argue that when answering surveys about WTP for environmental goods, people think as citizens and not as individual consumers. Krutilla (1967: 785), long ago expressed the idea that a sense of public responsibility influences choices concerning the passive use values. His exact words were:

> It will not be possible to achieve a level of well-being in the future that would have been possible had the conversion of natural environments been retarded. That this should be of concern to members of the present generation may be attributable to the bequest motivation in private economic behaviour as much as to a sense of public responsibility.

Although we are in agreement with the idea that passive use values belong to a different level of decision, we are less convinced by the possibility that respondents can respond to the same evaluation question as citizens and as individual consumers.

The concept of passive use value coincides with the description of public good. Therefore, without public intervention, resources encompassing these values will not be supplied in the market at the optimum quantity. Nevertheless, in some circumstances, e.g. when resources are considered unique, irreplaceable or endangered, this fraction of TEV may be quite important and accurate decisions can only be taken if TEV is considered.

Once the components of TEV are clarified, we have to ask: How can those values be incorporated into the economic analysis? Is it possible to express them in monetary units? The ideas of environmental values' incommensurability (Martinez-Alier *et al.*, 1998) and that monetization of some non-market environmental values is socially unacceptable (Kumar and Kant, 2007: 517) differ from our view. On the contrary, we accept the use of a common unit of account, specifically a monetary unit, through which values can be traded off. Along with Tacconi (1995: 229), we consider that it "is indisputable that valuation is a necessary step in the decision making process regarding the use of resources". At the same time, we recognize that it is harder to evaluate resources used passively regardless of the metric chosen (which could be monetary or non-monetary). But, this does not mean that individuals are unable to establish a maximum for their WTP or a minimum for their WTA.

Monetization is possibly not the perfect way of measuring the worth of things, but probably it is for now the best measure available and "better than nothing". We are not arguing that *any number is better than no number,* but that monetary estimates obtained through rigorous studies must be taken into account in decision-making processes. However, these values should be integrated in a more pluralistic and multi-scalar theory of valuation, as advocated, e.g., by Norton and Noonan (2007).

This discussion would be incomplete without making reference to three important additional aspects. First, the appropriate context for economic valuation is conditioned by the scale of environmental changes. Valuation is most meaningful when changes in environmental quantity/quality are small or marginal and keep the asset above some critical level (Turner *et al.*, 2003). Second, the above discussion about components of value adopted an anthropocentric and utilitarian stance. It assumes that valuation is anchored on human preferences and focused on instrumental values. However, some voices claim that the worth of an entity for its own sake, independent of human preferences, i.e., its intrinsic value, must also be considered. Accordingly, some environmental resources should be preserved because they have value in their own right. We consider that intrinsic worth cannot be part of the economic value. At the same time, we do not defend a "monistic" approach. We recognize that several motives, e.g., pure self-interest, preservation, environmental stewardship, altruistic, cultural, spiritual and ethical, can act as determinants of human preferences. Third, from a theoretical and schematic perspective, the components of TEV are typically presented as additive parts. In practice, however, values with respect to each motivation are not strictly separable and additive. Some empirical attempts have been made to estimate the different components of value, but it has proved to be a very hard task (Cummings and Harrison, 1995).

2.3. ECONOMIC VALUES AND NON-MARKET VALUATION METHODS

Freeman (2003b: 146) argues that instead of use values and non-use values, the distinction would be more meaningful if it is made between "values revealed from market behaviour" and "values not revealed from market behaviour". Indeed, the categorization of methodological approaches used in non-market environmental valuation follows from this distinction. The techniques are usually categorized into two major groups, SP and RP methods.

Use values relate to some use, activity or traceable economic behavioural trail, so they can be estimated using RP techniques, such as the TCM and the hedonic prices method. Passive use values are independent of any actual use of resources by the person evaluating them, so they have no clear behavioural footprint. Because of this, these values can only be estimated using SP techniques, such as the CVM and the CM. Accordingly, the estimation of both use and passive use values is only possible using SP techniques.

Theoretically, SP techniques enable non-market valuation in a variety of contexts. There are four main possibilities. The evaluation can relate to:

a) Direct and indirect use values accruing to the current users
b) Actual use and passive use values accruing to the current users
c) Passive use values accruing to the current users
d) Passive use value accruing to non-users.

Several studies have been conducted with the aim of disaggregating into use and non-use values or into values held by users and non-users. Three main procedures have been adopted to address this. One procedure consists of using an SP method to obtain an estimate of TEV and then deducting the actual use value estimated using an RP method (Bockstael and Freeman, 2003: 564; Eom and Larson, 2006). Another possibility involves splitting the sample into users and non-users and assuming that users are the only people that can have direct and indirect use values, while passive use values accrue only to non-users (León, 1996; Bateman and Langford, 1997; Sattout *et al.*, 2007). In the third option, people are asked about their WTP/A and requested to allocate the value to the different motivations for their subjective total value (Walsh *et al.*, 1984).

Each approach raises some problems. In the first one, the estimation is based on two measures, both estimated with unknown errors. In the second one, the value assigned by non-users may be an option value, a non-use value or both and users can hold non-use values. In the third one, respondents may be unable to make the allocation in the manner required. Even if they can identify their distribution among categories, the extrapolation of non-use values derived from samples of users to non-users may be incorrect. The difficulties encountered in empirical studies are discussed, e.g., in Cummings and Harrison (1995).

As mentioned earlier, there is no consensus on the classification of the option value. The reason for this divergence becomes clearer in the empirical studies. When the procedure chosen is to separate into users and non-users the option value tends to be included in the value assigned by non-users and is implicitly considered a non-use value. But users can also have an option value, which is ignored. Some authors suggest that empirical studies must measure the total WTP/A and that passive use values should be assessed as a residual, once estimated total direct and indirect use value have been subtracted. We recognize that this would avoid dual counting, but it does not overcome the other difficulties.

In brief, non-users' WTP/A may be justified by the willingness to preserve the resource for one of the following reasons:

a) To have the chance of using it in the future, even if this never happens – option
b) So that others do have the chance of using it at the present – altruism
c) So that others do have the chance of using it in the future – bequest
d) To assure its continuation, because it is that associated with some kind of well-being – existence.

In addition to those motivations users may have a direct use value and/or an indirect use value. We agree that all of these motivations exist and are consciously or unconsciously considered by respondents when answering the surveys. Our scepticism is related to the possibility that these motivations are distinct, separable and on the possibility of estimating TEV by means of an additive process.

2.4. WILLINGNESS TO PAY AND WILLINGNESS TO ACCEPT

The analysis of the impact of an environmental-change on welfare requires the estimation of how consumers and/or producers are affected. In this section we examine the monetary measures generally used in the evaluation of consumers' welfare change, mainly following Freeman (2003b: Chapter 3), Perman *et al.* (2003: Chapter 12) and Hanemann (1991). The theory of individual choice is the background of this analysis. The monetary measures of interest may be expressed as the maximum WTP or the minimum WTA.

2.4.1. INDIVIDUAL PREFERENCES

Economic theory very often works on the assumption that people have well-defined preferences between alternative bundles of goods and services. It is assumed that consumers are able to compare and to state a hierarchy for the different bundles, which can contain both market and non-market goods. Non-market goods include environmental goods and services.

A usual assumption in this context is that individual preferences are correctly described by a well-behaved utility function $U(\cdot)$[4], in which welfare directly depends on the quantity consumed (x). Hence, $x = [x_1, x_2, ..., x_n]$ denotes a bundle of n goods and services or, in mathematical terms, a vector. In the same context, one essential property of preferences is the substitutability between the components of the bundle. Considering only the case of "goods" (goods and services that consumers enjoy), when the quantity of a good in the bundle decreases, the consumers' utility will remain unchanged if the quantity of other good(s) rises sufficiently. The substitutability property makes it possible to establish a trade-off relation between pairs of goods. This trade-off is a substitution rate which gives

---

[4] For purposes of mathematical modelling, this function is usually assumed to be continuous, convex and twice differentiable.

information about the relative value that people place on goods. Therefore, if environmental goods and services are part of the bundle, it is possible to estimate their relative values.

In determining the substitution rate, one of the dual problems is usually considered. In the first one, the consumer's objective is utility maximization for a given level of income (M) and prices (P). Considering that x includes n private goods and the corresponding price vector is $P(p_1, ..., p_n)$, the problem is expressed mathematically as:

$$\underset{x_i \geq 0}{Max} U(x) \quad subject\ \ to\ (s.t.) \quad M = \sum_{i=1}^{n} x_i p_i \ . \tag{2.1}$$

The solution to this problem is given by Marshallian/ordinary demand functions, expressing the optimal quantity as a function of income and prices:

$$x_i = x_i(M, P) \ . \tag{2.2}$$

The indirect utility function is obtained by substituting the expression for x_i into the utility function:

$$U = V(M, P) \ . \tag{2.3}$$

The indirect utility function represents the highest level of utility obtainable when facing prices P and income M.

The dual to utility maximization is the expenditure (e) minimization problem. This consists of finding the minimum expenditure required to achieve a pre-defined level of utility ($\bar{U}$), and can be expressed as:

$$\underset{x_i \geq 0}{Min}\ e = \sum_{i=1}^{n} x_i p_i \quad s.t. \quad U(x) = \bar{U} \ . \tag{2.4}$$

The solution to this problem is given by the Hicksian/compensated demand functions, in which optimal quantities depend on prices and utility:

$$h_i = h_i(P, \bar{U}) . \tag{2.5}$$

By substituting them into the expression for the total expenditure, we find the expenditure function, given by:

$$e = e(P, \bar{U}) . \tag{2.6}$$

The inverse of the indirect utility function with respect to M is the expenditure function, because of the duality of the problems.

Very often the utility function cannot be determined and the observation of individual behaviour may be a way of estimating the demand function. For this function to satisfy integrability conditions it must contain all the information about underlying preferences. Hence, exact measures of welfare change resulting from modifications in economic conditions can also be estimated.

2.4.2. WELFARE MEASURES FOR A PRICE CHANGE

Five measures to examine the effect of a price variation on welfare level have been analysed in the literature, based on the demand functions mentioned above. These measures are the change in the Marshallian consumer surplus (MCS), and the four Hicksian welfare measures, namely: the compensating variation (CV), the equivalent variation (EV), the compensating surplus (CS) and the equivalent surplus (ES).

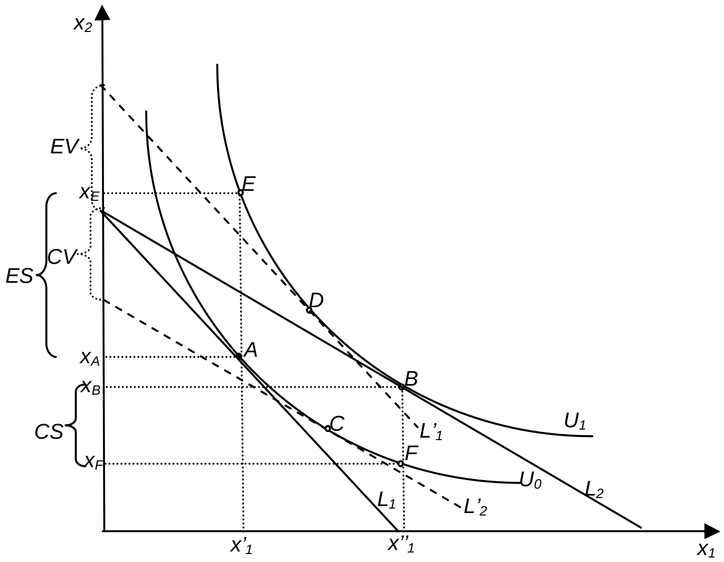
The MCS is, for each unit consumed, the difference between the maximum price a consumer is willing to pay and the price that they actually pay. Graphically, the MCS is the area under the Marshallian demand curve and above the horizontal price line. Consequently, a price change alters this area and modifies consumer well-being. The MCS change is the area under the Marshallian demand curve between the two price levels.

The MCS was the first measure developed to quantify the welfare change due to a price alteration. However, theoretically it is not the correct welfare measure because it holds income constant instead of well-being/utility. The MCS can only be correctly used as the welfare measure if some important conditions are satisfied. One condition is path independence in relation to the order in which prices and income change. The other condition is that the marginal utility of money is constant (consumer gets the same welfare from an extra unit of income regardless of income). These conditions ensure that the monetary measure is unique (Willig, 1976). The four Hicksian measures are a refinement of the original measure and are theoretically correct, as utility is held constant.

In the words of Seller *et al.* (1985: 157), “the compensating measures are defined as the amount of compensation paid or received, which would keep the consumer at the initial welfare level after the change had taken place” and “the equivalent measures are defined as the amount of compensation, paid or received, which would bring the consumer to his subsequent welfare level if the change did not take place”. In CV and EV it is implicit that the consumer can freely adjust the quantities of the commodity whose price changed. Conversely, it is implicit in CS and ES that the good whose price changed must be consumed in fixed quantities (Bockstael and McConnell, 1980: 56). Compensating and equivalent measures are illustrated in Figure 2.2. The MCS cannot be graphically determined using indifference curves so it is illustrated later in Figure 2.3.

In order to make the explanation simpler, we consider only two private goods, x_1 and x_2. Good x_2 can be regarded as a composite good whose units are chosen so that the price is equal to one. This good is taken to be the numeraire good. Thus, the maximum quantity of x_2 the consumer can buy is equal to income.

Figure 2.2: Hicksian welfare measures from a price change



Adapted from Freeman (2003b: 50)

Figure 2.2 shows two indifference curves and six bundles that would be chosen in specific circumstances. In the initial situation, prices are p_1 and p_2, income is M and the bundle A is the optimal choice (in the tangency of the original budget line (L_1) with the indifference curve). Next, we assume that an environmental improvement reduces the cost of producing x_1, so that its price drops from p_1 to p'_1. Consequently, the budget constraint moves to the right in the horizontal axis, to L_2. Consumer moves to B, with x''_1 units and a higher level of utility. The welfare benefit resulting from the price reduction may be calculated using four alternative measures and with x_2 as the numeraire.

The CV shows what compensating payment would be necessary so that the consumer returns to the level of utility achieved in the original situation, at the new price set. Since we are dealing with a price reduction, CV is the amount that must be taken from the consumer. The new budget line would be L'_2 (parallel to L_2) and the new optimal choice would be bundle C. This bundle has the same utility as bundle A, but the new implicit price. Hence, CV is the maximum amount that the individual would be willing to pay for the opportunity of consuming at the new price set. For a price increase, CV must be paid to the individual to compensate him/her for the reduction in the purchasing power.

Consequently, CV would define the amount that must be given to the individual to prevent a utility decrease and would be the minimum WTA.

The EV shows what change in income would be equivalent to the price change, in terms of the consumer's welfare. Thus, the EV is the income which, if given to the individual, without the price fall would bring them to the same level of utility that they would attain if the price fall had occurred. In Figure 2.2, if the price remains unchanged, income must increase EV, giving rise to L'_1 (parallel to L_1). The optimal bundle would be D, where the consumer would also be at point B, but facing the new price set. When considering a price increase, EV is the maximum sum of money that could be taken from the individual, yielding a loss of utility equivalent to that caused by the price increase. Therefore, this would be the maximum WTP.

We assume now that the quantity of the good whose price changes cannot be freely adjusted and the possibilities are x'_1 and x''_1. The CS shows what compensating payment would make the individual indifferent between the initial optimal solution (bundle A) and the consumption of good x_1 after the price change. Bundle F, with x''_1 units and the initial level of utility, would be chosen. Hence, the CS is given by $x_B - x_F$ units of x_2 and is the maximum amount the individual would be WTP for the price reduction.

The ES shows what change in income would be required to make the individual as well off as they would be with the new price, but keeping the consumption of x'_1 units and the original price. Bundle E, with x'_1 units and the final level of utility, would be chosen. The ES is given by $x_E - x_A$ units of x_2 and corresponds to the minimum WTA.

Variation measures may be defined analytically based on the indirect utility function or on the expenditure function. We next illustrate the analytical solutions, assuming a price fall. In terms of the indirect utility function, CV is the solution to:

$$V(P,M) = V(P',M-CV) = U_0 . \qquad (2.7)$$

In terms of the expenditure function, CV is the difference between the two expenditure levels:

$$CV = e(p_1, p_2, U_0) - e(p'_1, p_2, U_0) = M - e(p'_1, p_2, U_0)\,. \qquad (2.8)$$

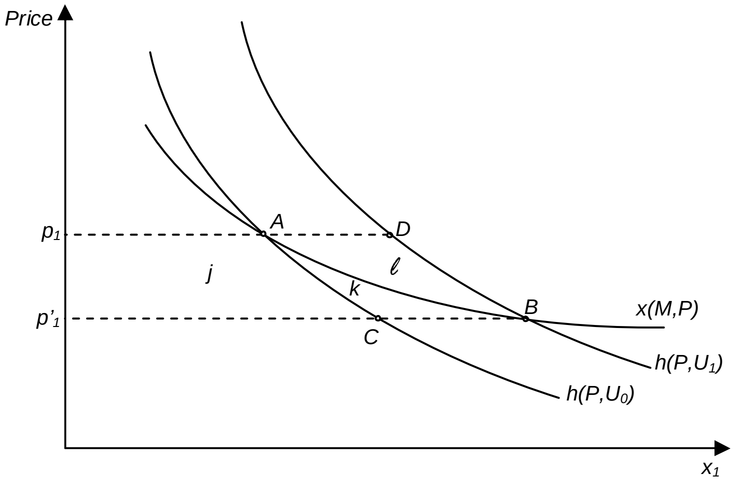
The analytical expressions for the EV are similar to those used for the CV. In terms of the indirect utility function, EV is the solution to:

$$V(P, M + EV) = V(P', M) = U_1\,. \qquad (2.9)$$

In terms of the expenditure function, EV is the difference between the two expenditure levels:

$$EV = e(p'_1, p_2, U_1) - e(p_1, p_2, U_0) = e(p'_1, p_2, U_1) - M\,. \qquad (2.10)$$

Figure 2.3: Demand curves and welfare measures



Adapted from Freeman (2003b: 54, 58)

Demand functions are an alternative way of representing the welfare measures. Three demand curves are represented in Figure 2.3: an ordinary demand curve, identified by $x(M,P)$; a compensated demand curve, corresponding to the level of utility achieved in the initial situation, identified by $h(P,U_0)$; and a compensated demand curve, corresponding to the level of utility achieved after the price decrease, identified by $h(P,U_1)$. The four points signalled in the quantity-price space, correspond to those presented in Figure 2.2. Points A and C are optimal to different implicit prices, but have the same level of utility, (U_0), so they are on the same compensated demand curve. The same happens with B and D, with a level of utility, U_1. The ordinary demand curve passes through A and B, which have different levels of utility but the same income.

The CV welfare measure is the area, j, between the two prices on the left of the compensated demand curve that passes through the initial bundle. The EV measure of welfare change is the area, $j+k+\ell$, between the two prices, on the left of the demand curve that passes through the final bundle. The MCS measure is the area, $j+k$, between the two prices, on the left of the ordinary demand function. The corresponding mathematical formulations for those areas are:

$$CV = e(p_1,p_2,U_0) - e(p'_1,p_2,U_0) = \int_{p'_1}^{p_1} \frac{\partial e(P,U_0)}{\partial p_1}\, dp_1 = \int_{p'_1}^{p_1} h(P,U_0)\, dp_1 = j\,. \qquad (2.11)$$

$$EV = e(p_1,p_2,U_1) - e(p'_1,p_2,U_1) = \int_{p'_1}^{p_1} \frac{\partial e(P,U_1)}{\partial p_1}\, dp_1 = \int_{p'_1}^{p_1} h(P,U_1)\, dp_1 = j+k+\ell\,. \qquad (2.12)$$

$$MCS = \int_{p'_1}^{p_1} x(M,P)\, dp_1 = j+k\,. \qquad (2.13)$$

Considering a positive income elasticity of demand, i.e., that x_1 is a normal good, for a fall in the price the following relation between the three measures holds: EV/WTA≥MCS≥CV/WTP. It can be seen in Figure 2.3. The relation will be symmetrical for a price increase, CV/WTA≥MCS≥EV/WTP (Bockstael and McConnell, 1980). The difference between the three measures increases with the income elasticity and when the income elasticity of demand is zero, the three measures are equal (Willig, 1976).

Although the above relations provide only limited information, they may be useful when knowledge of the absolute values is not essential. For example, if a policy measure that lowers the market price of a normal good has a cost higher than MCS, this cost will certainly be higher than CV. Therefore, it would not be strictly necessary to determine the theoretical correct measure to make a correct decision.

CV and EV have implicit different property rights associated with the price levels. In the CV context, the individual has the right to the original prices. But in the EV the presumption is that the individual has the right to the new price. The correspondence of the Hicksian measures for price change with the WTP/WTA are summarized in Table 2.1.

Table 2.1: Compensating and equivalent variation

Welfare measure	Price Increase	Price Decrease
CV	WTA to allow	WTP to obtain
EV	WTP to avoid	WTA to forgo

Adapted from Flores (2003: 38)

2.4.3. WELFARE MEASURES UNDER QUANTITY CONSTRAINTS

As a rule public programmes relate to changes in the quantity/quality of non-market environmental goods/services rather than changes in the prices of market goods. Furthermore, some of these environmental goods are indivisible (changes occur only in fixed quantities), non-exclusive and unpriced. Welfare measures therefore also have to be studied in this context. In order to do this, we have adapted the analysis of the preceding sub-section to the new context.

The utility function is now: $U(x,E)$, where E is a vector of environmental goods/services and is given to the individual. Considering that x, M and P have the same meaning has before and that R is the environmental good unit price, the budget constraint is given by:

$$P \cdot x + R \cdot \bar{E} = M . \tag{2.14}$$

The maximization problem now has an additional constraint:

$$\underset{x_i, E \geq 0}{Max}\, U(x,E) \quad s.t. \quad M = \sum_{i=1}^{n} x_i p_i + R \cdot E, \quad E = \bar{E} . \tag{2.15}$$

In the utility maximization context the optimal solution for private market goods is given by Marshallian demand functions of the type:

$$x_i = x_i(P, M - R \cdot \bar{E}, \bar{E}) . \tag{2.16}$$

These functions are conditional demand functions because E is quantity constrained. The optimal quantity of x_i depends on the quantity of E. The corresponding conditional indirect utility function will be:

$$V = V(P, M - R \cdot \bar{E}, \bar{E}) . \tag{2.17}$$

Inverting the conditional indirect utility function in order to $(M - R \cdot \bar{E})$ gives rise to the conditional expenditure function ($e*$). This function presents the minimum expense in private market goods required to achieve a specific level of utility, given P and E:

$$e* = M - R \cdot \bar{E} = e*(P, E, \bar{U}_0) \ . \tag{2.18}$$

Expenditure minimization subject to a utility constraint gives rise to the compensated demand functions:

$$h = h(P, R, \bar{E}, U_0) . \tag{2.19}$$

Compensated demand functions include a quantity constraint, in addition to the usual utility constraint. By substituting the optimal solution into the objective function the constrained expenditure function is obtained:

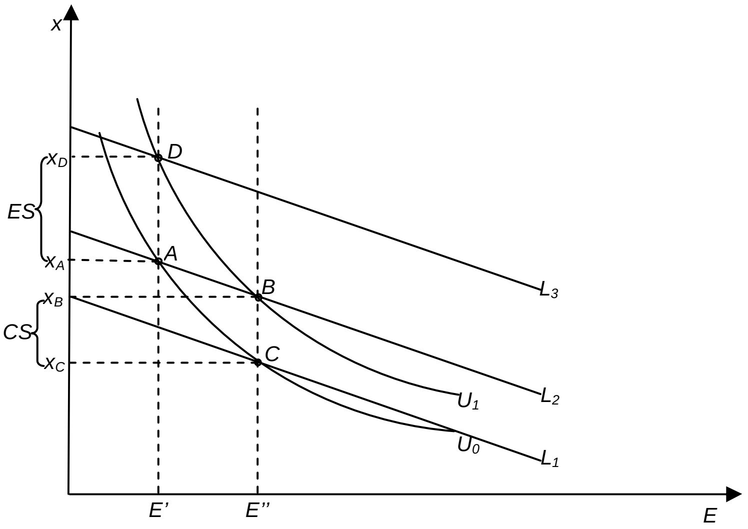
$$e = e(P, R, \bar{E}, U_0) . \tag{2.20}$$

This function gives the minimum expenditure necessary to achieve a level of utility U_0, given the $\bar{E}$ and the prices of the environmental and market goods. The conditional expenditure function refers only to expenditure related to private goods because E is pre-defined. Conversely, the constrained expenditure function includes the expenditure in both types of goods. Hence, expenditure functions are related by the following expression:

$$e = e(P, R, \bar{E}, U_0) = e* + R \cdot \bar{E} \ . \tag{2.21}$$

The presence of quantity constraints prevents the conventional optimizing conditions between marginal rates of substitution and price ratios from being satisfied by the adjustment in quantities. Consequently, relevant welfare measures are CS and ES and marginal rates of substitution are not revealed directly.

Figure 2.4: CS and ES for a quantity constrained good



Adapted from Freeman (2003b: 79)

For simplicity, in the graphical analysis it is considered that E is a vector of one element and x is the numeraire. In Figure 2.4 the two measures for an increase in quantity are presented, making use of an indifference curves map. Budget constraints are represented by negative sloped lines (L_1,L_2 and L_3), meaning that $R \geq 0$. Bundle A is defined as the initial solution, where E' and x_A units are consumed and a level of utility U_0 is achieved. The increase in E' to E'' raises the level of utility to U_1.

The CS corresponds to the money that must be taken from the consumer to bring them back to the original level of utility (U_0), but keeping the consumption of E''. The original level of utility and a consumption of E'' means that the consumer is at point C. Thus, CS is the distance between x_B and x_C, corresponding to the maximum WTP for the increase in the environmental good.

The ES corresponds to the money that must be given to the individual instead of augmenting E, so that they can achieve the new level of utility (U_1). A level of utility U_1 and a consumption of E' puts the consumer at point D. Therefore, ES is the distance between x_D and x_A and corresponds to the minimum WTA to forgo the increase in the environmental good.

Just as for CV and EV, there are three possible ways of determining CS and ES analytically. One possibility is based on the conditional indirect utility function; another is based on the expenditure function, and the third is based on the integral of the compensated inverse demand function.

Considering the notation used in Figure 2.4 and the conditional indirect utility function, the CS is the solution to:

$$V(P,M-R\cdot E',E')=V(P,M-R\cdot E''-CS,E'')=U_0\,. \tag{2.22}$$

In terms of the expenditure function, the CS is the difference between the two levels:

$$CS=e(P,R,E',U_0)-e(P,R,E'',U_0)=M-e(P,R,E'',U_0)\,. \tag{2.23}$$

Similarly, the ES is the solution to the equality between two conditional indirect utility functions, with different levels of E:

$$V(P,M-R\cdot E'+ES,E')=V(P,M-R\cdot E'',E'')=U_1\,. \tag{2.24}$$

ES is also given by:

$$ES=e(P,R,E',U_1)-e(P,R,E',U_0)=e(P,R,E',U_1)-M\,. \tag{2.25}$$

In the third possibility, CS and ES are given by the integrals:

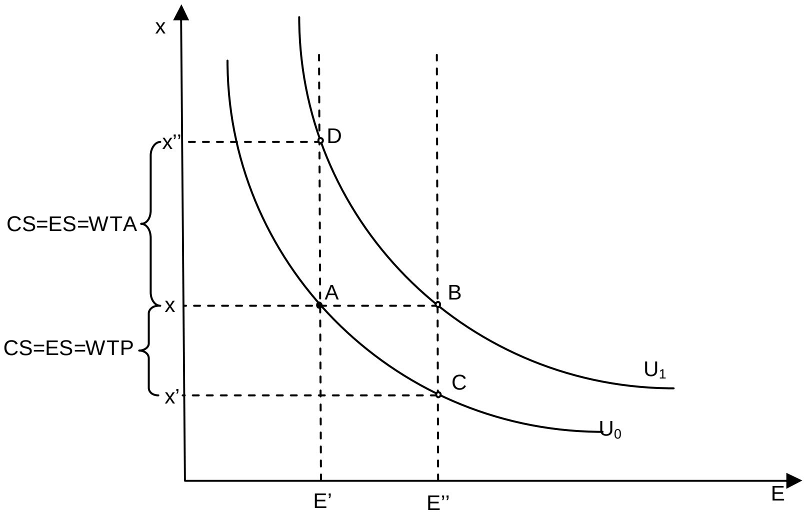
$$CS=-\int_{E'}^{E''}\frac{\partial e(P,R,E,U_0)}{\partial E}dE=-(e(P,R,E'',U_0)-e(P,R,E',U_0) \tag{2.26}$$

$$ES=-\int_{E'}^{E''}\frac{\partial e(P,R,E,U_1)}{\partial E}dE=-(e(P,R,E'',U_1)-e(P,R,E',U_1)\,. \tag{2.27}$$

We earlier compared CV, EV and MCS measures for a change in the price of a good whose quantities can be freely adjusted. However, as demonstrated above, for a change in the quantity of a quantity-constrained good, it is harder to use MCS as an approximation for the welfare monetary measure because Marshallian demand curves cannot be directly estimated. Nevertheless, the relation between CS and ES measures does not change. For an increase in the quantity of the restricted normal good: ES/WTA>CS/WTP. The relation will be the opposite for a reduction: CS/WTA>ES/WTP.

The assumption that the environmental good price is zero ($R = 0$) is a simplification of the above framework. This assumption gives horizontal budget lines, so they were omitted in Figure 2.5. This simplification is frequently used as a reference point for analysing the concepts of CS and ES, from the standpoint of property rights assignment. Whether one should consider CS or ES depends on the situation.

Figure 2.5: CS and ES for a quantity-constrained good with null price



Adapted from Freeman (2003b: 80)

We begin considering that the consumer is at point A, which is the initial endowment and the consumer has the right to it. If the policy action increases the environmental good from E' to E'' and everything else remains unchanged, the consumer will be better off, consuming at B. So, $x - x'$ is the maximum amount that the consumer would be willing to pay for this increase. This is a CS measure since the reference point is the initial situation.

In the second possibility, A is the initial endowment but the consumer has the right to B. The minimum amount the consumer would be willing to accept to forgo the improvement from E' to E'' would be $x''-x$. If this amount is paid, the consumer will be at point D with a level of utility, U_1, so it is the ES measure.

In the third possibility, B is the initial endowment, the consumer has the right to it, but policy action decreases the environmental good. The minimum amount the consumer would be willing to accept to allow the deterioration from E'' to E' would be $x''-x$. If this amount is received, the consumer will maintain the initial utility, but will move to point D. Now, the CS measures the WTA.

Finally, B is the initial endowment, policy action lowers the environmental good to E' and the consumer has the right to the new quantity. The maximum amount that the consumer would be willing to pay to avoid the deterioration would be $x-x'$, which is the ES.

Table 2.2 summarizes the possibilities that may be considered for a change in the quantity of a quantity-constrained good. The differences in relation to Table 2.1 are the designation of the welfare measures and the variable changing, which is now the quantity instead of the price.

Table 2.2: Compensating and equivalent surplus

Welfare measure	Quantity Increase	Quantity Decrease
CS	WTA to allow	WTP to obtain
ES	WTP to avoid	WTA to forgo

Adapted from Flores (2003: 38)

When an SP method is used, the way the valuation question is formulated determines which valuation measure is to be used. Depending on the situation, CS and ES are the correct welfare measures for a change in an indivisible or lumpy environmental good. Each may represent the maximum WTP or the minimum WTA. Let us give an example, using the context of forest policies. When the question is how much would be the respondent willing to pay to restore a woodland area, the WTP is the CS. When the question is how much he/she would be willing to accept as compensation for the forest's degradation, the WTA also corresponds to the CS. But if the question is how much they would be willing to pay to prevent the forest's degradation, the WTP is the ES. Finally, if respondents are asked about the minimum amount they would be willing to accept to renounce a forest improvement, the measure is the ES.

In order to analyse the WTP/WTA disparity, Hanemann (1991) relaxes the hypothesis of fixed quantities for environmental goods. In doing so, he provides an alternative analytical framework for estimating the welfare measures. We present here an overview and show that the integrals computed in (2.26) and (2.27) can be written in another way. To do that, we must go back and redefine the minimization expenditure problem. Thus, if we assume that E is no longer fixed and it is part of a consumer's problems, the minimization problem can be written as:

$$\underset{x_i,E\geq 0}{Min}\ e=\sum_{i=1}^{n} x_i p_i + R\cdot E \quad s.t. \quad U(x,E)=\bar{U}\,. \tag{2.28}$$

The solution for this problem is given by the compensated demand functions for private goods and for the environmental good which would respectively be written as:

$$x_i = h_i(P,R,U), \tag{2.29}$$

$$E = h_E(P,R,U). \tag{2.30}$$

The expenditure function is now written as:

$$e(P,E,U)=\sum_{i=1}^{n}\left(p_i\cdot h_i(P,R,U)\right)+R\cdot h_E(P,R,U)\,. \tag{2.31}$$

Equation (2.30) may be solved in order to R to obtain the inverse compensated demand function of the environmental good:

$$R=R^f(P,E,U). \tag{2.32}$$

$R^f(\cdot)$ is the hypothetical price of the environmental good that would induce the individual to buy E units in order to obtain a level of utility of U, given the private goods' prices. For each bundle, an R that turns that bundle into an optimal one must be estimated. For example, $R_0 = R^f(P, E_0, U_0)$ and $R_1 = R^f(P, E_1, U_1)$ denote the prices that would have supported the choice of E_0 and E_1, respectively.

Equation (2.21) can be re-written as:

$$e(P,E,U) + E \cdot R^f(P,E,U) = e\big(P, R^f(P,E,U), U\big). \tag{2.21'}$$

Equivalently,

$$e(P,E,U) = e\big(P, R^f(P,E,U), U\big) - E \cdot R^f(P,E,U). \tag{2.21''}$$

Equation (2.21') shows that the expenditure in private goods plus the expenditure in the environmental good equals the total expenditure, defined when the environmental good has a hypothetical price defined.

From (2.21") it is possible to relate changes in expenditure to changes in the environmental good quantity/quality:

$$e_E = \frac{\partial e(P,E,U)}{\partial E} = -R^f(P,E,U). \tag{2.33}$$

Based on the previous equation, CS and ES measures may alternatively be given by the integrals:

$$CS = \int_{E'}^{E''} R^f(P,E,U_0)\,dE\,, \tag{2.34}$$

$$ES = \int_{E'}^{E''} R^f(P,E,U_1)\,dE\,. \tag{2.35}$$[5]

---

[5] Since the expenditure associated with E' is higher than the expenditure associated with E'', $e(E'') - e(E') < 0 \Rightarrow -(e(E'') - e(E')) > 0$

The hypothetical price may also be found by estimating the ordinary inverse demand function. Given the level of E , we can ask what market price would induce the individual to buy that quantity of the environmental good if it was available in the market, but still buying the same quantity of x_i, facing P and M. In this case the individual must have an increase in income. For (P,E,M), R must satisfy:

$$E = x_E(P,R,M + R \cdot E). \qquad (2.36)$$

The solution will be given by $R = R^f(P,E,M)$. This inverse function is related to the inverse compensated demand functions through the identities:

$$R^f(P,E,M) \equiv R^f[P,E,V(p,E,M)], \qquad (2.37)$$

$$R^f(P,E,U) \equiv R^f[P,E,e(p,E,U)]. \qquad (2.37')$$

From equations (2.33) and (2.37'),

$$e_E = \frac{\partial e(P,E,U)}{\partial E} = -R^f[P,E,e(P,e,U)]. \qquad (2.38)$$

Finally, differentiating equation (2.37') and using (2.38), an estimation of the MCS is obtained by:

$$MCS = \int_{E'}^{E''} R^f(P,E,M)dE. \qquad (2.39)$$

2.4.4. COMPARISON OF WELFARE MEASURES

In this sub-section we aim to address two questions: i) What does the literature tell about the use of MCS instead of Hicksian variation and surplus measures? ii) What factors explain the disparity between WTA and WTP?. The first issue emerged because, for priced goods, Hicksian demand functions are not observable from market data whereas ordinary demand functions are, and so the MCS measure is easier to obtain. None of the measures is readily available for unpriced goods, but the way information is gathered makes it easier to estimate one of them. The second issue is important to understanding the differences found in empirical applications. When a cost-benefit analysis is conducted for an environmental programme/project researchers need to be aware of theory predictions.

Hicksian measures are the correct ones for estimating welfare changes. However, it has been advocated that in particular circumstances the MCS may be a very good approximation and may be used “without apology” (Willig, 1976). For the specific case of a price change, Willig (1976) demonstrated that the error resulting from the use of MCS is proportional to the income elasticity of demand and to the relation between consumer surplus and income. In the same work it is argued that in most empirical studies errors resulting from the approximation are small and overshadowed by the errors involved in estimating the demand curve. In these circumstances the areas ℓ and k in Figure 2.3 would be insignificant.

Randall and Stoll (1980) extended Willig’s work to the analysis of changes in the quantities of indivisible or lumpy goods. They show that in certain circumstances the CS and ES may also be deducted from the MCS. The approximation is possible if the proportion of total budget spent on the good whose quantity/quality changed is small and the individual’s income and the price flexibility of income[6] are known. However, errors are likely to be significant and the information needed to approximate MCS is hard to obtain if at least one of the following conditions holds: the change is large; the good is highly valued, or price flexibility of income is high and rises with income.

While the consumer’s WTP is always finite because income is limited, WTA may be infinite. Figure 2.3 shows an example in which the WTP (given by the CV) is smaller than the MCS which in turn is smaller than the WTA (given by the EV). Nevertheless, for some time it was generally accepted that differences between WTA and WTP would be very

---

[6] Price flexibility of income is the income elasticity of demand if post-change quantity adjustments are not possible. It also addresses the question of how changes in income affect the amount the consumer would spend to enjoy a given unit of the good. For normal goods the price flexibility of income is greater than zero.

small and without practical relevance[7]. Therefore, in cost-benefit analysis it would scarcely matter which measure is chosen.

However, several empirical works have been proving the opposite, some of them showing quiet large the differences between WTA and WTP. A number of economic and psychological reasons have been proposed to justify the disparity. Among the main explanations are: income effect, substitution effect, loss aversion, loss aversion combined with endowment effects, uncertainty combined with decision reverting costs and survey biases (Hanemann, 1991; Boyce *et al.*, 1992; Shogren *et al.*, 1994; Brown and Gregory, 1999; Zhao and Kling, 2001; Horowitz and McConnell, 2002). The WTA-WTP disparity is commonly identified as one of the behavioural anomalies which calls into question the principle of rationality assumed in neoclassical economic theory (Hanley and Shogren, 2005; List, 2005; Sugden, 2005; Shogren and Taylor, 2008; Venkatachalam, 2008).

The primary explanation for the WTA-WTP gap is the income effect. On the one hand, consumers' income constrains the payment but not the compensation required (Randall and Stoll, 1980). On the other hand, real income changes in a different way, depending on whether the individual has to pay or receive compensation (Pearce *et al.*, 2006: 161). Hanemann (1991) demonstrated that for changes in indivisible goods, large disparities are also explained by substitution effects. Even without income effects, the smaller the possibility of substituting environmental public goods for private market goods the greater the difference[8]. This means that when a public good has a small number of substitutes, high levels of compensation are required to tolerate a reduction in the quantity (see, e.g., Shogren *et al.*, 1994).

The notions of reference dependence and loss aversion of Tversky and Kahneman (1991) also help to explain the difference. Accordingly, reference levels are important in the determination of preferences. The initial endowment does matter in the valuation and

---

[7] Randall and Stoll (1980) show that this is true only under very restrictive conditions CS=MCS=ES.

[8] Hence, in the opposite case, if at least one private good is a perfect substitute for the public good, then CS=ES.

determination of an exchange rate between goods. Loss aversion translates into different value functions for losses and gains. Consequently, the loss in utility from giving up on a good is greater that the utility gained when receiving it. This is an endowment effect because the value is different when the asset is part of an individual's endowment and when is not. Another consequence of loss aversion is the induction of a bias that favours the maintenance of the *status quo*. Therefore, when the reference endowment point is the same for losses and gains, and both vary in the same amount, asymmetric valuations of losses and gains will be frequent. This gives rise to higher WTA values.

In Zhao and Kling's (2001) view, uncertainty and irreversibility combined with the opportunity of learning also help to explain the difference. When the consumer is uncertain about the value of a good, he may want to delay the decision of buying or selling in order to obtain additional information. If the decision has to be taken immediately, without extra information, some compensation is demanded. For this reason, individuals tend to require high values for WTA. The existence of *commitment costs* is enough to generate the disparity of values. After analysing several experimental studies, Zhao and Kling (2001) concluded that the conditions and products used influence the magnitude of the WTA-WTP gap. The type of product, the degree of uncertainty about the value of the good, the possibility of learning about the characteristics of the good through an experiment, the urgency to have the good and the reference point are among those conditions. Ordinary goods, lower levels of uncertainty, greater possibility of learning, less urgency to have the good and better references to evaluate the good, are all factors contributing to small WTA-WTP disparities. However, none of the experimental studies in which these conditions were analysed used environmental goods.

The difference between the two measures was analysed by Horowitz and McConnell (2002) based on 45 studies examining a diverse range of goods. The ratio of WTA/WTP was regressed on several explanatory variables. The incorrect construction or administration of the survey was included but was not significant as incentive compatible formats do not yield lower ratios and the disparity does not seem to result from experimental artefacts. The authors also concluded that real experiments were no different from the hypothetical ones. Furthermore, the analysis showed that ordinary goods have lower ratios than non-ordinary ones. This conclusion is in line with theoretical predictions. Perfectly divisible goods, transacted in large markets with null transaction costs confirm the equality ES=MCS=CS. The higher ratios found for public and non-market goods bear out the idea that the measure chosen in environmental valuations must be selected with special care.

Since differences between WTP and WTA are acknowledged, and to some extent explained, it is clear that the choice matters in empirical studies. The measure chosen will influence the results. Furthermore, whenever the discrepancy between the two measures is large, the MCS will not be a good approximation to at least one of these measures.

2.5. CONCLUDING REMARKS

This chapter discussed fundamental concepts and issues which, either explicitly or implicitly, are always present in non-market environmental valuation. A general review of the underlying theoretical and analytical frameworks has also been presented. Our interpretation of concepts and our opinion on methodological issues have been made clear on two particular topics.

First, we discussed which components make up the TEV and examined their meanings. There is without doubt less agreement on the notion of passive use value. In line with Adamowicz *et al.* (1998: 64), we define a passive use value as an economic value derived from a change in environmental attribute(s) that is not reflected in any observable behaviour. Therefore, it includes the option value and all non-use values. Moreover, we agree that different levels of decision lie beneath different components of value, but we distance ourselves from the possibility that each person is able to evaluate environmental goods as a citizen and as an individual.

Second, we agree that environmental goods are often multidimensional, complex, likely to involve a broad range of aspects (e.g., ecological, scientific, recreational, aesthetic, life support and spiritual) and that environmental goods are sometimes unfamiliar to people asked to evaluate them. Yet we do not think that converting environmental values into monetary units is impossible or an affront. Quite the contrary, we consider that when properly done and understood, valuation helps society to make better-informed choices about the trade-offs that are inherent to the scarcity restrictions of our daily decisions. This does not mean that we consider economic valuation perfect or that it should be used as the sole source of data for policy makers.

A similar point of view seems to be shared by the researchers involved in the TEEB (The Economics of Ecosystems and Biodiversity) initiative hosted by the United Nations Environment Programme[9]. This project brings together contributions from all around the world in order to "evaluate the costs of biodiversity loss and ecosystem degradation" and "to sharpen awareness of the value of biodiversity and ecosystem services". Hence, environment commensurability and monetization are accepted, but the limits of this kind of exercise are also acknowledged.

---

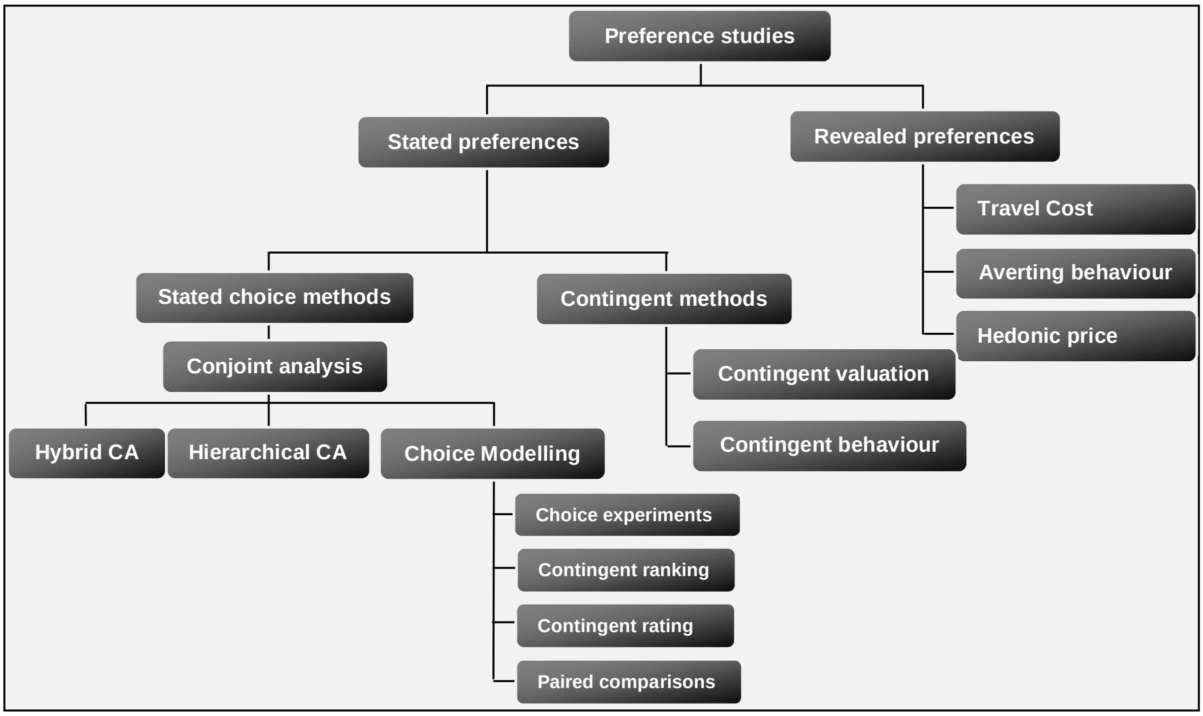
[9] See http://www.teebweb.org/ for further details and materials on this UNEP initiative.

CHAPTER 3 – NON-MARKET ENVIRONMENTAL VALUATION: AN APPRAISAL FOCUSED ON PORTUGAL[10]

3.1. INTRODUCTION

Methodological approaches used in non-market environmental valuation are usually categorized into two major groups, SP and RP methods. RP methods recover people's preferences from actual behaviour and this information is used to work out monetary welfare measures. Values are inferred from the observation of behaviour in related markets, thus they are also called indirect methods. SP methods use data derived from what people state when directly asked to declare their choices, evaluations or (dis)agreement. Hence, they are also called direct methods. Figure 3.1 summarizes diagrammatically the main RP and SP techniques used in environmental valuation.

Figure 3.1: Non-market valuation techniques



Adapted from Alriksson and Öberg (2008: 246)

---

[10] An initial version of this Chapter was presented at the Coimbra Conference – The Revival of Political Economy, in Coimbra, 21th October 2010. We are grateful to the participants for their helpful comments and suggestions. This Chapter was submitted to a scientific journal and is currently under review.

The group of the RP methods includes three main techniques conceptually different: the averting behaviour method[11], the hedonic price method[12] and the TCM. The group of SP methods has two main ramifications, one made up of the contingent methods and the other of stated choice models derived from the conjoint analysis (CA). Among the RP methods, the TCM is the most widely used, while the CVM is the most well-known among the SP techniques. Application of the CM has been growing rapidly, such as the combination of TCM with contingent methods (CB and CVM). In Section 3.2 we review the main theoretical aspects of these four methods. Section 3.3 focuses on their use in the context of non-market valuation of Portuguese environmental goods. Section 3.4 concludes.

3.2. VALUATION METHODS

3.2.1. CONTINGENT VALUATION

The CVM was originally suggested by Ciriacy-Wantrup in 1947, but its first empirical application was made by Davis in the 1960’s to estimate the economic value of big game hunting in Maine backwoods (Mitchell and Carson, 1989: 9). The CVM is a survey-based methodology which involves the construction of a hypothetical market where a proposed environmental program would be transacted. After the description of the hypothetical scenario, people are asked directly or indirectly how much they would be WTP/A to guarantee/avoid the proposed action[13]. The method is based on the assumption that

---

[11] The averting behaviour method is also known as defensive behaviour method. This method is based on the recognition that people is willing to make expenditures which protect themselves from risks, namely from the environmental ones (Whitehead *et al.*, 2008: 874). It is assumed that rational persons will take defensive behaviours as long as the value of the damaged avoided exceeds the costs of the protective action. The most common application of the averting behaviour method involves health valuation (Dickie, 2003: 396).

[12] The hedonic price method estimates the value of a non-market good by observing behaviour in the market for a related private good (Pearce *et al.*, 2006: 93). The non-market good is implicitly traded in that market as it is a characteristic of the transacted good. The most common application of the hedonic theory to environmental valuation has been in housing markets. However, the method has been also used to analyze other markets, such as the labour market (Taylor, 2003: 333).

[13] Mitchell and Carson (1989), Hanley (1989), Arrow *et al.* (1993), Portney (1994) and Carson *et al.* (2001) are among the most important references on this method.

individuals are able to identify the amount they would be WTP/A and that they will report the true value if the questionnaire is correctly designed.

There is general agreement that the CVM is the most versatile and powerful methodology for estimating non-market environmental values (Pearce *et al.*, 2006: 126). Several years of research and empirical application created the necessary space so that many methodological issues could be raised and discussed. Here we focus on three of them: the types of bias, the elicitation format and the treatment of uncertainty. They are presented as separated topics, but are clearly interrelated.

3.2.1.1. Types of bias

The biases likely to affect CVM responses and results have different natures. We structure them into two main groups. One is related to respondents' attitudes and the other to the questionnaire design and administration. While biases in the latter can be mitigated or even avoided, those in the former are more difficult to control and some of them can only be tested. Table 3.1 identifies the main biases and their sources.

Table 3.1: Biases likely to affect CVM

Bias	Source
Respondents' attitudes	
Hypothetical	Hypothetical nature of the CVM market; "*ask a hypothetical question and you get a hypothetical answer"*.
Strategic	Incentive to misrepresent true preferences, over or understating the true value in order to achieve a more desired outcome.
Compliance ("*yea-saying"*)	Respondents look for social approbation and respond in a way they consider socially desirable.
Embedding	WTP/A is irresponsive to the program extent and/or is responsive to the valuation order.
Scope effect	WTP/A is not sensitive to the program extent.
Sub-additivity effect	WTP/A is affected by the way the bundle is valuated (at once or each good separately).
Sequencing/order effect	WTP/A depends on the order of the items in the valuation exercise. Those presented earlier tend to be more highly valuated.
Payment vehicle	Respondents are responsive to the way by which the program is to be paid.
Questionnaire design and administration	
Information	The framing of the question influences the answers.
Operational	Respondent unfamiliarity with the good to be evaluated.
Interviewer	Different interviewers obtain answers which are statistically different.
Anchoring/starting point	Some value suggested in the questionnaire influences the answers.
Sample/self-selection	Respondents more interested in a subject are more likely to answer.

The identification of these various sources of bias is the result of extensive research. This does not invalidate the CVM, rather it serves to highlight the importance of thinking about these issues carefully when designing the survey and choosing the elicitation format.

3.2.1.2. Elicitation format

The most widely used CVM elicitation formats are reviewed in Table 3.2, but the list does not exhaust all the possibilities. Hybrid designs can also be found in the literature (see, e.g., Ready *et al.*, 2004). In any of them, the first step in the evaluation exercise is the accurate description of the proposed program and of the hypothetical market where it would be transacted. Depending on the design of the contingent market, the WTP/A will provide a measure of the Marshallian surplus or of the Hicksian welfare measures.

Table 3.2: CVM elicitation formats

Format	WTP/A question	Earlier applications
Iterative bidding (IB) [a)]	Various offers, each depending on the answer to the preceding one.	Davis (1964 *apud* Mitchell and Carson, 1989: 9)
Open ended (OE)	No value is suggested, respondents must state their maximum WTP or minimum WTA.	Various studies published in 1973 [b)]
Dichotomous choice (DC)	One offer. Possible answers: accept or not accept.	Bishop and Heberlein (1979)
Payment cards (PC)	Respondents choose a value from a set of possibilities.	Mitchell and Carson (1981)
Double bounded dichotomous choice (DBDC)	Two offers, the value of the second offer depends on the answer (accept or not accept) to the first one.	Carson *et al.* (1986)
Multiple bounded dichotomous choice (MBDC)	A list of bids is offered and individuals asked whether they would be WTP/A each of them.	Welsh and Poe (1998)
One-and-one-half-bound dichotomous choice (OHBDC)	Minimum and maximum values set upfront; the answer to the first offer dictates whether a second offer is made.	Cooper *et al.* (2002)

[a)] Also called bidding game and sequential bid.
[b)] Mitchell and Carson (1989: Appendix A) list three studies published in 1973 using the OE format.

Iterative bidding (IB) was the first format ever used in the empirical application of the CVM, but soon some shortcomings were identified motivating the development of alternatives. Successive new formats have been suggested in order to overcome the weaknesses of preceding ones, but experience has shown that each format is subject to its own problems.

There is some evidence that different formats are associated with specific advantages and problems/flaws. For example, IB is known to be more prone to starting point bias and OE lacks similarities with the "*take it or leave it*" marketplace situation, making the evaluation task more complex. On the other hand, DC is incentive compatible, but compared to OE and PC tends to produce higher WTP and these differences are statistically significant (Welsh and Poe, 1998: 170). This difference has been attributed to the compliance bias, which mainly affects DC formats. Therefore, there is no consensus on the most appropriate alternative and this tends to be dictated by the circumstances of each particular study.

3.2.1.3. Uncertainty in the CVM

The standard approach assumes that respondents know their utility function with certainty and therefore know their WTP/A for an environmental measure (Hanemann, 1984). However, there is evidence that actual payments can differ from the stated ones (Champ *et al.*, 1997). The difference has been ascribed to the compliance bias, typical of the DC formats, and to respondents' uncertainty when answering the hypothetical valuation question. The "*yes*" response seems to be the most affected. Several sources of uncertainty have been identified, the main ones are: (i) Unfamiliarity with the alternatives in the valuation exercise; (ii) Incomplete knowledge of the true value of environmental the resource/servive; (iii) True uncertainty about preferences; (iv) Misunderstanding of the contingencies in question; and (v) Inability to evaluate the proposed program in monetary units (Li and Mattsson, 1995; Shaikh *et al.*, 2007; Hanley *et al.*, 2009).

In order to explicitly consider the degree of respondents' certainty, additional questions have been introduced in CVM surveys, especially in those using DC formats. Certainty levels have been measured by means of a percentage scale (Li and Mattsson, 1995), a numeric scale, e.g., a 10-point or 5-point scale (Loomis and Ekstrand, 1998; Welsh and Poe, 1998; Samnaliev *et al.*, 2006; Shaikh *et al.*, 2007), or a descriptive word scale (Ready *et al.*, 1995). The "*do not know*" or "*not sure*" statements have also been interpreted as uncertainty indicators (Wang, 1997). The certainty levels assigned are then used to recode answers or to adjust monetary values. According to Loomis and Ekstrand (1998: 32), two conditions need to be satisfied if the predictive accuracy is to be improved when uncertainty information is considered. Respondents must be able to assess the level of certainty of their valuation with some degree of accuracy and all respondents must interpret the certainty scale equivalently.

Naturally, different calibration procedures embrace conceptual differences and lead to divergent results (Samnaliev *et al.*, 2006). There are also contradictory results concerning the effect of calibration on the goodness of fit, on the precision/efficiency and on the magnitude of the welfare estimates (Loomis and Ekstrand, 1998; Shaikh *et al.*, 2007). Yet, the certainty assessment is appealing because the supplementary question(s) brings additional data at a low cost and standard data remains available (Champ *et al.*, 1997: 161).

Biases and signs of uncertainty affecting the surveys' answers have been diagnosed mainly in the CVM framework. However, most of the biases are not exclusive to this method. Those related to respondents' attitudes are likely to affect any SP study, whereas interviewer and sample/self selection biases can affect RP studies as well. Furthermore, uncertainty is likely to be involved in responses to any hypothetical scenario, namely in the CM framework.

3.2.2. CHOICE MODELLING

CM is a family of survey-based methodologies which has its roots in conjoint analysis[14] (Adamowicz *et al.*, 1999: 461). It models preferences for goods described as sets of attributes, which can be quantitative or qualitative in nature and have different levels. Each combination of attributes is an alternative in the consumer's choice set. The inclusion of price as one of the attributes and the *status quo* situation as one of the alternatives enables the indirect estimation of the WTP/A and the relative values of different attributes. The CM method is consistent with Lancaster's characteristics theory of value which assumes that the utility consumers receive from the consumption of a good can be decomposed into the utilities from the component characteristics (Hanley *et al.*, 2001b: 436).

In a CM valuation exercise respondents are presented with various alternative descriptions of a good, distinguished by variations in the levels of the underlying attributes, and must choose one of the alternatives, rank or rate them. These different ways of measuring preferences correspond to the different variants of the CM method.

---

[14] For a more detailed review of conjoint analysis, see, e.g., Hensher *et al.* (1999).

The four main variants are listed in Table 3.3. These techniques differ in the degree of complexity, in the information provided and in the ability to produce WTP/A estimates consistent with welfare measures.

Table 3.3: Variants of CM

Format	Respondent have to...	Some earlier applications
Paired Comparisons	Score pairs of alternatives	Sinden (1974)
Contingent Ranking (CRk)	Rank the alternatives	Rae (1981) Desvousges *et al.* (1983)
Contingent Rating (CRt)	Score the alternatives	Mackenzie (1990)
Choice Experiment [a] (CE)	Choose one among two or more alternatives	Adamowicz *et al.* (1994)

[a] The CE version may also be seen as a generalization of the CVM (Adamowicz *et al.*, 1998).

The inclusion of several attributes and levels in the CM exercise makes the number of possible choice sets enormous. In general, it is not possible to analyse all of the alternatives and an orthogonal fractional factorial design which treats all attributes as independent is typically used (Mackenzie, 1993: 597). This design determines the minimum number of combinations necessary to compute estimations with precision and reduces the complexity of the choice task for the respondents (Álvarez-Farizo *et al.*, 2001: 687). Increasing the number of alternatives increases the probability that respondents can find the option that better matches their preferences, leading to a more precise selection (Caussade *et al.*, 2005: 624). On the other hand, the reduction in the number of choices will diminish the complexity and the cognitive burden, improving consistency (Rolfe and Bennett, 2009: 1147). Researchers always attempt to balance the data richness with the minimization of the choice complexity.

CM techniques provide a natural way of analysing environment multidimensionality, but unlike the TCM and the CVM, these techniques were not developed in the context of environmental economics. The earlier applications were made in the fields of psychometrics, marketing and transport (Mackenzie, 1990). The earliest application of CM in the environmental field we were able to find was conducted by Sinden (1974), who applied the TCM and paired comparisons in the valuation of recreational and aesthetic experiences. About a decade later, Rae (1981 *apud* Desvousges *et al.*, 1983: Chapter 6, p.9) used contingent ranking to evaluate air quality improvements in Mesa Verde National Park. The application of CM in the environmental field has been expanding rapidly since the beginning of the 1990s. In the majority of the earlier empirical analysis on environmental valuation, CM was applied in parallel with the CVM or the TCM. It seems that comparing the results of the different methods (that is, assessing convergent validity) was one of the aims of the researchers.

3.2.3. CONTINGENT BEHAVIOUR

The CB method refers to the use of hypothetical questions to get knowledge about behaviour in constructed scenarios. In the context of environmental valuation of resources used for recreational purposes, respondents have been asked about their intended visitation behaviour given a proposed change in price, quality or access conditions (Cameron, 1992; Grijalva *et al.*, 2002; Lienhoop and Ansmann, 2011).

Two main formats have been used. They differ in relation to the reference period for the CB question. In one format respondents are asked about how they would have behaved in the past if some hypothetical change in price or in quality has been observed. Respondents can, for example, be asked whether they would still have done the same number of trips if visits have occurred in different conditions. We have called this format *reassessed contingent behaviour* (RCB) because respondents are asked to reassess their trip behaviour in same previous period. Alternatively, respondents might be asked about their intended future trip behaviour in reaction to some proposed change[15]. Instead of reassessing their previous behaviour, respondents are asked to predict how they would behave in future proposed conditions (Christie *et al.*, 2007). Therefore, we have called this format *intended contingent behaviour* (ICB).

---

[15] This method has been named the hypothetical travel cost method by Layman (1996) and the contingent trip model by Betz *et al.* (2003).

The majority of the biases listed in Table 3.1 are likely to affect CB responses, namely the hypothetical bias, the strategic behaviour and the payment vehicle bias. The autonomous application of this method is rare (Betz *et al.*, 2003 is one of the few examples). The CB method has been applied jointly with a RP method, mostly TCM, to take advantage of the strengths of both techniques. Among the earliest applications there is Ribaudo and Epp (1984).

Future research should focus on analyzing the influence of the CB question format on answers. This aspect has not been explored in the literature, but we believe that some differences are likely. First, the RCB only requires one hypothetical question because current conditions represent the *status quo* situation. Conversely, when the researcher opts for the ICB format, the observed conditions no longer represent the *status quo* situation. Hence, respondents must answer at least two hypothetical questions, one assuming that current conditions will remain unchanged and other presenting hypothetical changes. Second, the choice conditions are not identical. Respondents are sure about their current income, while the future one is always linked to some uncertainty degree. Third, when visitors are asked to reassess their trip behaviour, current site conditions are already known but, especially in the case of the first visit, before the visit their knowledge was different. Hence, when behaviour is reassessed, the decision is made under different conditions. In the ICB format, available data on the *status quo* and conditions after the proposed change are the same.

3.2.4. TRAVEL COST METHOD

The foundation of the TCM is ascribed to Hotelling, who in 1947 suggested the use of the zonal version of the model in a letter to the director of the United States National Park Service. The TCM establishes a site demand curve by associating the number of trips, or visit rates, to a recreational site with the implicit trip price. Economic benefits are derived by the area under this demand curve between the current price and the choke price. The implicit price (or travel cost) is given by travel expenditures.

The method is based on the premises that visit frequency to a recreational site declines with increasing travel distances (due to higher costs) and that people consider travel costs similarly to entrance fees. The idea is that the observation of the travel cost that individuals bear to gain access to recreational sites makes it possible to infer how much people value each site.

Table 3.4: Versions of the TCM

Version			Seminal work	Choice	
				Number of sites	**Time**
SINGLE SITE	**Zonal**		Hotelling (1947)*	One site, several origins	Number of trips during a period of time.
	Individual		Brown and Nawas (1973)*	One site, several individuals or households	
MULTIPLE SITE	Regional Recreation Demand	Demand system a)	Burt and Brewer (1971) Brown and Hansen (1974)	Several sites, several origins or individuals	
		Gravity	Cesario and Knetsch (1976)		
	Varying parameter b)		Vaughan and Russell (1982) Smith and Desvousges (1985)		
	Pooled				
	Hedonic Travel Cost		Brown and Mendelsohn (1984)		
	Random Utility	Site choice	Hanemann (1978 *apud* Kling and Crooker, 1999)	Several sites, several individuals	Selection made at a choice occasion.
		Linked	Bockstael *et al.* (1987a) Bockstael *et al.* (1989a)		Selection made at a choice occasion + Number of trips during a period of time.
		Repeated choice	Bockstael *et al.* (1989a) Morey (1993)		
	Kuhn-Tucker		Phaneuf *et al.* (2000)		

* These authors presented the original ideas of the model, but did not perform the first empirical application.

a) Also called “multiple equation”, “partitioning” (Mendelsohn, 1985) and “site-specific multiple site model” (Ward and Beal, 2000: 135).

b) Also known as generalized travel cost model (Smith and Desvousges, 1985).

Table 3.4 summarizes the most widely used versions of the TCM, highlighting the aspects which make the versions distinctive. This table shows the distinction between two broad groups of models: the single site and the multiple site versions. Seminal works which settled the roots or become references of each version are presented following the chronological order. The table also shows that two main frameworks can be identified concerning the time. One analyzes the number of trips made during a period of time.

Another examines the choice of a site from a set of substitute sites at a moment in time. Models in the three final lines are mixtures of two time frameworks.

3.2.5. CLOSING NOTES

The CVM and CM are both SP methods and as such, theoretically, are both able to deal with any component of TEV. However, these methods have been approached differently in literature. Much research on the CVM has been devoted to the analysis of its main biases and ways of overcoming them. The discussion regarding the application of CM in environmental non-market valuation has been more focused on its advantages relative to the CVM and on the comparison of results across techniques. The most widely emphasized advantages of CM regard the fact that respondents are made aware that different amounts of each attribute might be available, and that price is treated simply as one of the attributes, without being the focus of the survey (Mackenzie, 1990). However, the flipside of each of these advantages is a disadvantage. The most obvious is the higher degree of complexity in comparison with the CVM. For example, Madureira *et al.* (2011: 402) report the excessive cognitive burden noticed in the pilot survey using CE as the reason for applying CVM in the main survey instead.

In comparison with the SP methods, the TCM suffers from some limitations. First, like any RP method, it cannot be used to estimate any component of passive use value. Second, the welfare measure directly obtained is the MCS[16], while with CM and the CVM one of the Hicksian welfare measures can be recovered. Finally, it is based on historical data and thus does not enable the estimation of values for quantity/quality levels that have not been experienced. The TCM has also strengths. The TCM is based on observed behaviour and not on the answers to hypothetical questions, thus avoiding all the bias associated with the hypothetical scenario, which is an important advantage. Furthermore, Hicksian welfare measures can be indirectly derived. Hence, we argue that the TCM should be preferred

---

[16] As explained in Chapter 4, with the exception for the random utility model (RUM).

when the research aims to compute actual use values only.

3.3. EMPIRICAL RESEARCH FOCUSING ON PORTUGUESE RESOURCES: A REVIEW

The survey of empirical research which focuses on Portuguese resources presented in this section aims to answer three main questions. We begin by asking: "*What has been done in the domain of non-market environmental valuation in Portugal?*". The following logical question is "*What common features can be observed across different studies?*". The last question is "*What do we know about the validity/reliability of the monetary values obtained?".*

3.3.1 What has been done in the domain of non-market environmental valuation in Portugal?

Table 3.5 provides a partial answer to this question as it lists the main studies and for each one identifies: i) the type of publication; ii) the resource involved and the year the survey was administered; iii) the policy measure (when applicable); iv) the method, the question format and the sample size; v) the population surveyed; vi) the components of TEV under evaluation and viii) the payment vehicle (when applicable).

Table 3.5: Studies applying non-market valuation methods to Portuguese resources

Study	Type of publication	Resource and year of the survey	Policy measure	Method, format and sample size	Population surveyed	Values	Payment vehicle
Perna (1994)	Master's thesis	*Culatra* Island beaches, Ria Formosa Natural Park (1992)	n.a	Single site ZTCM, n=406	Visitors	Recreational use	n.a
				CVM-DC, n=406			
Machado and Mourato (1998)	Conference paper	Estoril Coast beaches (1997)	Water quality improvement	CVM-MBDC, n=401	Region visitors	Health benefits	Fee
			n.a	CM-CRk, n=195		Use value	Access cost
Santos (1998)	Book [a]	Agricultural landscape of the PGNP (1996)	Landscape conservation	CVM-DC, n=704	Visitors	Recreational use + Passive use	Household income tax increase
Perna (2001)	Doctoral thesis	*Culatra* Island beaches, Ria Formosa Natural Park (1997)	Conservation	CVM – OE, n=604	Visitors	Recreational use	Environmental fee
				CVM – DBDC, n=577			
Santos *et al.* (2001)	Report	Sportive Fishing (2000-2001)	n.a	Multiple site ZTCM n=905	Users	Recreational use	n.a
Madureira (2001)	Doctoral thesis	Traditional landscape of almond trees (1998-1999)	Landscape conservation	CVM – DC, n=1 027	Visitors and region residents	Recreational use + Passive use	Increase in the annual income tax
				CM – CE, n=796			
Ribeiro (2002)	Master's thesis	Sportive fishing in Lagoons in the Alentejo region (2000)	Access restriction	CVM –DBDC, n=223	Users	Recreational use	Entrance fee
			n.a	Multiple site ZTCM, n=497 and ITCM, n=325			n.a
Nunes (2002a, 2002b)	Journal article	*Sudoeste Alentejano* and *Costa Vicentina* Natural Park (1997)	Protection programs	CVM – DBDC, n=1 678	General population	Recreational use + Passive use	One-time donation
Marta-Pedroso *et al.* (2007)	Journal article	Cereal Steppe of *Castro Verde* (2001)	Landscape preservation	CVM – OE, n=422	General population	Passive use	Annual tax increase or One-time donation
Madureira *et al.* (2011)	Journal article	Forestry perimeter of *Cantão das Hortas* (2003)	Management strategies	CVM – DC, n=900	Region residents	Recreational use + Passive use	Household annual income tax increase
Mendes and Proença (2011)	Journal article	PGNP (1994)	n.a	Single site ITCM, n=243	Visitors	Recreational use	n.a
Cunha-e-Sá *et al.* (2012)	Journal article	Traditional landscape in the Douro Region (2006)	Landscape conservation	CVM – DC, n=706	Visitors	Recreational use + Passive use	Household annual income tax increase

[a] Results of the parallel study for the Pennine Dales (United Kingdom) are not reported here. (n.a. – not applicable).

Research in the field began about twenty years ago. The earliest studies we could find date from the 1990s. The earliest is Perna's (1994) Master's thesis, where the CVM and TCM were used in evaluating the recreational use value of the Culatra Island beaches. Given the number of studies and articles published in journals, interest in environmental non-market valuation seems to have been increasing since the late 1990s, following the general trend of research in the economic field in Portugal (Guimarães, 2002: 8). Research results have been made available through different channels, namely, academic theses, books, technical reports, working-papers, conference papers or/and journals articles. When one main piece of research gave rise to different publications, only the main work is included in Table 3.5[17].

As a general overview, we emphasize five main features. First, the CVM is the prevailing method, probably because of its ability to estimate any component of economic value and its lower degree of complexity in comparison with CM. Second, in SP studies visitors have been the population most often surveyed; accordingly, assessing recreational use value seems to be of particular interest. Third, the Peneda-Gerês National Park (PGNP) has received special attention, most likely because of its features, which make it the only national park in the country. Fourth, the loss of positive externalities as a result of the abandonment of traditional agricultural activities with impacts on fauna and/or on landscape conservation seems to concern the researchers in this area. Finally, interest in the subject seems to be shared equally by economists and agronomists/biologists.

---

[17] Intentionally, we leave out of the survey the following studies: i) the CVM results of Perna's Master's thesis because the same method was later applied in more depth by the author for the same resource. ii) A report by Cruz and Royuela (2009) concerning the estimation of the socio-economic benefits of the Special Protected Area of Pico da Vara/Ribeiro do Guilherme in the S. Miguel Island (in the Azores archipelago). Though estimations based on the TCM and the CVM are referred to, methodological details are not provided. iii) Figueira (1994) who tried to apply the CVM to estimate the WTP for water quality improvement in the public supply system. The research was conducted in an unfavourable social context: the population was elderly, had a low level of education and participation was low. The WTP could not be asked directly and only 41 persons took part. iv) Pereira (2004) as we place some reserve upon the valuation scenario.

3.3.2 What common features can be observed across different studies?

Concerning the CVM, three elicitation formats have been used in the WTP question. The National Oceanic and Atmospheric Administration – NOAA panel (Arrow *et al.*, 1993) recommended the use of DC because it better mimics the market *take it or leave it* situation, characteristic of private goods' markets. Instead of DC, its variants are frequently preferred because the additional question(s) improves the efficiency of estimates. In the studies surveyed, the DC formats are indeed dominant but only Machado and Mourato (1998) assessed the degree of certainty in responses. OE has been used as well. The preference for this format is usually justified by its more conservative estimates. This result is corroborated by Perna (2001), whose estimates using DBDC are 1.57 higher than using the OE format.

There are a few variables which seem to be globally important to explaining the WTP, as they are statistically significant across studies. In SP models, the past and current use of the site being studied is associated with higher WTP levels. WTP is also positively related to income and formal education, while age seems to exert a negative influence. Furthermore, Nunes (2002b) and Santos (1998) concluded that the WTP of urban populations is significantly higher than that of rural ones.

Madureira (2001) used the CE in addition to the CVM to assess the value of different landscape attributes. Two important conclusions are that the order of preference concerning different combinations of landscape attributes does not differ among methods and that the CVM produced the most conservative estimates. The author attributes the difference to a flawed focus of the respondents on the price in the CM exercise, which counters the idea that treating price as one among the many attributes is an advantage.

The internet is the most recent channel used in questionnaire administration and was used by Marta-Pedroso *et al.* (2007) in parallel with in-person interviews. This seems to be a promising option in Portugal as well because in spite of the very low response rate, no major differences were found between the sub-samples. This channel is advantageous in what concerns to time and budget constraints, which are always important obstacles to obtaining larger samples.

The TCM was used in the estimation of the recreational values of three quite different resources/activities: a beach, used mainly for bathing; a set of lagoons, used for fishing activities; and the PGNP. The majority of the studies applied the zonal version and regressions were always performed using administrative zones as origins due to the difficulties in obtaining data to deal with concentric rings. Mendes and Proença (2011) opted for the individual version of the model which currently dominates the literature. In both versions, besides the travel cost, some measure of income (household income, income available for recreational activities and purchasing power) proved to be significant in explaining the demand level. The effect of the travel cost is always negative (as expected), while the influence of income on demand differs across studies. All of the authors considered the opportunity cost of time as a component of the total travel cost. The percentage of the wage rate used as *proxy* of the opportunity cost was not uniform across studies, which is evidence of the lack of consensus among researchers.

3.3.3 What do we know about the validity/reliability of the monetary values estimated?

In preference studies, researchers are unable to observe true economic values. Hence, one of the main areas of concern regards the ability of valuation methods to produce reliable and valid estimates. Reliability concerns the replicability of the measurements and validity is about the correspondence between what one wishes to measure and what is actually measured (Carson *et al.*, 2001: 193).

Three main types of validity can be assessed: content, criterion-related and construct (Mitchell and Carson, 1989: 190). Content validity "refers to the extent to which design and implementation of the survey conform to the generally recognized best practice or state of the art" (Freeman, 2003b: 178). Criterion validity is confirmed when the welfare measure estimated is not statistically different from a value known to be the truth or close to the theoretical construct under investigation (Carson *et al.*, 1996: 80). Construct validity includes convergent and theoretical validity (Bishop, 2003: 543). Theoretical validity is verified when results conform to the economic theory. Convergent validity is confirmed when different methods yield measures that are not statistically different, without any presumption about which method is the most correct one. In the words of Bishop (2003: 543), “the measures have roughly equal status”, otherwise it would be a criterion test.

The research contexts underlying the studies surveyed are conducive to content validity as these studies were produced in the context of supervised academic research or evaluated by peers before publication. There is also evidence of theoretical validity since price and income are significant explanatory variables of demand and sensitivity to scope is verified. Convergent validity can only be assessed when more than one method is used in a similar evaluation exercise. That is, when the resource and the components of value involved coincide. Convergent validity was not confirmed by Ribeiro (2002) who compared the results derived from the CVM and the TCM. In Madureira (2001), after correcting for the *yea-saying* bias in the CVM, welfare measures were not statistically different from the ones obtained through the application of CE.

Reliability involves the extent to which a survey will yield statistically equivalent estimates in repeated trials. Test-retest procedures and temporal stability tests have been used to assess reliability. Temporal stability is tested by comparing monetary values obtained interviewing two different samples using the same survey instrument, at two different points in time (Carson *et al.*, 2001: 195). Test-retest procedures are even more demanding as they require the same sample to be re-interviewed using an identical survey instrument (Loomis, 1993: 184). These tests are rare, mainly due to the high costs involved, and this is probably the reason why none of the studies listed above has conducted such tests.

3.4. CONCLUDING REMARKS

This chapter has provided an overview of four non-market valuation methods widely used. The CVM and the TCM have a long tradition in the field, while the use of CM and CB is growing rapidly. Whereas CM is seeing as an alternative to CVM, the CB has been used jointly with the TCM. From a global perspective, we can say that research relying on these methods has been intense. In Portugal, the decade of the 1990s can be identified as the turning point.

Our survey shows that past and current use of natural areas and similar sites for recreation purposes positively influences the WTP. In general, income has been shown to have a significant and positive effect on the value of resources. Furthermore, evidence that higher levels of formal education are associated with higher demand for outdoor recreation sites and with a higher WTP for conservation (Madureira, 2001; Nunes, 2002b), lead us to expect that, as the level of formal education improves in Portugal, values assigned to natural resources will tend to increase. The accelerated urbanization of the country is likely to act in the same direction.

Assessing the preferences, perceptions and concerns of Portuguese citizens regarding natural areas is particularly relevant for policy-makers. Abroad, public agencies recognize the usefulness of the estimates obtained by the application of these techniques for deciding among alternative policies (List, 2005; Sugden, 2005) and studies have been conducted in order to meet the needs of public agencies (Cameron *et al.*, 1996). However, it seems that in Portugal, in ten years, the state of affairs has not changed significantly. As observed by Perna (2001: 254), “in our country, out of the academic circle, there is not yet enough knowledge and/or trust to use results of non-market valuation as data sources for public decisions”.

The survey also shows that different types of natural resources have captured the attention of researchers. However, in spite of the number of national forests in the country and their probable high non-timber value, not enough attention has been devoted to their non-market valuation.

CHAPTER 4 – THE TRAVEL COST METHOD: A REVIEW[18]

4.1. INTRODUCTION

The TCM is the oldest and most widely used RP technique to assess the recreational benefits derived from the use of natural resources. It has been applied to different natural resources and recreational activities with several specific objectives lying behind its application. An obvious one is to assess the value of current benefits in order to know the extent of the loss if the resource were to be employed for other purposes[19]. As described by Hotelling (1947):

> It is this consumers' surplus cost (calculated by the above process with deduction for the cost of operating the park) which measures the benefits to the public in a particular year. This, of course, might be compared directly with the estimated annual benefits on the hypothesis that the park area was used from some alternate purpose.

A second reason for using the TCM is to predict the benefits accruing from the creation of a new site similar to other(s) already used for recreation purposes. A third classical motivation is the need to understand how different characteristics add to the resource's economic value. Related to the latter goal is the need to assess how a change in sites' characteristics (e.g., a quality change) affects users' well-being.

Several variants of the TCM have been conceived over the years directly related to these motivations. Given the multiplicity of versions of this method, it is helpful and enlightening to provide an organized overview of the developments usually grouped under the umbrella of the TCM. This is done in Section 4.2. Section 4.3 reviews the main issues of the method, discusses the empirical solutions adopted to deal with them and presents the main aspects of the theoretical framework underlying those solutions. Section 4.4 contains the concluding remarks.

---

[18] A review article containing an in-depth literature review on the use of the several versions of the TCM (whose content is predominantly taken from this chapter), as well as of the most recent trends and efforts to combine this technique with contingent methods (whose content is predominantly taken from section 8.2) was submitted to a scientific journal and is currently under review

[19] A recent example is provided by Hynes and Hanley (2006) concerning the management of Irish rivers.

4.2. THEORETICAL BACKGROUND

As highlighted before, the TCM belongs to the group of RP techniques. The formal definition of revealed preference, first described by Samuelson, can be enunciated in the following way, quoting Varian (2006: 2):

> *Given some vectors of prices and chosen bundles* (p^t, x^t) *for* $t = 1, \ldots, T$ *, we say* x^t *is directly revealed preferred to a bundle* x *if* $p^t x^t \geq p^t x$ *(written* $x^t R_D x$*). We say that* x^t *is revealed preferred to* x *(written* $x^t R x$ *) if there is some sequence* $r, s, t, \ldots, u, v$ *, such that* $p^r x^r \geq p^r x^s,\ p^s x^s \geq p^s x^t, \ldots, p^u x^u \geq p^u x$*. In this case, we say the relation R is the transitive closure of the relation* R_D *.*

Economic valuation of non-market environmental resources using the TCM is based on the assumption of weak complementary between the resource and some market good. Weak complementarity between a non-market environmental good (*E*) and a market good (*W*) requires that both are consumed together and is based upon two assumptions (Freeman, 2003b: 112).

First, if the demand for the private market good is zero, then the marginal value of a change in the environmental resource is also zero. This relation is expressed by means of the utility function. Hence, if *Y* denotes the demand for all the other goods and $U(\cdot)$ is the individual's utility function:

$$\left.\frac{\partial U(W,E,Y)}{\partial E}\right|_{W=0} = 0, \tag{4.1}$$

where *E* represents a specific level of a characteristic of the private good and its level matters only if the individual consumes *W*.

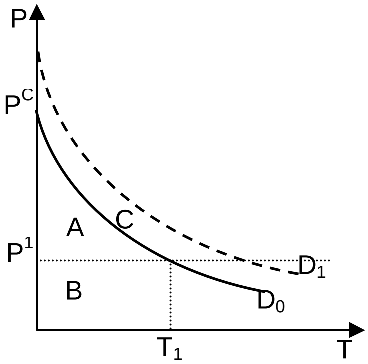
Second, the market good must be non-essential. A good is non-essential if there is some choke price where the price is so high that quantity demanded falls to zero. When the price is at or above its choke price, changes in *E* have no influence on expenditure and have no value to the consumer. Considering P_w^C the choke price for *W*, P_Y the vector of prices of all the other goods and $e(\cdot)$ the expenditure function, that relation is expressed by:

$$\frac{\partial e\left(P_Y, P_W^C, E, U^0\right)}{\partial E} = 0. \tag{4.2}$$

These conditions are the starting point for deriving the value of environmental resources. These values are implicit in the expenditure on market goods because changes in environmental quality will be reflected in the demand for the market good which is weak complementary. The CS for a change in *E* might be measured by the area between two compensated demand curves for *W*.

Another implicit assumption is that, for the representative visitor, the utility function is separable in the recreational activity being modelled. This means that the demand for the recreational activity can be estimated independently of demand for alternative leisure activities and alternative non-leisure market goods (Hanley and Spash, 1993: 84).

Figure 4.1: Demand curve and consumer surplus for trips



Since RP methods are based on market information, they typically provide estimations of the Marshallian and not of the Hicksian surplus. The travel demand estimated is a Marshallian demand function, so the MCS is an approximation to the true economic value of nature-based recreation (Willig, 1976). Graphically, the MCS is the area under the demand curve between the current travel cost (P^1) and the choke price (P^C). The MCS is the area A depicted in Figure 4.1. Generically, the area is given by:

$$MCS = \int_{P^1}^{P^C} f(x,P)dP \ . \qquad (4.3)$$

This is a measure of the social benefits accruing from the access to the environmental resource. If the site were closed, area A would be the value of the loss associated with the actual use. The area under the demand curve ($A + B$) is the WTP for trips. A change in price or in one of the demand determinants would change the MCS. For example, an improvement in environmental quality would increase the MCS by the area C.

4.3. THE TCM MODELLING FRAMEWORK

As noted in Section 3.2.4, the origins of the TCM are attributed to Harold Hotelling (1947). The earlier published empirical studies that apply the method include the works by Trice and Wood (1958) and Clawson (1959 *apud* Ward and Beal, 2000: 33). Earlier applications refer mainly to the estimation of the monetary value of actual users' benefits derived from water based recreational activities (see, e.g., Trice and Wood, 1958; McConnell and Strand, 1981; Vaughan and Russell, 1982; Desvousges *et al.*, 1983). The TCM is now applied to an extensive spectrum of recreational sites, such as forests, parks, lakes, rivers, beaches, heritage sites and related activities (e.g., fishing, kayaking, rock and ice climbing). These sites and activities have two main common features: users must travel to the site to enjoy it and access is free or only a nominal entrance/licence fee is charged.

The implicit price (or travel cost) is given by travel expenditures. The method is based on the premises that visit frequency to a recreational site declines with increasing travel distances (due to increasing costs) and that individuals respond to changes in travel costs in a similar manner to changes in site entrance fees[20]. The idea is that the observation of the travel cost that individuals bear to gain access to recreational sites makes it possible to infer how individuals value each site.

---

[20] For an interesting theoretical treatment of this aspect, see Bowes and Loomis (1980: Section II).

Travel costs may include several components, such as travel expenditures, entrance fees, the opportunity cost of time, on-site expenditures and expenditure on equipment. A number of factors, such as substitution possibilities and socio-demographic characteristics act as demand determinants and help in explaining visitors' recreation behaviour. These factors are believed to explain the demand for trips as visitors with particular characteristics travel to specific sites with preferred attributes to attain the desired recreation experience (Shrestha *et al.*, 2007).

As shown in Table 3.4, recreation demand analysed in the TCM framework may refer to a single site or to several sites. In the first case, a single site model is applied. A multiple site model is usually used to estimate recreation demand considering various substitutes sites. There are several versions of the multiple site model which has evolved from the earlier demand system into other sophisticated models based on a discrete choice framework. Regarding time, choices were originally modelled following one of two possible perspectives: the total number of visits was measured over a specific time period, such as a year or a season; or the decision was made at a particular moment regarding the choice of a recreational site from a set of sites. The development of hybrid approaches has also been common.

4.3.1. SINGLE SITE MODELS

In single site models the quantity demanded is the number of trips to a recreational site or the visitation rate during a certain period. The price is given by the (travel) cost paid to reach the site. Since visitors live at different distances from the recreational site and make a different number of trips, it is possible to observe different quantities corresponding to different price levels. A negative relationship between quantity and price is expected.

This approach is useful when the objective is to estimate the total use/access value of the site or the value associated with changes in the access cost. It requires less data than multiple site models and is particularly suitable when the number of substitutes sites is small (Parsons, 2003: 324). Single site models can be applied in three possible ways: as a zonal travel cost model (ZTCM), as an individual travel cost model (ITCM) or using a hybrid structure.

4.3.1.1. Zonal Travel Cost Model

The ZTCM, also known as the Clawson-Knetsch method, was the first version of the TCM to be employed (Sutherland, 1982a: 231; English and Bowker, 1996: 80). This version relies on the definition and use of a number of geographic zones of origin, at varying distances from the recreational site. Zones of origin may be concentric rings radiating from the site, as suggested by Hotelling, or administrative zones. In both cases, the area surrounding the recreational site is portioned into distinct areas and the number of visits from each is divided by the corresponding population, giving rise to visit rates (Brainard *et al.*, 1997). Measuring the visits on a *per capita* basis accounts for the number of visits and the probability of visiting the site as function of the distance (Oster *et al.*, 1987: 30).

Individuals from each zone are assumed to have the same socio-demographic characteristics, to support equal travel costs and to find the same quantity and quality of substitute sites available. Variations in the visitation rates are explained by variations in these variables among zones. Collecting demographic, social and economic data is easier when administrative zones are used, so this option is more likely to be selected (Fleming and Bowden, 2009)[21].

The classic ZTCM relates visits *per capita* from each original geographic zone with the costs incurred travelling. A single-destination trip generating function can be described as:

$$T_i / N_i = f(C_i, X_i), \quad i = 1, \ldots, I, \qquad (4.4)$$

where T_i is the number of visitors from zone *i* in a given period of time; N_i is the population of zone *i;* C_i is the travel cost borne by visitors from zone *i* to the recreational

---

[21] Smailes and Smith (2001) is the most recent study we found using concentric rings. They used nine distance bands at 5km intervals.

site; X_i are other explanatory variables that differ among zones and determine the probability that any individual visiting the site came from zone *i,* e.g., average zonal values for income, education, unemployment (Willis and Garrod, 1991: 515); T_i / N_i is the visit rate and hence is a continuous variable. The trip generating function can take a variety of alternative functional forms (linear, quadratic, semi-log or double-log). However, there is no recognized theoretical superiority for any of them (Brainard *et al.*, 1997).

The ZTCM develops in several stages. First, a trip demand curve is computed using information from all zones of origin. Second, the coefficients of the first demand curve are used in estimating the visitation rates from each zone at varying travel costs. Third, the aggregate demand curve is generated by multiplying the population in each zone of origin by the visitation rate at each hypothetical travel cost and then summing the number of visits at each price (Dwyer *et al.*, 1977).

Once a trip generating function has been estimated, the researcher has established a relation between the cost and the number of visits, for a given level of the other explanatory variables. The trip generating function may be used to predict what would happen to visits from each zone as the explanatory variables change. The area above the price and under the demand curve is an assessment of the MCS and differs among zones of origin at varying distances.

Although less common, a reverse gravity model has also been used to estimate the demand and the economic value of a recreational site receiving visitors from a set of origin zones. As presented by McCollum *et al.* (1990), origins are weighted in inverse proportion to the cost of reaching the recreational site. This specification comprises two multiplicative parts, a trip generation model and a trip distribution model. The trip generation model predicts the total number of trips received by a recreational site as a function of the site's atractiviness and acessibility. The trip distribution model estimates the relative proportion of trips coming from each origin in the market area. This relative proportion depends on origin characteristics, the travel cost and the price of substitute sites.

The ZTCM offers two main advantages over the ITCM. It is less data intensive since the home post code is the only individual information that is strictly necessary. Consequently, data collection is easier and less expensive. However, the literature reports some problems with applications of the ZTCM (Sutherland, 1982b; Ward and Loomis, 1986; Rosenthal, 1987; Garrod and Willis, 2001: 59; Parsons, 2003: 295). The most widely highlighted are:

a) The difficulty in accurately defining the zones of origin;

b) The use of origin zones works with the restrictive assumption of homogeneous populations in each zone, hence heterogeneity among respondents cannot be accounted for;

c) The aggregation into zones leads to multicollinearity and efficiency losses in the estimation of statistical parameters. Consequently, larger samples are required;

d) The model suffers from a lack of consistency with basic consumer theory;

e) This version is unsuitable for estimating the benefits of recreational areas which are linear in topology (e.g., rivers).

As a consequence of these limitations, the ITCM has been gradually favoured in applied research. However, the ZTCM continues to be used (Fleming and Cook, 2008; Grossmann, 2011) specially due to data constraints (see, e.g., Rosenthal, 1987: 829). The zonal version is the most suitable alternative, for example:

a) When a site receives a high number of first-time or one-off visitors;

b) When visitation rates are low;

c) If it is not possible to collect individual data with the necessary level of detail.

The comparison of the results obtained from the two versions is another motivation for the use of the zonal version. When the recreational site is visited mainly by a local population who bear very low travel costs, neither of the versions can be used (Bishop, 1992).

4.3.1.2. Individual Travel Cost Model

Concerns about the zonal version led Brown and Nawas (1973) to suggest that individual users might be a better unit of observation than zone population. This contribution was an important stimulus to the use of the ITCM. The main arguments against the use of the ZTCM are advantages of the ITCM that, in addition, make it possible to disaggregate visitors according to the type of recreational activity and the type of visit undertaken. This information can be used to split respondents into sub-groups and to analyse them separately (Willis and Garrod, 1991: 515).

In the ITCM, the dependent variable is the number of trips each individual visitor, or household, *i* made to a site *j,* during some time period. Variation in price is obtained by observing the behaviour of people living at different distances from the site. As usual, utility maximization is the individual's objective. The individual's utility function may be defined as:

$$U_i = U(T_i, Y), \quad i = 1, \ldots, I, \tag{4.5}$$

where T_i is the number of trips made by individual *i* to the recreational site *j* during a season; Y is a vector of all the other goods.

If C_i is the cost *per* person, *per* trip and the individual's income (M_i) is spent on recreational trips and on all the other goods with unit price P, the budget constraint is given by:

$$M_i = P \cdot Y + C_i \cdot T_i. \tag{4.6}$$

Utility maximization, subject to the budget constraint, leads to Marshallian demand functions of the type:

$$T_i = f(C_i,, M_i) \ . \tag{4.7}$$

As before, the MCS of trips to site *j* is the area below the demand function and above the implicit price (C_0). The MCS is an approximation to the WTP of visitors to recreation sites and is given by:

$$MCS_i = \int_{C_0}^{C_C} T_i(\cdot)dC_i \ . \tag{4.8}$$

The set of explanatory variables is usually enlarged to include additional data which helps to explain different levels of demand by different individuals. These variables include: the travel cost to substitute recreational sites, respondents' individual characteristics (such as age, gender, family size, occupation, income, level of education), the major recreation activities performed at the site, years of experience in visiting the recreational site, attitudinal variables and rating variables, representing the classification of site attributes. Unobservable individual factors that influence recreation demand will be incorporated into the error term in the econometric model.

4.3.2. MULTIPLE SITE MODELS

This sub-section presents an overview of the main variant of the TCM developed in order to deal with sets of sites. The designation *multiple site models* is used here to identify the models that deal with the choice between substitute sites, which may or may not take into account the number of trips made to each of the recreational sites. The main versions are: the regional recreation demand model (RRDM), which includes the demand system model and the gravity model; the varying parameter model, the pooled model[22], the hedonic travel cost method (HTCM) and the random utility model (RUM).

In this framework, the prices and characteristics of substitute sites are important determinants of demand and substitution possibilities between recreational sites are explicitly recognized. Whatever the multiple site model chosen, the selection of the sites to

---

[22] See, e.g., Parsons (2003: 296) for a different structure.

be considered as substitutes and the measurement of the attributes are always complex and crucial tasks.

4.3.2.1. Regional recreation demand model

The RRDM is a generalization of the single site model. Early developments date back to the beginning of the 1970s (Brown and Hansen, 1974). These models were widely used to assess the recreational value of water reservoirs managed by the US Army Corps of Engineers and the main developments were achieved in this context (Ward *et al.*, 1996: 10-11). The RRDM typically identifies the determinants of recreation use with the ultimate purpose of using the quantitative relation estimated to forecast the effect of changes in the explanatory variables on the number of visits (or visit rates), as well as the impacts of new projects.

These models combine visitor data from several origins with data from several recreation sites within a region to estimate a recreation demand and the benefits for a group of sites (Sutherland, 1983). In the most recent versions, four main categories of factors determining recreation use are usually considered, specifically: travel cost, qualitative characteristics of natural resources (e.g. size and facilities), user characteristics (e.g. age, gender and income), and availability and price of substitutes (Henderson and Allen, 1994). The dependent variable is the number of trips, or *per capita* trips. Hence, a possible general specification of the demand equation is of the kind:

$$\frac{T_{ij}}{N_i} = \beta_0 + \beta_1 \cdot C_{ij} + \beta_2 \cdot S_i + \beta_3 \cdot Z_j + \beta_4 \cdot E_i + \varepsilon_{ij}, \tag{4.9}$$

where T_{ij} is the number of trips made by individual i to the site j; N_i is the population of zone i; C_{ij} is the trip cost borne by individuals in zone i to access site j; S_i is the vector of demographic variables; Z_j is the vector of site j characteristics; E_i is the vector of attributes of substitute sites available for origin i visitors; and β are the coefficients to be estimated.

More than thirty years ago, when many of the versions were yet to be developed, Dwyer *et al.* (1977: 105) structured the published work concerning multiple site models into two broad classes of regional models. One class is based on a linear system of interrelated demand equations. The other class is based on a single equation gravity model. Both models require sites to be distinct due to observable attributes and users' preferences and a survey of individual or household recreation behaviour to be used in collecting data.

Demand system model

The models of Burt and Brewer (1971) and Cicchetti *et al.* (1976) belong to the first class identified by Dwyer *et al.* (1977). They are among the first works to explicitly specify multiple site demand. Authors extended the single site to a multiple site model using a system of interrelated demand equations and household data (Bockstael *et al.*, 1985). Similar sites are grouped into homogeneous categories and substitution across categories is explicitly modelled based on price. Sites within a category are considered to be perfect substitutes; hence it is assumed that each individual would visit only the one nearest to his/her home.

These models are formally represented by a system of *K* interrelated equations, such as (Dwyer *et al.*, 1977: 112):

$$T_{ij} = \alpha_j + \beta_{jk} \cdot C_{ik} + \gamma_j \cdot S_i + \varepsilon_{ij}, \quad j,k = 1,\dots,K; \quad i = 1,\dots,I; \tag{4.10}$$

where T_{ij} is the number of trips made by individual *i* to the nearest site of type *K*; C_{ki} is the trip cost borne by individual *i* to access the closest site of type *K*; S_i are the demand shift variables; $\alpha_j, \beta_{jk}, \gamma_j$ are the coefficients to be estimated.

Cross price effects must be symmetric; hence the constraint $\beta_{jk} = \beta_{kj}$ must hold. The first and last equations are:

$$\begin{cases} T_{i1} = \alpha_1 + \beta_{1k} \cdot C_{ik} + \gamma_1 \cdot S_i + \varepsilon_{i1} \\ (...) \\ T_{ik} = \alpha_k + \beta_{kk} \cdot C_{ik} + \gamma_k \cdot S_i + \varepsilon_{ik} \end{cases} . \tag{4.11}$$

The *K* equations are estimated jointly in order to reflect the interrelation between the demand for each site category. As in the single site version, benefits are calculated as the area under the demand curve previously estimated. If C_1^0 is the travel cost to site 1 and that C_1^C is the choke price, the MCS for site 1 is given by (Seller *et al.*, 1985):

$$MCS = \int_{C_1^0}^{C_1^C} T_1(C_1,...)dC_1 . \tag{4.12}$$

The main novelty of this model is that all of the sites considered to be part of a recreational site system are analysed together. Thus, consistently with economic theory, the substitution effect in demand is captured by the inclusion of the price of substitute sites in demand for trips to site *j*. The main limitations concern the definition of homogeneous sites and the fact that the method becomes cumbersome as the number of categories increase (Mendelsohn, 1985). Furthermore, although the existence of substitutes is explicitly accounted for by the model, the role of the attributes of sites attributes is ignored.

Gravity models

The second class of multiple site models identified by Dwyer *et al.* (1977) are the gravity models which predict the distribution of recreational trips among several sites. Recreational sites have to be homogeneous entities assumed to be part of the same trip demand function, such as sets of lakes, forests or state parks. Sites within each set are assumed to differ only in attributes and accessibility. An attractiveness index is often incorporated into the specification, distinguishing the various recreation sites. This index reflects aspects such as the activities available, the quality of facilities and users' satisfaction. The number of visits received by each site is explained by its relative attractiveness and by the distance from the origins of all the recreational sites.

The gravity models applied in the multiple site context evolved from an earlier version where the number of trips to a site was exogenous to one that estimates the number of trips and their allocation across sites (Mendelsohn, 1985). Considering *K* recreation sites, a general model may be defined as follows (Buhyoff *et al.*, 1981):

$$T_{ij} = \left[\theta \cdot S_i \left(\sum_{k=1}^{K} A_k \cdot \exp(\beta C_{ij}) \right)^{\alpha+1} \right] \cdot \left[\frac{A_j \cdot \exp(\beta C_{ij})}{\sum_{k=1}^{K} A_k \cdot (\exp \beta_{ik})} \right], \qquad (4.13)$$

$$j,k = 1, 2, \ldots, K; \quad i = 1, \ldots, I;$$

where T_{ij} is the number of trips from origin *i* to site *j*; S_i are the population characteristics of zone *i*; A_j is the attractiveness index of site *j*; C_{ij} is the travel cost from origin *i* to site *j*; θ, α, β are the coefficients to be estimated.

The first term is the trip generation function which determines the aggregate number of trips. The second term is the trip distribuition function which allocates trips across sites. Slightly different versions of this type of recreational model were applied, e.g. by Buhyoff *et al.* (1981) to assess the benefits of visits to national parkways; by Sutherland (1982a) to camping, fishing, boating and swimming in 179 sites in the Pacific Northwest; and by Rosenthal (1987), who additionally estimated a system of interrelated demand equations based on the number of visits to eleven reservoirs.

One of main criticisms of the gravity models is the need to define an attractiviness index, which implies some subjectivity on the part of the researchers (Desvousges *et al.*, 1983: 7-11; Seller *et al.*, 1985). Additionally, Mendelsohn (1985) argues that the gravity model simply predicts participation and for this reason it is not sufficient to estimate the value of a site. In a critique of Sutherland´s work (1982b), who added a valuation part, he notes that if this valuation part is added, the choice process must be the same one that determines trip choice. These models had also been accused of merely being statistical allocation models, which were not based on economic principles.

A merit shared by RRDMs is that the demand for a new site and the resulting benefits can be evaluated. However, while Burt and Brewer (1971) were motivated by their interest in measuring the value of introducing a new site, the original gravity models were not designated for this purpose (Dwyer *et al.*, 1977: 114). The new site must be assigned to one of the site categories and in the gravity model the attractiveness index must be computed. Then, the market area must be defined and travel time estimated. The benefits and demand for the new site can then be estimated as before.

Enthusiasts of the most recent developments in RRDMs highlight two main advantages. First, the possibility of evaluating the impact on demand resulting from: new similar projects, changes in site attributes, or changes in demographic characteristics. Hence, results can be generalized to a wide range of management actions, site locations, visitor populations and substitute opportunities, as well as to a wide range of future on-site conditions. Second, results can be used to accurately transfer predicted visits or benefits to unstudied sites in the study region or outside, as long as user surveys are available that provide individual post codes, and regional recreation demand conditions are comparable (Seller *et al.*, 1985: 157; Sorg *et al.*, 1985; Ward *et al.*, 1996; Ward and Beal, 2000: 145). However, Phaneuf and Smith (2005: 696) argue that the RRDM, as well as the varying parameter model, do not provide a consistent or utility-based theoretical link from choice to empirical demand analysis.

4.3.2.2. Varying parameter and pooled models

The varying parameter model proposed by Vaughan and Russell (1982) was meant to account for the impact of site quality characteristics on aggregate visitation rates (Samples and Bishop, 1985). This version assumes that the parameters of individual site demand models are functions of site attributes. The visitation functions are usually estimated in two stages. In the first stage, a separate travel demand function to each recreational site is computed using data on travel and individual characteristics. In the second stage, there are as many equations as parameters from the first stage (Bockstael *et al.*, 1989c). The parameters are then the dependent variables and are explained by the differences in site characteristics. Smith *et al.* (1983) proposes a theoretical basis for the varying parameter model based on the household production framework[23]. In this framework, attributes determine the relative productivity of each site.

Analytically, the model develops as follows. In the first stage, *K* independent demand equations, one for each recreation site *j,* are computed. These equations may be expressed as follows (Vaughan and Russell, 1982; Bockstael *et al.*, 1989b):

$$T_{ij} = \beta_{0j} + \beta_{hj} \cdot X_i + \varepsilon_j, \qquad i = 1,...,I; \quad j = 1,...,K; \quad h = 1,...,H \, , \tag{4.14}$$

where T_{ij} are the visits *per* season to site *j* measured *per* travel zone or *per* person *i*; β is a vector of $H+1$ parameters to be estimated for each site/equation; X_i is a vector of *L* explanatory variables, which vary across individuals or origin zones, including variables such as travel cost, zonal or individual income and dummies; ε_j is the stochastic disturbance.

The first and the last equations would be:

$$\begin{cases} T_{i1} = \beta_{01} + \beta_{h1} \cdot X_i + \varepsilon_1, \\ (\cdots) \\ T_{iK} = \beta_{0K} + \beta_{hK} \cdot X_i + \varepsilon_K \end{cases} . \tag{4.15}$$

In the second stage, $H+1$ independent equations are computed, one for each parameter. L sites' characteristics are considered and there are *K* observations for each equation, such that:

$$\beta_h = \gamma_{0\ell} + \gamma_{b\ell} \cdot Z_j + \eta_\ell, \quad \ell = 1,...,L; \quad b = 1,...,H; \tag{4.16}$$

---

[23] The econometric model used was later revised by Smith and Desvousges (1985).

where Z_j is a vector of site characteristics that vary across sites; γ is a vector of $(L+1)\cdot(H+1)$ parameters to be estimated and that are invariant across sites; η_ℓ is the stochastic disturbance. The first and the last equations would be:

$$\begin{cases} \beta_0 = \gamma_{00} + \gamma_{b\ell} \cdot Z_j + \eta_0 \\ (\cdots) \\ \beta_H = \gamma_{0L} + \gamma_{bL} \cdot Z_j + \eta_L \end{cases} . \tag{4.17}$$

If C_1^0 is the current travel cost to site 1 and C_1^C is the choke price, the MCS for site 1 is given by (Brox and Kumar, 1997):

$$MCS = \int_{C_1^0}^{C_1^C} T_1(X_i)\, dC_1 . \tag{4.18}$$

The novelty of this model is that it accounts for sites' attributes and evaluates their changes without imposing constant demand parameters for all the sites. However, the varying parameter model requires a considerable number of observations for each site and the inclusion of several sites (Bockstael *et al.*, 1989b: 168). When there is a limited number of observations for each destination, a pooled model can be used (Caulkins *et al.*, 1986; Kling, 1988; Agnello and Han, 1993).

In the pooled model the two stages collapse into one by pooling the data from the various sites. Visitation is estimated as a function of travel cost, individual/zonal characteristics and site attributes from a pooled sample of all available observations (Vaughan and Russell, 1982). The number of coefficients to be estimated will be $L + H + (L \cdot H) + 1$. With two explanatory variables $(H = 2)$ and two attributes $(L = 2)$, the equation (4.16) is inserted into equation (4.15) to obtain the following expression:

$$\begin{aligned} T_{ij} = {} & \gamma_{00} + \gamma_{01} \cdot Z_{1j} + \gamma_{02} \cdot Z_{2j} + \gamma_{10} \cdot X_{1i} + \gamma_{20} \cdot X_{2i} + \gamma_{11} \cdot Z_{1j} \cdot X_{1i} + \gamma_{02} \cdot Z_{2j} \cdot X_{1i} \\ & + \gamma_{21} \cdot Z_{1j} \cdot X_{2i} + \gamma_{22} \cdot Z_{2j} \cdot X_{2i} + \eta_0 + \eta_1 \cdot X_{1i} + \eta_2 \cdot X_{2i} + \varepsilon_{ij} . \end{aligned} \tag{4.19}$$

The equation above is potentially heteroscedastic because the error term $\left(\eta_0 + \eta_1 X_{1i} + \eta_2 X_{2i} + \varepsilon_{ij}\right)$ is a function of the explanatory variables. This is a potential limitation of the one step procedure.

4.3.2.3. Hedonic travel cost method

A distinctive feature of the HTCM is that it focuses on the valuation of site attributes. Therefore, while the versions presented above analyse the demand for sites, the HTCM examines the demand for attributes. In the HTCM framework a group of sites is considered and each one is seen as a different bundle of attributes. The cost that visitors are willing to bear when going from origin *i* to each site *j* is functionally related to the attributes of *j*. As defined by Brown and Mendelsohn (1984), this technique attempts to reveal how much users are WTP for specific characteristics/attributes of recreational sites. Accordingly, the extra cost supported when people travel to a more distant site is assumed to be justified by its higher quality.

Each hedonic market is composed by sites visited by people coming from the same origin. If the same person visits more than one site, each visit will be a separate observation in the data set (Bockstael *et al.*, 1989b: 168). The analytical framework can be expressed in the following way: each site is described as a bundle of *L* attributes entering directly into the consumer's utility function $U(\cdot)$; C_{ij} is the travel cost supported when travelling from the origin *i* to site *j* and is exogenously determined; Y represents all the other goods and P is their unit price. The optimization problem can be formulated as (Pendleton, 1999):

$$\max_{W,Z \geq 0} U\left(Z_j, Y\right) \quad s.t. \quad M_i = P \cdot Y + \sum_{j=1}^{K} C_i\left(Z_j\right), \quad i = 1,\ldots,I; \quad j = 1,\ldots,K\,. \qquad (4.20)$$

This model also develops in two stages. In the first stage, prices are obtained through the estimation of a hedonic price function, such as $C_{ij} = C_i(Z_j)$, which is used in the regression equation relating travel costs from the origin and site characteristics. Hence, the price of each site is decomposed into a set of implicit prices, one for each attribute, and each origin will be associated with a different vector of price coefficients. If λ is marginal utility of income, the marginal implicit cost of the ℓ^{th} attribute is given by the first order condition:

$$\frac{\partial U}{\partial Z_\ell}\lambda^{-1} = \frac{\partial C_i}{\partial Z_\ell} \, . \qquad (4.21)$$

Utility is maximized when the individual chooses a recreational site so that the marginal value of each attribute equals the implicit marginal cost of experiencing it. This marginal cost is the hedonic price which varies across zones of origin for a given attribute (Hanley and Spash, 1993: 92). From all the available recreational sites, consumers will choose to visit only those which lie along their hedonic cost frontier. If the price gradient remains unchanged, the hedonic price function is a measure of the value of the change in the quality of a single site.

In the second stage, a demand curve for each attribute is estimated by regressing site attributes against its own implicit prices (estimated in the first stage) and other related variables (e.g. socio-demographic) from each origin zone. If $MC_\ell = \partial C / \partial Z_\ell$, and S_i is the vector of socio-demographic variables, the L demand functions are given by[24]:

$$\begin{array}{l} Z_1 = f\left(MC_1, S_i\right) \\ (\cdots) \qquad\qquad . \\ Z_L = f\left(MC_L, S_i\right) \end{array} \qquad (4.22)$$

These demand functions can be used to calculate the consumer surplus associated with each attribute. Using these functions it is also possible to evaluate the change in welfare associated with the complete elimination of an attribute at all sites.

If the price gradient remains unaffected, the welfare change associated with modifications in an attribute over the whole system can be computed in the same way. On the other hand, if a change in only one attribute at a specific site alters the price gradient, the

---

[24] For example Brown and Mendelsohn (1984) and Bockstael *et al.* (1989b: 169) estimated the inverse demand function.

welfare change will be much more difficult to estimate. A new functional relationship between travel cost and site attributes must be defined reflecting the choices made after the change. When it is not possible to know the sites that would be chosen after the change, the HTCM cannot be used to predict the effects of actions or events that alter the cost of attributes (Bockstael and McConnell, 1999).

The HTCM borrows some aspects from the hedonic price method applied to private markets (mainly to property value and wages), but the framework is not equivalent. The hedonic price function describes the set of prices that ensure the market equilibrium. Yet, in the recreation context there is no market mechanism ensuring that the price of each attribute corresponds to the equilibrium (Freeman, 2003b; Phaneuf and Smith, 2005: 715). In the HTCM it is chance or public intervention which provides sites and attributes, and not the market. In private markets more of some desirable attribute means higher prices, in travelling for recreation this is not necessarily true. Furthermore, a problem with the HTCM is the possibility of estimating negative hedonic prices (Smith and Kaoru, 1987).

4.3.2.4. Random Utility Model

The application of RUMs to recreation is a step further in the successive attempts to develop a model which accounts for the choice among multiple substitute sites as a function of the costs and characteristics of the alternatives. In the TCM context, the RUM seeks to explain individual’s discrete choice of a site for a recreational trip, from a set of many mutually exclusive alternative sites. Each site has a price (travel cost) and is described as a bundle of attributes, such as environmental quality and congestion. The selection of any given site depends on the attributes of that site as well as on those of all other sites in the choice set. The choice indirectly reveals how people trade off one site characteristic for another.

In each choice occasion, recreationists must compare a set of K sites and choose to visit the site j which maximizes their random utility. The optimal solution corresponds to an extreme corner solution. The utility associated with each alternative j may be express as:

$$U_{ij} = V_{ij}\left(\left(M_i - C_{ij}\right), Z_j, S_i\right) + \varepsilon_{ij}, \quad i = 1,\ldots,I; \quad j = 1,\ldots,K; \tag{4.23}$$

where, Z_j is the vector of attributes that describes a recreational site j; S_i is a vector of individual-specific or shift variables; C_{ij} is the travel cost borne by individual i to access the recreational site j; M_i is the consumer's income available for recreation when the choice is made. Accordingly, V_{ij} is the deterministic part and ε_{ij} is the random component which occurs due to the omission of explanatory variables, random preferences and errors in measurement.

Consumers will choose the site that maximizes their random utility, hence site j will be chosen if:

$$U_{ij} > U_{ik} \quad \forall k \neq j. \tag{4.24}$$

The conditional probability that individual i, who has decided to visit some site on a specific occasion, will choose site j is given by:

$$\Pr_i\left(choose\, j\right) = \pi_{ij} = \Pr ob\left[U_{ij} > U_{ik}, \quad \forall k \neq j\right]. \tag{4.25}$$

When errors are assumed to be independent and identically distributed (iid), following a type I extreme value distribution, a multinomial logit[25] model will give π_{ij} :

$$\pi_{ij} = \frac{\exp\left(V_{ij}\left(\left(M_i - C_{ij}\right), Z_j, S_i\right)\right)}{\sum_{j=1}^{K} \exp\left(V_{ij}\left(\left(M_i - C_{ij}\right), Z_j, S_i\right)\right)}. \tag{4.26}$$

The associated log-likelihood function maximized across N individuals and K sites is:

---

[25] These models have also been named Logit models due to the assumption concerning the error distribution (Kling, 1988).

$$\ln L = \sum_{i=1}^{I} \sum_{j=1}^{K} d_{ij} \ln(\pi_{ij}), \tag{4.27}$$

where d_{ij} equals one if site *j* is chosen by individual *i*, and zero otherwise. Once the functional form for V_{ij} is specified, its parameters can be estimated using maximum likelihood methods. Different functional forms can be considered for the utility function (Pendleton, 1999).

RUMs have been used not only to value changes in site attributes, but also to value economic losses associated with the loss of access to a site (Montgomery and Needelman, 1997; Grijalva, 2000; Parsons *et al.*, 2009). Assuming that the marginal utility of income is constant and that there is a change in an attribute of *j*, the CS may be expressed as:

$$V_{ij}\left(\left(M_i - C_{ij} - CS\right), Z_j^1, S_i\right) + \varepsilon_{ij} = V_{ij}\left(\left(M_i - C_{ij}\right), Z_j^0, S_i\right) + \varepsilon_{ij} . \tag{4.28}$$

Assuming that β_M is the coefficient on income, the CS resulting from a change in an attribute is given by:

$$CS_{ij} = \frac{\ln \sum_{j=1}^{K} \exp\left(V_{ij}\left(\left(M_i - C_{ij}\right), Z_j^1, S_i\right)\right) - \ln \sum_{j=1}^{K} \exp\left(V_{ij}\left(\left(M_i - C_{ij}\right), Z_j^0, S_i\right)\right)}{\beta_M} . \tag{4.29}$$

And the CS resulting from the loss of access to a site is given by:

$$CS_{i,K-1} = \frac{\ln \sum_{j=1}^{K} \exp\left(V_{ij}\left(\left(M_i - C_{ij}\right), Z_j, S_i\right)\right) - \ln \sum_{j=1}^{K-1} \exp\left(V_{ij}\left(\left(M_i - C_{ij}\right), Z_j, S_i\right)\right)}{\beta_M} . \tag{4.30}$$

Like the varying parameter model and the HTCM, the RUM considers a set of substitute recreation sites, which differ in specific attributes. Indeed, RUMs belong to the same class of models as the HTCM which consider environmental quality as a vector and where each element of the vector is a realization of quality at a different location (Bockstael and McConnell, 1999). Concerning time, the RUM departs from the standard framework since decisions are considered to be taken on a particular choice occasion, independent of the other occasions, and not during a period of time (Smith, 1989).

An advantage of RUMs over the remaining versions is their ability to deal with substitution in consumption and non-market quality attributes in ways that offer measures of the Hicksian welfare change (Phaneuf and Smith, 2005: 692). However, it suffers from three important limitations. Firstly, it is not appropriate for the analysis of attribute changes which alter the total number of recreational trips instead of re-allocating trips across sites (Buchli, 2004: 35). Secondly, the assumption that each trip occasion is a decision independent of past or planned visits is not completely plausible (Caulkins *et al.*, 1986). Thirdly, in RUMs the ability to explain the choice among sites as a function of their attributes comes at the expense of the inability to predict the total number of recreational trips (Freeman, 2003b: 433). This limitation has been overcome by considering it as separate problem or by extending the discrete choice framework. Two main approaches have been used in dealing simultaneously with the random utility framework and seasonal participation: the linked model and the repeated choice model.

The linked model has been used in the estimation of the participation and site-choice decisions, using two stages in the estimation (see, e.g., Kling, 1988; Parsons *et al.*, 2009; Bestard and Font, 2010). The first stage refers to the site selection component. The choice of the site is explained on the basis of site attributes, individual characteristics and travel cost, applying a RUM framework. In the second stage, the trip frequency is modelled by means of a separate participation equation using, e.g., count data models.

In repeated choice models, trips are explained as the outcome of repeated site choice decisions and not making a trip must also be considered as an alternative. The three-level nested logit model of participation and site choice proposed by Morey (1993) is among the seminal works. In this model, in each of the T periods the individual decides whether or not to go, chooses the region to visit and the site j of that region that provides the greatest utility.

The Kuhn-Tucker demand system is an alternative way of modelling the choice among substitute sites based on consumer's choice problems, while at the same time allowing for corner solutions. In this framework, individuals simultaneously decide which sites to visit and how many trips to make to each site over the course of a season. Behaviour regarding site selection and trip frequency stems from a single utility maximization problem; hence there is a utility-based theoretic link between the choice of a recreation site and its attributes (Whitehead *et al.*, 2010: 99). In this model integrability conditions are usually imposed on the functional form of the demand system. As recognized by its authors, the main disadvantage of this model is that it is computationally demanding (Phaneuf and Smith, 2005).

The discussion provided in this section is summarized in Table 3.4, presented in the preceding chapter. The table provides a summary of the main TCM versions but does not exhaust all the possibilities. Hybrid models, which contain features of more than one of the models analysed above, have been as well applied in empirical analysis. Models combining the best features of the zonal and individual observations are classic examples (Ward and Loomis, 1986: 169). In this line, Brown *et al.* (1980) maintained individual observations but divided each individual's trip by his share his zone's population. More recently, the low individual visitation rate lead Grossmann (2011) to adopt a similar procedure. Its starting point was data on individual observations collected on-site. The dependent variable was the aggregate demand by zone *j*, which was computed as the number of potential users times the demand by a representative individual.

4.4. MAIN ISSUES IN THE TCM

There are a set of methodological issues which are inherent to most of the TCM versions covered above. The difficulties are mostly related to: i) the valuation of time devoted to recreation, which is likely to be an important fraction of the travel cost; ii) the treatment of substitute sites; iii) the adjustment of costs for multiple purpose/destination trips. These issues have been addressed in the literature at a theoretical and practical level. An overview of these issues is provided in this sub-section (see, e.g., Bishop and Heberlein, 1979; Randall, 1994; Ward and Beal, 2000: 36-48).

4.4.1. TIME

Economists have long recognized that time is a scarce resource which plays an important role in the demand for outdoor recreation (Cesario, 1976). When time is allocated to outdoor recreational activities some other end will not be accomplished and some opportunity cost is likely to be involved. The proper measure to evaluate this opportunity cost, the parcels of time to be included in the model and the econometric difficulties involved are some of the issues in the long-running debate which has been a feature of TCM analysis since its earliest developments.

One of the sources of difficulties is the fact that two different time horizons are involved in the decision-making process. Decisions about recreation are mainly short-run choices conditioned on long-run labour supply decisions. Furthermore, the appropriate value of time depends on complex institutional, social and economic relations (McConnell and Strand, 1981; Feather and Shaw, 1999).

In the more common framework, the total time available (H) is considered to have two competing allocations. It may be spent working (h_w) or in recreational activities (h_r), such that the time constraint is given by[26]:

$$H = h_w + h_r \,. \qquad (4.30)$$

The corresponding budget constraint is expressed as:

$$M = w \times h_w + \bar{M} \,, \qquad (4.31)$$

---

[26] Palmquist *et al.* (2010) developed a model that included a third component – household maintenance – which competes with recreational time.

where w is the wage rate *per* time period, $\bar{M}$ is exogenous income and M is total income. The optimal solution is obtained by solving a utility maximization problem subject to time and budget constraints. The basic idea is that time spent on recreation activities cannot be spent working, thus it is likely to have an opportunity cost. This cost will depend on an individual's labour market situation (Freeman, 2003b: 420).

Employed workers are considered to be in one of two situations. They may face a "take it or leave it" situation, in which they must work during a fixed number of hours. Or, they may have a flexible work schedule where working hours may be adjusted until equilibrium is reached. For people facing a flexible work schedule, information contained in the two constraints can be collapsed into one and the marginal wage rate will be the correct measure for the opportunity cost of time. These individuals will adjust working hours until the wage rate equals the marginal value of time. The optimal solution is an interior solution. For people over or under employed and facing a fixed work schedule the two constraints cannot be collapsed. Workers are in a corner solution where the wage rate is not the correct measure for the opportunity cost of time. It is a non-observable parameter (Bockstael *et al.*, 1987b; McKean *et al.*, 1995; Feather and Shaw, 1999).

Some practical difficulties arise when applying this theoretical structure. On the one hand, it is very hard to classify a worker's situation. On the other hand, even when this is possible, it is not clear what marginal wage rate to consider for those in a corner solution. Furthermore, the valuation of the recreational time of students, unemployed and retired workers is not possible in this context.

The estimation of hedonic wage rates using survey data, or data from a secondary data source (Hynes *et al.*, 2009), and the computation of shadow wages rates (Feather and Shaw, 1999) are some of the attempts that have been made to improve time valuation and the accuracy of consumer surplus estimates. However, due to the numerous difficulties in obtaining a reliable measure for the wage rate and based on earlier studies, such as Cesario (1976), in the majority of the studies some proportion of the respondent's wage rate is considered as a *proxy* for the opportunity cost of time. That proportion varies between one-quarter and the full wage rate. The use of one third of the wage rate seems to be the most widely accepted measure (see, e.g., Englin and Cameron, 1996; Chakraborty and Keith, 2000; Hanley *et al.*, 2001a; Bergstrom *et al.*, 2004; Anderson, 2010, among others). It is also a common procedure to test several percentages and choose the “best one” in accordance with statistical results.

The evaluation of the different parts which make up recreational time is another matter for discussion. Time spent on recreation activities may be divided into two additive parts, travel time (h_{rt}) and on-site time (h_{rs}), thus:

$$h_t = h_{rt} + h_{rs} \,. \tag{4.32}$$

Typically, the two fractions have been subject to different treatments. While travel time is always considered, on-site time is very often ignored. There are at least three main reasons for this. First, the estimation is simpler if it is assumed that on-site time and the opportunity costs of time are the same for all individuals. These were the assumptions made in the earlier studies where differences in travel costs were explained by out-of-pocket expenses and by the differences in travel time (McConnell, 1992). Furthermore, when secondary data is used, it is a necessary assumption due to the lack of information. Second, on-site time is both a choice variable and a component of the trip cost, while travel time is only a component of the cost (Smith *et al.*, 1983). Third, whereas most authors consider that travel time represents only a cost, it is of general agreement that on-site time includes two contradictory effects, a benefit and a cost. This latter reason is related to two arguments favouring the inclusion of on-site time. Concerning the cost perspective, travel and on-site time are allocations of the scarce resource and for this reason both should be included in the time constraint (McConnell, 1975; Bockstael *et al.*, 1987b: 294). In accordance with McKean *et al.* (1996), from the benefit perspective, on-site time is essential because of the principle of weak complementarity upon which measurement of benefits from the TCM is founded. Without on-site time there is no utility.

Some claim that travel time may also be associated with the two contradictory effects: the opportunity cost of time and the recreation benefit. The trip may be perceived as pleasant, for example, when it is made along a scenic route or the recreationist likes to drive. In view of the fact that time spent on a trip may be a benefit or a cost and especially because often work time is not freely adjusted by workers, some authors prefer not to consider the cost of time (e.g., Beal, 1995; Smailes and Smith, 2001; Bhat, 2003; Hanley *et al.*, 2003; Hynes and Hanley, 2006; Alberini *et al.*, 2007; Fleming and Bowden, 2009). Nevertheless, ignoring the opportunity cost of time is likely to result in a downward bias in the estimate of consumer surplus, which should therefore be regarded as a lower bound. The cost is an opportunity cost, such as that incurred with travel time. The benefit accrues from the recreation experience. Hence, time on-site is likely to be an argument in the utility function. Ward and Beal (2000: 39) consider that it must be presumed that the benefits of on-site time are at least equal to the time cost, and probably exceed it.

When the cost of time is included, typically the total travel cost is the sum of two additive parts, the direct and indirect out-of-pocket expenses plus the opportunity cost of time. The two components of cost may instead be included in the model separately and thus two coefficients are estimated (Brown and Nawas, 1973; Gum and Martin, 1975). Shaw and Feather (1999) also favour this practice, especially if workers are in a corner solution in the labour market. However, this alternative has seldom been applied because both variables are calculated as functions of distance, creating multicollinearity problems (Cesario, 1976; Ribaudo and Epp, 1984; Bockstael *et al.*, 1987b; Englin and Shonkwiler, 1995). Due to the data aggregation, in the ZTCM multicollinearity problems seem to be even stronger (Brown and Nawas, 1973).

In order to avoid the imposition of an opportunity cost of time, in some works travel time, and not the cost of travel, enters into the model as an independent variable. This procedure, which follows the pioneering work of McConnell and Strand (1981), allows the cost to be established from the data. Empirical works produced mixed results (see, e.g., Fix and Loomis, 1998; Loomis *et al.*, 2000; and Azevedo *et al.*, 2003).

Another alternative in terms of econometric modelling, especially when the ITCM is used, is to include on-site time separately from the remaining costs. If spending more time on-site positively affects the value of the trip and increases the number of visits, the sign will be positive. Conversely, if on-site time is substituted for trips, the sign will be negative. The nature of this relation remains an empirical question (McKean *et al.*, 1996).

The knowledge that the estimates of consumer surplus are very sensitive to time valuation continues to stimulate interesting discussions about how to value recreational time and how to incorporate it into empirical applications. There are a variety of approaches to handing time costs, and no established approach, and researchers continue to "improvise". What seems clear is the need to take into account both components of time in the analysis.

4.4.2. SUBSTITUTE SITES

Economic theory suggests the price of related goods, whether complements or substitutes, is an important determinant of demand. There is no reason to believe that demand for outdoor recreation should be modelled differently. Thus, when modelling behaviour regarding the demand for recreational trips, the price of substitute sites should be included. The HTCM and RUMs deal explicitly with the choice between similar sites and substitutes need to be identified. In these models, the choice of the set of sites to be included is not straightforward and involves some degree of subjectivity, but the difficulty is faced *a priori* by the researcher. Hence, our focus is on difficulties in the identification of substitutes noted in the ITCM framework.

There are a number of practical difficulties that preclude or make it hard to include the price of substitutes in the ITCM analysis (Bowker *et al.*, 1996). First, different preferences and possible substantial differences in the access to substitutes make it difficult to identify and include all the potential alternative sites. Second, the multicollinearity usually verified between the own price and substitute prices may compel the researcher to drop the substitute site price (Bergstrom *et al.*, 2004). These difficulties lead to the omission of this explanatory variable in much applied work (e.g., Englin and Cameron, 1996; Hesseln *et al.*, 2004; Hynes and Hanley, 2006). However, the omission of substitute prices is likely to have some undesirable consequences, namely, specification bias and a decrease in the variables' explanatory power (Caulkins *et al.*, 1986). For example, Rosenthal (1987) reported a bias in the own price coefficient due to the omission of substitute prices. This caused a reduction in the price elasticity of demand, an increase in the consumer surplus and a reduction in the explanatory power of the models. In response to these difficulties, several alternatives had been experimented with in empirical work.

One commonly used procedure is the identification by the researcher of one or more sites that he considers to be substitutes for the recreational site being studied (Fix and Loomis, 1998; Whitehead *et al.*, 2000; Alberini *et al.*, 2007). Another option is to ask the respondent to identify the substitute. This may be, for example: a site yielding comparable attributes (Rosenberger and Loomis, 1999); the closest to the home residence from the set of comparable recreation sites (Anderson, 2010); or the site most often visited (Liston-Heyes and Heyes, 1999). A third option is to consider the site nearest to the destination and which has similar characteristics (Hellerstein, 1993). In any case, the distance from respondent's home to each of the places is used in designing the *proxy* for substitute prices. Unavoidably, some degree of subjectivity always remains and in the ITCM, when data refers to a large set of substitute sites, the inclusion of the price of all of them is not possible (Creel and Loomis, 1990).

4.4.3. TRIP

A standard assumption in the TCM framework is that the travel cost is incurred to visit one recreational site exclusively. When this assumption is verified, all the travel costs are incurred to access a specific recreational site and activity and cost imputation is straightforward. However, often a trip has more than one destination and/or purpose. For trips with multiple destinations, travel cost represents joint costs. In empirical analyes several options have been suggested to deal with this issue, but there is no established theoretical procedure.

A common empirical practice has been to exclude visitors engaged in multiple destination trips. However, these visits are likely to have a high use value. Hence, benefits *per* person would be overestimated and total benefits underestimated (Fleming and Cook, 2008). Furthermore, excluding multiple purpose and multiple destination visitors may bias the sample significantly, especially in terms of the socio-demographic characteristics (Martínez-Espiñeira and Amoako-Tuffour, 2009). On the other hand, treating all the observations in the same way is likely to overestimate the benefits and so the economic value of the site (Haspel and Johnson, 1982). Because of the problems related to this more drastic option, some other alternatives have been proposed and discussed.

One option is to use the cost share of each destination. This can be computed by directly asking visitors about: a) the influence of the site on the decision to take the trip (Martínez-Espiñeira and Amoako-Tuffour, 2009); b) the length of the stay at each destination (Yeh *et al.*, 2006); c) the number of sites visited (Beal, 1995: 297); d) the proportion of the travel cost related to each destination (Loomis *et al.*, 2000: 184). Haspel and Johnson (1982) suggested a fifth alternative: an itinerary data approach that involved the assignment of the total travel cost to each destination in accordance with the marginal distance travelled from the previous site visited. A related issue concerns the treatment of incidental visits/activities (Parsons and Wilson, 1997).

The treatment of accommodation and related costs in visits that encompass more that one day is another issue related to multiple destination and purpose visits. Although there is no uniform treatment, some authors excluded information regarding holiday makers. Hanley and Spash (1993: 88) suggest including these visitors in the aggregation, but using the imputed costs of day-trippers. Another option is to consider only the daily travel costs from the place of accommodation to the recreation site. But this option is likely to ignore significant costs incurred by those travelling significant distances and to provide excessively conservative estimates (Liston-Heyes and Heyes, 1999).

4.4.4. OTHER ISSUES

Conventional TCM practice treats the distance from home to the recreation site as an exogenous variable. However, recreation preferences may have some influence on the choice of the residential location. Households may choose to live near the site to avoid time and transport costs and this information is usually not included in the model.

Others gaps in the TCM framework have received less attention. According to one of the basic premises, admission fees are usually treated by researchers in the same way as the travel expenditures, but this may not be the recreationist’s perception. The existence of substantial differences in the recreationist’ tastes/preferences and in income levels at varying distances from the recreation site are potential sources of bias which increase the complexity of the analysis. Furthermore, the number of other visitors simultaneously visiting the recreational site is likely to affect the visit quality. However, not much attention has been devoted to the analysis of the effects of congestion/crowding on demand and on consumer surplus.

4.5. CONCLUDING REMARKS

This chapter presented an overview of the TCM following two complementary routes. The first described the main variants of the model, providing a chronological overview. The main objectives and shortcomings of each alternative were emphasized. The second discussed the main empirical issues that researchers must deal with.

The TCM is the most widely used RP methodology for evaluating the benefits users derive from outdoor recreation. The initial roots and the earlier versions and refinements of the method were developed in the USA, many of them in the context of recreation studies sponsored by government authorities. The predominance of water-based recreation activities in the earlier empirical analyses is probably due to this influence of the American public authorities. The method is now applied in many developed and developing countries and to a wide range of recreational areas.

Nowadays, the RUM, for multiple site models, and the ITCM with count data models, for a single site model, seem to be dominant in the literature. RUMs have the virtue of directly producing estimates of the Hicksian welfare measures, while in the remaining approaches the measure directly obtained is the MCS. In spite of its high degree of complexity, the Kuhn-Tucker demand system model seems to be a promising option due to its ability to simultaneously deal with the choice among various sites and the number of visits using a framework that is consistent with choice theory.

Research on the subject is expected to continue to intensify as none of the approaches is entirely satisfactory. One of the most recent lines of research is the joint use of the TCM with a SP technique. Several studies combining the TCM with CB or CVM have been applied more recently. A summary of the relevant literature that focuses on this type of technique combination to analyze outdoor recreation is presented in Chapter 8, where an empirical analysis combining TCM and CB is offered.

PART II: EMPIRICAL ANALYSIS

CHAPTER 5 – NON-MARKET RECREATIONAL VALUE OF A NATIONAL FOREST: SURVEY DESIGN AND RESULTS[27]

5.1. INTRODUCTION

Several developed countries have been collecting outdoor recreation data in the context of national recreation surveys. USA[28] is pioneer in this task, but Canada and several European countries, such as Denmark, Finland, Holland and United Kingdom have also been conducting large scale surveys. In a few countries, visitation monitoring systems to (semi)natural areas have also been implemented (Bell *et al.*, 2007: 13). Datasets resulting from these efforts provide valuable information for people involved in the management of these areas. Moreover, analysis of these data reveals that demand for outdoor recreation in general is rising and projections point towards continuing increase (Bell *et al.*, 2007: 35; Carpio *et al.*, 2008: 430).

In Portugal, it is probable that outdoor recreation demand follows a similar trend. Nevertheless, as far as we know no national survey has been conducted and no data is currently available concerning recreation in Portuguese (semi)natural areas. Our research contributes to partially fill this gap. A questionnaire focussed on forest recreation demand was developed and administered. The questionnaire brings together data on trips' characteristics, socio-demographic features, perceptions and attitudinal data and intended behaviour in hypothetical circumstances. The Bussaco National Forest is used as case study, but the approach and the conclusions can be further extended to other similar resources.

---

[27] This Chapter is available as the Discussion Paper number 09-2012 by GEMF, FEUC (Simões *et al.*, 2012).

[28] Phaneuf and Smith (2005) mention three illustrative examples of the work made and people involved: i) the Public Area Recreation Visitors Survey coordinated by H. Cordell of the US Forest Service to collect on-site recreation surveys for forest service areas and in coordination with NOAA, for beaches around the US; ii) V. Leeworthy has been especially active in developing well-documented recreation databases relevant to NOAA's activities; and iii) D. Hellerstein has focused efforts at the Economic Research Service on related activities for recreation sites relevant to agricultural policy.

This chapter is organized as follows. In Section 5.2 the study area is described. Section 5.3 explains the main elements of the survey. It includes the questionnaire description and administration, the explanation of the objectives behind each question and a brief presentation of the descriptive statistics. Section 5.4 provides the concluding remarks.

5.2. CASE STUDY AREA: BUSSACO NATIONAL FOREST

Bussaco is an ancient national forest. It covers an area of 105 hectares surrounded by a wall with 5 300 meters of extension and an average of 2.5 meters high. It is located in Baixo Vouga in the Centro Region of Portugal. The nearest medium size city is Coimbra, at about 30 kilometres (km) distance. The forest has many walking trails, fountains and pools, several picnic areas and also religious heritage. In the centre of the forest there is a hotel working in an old king's summer palace built in the beginning of the 20th century (Santos, 2002).

The forest was first settled by Benedictine monks in the 6th century. Until the 16th century arborisation was made with autochthones species. In the following century, it was donated to an Order of barefooted monks, the Discalced Carmelites, who established here their sacred wilderness in Portugal. This Order built a monastery and the wall surrounding the forest. Since the 19th century it has been managed by the Portuguese central government.

Over the centuries, monks and governmental forest authorities have planted many varieties of trees, shrubs and flowers. This continuous action resulted in the construction of a rich and diversified natural heritage, which is now considered to be one of the best dendrological collections in Europe. This natural heritage boats exotic species brought from all over the world standing alongside natives species. The Mexican Cedar (*Cupressus Lusitanica*) is probably the most emblematic one.

The forest of Bussaco is now a semi-natural public area used mostly for recreation purposes. This area is demanded mainly because of its unique natural characteristics. Visitors come all year round, although the vast majority of visits are made in spring and summer when the weather is warmer and drier. Currently, motorized vehicles must pay an entrance fee during the lighting hours. In the spring and summer this period it goes from 8a.m to 8p.m. In the autumn and winter period it goes from 8a.m to 6p.m.

Several reasons justify the choice of this place for this research economic assessment. Firstly, it is a space biologically rich. Secondly, it is far enough from any city so it cannot play the role of a city park. Thus, the majority of the visitants cannot get there walking[29]. Thirdly, it attracts mainly day-trip visitors; hence the basic assumption of single purpose recreational trip is more likely to be verified. Fourthly, it is too far from the coast line so that it can be visited in the very warm hours of the summer days when it is not pleasant to be on the beach, as it happens with other Portuguese national forests. Fifthly, it is not a very large resource as it is the case of national and natural parks and estuaries and there are not similar resources in proximity. To sum up, this space can be considered a well-defined entity (Smith and Kaoru, 1990). Furthermore, we are not aware of any TCM study applied to forest recreation in Portugal.

5.3. THE SURVEY

5.3.1. QUESTIONNAIRE ADMINISTRATION

Three main traditional channels have been used in the questionnaire administration: mail, telephone and in-person (typically on respondents' home or on-site). The internet is a fourth channel which use is more recent but is growing rapidly. Each of these survey modes embraces advantages and disadvantages (see, e.g., Champ, 2003, for a discussion).

In the context of our research, the most feasible option was the on-site questionnaire administration. A number of reasons justify this choice. Firstly, the questionnaire is excessively detailed to be administered by telephone. Secondly, administration by mail or

---

[29] As pointed by Bishop (1992), it would make impracticable the application of the TCM.

in-person interviews on respondents' home comprising a representative sample of the general population would be prohibitively expensive. Thirdly, the administration of non-market valuation questionnaires via the internet is still in its infancy and offers no guaranties of representative samples[30] (Lindhjen and Navrud, 2011). Finally, although on-site sampling does not allow the data to be drawn from a random sample of the general population, this sample scheme has the advantage of hitting the target population directly (Parsons, 2003: 276)[31].

Visitors were thus approached at the three main entrances, asked to complete the questionnaire by the end of the visit and to return it before leaving the forest. The use of reminders was not possible as no individual information was requested when the questionnaire was handed over. As a result, the questionnaire was self-administered. It is known that in-person interviews may cause some undesirable interviewer effects because commonly respondents try to be pleasant and socially correct (Loureiro and Lotade, 2005). In addition, in self-administered questionnaires respondents are likely to use more time to think about their preferences and to remember past behaviour because there is no one waiting for an immediate answer (Carlsson, 2010: 174). Furthermore, it has been argued that if questionnaires are answered in respondent's own environment and time, the sense of anonymity is reinforced and the reflection is favoured (Ward and Beal, 2000: 162).

Two groups of visitors were deliberately not surveyed: foreign visitors and people participating in commercial organized excursions. Concerning the first group, it is not realistic to consider that these visitors are likely to make the trip since home to visit only Bussaco forest. Cost imputation in these circumstances is particularly troublesome. The second group was left out because typically many sites are visited in the same trip and

---

[30] When the survey began, the site on the internet of the *Fundação Mata do Bussaco* (www.fmb.pt) was not available yet. This foundation was created by the Portuguese government in May 2009 and manages the forest since then.

[31] Due to similar reasons, this sample scheme has been followed by numerous researchers (e.g., Heberling and Templeton, 2009; Martínez-Espiñeira and Amoako-Tuffour, 2009; Lienhoop and Ansmann, 2011).

time spent in each site is pre-defined by the trip organizer. Consequently, data on these visitors could hardly be used in the TCM analysis.

A pilot survey to ensure clarity and ease of answering was conducted in June 2010 using face-to-face, intercept interviews. Following the pilot survey, few corrections were introduced. A total of 1 055 questionnaires were distributed from July 2010 to June 2011. Of those, 311 were returned, corresponding to a response rate of 29.5%. This is a low response rate typical for long questionnaires (Martínez-Espiñeira and Amoako-Tuffour, 2008). It is, however, similar to the one obtained by Mendes (1996) who used a analogous administration procedure. The in-person administration of surveys is likely to provide higher responses rates but is not exempt of difficulties. For example, in Madureira *et al.* (2011) 40% of the answers were classified as protests and were excluded from the econometric analysis.

Concerning this research sample, 35 questionnaires had to be excluded following the conclusion that critical questions were not answered (e.g., the number of trips or the home zip code) or the respondent was younger than 16. Additionally, some questionnaires considered were not fully completed, so the total number of answers is not the same for all the questions and sections. It varies between a minimum of 212 observations regarding trip expenses and a maximum of 276.

According to media, the forest received around 200 000 visitors in 2011, half of them being foreign visitors[32]. Therefore, the population of interest is of about 100 000 individuals and can be classified as a large population. According to Champ (2003: 68), for a population above 10 000 individuals, a sample size of around 380 is needed, so that the sample error is of ±5%. The number of usable questionnaires we could obtain is below this number and imposes a larger margin of error. Even though below the desirable, our sample size is comparable to other studies (Hynes and Hanley, 2006, n=144; Alberini *et*

---

[32] http://ecosfera.publico.pt/ (02/01/2011)

al., 2007, n=269; Shrestha *et al.*, 2007, n=263; Donovan and Champ, 2009, n=143; Lienhoop and Ansmann, 2011, n=187).

Two important factors determining the sample size are the level of precision (or sampling error) and the statistical power required (the probability of correctly reject the null hypothesis when it is indeed false). The level of precision is measured by the statistical significance level (the probability of wrongly reject the null hypothesis when it should be accepted) and by the confidence interval (the interval where the mean of the population is with a given confidence level). Although the relation is not linear, as the sample size increases, the statistical power improves.

5.3.2. DATA

When the researcher intends to apply the individual version of the TCM in the econometric analysis, the questionnaire must include a set of mandatory questions (e.g., regarding socio-demographic data and trips characteristics). Other questions, which could be left out without compromising the TCM application provide important information likely to be useful not only to site managers but also to the decision makers responsible for the management of others (semi)natural areas. These questions can also produce complementary data which assist modelling individual preferences. The two sets of questions were included.

The questionnaire (available in Appendix A) comprised thirty three questions structured into four sections plus an introductory text. The introductory text explained the questionnaire´s objective and scope, ensured anonymity and confidentiality of individual data and provided instructions for completion. Section I – *Perceptions and preferences* – was meant to collect data on respondents’ perceptions about environmental protection in Portugal and their recreational preferences. Section II – *Bussaco forest* – collected data regarding the number of visits, the main motives justifying visits, the perceptions about the forest and about its current conservation conditions. Section III – *Trip and related* *expenses* – was about the characteristics of the trip undertaken to visit Bussaco and to respondents' behaviour in hypothetical circumstances. Finally, Section IV – *Personal data* – was aimed to collect respondents' socio-demographic data. Accordingly, the usual structure was adopted. It began with unproblematic questions of a general character, while personal data and other questions potentially more intrusive were left for the last part. Questions were grouped in a manner that should make sense to the respondent and not attending to research aims. For this reason, the order followed here may differ from the order in which questions were presented in the questionnaire.

Data analysis begins with the socio-demographic characterization. Along this text, the number of the question is identified by using the abbreviation Q#.

5.3.2.1. Socio-demographic data

In order to carry out visitors' socio-demographic characterization, the questionnaire asked for respondent gender (Q23), age (Q24), formal education (Q25), number of persons in the household (Q27), number of persons contributing to the household income (Q28), household monthly income after taxes (Q29) and occupation (Q30).

Concerning respondents' gender, females represent 57% of the sample and males the remaining 43%. This structure shows a slightly higher percentage of females than at the national level, where females represent 52% of total population. The age structure is as follows: 16% of the respondents are between 15 and 25 years of age, 34% are in the range of 25 and 34 years, 40% belong to the range of 35 to 54 years and 10% are over 54 years. This distribution does not reflect the national one – in the sample the two younger cohorts are over represented and the eldest is under represented (INE, 2011a). The sample mean age is of 36.5 years old with a standard deviation of 12.2.

Regarding formal education, 3% of the respondents completed the elementary school or less, 3% the 2nd cycle, 10% completed the lower secondary school, 33% the upper secondary school, 34% have a bachelor or university degree and 17% have a post-graduation or a higher level of education. This distribution is very different from the Portuguese one given that only 11% of the population above 15 years old have a bachelor/university degree or higher (INE, 2011b). The sample mean number of years of formal education is 13.6, with a standard deviation of 3.6. The over-representation of persons with higher school levels has also been noted in other studies analysing recreation in (semi)natural areas (e.g., by Egan and Herriges, 2006: 575).

As Hynes *et al.* (2009), this research questionnaire asked for household disposable income and not for gross earnings because in allocation decisions people will consider what they really can afford to pay. Instead of an open question, several brackets to the monthly household disposable income were provided in order to motivate respondents' answer[33]. The two lower brackets account for 20% of the sample, the two middle brackets represent 40% and the two highest for the remaining 40%. The uppermost income category is the one counting the greatest percentage of respondents. It is probably related to the high representativeness of people with a university degree. Table 5.1 lists frequency and cumulative distribution in percentage.

Table 5.1: Household monthly disposable income

Categories (€)	Percentage	Cumulative
< 500	2.27	2.27
[500; 1 000[	17.42	19.70
[1 000; 1 500[	23.48	43.18
[1 500; 2 000[	16.67	59.85
[2 000; 2 500[	14.39	74.24
≥ 2 500	25.76	100.00

On average, households were composed by 3.1 persons and 2.0 persons contributed to the household income. Accordingly, the average monthly disposable income *per* worker in the sample was approximately €997, well above the €754 reported in national statistics (INE, 2011b). Concerning the labour situation, 76% of the respondents were employed

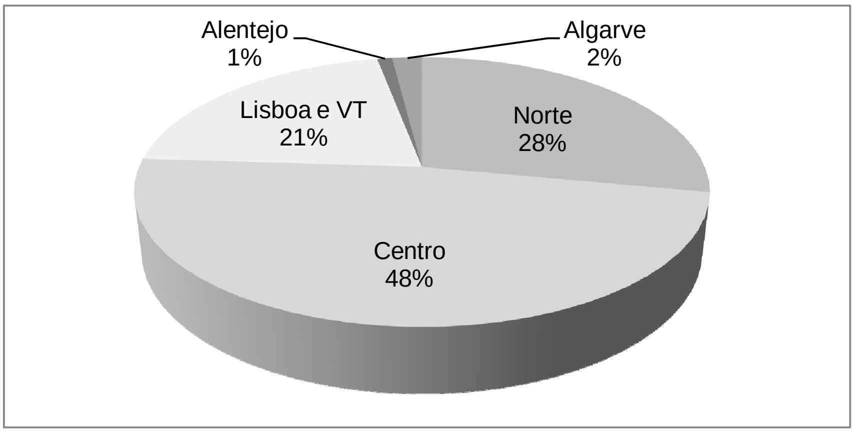
---

[33] Albeit the efforts, results are in line with preceding experiences since this question was answered in only 80% of the returned questionnaires.

workers, 14% were students, 6% were retired workers, 3% were unemployed workers and 1% were housekeepers. Retired and unemployed individuals are expected to have a lower opportunity cost of time because they have more leisure time available.

The distribution of respondents by geographical origin is displayed in Figure 5.1. Almost half of the respondents live in the Centro Region, where Bussaco forest is also located. Norte is the next most important respondents' origin, followed by Lisboa e Vale do Tejo. Algarve is the most distant region and Alentejo the less populated in mainland. These regions are the origins for 2% and 1% visitors, respectively.

Figure 5.1: Respondents' origin



5.3.2.2. Preferences and perceptions

Preferences and perceptions were scrutinized in Q1 to Q5. In Q1 a set of attitudinal questions about the environment were presented. Q2 and Q3 were aimed to characterize visitors' recreation preferences. In Q4, respondents were asked whether they were members of any environmental organization and in Q5 if they belong to any nature-based sports group.

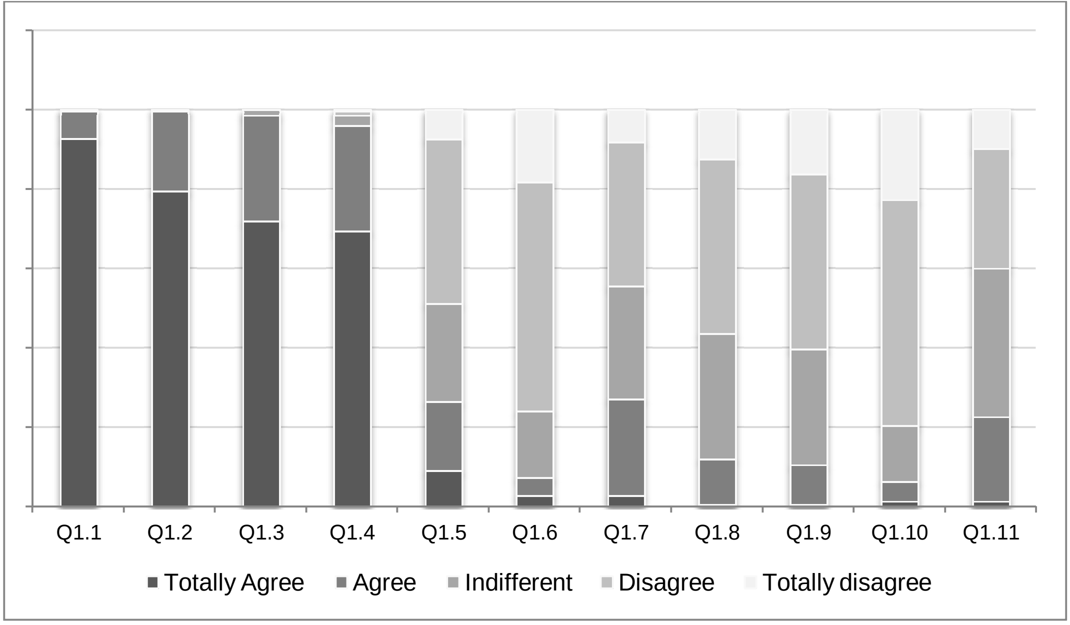
Q1 asked for respondents' opinion about the importance of: environmental protection, preservation of natural spaces in Portugal and supply of green public spaces in Portuguese cities. Eleven statements were presented and respondents should answer making use of a five point Likert scale, where a score of 1 means "*totally disagree*" and a score of 5 means "*totally agree*". Table 5.2 reports mean values, while Figure 5.2 illustrates the answers.

Table 5.2: Perceptions concerning environment

		Mean
1.	Environmental protection is very important	4.92
2.	As an individual, you can play a role in protecting the environment	4.79
3.	A good quality of life requires high environmental quality	4.70
4.	Environmental problems have a direct effect on your daily life	4.63
5.	Environmental issues are among the main concerns in Portugal	2.79
6.	Environmental protection measures adopted in Portugal are adequate	2.15
7.	Natural spaces in Portugal are sufficient	2.76
8.	Natural spaces in Portugal are well preserved	2.43
9.	Wildlife in Portugal is well preserved	2.34
10.	Green spaces in Portuguese cities are sufficient	2.04
11.	Green spaces in Portuguese cities offer good quality	2.73

Statements 1 to 4 were adapted from a questionnaire used by the Eurobarometer (2008). According to this report, environmental protection was considered to be "*very important*" by 67% and "*fairly important*" by 30% of the Portuguese respondents. Although the comparison is not straight, the results of the sample of Bussaco forest visitors seem even more meaningful. Indeed, 93% "*totally agree*" with the statement "*Environment protection is very important*" and the remaining 7% just "agree".

Figure 5.2: Environmental perceptions



A list of ten classes of recreational activities was presented in Q2. Respondents should signal the regularity of engagement in each one making use of a five levels' scale, where 1 means "*never*" and 5 denotes "*very often*". The answers led to conclude that recreation activities in which Bussaco visitors take part with greater regularity were: visits to the beach, visits to historic places, cinema sessions and the practice of open air sports.

In Q3 visitors were asked to point out by order of preference their three favourite recreational activities. The objective was to check out the position of the open air nature based recreation activities in preferences. The answers showed that open air sports, such as walking, cycling and jogging, were the most favourite ones. Visits to the beach and sports requiring specific infrastructures, such as swimming, football and tennis, occupied the second and third positions, respectively.

In addition to its general interest when environmental issues are under analysis, data on environmental protection association membership is usually collected in non-market valuation questionnaires as it captures involvement and interest in environmental issues (Garrod and Willis, 1992; Nunes, 2002b; Parsons, 2003: 279). In the sample, 4.0% of the respondents affirmed to belong to some environmental association, which is much above the 1.8% recorded at the national level (INE, 2011c), or the 0.1% indicated by the *Liga Protectora da Natureza*[34]. The percentage of respondents affirming to belong to a nature-based sports association is also low (only 2.2%), but we could not find national figures to compare with.

5.3.2.3. Perceptions about Bussaco forest

The perceptions about Bussaco forest were enquired in Q10 which was structured on the five point Likert scale used before in Q1. This question was divided in two parts. The first respected to the visitors general opinion of Bussaco features and to the current conservation conditions. It pursued four main objectives, namely: i) to ascertain whether

---

[34] *Liga Protectora da Natureza* is a Portuguese non-governmental association of environmental protection. This data is available in http://www.lpn.pt (03-05-2012).

visitors consider that there are Bussaco's substitutes at regional or national levels; ii) to assess the importance of different heritage dimensions (natural, constructed and religious); iii) to know the perceptions of recreationists regarding current conditions; and iv) to notice if recreationists believe that this space is under threat. The second part enquired whether visitation frequency would be affected if some proposed changes occur. The results are summarized in Table 5.3.

Table 5.3: Opinions about Bussaco

Bussaco forest...	**Mean**	**% of totally**	
		agree	**disagree**
Is an excellent recreational space	4.49	56	0
Is a unique space in the Centre Region	4.07	40	1
Is a unique space in the country	3.36	18	5
Has remarkable environmental conditions	4.07	26	0
Has remarkable architectonic heritage	4.02	26	0
Has remarkable religious heritage	3.27	10	5
Provides the necessary information to the visitors	2.72	2	15
Is overcrowded	2.31	2	20
Has good parking conditions	2.91	7	8
Is a safe place	3.83	16	0
Is currently threatened by forest fires	3.44	13	4
Is currently threatened by biotic agents	3.21	11	3
Is currently threatened by human action	3.39	12	4

We can conclude that Bussaco forest was regarded as a very good recreational space, but that the majority of the respondents did not considered it as a unique space at national or regional levels. Moreover, while the constructed dimension was sought as important, the natural was considered as the most remarkable one.

According to Bell *et al.* (2007: 22) "forests have a relatively large social capacity *per* hectare. Because of the screening effect of trees there can be many people present without the area feeling crowded". In fact, it seems that in visitors' perception Bussaco forest was not overcrowded and that parking conditions were acceptable.

Effects of fear and feeling unsafe have been reported as barriers to the use of some natural spaces to recreational activities (Bell *et al.*, 2007). This seems not to be the case in Bussaco since no one disagreed that this forest can be defined as a safe place. Regarding factors potentially threatening conservation conditions, visitors did not manifested a strong opinion.

In the second part of Q10, respondents showed relative indifference to the changes proposed as the mean value is very close to three in all of them. At the same time, the final open-question, "*other motives*" was chosen by several respondents who point out that they would visit Bussaco forest more often if they live closer to it. A closed-ended question was chosen believing that it would be less burdensome to the respondents (Champ, 2003: 82). Provided that the proposed changes were not meaningful, we consider now that it would have been a better option to use an open question, such as it was done, e.g., by Smailes and Smith (2001).

5.3.2.4. Visits

Data on the number and duration of visits and their motivations was gathered in Q6 through Q9, plus Q15. Q6 inquired whether the current visit was the first one ever made to Bussaco forest; Q7 asked for the number of trips in the last 3 years; and Q8 for their distribution among seasons (spring/summer versus autumn/winter). Q9 enquired about visit motivations and Q15 about their duration.

Regarding the trip frequency, when the individual version of TCM is to be applied, two main routes have been followed. When the number of trips *per* season/year is expected to be high, a season/year have been commonly considered as the reference period (e.g., when trips are related to fishing or hunting activities as in Alberini *et al.* (2007) and Shrestha *et al.* (2002)). On the other hand, whenever trips are expected to be less frequent, the choice of a larger period of time (e.g., several years) offers two advantages. First, some econometric problems can be avoided as the likelihood of observing a high proportion of "one" is reduced[35]. Second, in a longer period demand is less sensitive to the specific conditions of a single season/year (Englin and Shonkwiler, 1995). For these reasons, e.g., Bhat (2003) and Martínez-Espiñeira and Amoako-Tuffour (2009) considered a five years period, while Chakraborty and Keith (2000) worked with a time horizon of ten years[36]. As the period length increases, the amount of data also increases. At the same time, it becomes harder for the respondents to report accurately the number of visits. Therefore, the period of reference chosen for this survey is of three years. A longer period is perhaps excessively demanding.

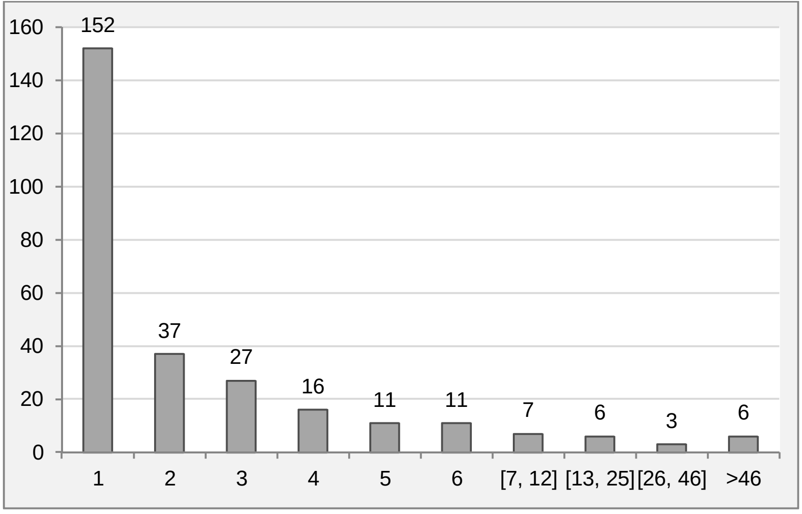
The total number of visits reported was 1 551, corresponding to an average of 5.6 during the three years and to 1.9 annual visits. Respondents making the first visit to Bussaco forest represent 34% of the sample and 6% of the visits. Excluding these novices, the average raises to 8.0 visits during the three years and to 2.7 annual visits. Assuming that the respondents' group in preceding visits was equal to the one in which the questionnaire was administered, the total number of visits by respondents' group along the three years was of 4 596. Answers to Q8 show that 61% of those visits were made during the spring/summer seasons. The average visits' duration was of three hours. Furthermore, the majority of the respondents (55%) made only one visit during the period. Two or three visits register the next high frequencies and together represent 23% of the sample. Four to six visits are as well frequent, totalizing 14%. Higher records are uncommon. Figure 5.3 displays visit counts.

---

[35] For example, Ribeiro (2002: 86) in the individual model excluded respondents reporting less than three visits.

[36] As alternative, Grossmann (2011) suggests the use of a variant of the zonal version when annual visitation rates are low, while Martínez-Espiñeira and Amoako-Tuffour (2008) considered the number of visits of respondents' group.

Figure 5.3: Visit counts



Q9 follows Shrestha *et al.* (2007), Hanley *et al.* (2003) and Willis and Garrod (1991). It was aimed to identify the main motives for visiting Bussaco, which are the determinants of its recreation use value. The answer to these questions was based on a five point scale, where 1 means "*not important at all*" and 5 means "*very important*". The most important motive indicated was "*contact with nature*", with an average value of 4.4. Table 5.4 presents mean values of the five main visit motives.

Table 5.4: Visits' purpose

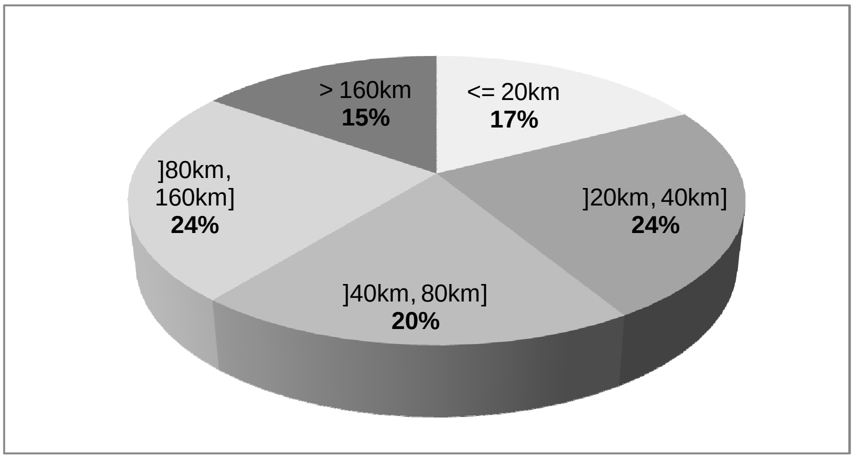
Motive	Mean value
To contact with nature	4.4
To know the heritage	4.0
Socializing with friends/family	3.9
Walking	3.8
To take photos	3.3

5.3.2.5. The trip

Trip' characteristics were known through Q13, Q14 and Q18 which enquired, respectively, about the mode of transport, the total number of vehicle's occupants and the parking choice. In order to compute the travel distance and time, the residential four digits zip code was asked in Q26 (in *personal data* section).

All the respondents travelled by car and the average group was of 3.2 persons. The distance and travel time were calculated using the *Google maps* software, assuming the recommended itinerary. The mean one-way travel distance was of 80.4 km and the minimum of 0.66 km. The mean travel time was of one hour and the minimum about one minute. The distribution of the travel distance categories are illustrated in Figure 5.4.

Figure 5.4: Distance categories



5.3.2.6. How did we dealt with the main issues in TCM?

The TCM is based on the premises that visit frequency to a recreational site declines with longer travel distances (due to higher costs) and that people consider travel costs to have the same kind of influence than entrance fees. That is, following the law of demand, as the price rises, *ceteris paribus*, quantity demanded falls. Although the extensive empirical literature, researchers had not established standard measures to be used in demand estimation. The definition of the accurate *proxy* for the price is one of the unsolved issues. The thorny subject here is which costs to include.

Concerning direct travel costs, there are two understandable options: including only fuel cost or accounting for the full costs, which additionally include depreciation and insurance costs. A frequent pragmatic procedure has been the use of the government reimbursement rate for automobile travels as the *proxy* of vehicle operating cost *per* kilometre (Parsons *et al.*, 2009). Once acknowledged the cost *per* kilometre, computation of direct cost is straight[37].

In our case study, direct costs must also include the entrance fee that motorized vehicles pay to be admitted in the forest precinct. Hence, Q18 distinguishes people who brought the vehicle inside the forest precinct and paid an entrance fee, from people who left their vehicles outside walls and entered walking without charge. The trip length is used in total travel cost estimation whenever the opportunity cost of time is to be accounted.

There are some other methodological issues inherent to all the versions of TCM which must be well thought-out when designing the questionnaire. These issues are structured here in four categories: i) the adjustment of costs for multiple destination trips (Q16); ii) the valuation of time devoted to the recreation trip, which includes travel time plus on-site time (Q31 and Q32); iii) trip expenses (Q22); and iv) the identification and treatment of substitute sites (Q11, Q12 and Q16). The following paragraphs justify the options made in the questionnaire design so as we can later deal with these issues.

i) Multiple destination trips

Contrary to one of the TCM basic premises, frequently recreation trips have multiple destinations. A few ways of dealing with this issue were already reviewed and discussed in Section 4.4. We believe that option d), consisting in the identification of the proportion of the travel cost related to each destination, is excessively burdensome and that Haspel and Johnson treatment may be quite erroneous[38]. Thus, in order to distinguish between multiple and single destination trips and to characterize the former, the questionnaire included alternatives from a) to c).

---

[37] Currently, the reimbursement paid by the Portuguese government is of €0.36/km, as defined by the Portuguese governmental order n.º 137/2010, published at 28/12/2010.

[38] An example illustrates why. Let's assume that the site *B* is the one motivating the trip. However, given their proximity, recreation sites *A* and *C* are as well visited. These visits are incidental. It is assumed, as well, that *R* means residence, visitation follows the sequence *R – A – B – C – R* and distances are: *RA=50km*, *AB=5km*, *BC=10km, CR=40km*. Since *B* is near *A*, and *C* near *B*, applying the itinerary approach recreation values of *B* and *C* would be low. This assignment would be incorrect because, by hypothesis, if *B* does not exists, the trip to *A* and C would have not taken place.

At first, in Q16 respondents were asked whether some other place was or would be visited in the same trip. If the answer was negative, no further questions about this topic were made. For people visiting other places, four other questions were presented. In the first (Q16.1), it was asked if the visit to Bussaco was the main purpose of the trip. It was a closed-ended question with two options "*yes*" and "*no*". In the following question (Q16.2), respondents are asked to list the other spaces visited besides Bussaco using an open-ended format. The third question (Q16.3) requested respondents to indicate the percentage of recreation time spent in Bussaco. The last question (Q16.4) asked for the weight the visit to Bussaco had on the travelling decision. Here the "*do not know*" option was also available.

Visiting Bussaco was the sole purpose of the trip to half of the respondents and the main purpose of the trip to 27%. That is, 55% of the people engaged in a multiple destination trip have Bussaco as the main purpose of the trip. In average, respondents in a multiple destination trip spent 46% of the recreation time in Bussaco and the visit had a weight of 59% in the trip decision. Furthermore, in addition to Bussaco an average of 1.5 sites were visited. Luso, which is a locality 2.1 km away from Bussaco, was the site more often cited (by 45% of people engaged in a multiple destination trip).

ii) Opportunity cost of time

Time is a scarce resource and thus its allocation to recreation competes with other activities. Consequently, recreation is likely to embrace an opportunity cost. The cost will depend on peoples' situation, namely in the labour market. Q31 and Q32 were directed to this analysis and follow from McConnell´s (1975: 331) suggestions. They are complementary as Q32 ascertained if it was possible to exchange recreation for working time, while Q31 checked whether that would has been actually the alternative chosen.

Q31 asked how the respondent would have occupied his/her time if he/she had not come to Bussaco. The questionnaire offered three options: "*Visiting another recreational site*", "*With another activity*" and "*Do not know*". If the respondent chooses the first option, he/she was then asked about the substitute site; if the second option was chosen the activity was then asked. This question was not answered by 6% of the respondents, 43% signed the "*Do not know*", 39% would have visited another place and 12% would have been occupied with some other activity.

In Q32 respondents were questioned directly about the possibility of substituting recreation by working time (Wegge *et al.*, 1986). When respondents recognized such possibility, the hourly wage was then asked. The majority (83%) considered that would not have the possibility of increasing the number of working hours. From people who declared that the substitution was possible, 21% considered that they would not earn any extra money. Concerning the remaining ones, the average payment expected was of €16.8/hour.

iii) Trip expenses

In the majority of the empirical studies, trip costs are computed indirectly by the researcher based on the data gathered in the survey. However, some authors (e.g., Cameron, 1992; Loomis, 1997) opted to ask directly people about the expenses related to the visit. For example, Layman *et al.* (1996) are enthusiasts of this procedure since they consider that it is the respondents' perceived cost (and not actual costs) that determines travelling decisions. Accordingly, in Q22 respondents were requested to indicate the expenses they made (or expect to make) related to the trip and the locality where the expenses occurred. The first aim was to produce two different proxies for the price of the trip, one based on respondent's estimation and the other based on our own construction of the travel cost. The comparison of the econometric results and welfare measures should then be possible. The second aim was to estimate the regional economic impact of expenses resulting from direct, indirect and induced effects (Bergstrom *et al.*, 1990). Nevertheless, neither of these objectives can be accomplished because this question was poorly answered or left in blank by the majority of the individuals surveyed. We believe that it was regarded as intrusive and/or burdensome.

iv) Substitute sites

Several difficulties have been acknowledged in the identification and econometric treatment of substitutes. Yet, it is consensual that their omission tends to biases the results (Rosenthal, 1987). Two main alternatives have been suggested in order to define the substitute: respondents identify the site(s) they perceive as having characteristics similar to the one in study (Rosenberger and Loomis, 1999); or indicate the most visited recreational site (Liston-Heyes and Heyes, 1999). Q11 and Q12 were used, respectively, to collect this data without a place being suggested.

In the answers to Q11 two places clearly showed up as the substitutes, Sintra forest and Gerês. Sintra was referred by 61% of the respondents and is 250 km distant from Bussaco, in a southern region. Gerês was referred by 36% of the respondents and is 230 km distant from Bussaco, in a northern region. In Q12 the three most quoted responses were urban green areas and natural parks, beach and Bussaco, accounting for 31%, 7% and 5% of the answers, respectively.

Crossing answers to Q10.3 and Q11 we found some inconsistencies. In Q10.3 there are 38 answers of "*totally agree*" to the statement "*Bussaco forest is a unique space in the country*", but surprisingly, 44% of those respondents declared in the Q11 that they know a similar place and all of them identified a Portuguese one.

5.3.2.7. Hypothetical questions

The combination of the TCM with the CB enables the researcher to take advantage of both methods. The CB data is meant to increase the data behind the range of historical and current observed conditions. Hence, hypothetical questions regarding respondents' behaviour in hypothetical circumstances have to be included in the questionnaire. In order to capture the net effect of the hypothetical changes, visitation behaviour must be evaluated comparing the predicted number of trips with and without the change (Huang *et al.*, 1997). As Whitehead *et al.* (2008: 889) notes, even if conditions remain unchanged, observed and intended trips are not necessarily equal. Jeon and Herriges (2010) corroborate this view as they show that maintaining the present conditions unchanged, the expected number of trips would be quite higher than the observed.

We are interested in how demand is affected by changes in the entrance fee or in the conservation conditions. Accordingly, the questionnaire included four hypothetical questions. Q17 asked for the number of trips anticipated in the following year if current conditions hold. Q20.1 and Q20.2 refer to the intended visit frequency if the entrance fee changes, holding constant the current conservation conditions[39]. Q21 asked for the intended number of trips in the following year if ¼ of the forest was damaged by a fire. Other scenarios were not included because an increase in the number of hypothetical questions is likely to cause respondent fatigue and related problems (Whitehead *et al.*, 2008). Figures available in Table 5.5 make clear that a forest fire would reduce the number of visits and that price and quantity move in opposite directions.

Table 5.5: Visitation levels

Visits *per* person/year	Mean	Min.	Max.	Obs.
Full sample				
Observed	1.87	1	104	276
If current conditions keep up	2.98	0	104	268
If 25% of the forest is affected by a fire	1.93	0	104	254
Visitors parking in the forest				
Observed	1.57	1	30	238
If current conditions keep up	2.10	0	30	233
If 25% of the forest is affected by a fire	1.33	0	30	221
If access fee decreases 50%	3.16	0	30	179
If access fee increases 50%	1.57	0	30	160
If access fee doubles	0.97	0	20	149

It would be meaningless to ask people entering in the forest as pedestrians or cyclists (whose entrance is free of charge) how they would react to an entrance fee increase.

---

[39] There were three versions of this question accounting for different variations. In version 1 there were an increase of 50% and another of 100%. In version 2 there was a decrease of 50% and an increase of 50%. In version 3 there was a decrease of 50% and an increase of 100%.

Instead, they were asked in Q19 whether they would park their vehicles in the forest if the entrance fee decreases to half. The majority, more precisely 66%, gave a negative answer.

5.3.2.8. Additional questions

The last question (Q33) asked whether respondents wish to have access to the final results of the research. People answering positively were asked to provide their electronic address. About 40% of the respondents showed interest in knowing the results. At the end of the questionnaire a few lines were made available for comments regarding the survey and/or the forest. A total of 36% of the respondents write down some comments. These comments were mainly about what should be improved in the forest and were promptly made available for the Bussaco managers. The questionnaire was regarded as excessively long by nearly 13.4% of the respondents who made some comment, corresponding to 5.4% of the sample.

5.4. CONCLUDING REMARKS

In this chapter the main aspects of a questionnaire used in surveying the visitors of the Bussaco National Forest were presented and discussed. The design was justified in detail and the main descriptive statistics were presented and analysed.

Comparing to the Portuguese population, we concluded that: the elderly generations were underrepresented in the sample; respondents had a higher degree of formal education; and had a higher disposable income. These features are probably related to the fact that younger people tend to have a higher level of instruction, which is often associated with higher salaries. Furthermore, the majority of the respondents classified the environmental protection as very important. Their answers denoted that they are not pleased with the current level of environmental protection and with the supply of green spaces in Portugal. Yet, globally Bussaco forest was regarded as being a good quality recreational space, currently not facing serious threats to conservation.

Visits have take place mainly in spring and summer seasons, with the main purpose of contacting with nature. Indeed, the natural dimension was classified as the most important one. In average, during the last three years each respondent visited Bussaco 5.6 times and all of them travelled by car, although some of them did not parked in the forest precinct. The average one-way distance travelled was of 80 km and the average travel time was of one hour. Data confirmed the main premise of TCM since people living closer to Bussaco visited it more often. Furthermore, intended behaviour in hypothetical scenarios shows that respondents are sensitive to changes in the entrance fee and in site conservation conditions. It is worth noting that it is not possible to verify whether this research results are representative of the entire population of visitors or if they are the outcome of self-selection bias. The choice of another survey mode would not have avoided this shortcoming, though.

Subsequently to the administration process three main failures in the questionnaire are recognize. First, the questionnaire could not identify the factors likely to contribute to increase the visitation frequency of current visitors. Hence, the choice of a closed-ended question seems now not to have been the best choice. Second, the question regarding travel expenditures was poorly answered and the response rate was low. Finally, the difficult to balance between the extension of the questionnaire and the response rate is evident. A shorter questionnaire would probably have a greater response rate, though at the expense of less richness of the data collected.

In spite of those shortcomings, important data was collected. A comprehensive characterization of the visitors is now available which is useful to forest managers. It includes socio-demographic data, perceptions, preferences and opinions. Furthermore, the questionnaire should be noted as successful in collecting the data necessary to the TCM and TCM-CB analysis and particularly in distinguishing and characterising multiple destination trips and in identifying Bussaco substitutes. Estimation of the WTP and consumer surplus *per* visit is also feasible. Moreover, as data is reported with detail, it will be possible to use it in benefit transfer and meta-analyses.

Thus, the work here presented reports to what should be considered as the initial steps of a broader research project directed towards the valuation of forests used mainly for recreational purposes. Although differing in dimension and features, in Portuguese mainland there are twenty seven other areas classified as national forests[40]. Worldwide there are much more requiring an efficient management.

---

[40] See http://www.afn.min-agricultura.pt/portal/gestao-florestal/regflo/stmatas

CHAPTER 6 – APPLICATION OF THE TRAVEL COST METHOD USING COUNT DATA MODELS ACCOUNTING FOR ENDOGENOUS STRATIFICATION[41]

6.1. INTRODUCTION

Efficient management requires benefits to be weight against costs. However, free access to wooded and natural areas makes the task quiet difficult. The public or semi-public character of these spaces precludes the existence of established markets ensuring an efficient allocation. Public authorities managing these areas are likely to be acquainted with current maintenance costs, but benefits must be estimated indirectly. This estimation is complex because a wide range of benefits is likely to be involved. Recreational use value is frequently among these benefits.

The TCM has been the RP technique favoured to assess the actual use value of natural and semi-natural areas visited for recreation purposes (Adamowicz *et al.*, 1994). As described in Chapter 4, different versions of TCM have been used depending on the research context. One version is the ITCM which has been applied whenever the observed behaviour concerns the number of visits a person makes to a site or group of similar sites within a period of time. The dependent variable in the demand model is thus non-negative and integer. As a result, since the late 1980s, the ITCM econometric analysis has been performed applying mainly count data models.

In this chapter the ITCM is applied to visits made to Bussaco National Forest during a period of three years. As noted before, data were collected by means of a questionnaire administered on-site to visitors to Bussaco. This sample procedure, adopted due to budget constraints, although common in outdoor recreation studies, had two direct consequences. There is no data on non-visitors and the most frequent visitors were more likely to be surveyed. As a result, counts are truncated at zero and the sample is

---

[41] A paper presenting the main empirical findings of this Chapter and of Chapter 7 was submitted to a scientific journal and is currently under review.

endogenously stratified.

The importance of accounting for endogenous stratification and truncation at zero in recreation demand models was first noted by Shaw (1988). His influential article proposed a correction for the Poisson regression model. The same kind of correction was later extended to the NB model. These are the econometric models used in data analysis.

Despite the fact that we follow the standard methodology applied in ITCM studies using on-site surveys, this chapter adds to the limited valuation literature applied to Portuguese environmental goods. The remainder of this chapter proceeds as follows. The next section reviews the main theoretical modelling aspects of Poisson and NB models corrected from endogenous stratification. In Section 6.3 the set of candidate explanatory variables is defined. Section 6.4 estimates the demand functions for recreation trips and discusses the empirical findings. Section 6.5 provides several welfare measures. Section 6.6 contains the concluding remarks.

6.2. MODELLING ASPECTS

In the ITCM analysis, the dependent variable is the number of trips made by a person or group of persons, during a pre-defined time period. The dependent variable is thus non-negative and integer. Accordingly, there are three basic candidate approaches for estimating regressions where the dependent variable has these features: linear regression; ordered models; and count data models. Linear regression has been used when the mean of the dependent variable is quite high (see, e.g., Bergstrom *et al.*, 2004; Alberini *et al.*, 2007). However, it is not entirely recommended because it is likely to give positive probability to fractional numbers and to produce meaningless negative values. On the other hand, ordered models may require some data aggregation if there are many counts, or few observations for a given count (Cameron and Trivedi, 1998: 88). Therefore, in this chapter the econometric analysis is performed using count data specifications based on the Poisson distribution.

The application of standard count data models is well established in economic and non-economic analysis. In the economic domain there is a wide spectrum of applications, namely, in patents registration and innovation intensity (Hausman *et al.*, 1984), in health economics (Lourenço, 2007), labour mobility (Winkelmann, 2003: 229) and in demand for outdoor recreation. Shaw (1988), Creel and Loomis (1990), Hellerstein (1991), Grogger and Carson (1991), Hellerstein (1991) and Englin and Shonkwiler (1995) are among the pioneer authors applying count data models in outdoor recreation demand analysis or proposing new developments to these models. Progresses have been constantly done in order to adjust the models to particular data features and researcher aims. These particular features result from the sampling plan, e.g., in the case of endogenous stratification, or from the data generating process, e.g., when sample includes a large number of zero observations for the dependent variable (Haab and McConnell, 1996).

6.2.1. POISSON REGRESSION MODEL

The Poisson regression model is usually considered the benchmark model for count data and so the fundamental departure point in count data analysis. This model specifies that each count (t_i will be used here representing the number of trips) is drawn from a Poisson distribution with parameter λ_i, which is related to the regressors (x_i). The primary equation of the model is (Greene, 2000: 880):

$$\Pr\left(T = t_i \middle| x_i\right) = \frac{\exp(-\lambda_i)\cdot \lambda_i^{t_i}}{t_i!}, \quad t_i = 0, 1, 2, \ldots; \quad \lambda_i > 0 . \tag{6.1}$$

To ensure the non-negativity of the mean, the usual specification is:

$$\lambda_i = \exp\left(x_i^{'}\beta\right), \quad i = 1, 2, \ldots, n . \tag{6.2}$$

Where $x_i^{'}$ is a $(1 \times k)$ matrix of observations on independent variables associated with the individual i; β is a $(k \times 1)$ vector of unknown parameters to be estimated.

Poisson is a one-parameter distribution since conditional mean and variance are equal.

They are given by:

$$E\left[t_i \middle| x_i\right] = Var\left[t_i \middle| x_i\right] = \lambda_i \ . \tag{6.3}$$

This property is known as equidispersion because it implies that the variance-mean ratio is unitary.

In β estimation, the maximum likelihood method is generally applied meaning that parameters values for which the likelihood is maximized have the highest relative likelihood of having generated the observed data (King, 1989). The likelihood based on the Poisson distribution is:

$$L(\beta) = \prod_{i=1}^{n} \frac{\exp(-\lambda_i) \cdot \lambda_i^{t_i}}{t_i!}, \quad t_i = 0,1,2,... \tag{6.4}$$

The corresponding log-likelihood is:

$$\ln L(\beta) = \sum_{i=1}^{n} \left[t_i \cdot x_i^{'}\beta - \exp\left(x_i^{'}\beta\right) - \ln\left(t_i!\right)\right]. \tag{6.5}$$

Often the equidispersion property makes the Poisson model too restrictive because in economic empirical applications data is typically over-dispersed (Cameron and Trivedi, 2005). In recreational demand context it is common that a small number of respondents make a large number of trips while the majority of respondents make only a few, causing overdispersion in data. It may be caused by unobserved heterogeneity or there may be interdependence between the occurrences of successive events. In these conditions the Poisson model tends to produce spuriously small standard errors of estimated coefficients, inflating the *t*-statistics. If there is overdispersion, the Poisson estimator is consistent (as long as the conditional mean is correctly specified), but biased. Consequently, the Poisson model must be abandoned in favour of a more general count data model. The most usual choice has been the NB model (Cameron and Trivedi, 1998: 71).

6.2.2. NEGATIVE BINOMIAL MODEL

The NB distribution can arise from a number of ways. The most common is the generalization of the Poisson distribution through the introduction of a term of unobserved heterogeneity (ε) uncorrelated with the explanatory variables (Winkelmann and Zimmermann, 1995; Cameron and Trivedi, 1998: 71). Individuals are assumed to differ randomly in a way that is not entirely accounted by the observed covariates. The error term could reflect a specification error, such as unobserved omitted exogenous variables or intrinsic randomness.

Unexplained randomness is introduced in the Poisson distribution parameter (we now term it μ) and (6.2) is replaced by (Cameron and Trivedi, 1986):

$$\ln(\mu_i) = x_i^{'}\beta + \varepsilon_i \Leftrightarrow \mu_i = \exp(x_i^{'}\beta + \varepsilon_i) \Leftrightarrow \mu_i = \exp(x_i^{'}\beta)\cdot\exp(\varepsilon_i) = \lambda_i \upsilon_i \,. \tag{6.6}$$

Considering $g(\upsilon_i)$ the density function of υ_i, the density of t_i, is obtained by integrating with respect to υ_i, as follows:

$$f(t_i|x_i) = \int f(t_i|x_i, \upsilon_i)\cdot g(\upsilon_i)\, d\upsilon_i \,. \tag{6.7}$$

It is further assumed that $g(\upsilon_i)$ follows a Gamma distribution (Hausman *et al.*, 1984; Cameron and Trivedi, 1986). This choice leads to a compounded data generating process following a NB distribution[42], whose density function is:

$$f(t_i|x_i) = \frac{\Gamma\left(t_i + \frac{1}{\alpha}\right)}{\Gamma(t_i + 1)\cdot\Gamma\left(\frac{1}{\alpha}\right)} \left(\frac{\frac{1}{\alpha}}{\lambda_i + \frac{1}{\alpha}}\right)^{1/\alpha} \left(\frac{\lambda_i}{\lambda_i + \frac{1}{\alpha}}\right)^{t_i} \,. \tag{6.8}$$

After a few transformations it takes the form usually presented in literature (see, e.g., Creel and Loomis, 1990; Shrestha *et al.*, 2002; Betz *et al.*, 2003; Hynes and Hanley, 2006):

---

[42] Mathematical demonstration is available in Appendix B.

$$f(t_i|x_i) = \frac{\Gamma\left(t_i + \frac{1}{\alpha}\right)}{\Gamma(t_i+1)\,\Gamma\left(\frac{1}{\alpha}\right)}(1+\alpha\lambda_i)^{-\left(t_i+\frac{1}{\alpha}\right)}(\alpha\lambda_i)^{t_i}, \quad t_i = 0,1,2,...; \quad i = 1, 2, ..., n; \tag{6.9}$$

where α is a nuisance parameter to be estimated; parameters α and λ are strictly positive; the specification for the mean usually adopted is the one presented before for the Poisson regression model, $\lambda_i = \exp\left(x_i^{'}\beta\right)$.

The mean of the random variable, t_i, is still:

$$E[t_i|x_i] = \lambda_i = \exp\left(x_i^{'}\beta\right). \tag{6.10}$$

And the variance is:

$$Var[t_i|x_i] = \lambda_i(1+\alpha\lambda_i). \tag{6.11}$$

Given that $\alpha,\lambda > 0$, the variance will always be greater than the mean. The variance-mean ratio is $(1+\alpha\lambda)$, so the degree of dispersion is a positive function of α and λ.

The likelihood based on the NB distribution is:

$$L(\alpha,\beta) = \prod_{i=1}^{n} \frac{\Gamma\left(t_i + \frac{1}{\alpha}\right)}{\Gamma(t_i+1)\,\Gamma\left(\frac{1}{\alpha}\right)}(1+\alpha\lambda_i)^{-\left(t_i+\frac{1}{\alpha}\right)}(\alpha\lambda_i)^{t_i}. \tag{6.12}$$

The corresponding log-likelihood is:

$$\ln L(\alpha,\beta) = \\ = \sum_{i=1}^{n}\left[\ln\left(\Gamma\left(t_i+\frac{1}{\alpha}\right)\right) - \ln(\Gamma(t_i+1)) - \ln\left(\Gamma\left(\frac{1}{\alpha}\right)\right) - \left(t_i+\frac{1}{\alpha}\right)\cdot\ln(1+\alpha\lambda_i) + t_i\cdot\ln(\alpha\lambda_i)\right]. \tag{6.13}$$

Contrary to the linear regression, the parameters estimated trough the Poisson and NB regression models do not indicate the marginal effect of a covariate on the dependent variable because these are non-linear models. The marginal effect of an explanatory variable, *k*, vary among individuals and is function of x_i and β_k. The individual effect for a continuous variable is computed as:

$$\frac{\partial E(t_i|x_i)}{\partial x_{ik}} = \beta_k \cdot E(t_i|x_i) = \beta_k \cdot \lambda_i = \beta_k \cdot \exp\left(x_i^{'}\beta\right). \tag{6.14}$$

Elasticities take the form of:

$$\xi_{E,x_{ik}} = \frac{\partial E(t_i|x_i)}{\partial x_{ik}} \frac{x_{ik}}{E(t_i|x_i)} = \beta_k \cdot E(t_i|x_i) \frac{x_{ik}}{E(t_i|x_i)} = \beta_k x_{ik}. \tag{6.15}$$

The overall response in population is typically of main interest but there is no single way of compute the marginal effect on mean. From the three alternatives of valuation pointed by Cameron and Trivedi (2009: 334), the average marginal effect (AME)[43] seems to be the most defensible one (Cameron and Trivedi, 1998: 80). The AME is given by:

$$AME = \frac{1}{n}\sum_{i=1}^{n}\beta_k \exp\left(x_i^{'}\beta\right) = \beta_k\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left(x_i^{'}\beta\right)\right] = \beta_k \frac{1}{n}\sum_{i=1}^{n} t_i = \beta_k \bar{t}. \tag{6.16}$$

The corresponding elasticity is given by:

$$\xi_{E,x_{ik}} = \sum_{i=1}^{n}\beta_k x_{ik} = \beta_k \sum_{i=1}^{n} x_{ik} = \beta_k \bar{x}_k. \tag{6.17}$$

If instead of a continuous variable, the effect of a change in an indicator variable is to be assessed derivates are not calculated. The impact is computed as the difference between the two expected values.

6.2.3. ENDOGENOUS STRATIFICATION

As indicated before, main data were obtained using an on-site survey. Hence, the dependent variable is truncated at zero because everyone reported at least one trip to

---

[43] One of the alternative measures (applied, e.g., by Curtis, 2002) is the marginal effect evaluated at the data mean. Instead of averaging the marginal impacts computed for each individual *i*, marginal impacts are computed at the explanatory variables average. The third possibility is to choose values of particular interest for the explanatory variables.

Bussaco forest. Furthermore, the probability of an individual being sampled is positively related to the number of trips made, i.e., the sample is endogenously stratified. In these circumstances, the standard Poisson and NB estimators are biased and inconsistent (Grogger and Carson, 1991). Hence, Poisson and NB estimators accounting for truncation and endogenous stratification must be applied. They are presented in the following sub-sections.

6.2.3.1. Endogenous stratified Poisson

As demonstrated by Shaw (1988), the Poisson probability function adjusted from truncation at zero and endogenous stratification is given by:

$$\begin{aligned} f_{ES}(t_i|x_i) &= f(t_i|x_i)\cdot\frac{t_i}{E(t_i|x_i)} = \frac{\exp(-\lambda_i)\cdot\lambda_i^{t_i}\cdot t_i}{t_i!\,\lambda_i} \\ &= \frac{\exp(-\lambda_i)\cdot\lambda_i^{(t_i-1)}}{(t_i-1)!}, \quad t_i = 0,1,2,...; \quad i = 1,2,...,n. \end{aligned} \tag{6.18}$$

The conditional mean and variance are given, respectively, by:

$$E[t_i|x_i, t_i > 0] = \lambda_i + 1; \tag{6.19}$$

$$Var[t_i|x_i, t_i > 0] = \lambda_i. \tag{6.20}$$

In model computation, the only additional requirement is the construction of a new dependent variable by subtracting one from the observed number of trips. Therefore, there is no need for supplementary programming to conduct the empirical work.

6.2.3.2. Endogenous stratified Negative Binomial

If the hypothesis of no overdispersion is rejected and there is endogenous stratification, the NB model corrected for endogenous stratification is the natural alternative. The corresponding density function was first deducted by Englin and Shonkwiler (1995) who extended Shaw's (1988) conclusions to the NB density function:

$$f_{ES}(t_i|x_i) = f(t_i|x_i)\,\frac{t_i}{E(t_i|x_i)} = \frac{\Gamma\left(t_i + \frac{1}{\alpha}\right)}{\Gamma(t_i+1)\,\Gamma\left(\frac{1}{\alpha}\right)}(1+\alpha\lambda_i)^{-\left(t_i+\frac{1}{\alpha}\right)}(\alpha\lambda_i)^{t_i}\,\frac{t_i}{\lambda_i}$$
$$= t_i\,\frac{\Gamma\left(t_i + \frac{1}{\alpha}\right)}{\Gamma(t_i+1)\,\Gamma\left(\frac{1}{\alpha}\right)}(1+\alpha\lambda_i)^{-\left(t_i+\frac{1}{\alpha}\right)}(\alpha)^{t_i}\,\lambda_i^{(t_i-1)},\ t_i = 0,1,2,...;\ i = 1,2,...,n. \tag{6.21}$$

The conditional mean and variance are now given, respectively, by:

$$E[t_i|x_i, t_i > 0] = \lambda_i(1+\alpha)+1; \tag{6.22}$$

$$Var\,[t_i|x_i, t_i > 0] = \lambda_i(1+\alpha+\alpha\lambda_i+\alpha^2\lambda_i). \tag{6.23}$$

The likelihood based on endogenous stratified NB distribution is given by:

$$L = \prod_{i=1}^{n} t_i\,\frac{\Gamma\left(t_i + \frac{1}{\alpha}\right)}{\Gamma(t_i+1)\Gamma\left(\frac{1}{\alpha}\right)}(1+\alpha\lambda_i)^{-\left(t_i+\frac{1}{\alpha}\right)}(\alpha)^{t_i}\,\lambda_i^{(t_i-1)}. \tag{6.24}$$

The log-likelihood function can be written as:

$$\ln L = \sum_{i=1}^{n}\left[\ln(t_i)+\ln\left[\Gamma\left(t_i+\frac{1}{\alpha}\right)\right]-\ln[\Gamma(t_i+1)]-\ln\left[\Gamma\left(\frac{1}{\alpha}\right)\right]-\left(t_i+\frac{1}{\alpha}\right)\ln(1+\alpha\lambda_i)\right.$$
$$\left.+\,t_i\ln(\alpha)+(t_i-1)\ln(\lambda_i)\right]. \tag{6.25}$$

Contrary to the Poisson model, when correcting for truncation at zero and endogenous stratification in the NB model, there are structural changes in the likelihood function. Therefore, a new maximum likelihood routine must be programmed. The AME and the elasticity are now given by different and more complicated expressions. Elasticity is given by:

$$\xi_{E,x_{ik}} = \frac{\partial E(t_i|x_i; t_i > 0)}{\partial x_{ik}}\,\frac{x_{ik}}{E(t_i|x_i; t_i > 0)} = \beta_k\lambda_i(1+\alpha)\frac{x_{ik}}{\lambda_i(1+\alpha)+1}. \tag{6.26}$$

It has been noted that correcting for on-site sample in a zero truncated context has a very low impact on estimates because correcting for zero truncation obviates the need for further correcting the endogenous stratification. A few studies seem to confirm this view (e.g., Martínez-Espiñeira and Amoako-Tuffour, 2008; Donovan and Champ, 2009). In general, results of both models are very similar, although fit statistics in models accounting for endogenous stratification are modestly better. Therefore, in this chapter only these models are used.

6.3. THE VARIABLES

In the single-site ITCM framework people are assumed to choose the optimal number of trips in the context of a utility maximization problem subject to budget and time constraints (Freeman, 2003a: 420). The optimal number of trips is the quantity demanded which, in accordance with economic theory, is explained mainly by the price of the good, consumer income, substitutes' price and consumer preferences. These preferences are a function of individual characteristics and are expressed through choices.

We shall briefly explain which proxies were defined for the set of candidate explanatory variables and discuss the expected sign in the econometric model. Table 6.1 summarizes main aspects. Some variables were defined to be used in alternative to each other and not together as they contain alike data expressed in different forms (e.g. D_{UR} and Tipau).

All the variables, dependent and explanatory, were constructed based on answers to the questionnaire. The quantity demanded is the dependent variable (t_i) which is defined as the number of trips each respondent made to Bussaco in the previous three years (including the trip in which the survey was answered). When the number of trips reported is higher than one, it is assumed that preceding trips have the same features as the one in which survey was completed. Furthermore, group size was not accounted when defining t_i.

Table 6.1: Potential explanatory variables

Variable	Definition	Expected sign
TC	Travel cost *per* person	(-)
M	Income	(+)
TC_S	Travel cost to the substitute site	(+)
D_{Sd}	=1 if distance to substitute <80km	(-)
D_{BSd}	=1 if distance to substitute < distance to Bussaco	(-)
Age	Age	(- / +)
Gen	Gender (1=female)	(- / +)
Educ	Years of formal education (min=4; max=18)	(- / +)
D_w	=1 workers; =0 if unemployed, retired or student	(-)
D_{ns}	=1 if the most visited site is a natural space	(+)
D_{nb}	=1 if preference for nature-based activities	(+)
DD	=1 if the most visited site is a natural space & respondent prefer nature-based activities	(+)
B_C	Opinion about Bussaco conditions	(+)
E_P	Opinion about environmental protection in Portugal	(+)
D_{BC}	=1 if vry good/good opinion of Bussaco conditions	(+)
D_{EM}	=1 if member of an environmental or sport nature group	(+)
D_{vm}	=1 if contact with nature & walking are very important visit motives	(+)
D_{UR}	=1 if urban zone origin	(- / +)
Tipau	1 if predominantly urban area; 2 if moderately urban area; 3 if predominantly rural area	(- / +)
V_w	% visits in autumn/winter periods	(- / +)
O_C	Group size	(- / +)
h_{rs}	On-site time	(- / +)

Price is critical in the estimation of any demand curve and a negative sign is always expected. In the TCM framework, travel cost is the *proxy* for the price and its definition typically involves some subjectivity. Several measures were therefore defined, all of them expressed in monetary terms. The first (TC_1) includes only out-of-pocket expenses and was computed based on: i) the round-trip distance in kilometres from the home zip code to the Bussaco forest; ii) the reimbursement rate *per* kilometre established by the Portuguese government for 2011 (€0.36[44]); iii) the entrance fee (*ef*), paid only by those taking a motorized vehicle into the forest area; iv) the group size; and v) a correction factor (*c*) accounting for multiple destination trips.

In the traditional TCM framework it is assumed that the travel cost is incurred to visit just one recreational site. This means that visitors on multiple destination trips are often

---

[44] Defined by the Portuguese Government order n.º 137/2010, published on 28/12/2010. This is also the average cost *per* km reported in Zandersen (2005: 23).

excluded (Englin and Shonkwiler, 1995; Chakraborty and Keith, 2000; Parsons, 2003: 280). However, these visits are likely to involve a high use value and if excluded, benefits *per* person will probably be overestimated and total benefits underestimated (Fleming and Cook, 2008). Furthermore, excluding multiple destination visitors may bias the sample significantly, especially in terms of the demographic and socioeconomic characteristics (Martínez-Espiñeira and Amoako-Tuffour, 2009). But, if all the observations are treated in the same way this is likely to overestimate the benefits and so the economic value of the site. In order to avoid these shortcomings, a correction factor was constructed. The correction factor equals one for single destination trips. For multiple destination trips it was defined as the average of two values: the proportion of the trip's recreational time spent in Bussaco and the visit weight in the trip decision. Percentages were specified in answers to the Q16.3 and the Q16.4. TC_1 is obtained from the following expression:

$$TC_1 = \frac{km \times 2 \times 0.36 \times c + ef}{group\ \ size}, \qquad 0 < c \leq 1, \tag{6.27}$$

where km is the one-way distance travelled from home to Bussaco.

TC_1 is the starting point for the other travel cost measures. The other measures, besides out-of-pocket expenses include the opportunity cost of travel time. Two alternative *proxies* for the opportunity cost of time were computed bearing in mind the Portuguese context. Therefore, fourteen months of payment, eleven months of work, with an average of twenty one working days and eight hours work *per* day were considered.

In the first alternative, the reference is the minimum monthly salary (MMS) in 2010, in Portugal, which was of €475. When the MMS is used, hourly salary (mw_h) is €3.6 $\left(mw_h = 3.6 = (14 \times 475)/(11 \times 21 \times 8)\right)$. We intended this to be a conservative choice as workers can seldom freely adjust the number of paid working hours. As confirmed by Palmquist *et al.* (2010), when free time is not used for recreation it is usually spent on domestic activities.

Defining h_{rt} as the one way travel time in hours, TC_2 was obtained by:

$$TC_2 = TC_1 + 3.6 \times 2 \times h_{rt} \times c\,. \tag{6.28}$$

In the next measure (TC_3), the opportunity cost of on-site time (h_{rs}) is included. As this cost is independent of the trip's nature (single or multiple destination) it is added to TC_2. This measure is given by:

$$TC_3 = TC_2 + 3.6 \times h_{rs}\,. \tag{6.29}$$

In the second alternative, the standard measure for the opportunity cost of time, i.e., one third of the hourly salary (w_h) was used as in the seminal contribution of Cesario (1976). It was computed as $\left(w_h = (14 \times M_w)/(11 \times 21 \times 8)\right)$, using the disposable monthly salary *per* worker (M_w). This is the possible and not the optimal measure because M_w is an average, as explained later. TC_4 and TC_5 were obtained respectively by:

$$TC_4 = TC_1 + 1/3 \times w_h \times 2 \times h_{rt} \times c\,. \tag{6.30}$$

$$TC_5 = TC_4 + 1/3 \times w_h \times h_{rs}\,. \tag{6.31}$$

Respondents were offered six income categories in the questionnaire. For regression purposes the midpoint of income ranges is taken, except for the first and last categories. In the first category the MMS was assumed and in the last one €3 000 was the amount chosen. Three measures of income were created: monthly household disposable income (M), monthly disposable income *per* worker (M_w) and monthly disposable income *per capita* (M_{pc}). In order to obtain M_w, M was divided by the number of people contributing to the household income (Bergstrom *et al.*, 2004 used a similar procedure). M_{pc} was computed by dividing M by the number of people in the household (Henderson and Allen, 1994).

Preceding work has arrived at mixed results regarding the statistical significance of income and the coefficient sign (see, e.g., Azevedo *et al.*, 2003; Shrestha *et al.*, 2007;

Heberling and Templeton, 2009; Anderson, 2010). Concerning TCM studies conducted in Portugal, while Mendes and Proença (2011) found a positive influence of income on demand, Perna's (1994) conclusion was the contrary. On the other hand, Santos *et al.* (2001) found opposite signs when modelling demand for fishing trips in two different samples in two regions in the country. The recreation space evaluated by Mendes and Proença (2011) is the most similar to Bussaco forest, hence a positive influence of income on demand is hypothesized.

The substitute site was defined as the place nearest the respondents' home that had characteristics similar to Bussaco. These places were identified by respondents in the Q11 and, based on that, three explanatory variables were defined. The first is a measure of the substitute's price (TC_S) and was defined as the round trip out-of-pocket cost to the substitute site $(km \times 2 \times 0.36)$. A positive relation between t_i and TC_S is thus expected. The second variable is a dummy (D_{Sd}) identifying whether respondent has a substitute at less than 80 km from home (about one hour of one-way travelling time, assuming like Englin and Cameron (1996), an average driving speed of $80\,km\ /\ hour$). The third variable is another dummy (D_{BSd}) which identifies whether the substitute is closer to the respondent's home than Bussaco. A negative sign is expected for both dummies.

The inclusion of gender (Gen), age, years of formal education $(Educ)$ and respondent's occupation (D_w) is always suggested (Parsons, 2003: 278), but there are no anticipated relations for these variables. Economic theory and empirical research do not provide guidelines for expected statistical significance and sign. However, these variables express individual and professional characteristics, therefore are likely to influence preferences and so the demand.

A three levels variable $(Tipau)$ identifying respondents' residence area type (INE, 2009) is also included in the explanatory variables set. Two CVM studies on Portuguese environmental resources found this distinction significant in explaining WTP (Santos, 1998; Nunes, 2002b). It has also recently been integrated in TCM studies (e.g., by Grossmann, 2011). *Tipau* was defined based on the *tipologia das areas urbana*s, which classifies Portuguese municipalities as: predominantly urban areas, moderately urban areas and predominantly rural areas. A dummy distinguishing only rural from urban visitors (D_{UR}) was included in the set of candidate explanatory variables.

A set of variables expected to express visitors' preferences was further defined. The first (D_{ns}), based on answers to Q12, indicates whether the site visited most often is a natural or semi-natural space. Included here are natural parks, mountains and large city parks. The second variable (D_{nb}) indicates whether visitors' favourite recreational activities involve contact with nature. It was defined based on answers to Q3. The third dummy (D_{vm}) indicates whether walking and contact with nature are very important reasons for the visit. The interaction of D_{ns} with D_{vm} yielded a new variable, DD. The influence of environmental or nature-based sports group membership (D_{Em}) was also considered.

Two variables accounting for respondent opinion are also included. One relates to environmental protection in Portugal (E_p) and the other to the Bussaco forest's environmental conditions (B_C). Both were primarily obtained using a five point Likert scale, where one signifies the worst opinion. B_C was then converted into a dummy (D_{B_C}) with one meaning a good or very good opinion. A positive coefficient is expected for the latter but there is no *a priori* expectation for the E_p.

Although Bussaco National Forest is not perceived to be crowded, we are aware that there are fewer visitors in autumn and winter. Hence, the percentage of visits in this period (V_w) was included in regression in order to ascertain if visit distribution during the year is related to the total number of visits.

Finally, the on-site time variable (h_{rs}) is intended to see if there is any link between the length and the number of visits. It would be plausible that people travelling longer distances and visiting Bussaco forest less often would stay longer in the park. However, other work does not suggest this (Shrestha *et al.*, 2007; Martínez-Espiñeira and Amoako-Tuffour, 2008). Table 6.2 displays the descriptive statistics of potential explanatory variables.

Table 6.2: Descriptive statistics of potential explanatory variables

Variable (n=264)	Mean	Stand. deviation	Minimum	Maximum
TC	31.97	28.42	0.22	148.66
M_{pc}	652	349	68	1815
TC_S	34.75	25.94	0.51	168.00
D_{Sd}	0.32	0.47	0	1
D_{BSd}	0.54	0.50	0	1
Age	36.62	12.24	16	71
Gen	0.56	0.50	0	1
Educ	13.57	3.58	4	18
Tipau	1.46	0.71	1	3
D_{UR}	0.87	0.34	0	1
D_w	0.77	0.42	0	1
D_{ns}	0.06	0.24	0	1
D_{nb}	0.42	0.49	0	1
D_{Bc}	0.72	0.45	0	1
Ep	2.47	0.59	1	5
D_{Em}	0.06	0.24	0	1
D_{vm}	0.28	0.45	0	1
DD	0.11	0.32	0	1
V_w	0.12	0.23	0	1
h_{rs}	2.97	1.47	1	7

Due to the difficulties of modelling large integers with count data models, especially when the distribution must be corrected because of truncation and/or endogenous stratification, a usual procedure is to exclude observations corresponding to large counts. For example, Englin and Shonkwiler (1995) and González *et al.* (2008) excluded visitors reporting more than 12 visits[45]. Following this procedure, visitors reporting more than 15 visits (more than five *per* year in average) were excluded from the analysis. This leads to the exclusion of 12 observations. Consequently, models were computed accounting for the other 264 observations. Mean sample figures were used for 27 respondents not reporting household income.

---

[45] Nakatani and Sato (2010) and Donovan and Champ (2009) also reported convergence problems in the NB models accounting for zero truncation and endogenous stratification.

6.4. ESTIMATIONS AND RESULTS[46]

Three alternative model specifications where computed, all accounting for zero truncation and endogenous stratification. The models are: the Poisson (ESP); the NB, where the dispersion parameter is estimated along with explanatory variables' coefficients (ESNB)[47]; and the NB where the dispersion parameter was settled *a priori* (ESNBF). Table 6.3 displays the parameter' coefficients, standard errors and income and price elasticity of demand. Econometric computations were performed using Stata software package version 11.2.

Table 6.3: Estimation results using count data models

	ESP		ESNB		ESNBF	
TC	- 0.0210	(0.0060)	- 0.0196	(0.0044)	- 0.0199	(0.0044)
TC_s	0.0067 $^{a)}$	(0.0029)	0.0065	(0.0032)	0.0066 $^{a)}$	(0.0032)
M_{pc}	0.0005 $^{a)}$	(0.0002)	0.0003 $^{c)}$	(0.0002)	0.0003 $^{b)}$	(0.0002)
DD	0.6667	(0.2054)	0.6644	(0.2518)	0.6663	(0.2518)
h_{rs}	0.2335	(0.0621)	0.1678	(0.0618)	0.1772	(0.0618)
D_{Bc}	- 0.4800	(0.1808)	- 0.3574 $^{a)}$	(0.1942)	- 0.3671 $^{a)}$	(0.1942)
V_w	1.9309	(0.3293)	3.0217	(0.4039)	2.8283	(0.4039)
Constant	- 0.7816 $^{a)}$	(0.3509)	- 2.7620 $^{a)}$	(1.3640)	- 1.6106	(1.3640)
α	---		6.0160	(9.1800)	1.2500	
$\xi_{t_i,TC}$	-0.6719		-0.3689		-0.3937	
$\xi_{t_i,M}$	0.3004		0.1217		0.1328	
Log-likelihood	-384.1324*		-338.2777		-340.0881	
LR	352.0020		443.7114		440.0906	
Pseudo-R^2	0.3142		0.3961		0.3928	
AIC	784.2649		694.5553		696.1762	
BIC	812.8725		726.7389		724.7838	

* Log pseudo-likelihood. n=264. Standard errors are reported in parentheses.
Significance: $^{a)}$ $0.01<p\leq0.05$; $^{b)}$ $0.05<p\leq0.1$; $^{c)}$ $p>0.1$; blank if significant at the 0.01 level or below.

The formal test of overdispersion suggested in Cameron and Trivedi (2009: 561) was

---

[46] See Appendix C for the procedures used in Stata and the corresponding outputs, and Appendix G for measures of goodness of fit.

[47] Several tries were made in order to parameterise the NB overdispersion parameter but none of the variables was found statically significant, so overdispersion does not seem to be explained by visitors' features. Results from generalized NB are not reported.

computed and the null hypothesis $(\alpha = 0)$ was strongly rejected[48]. The Poisson model provides consistent parameters coefficient but biased estimates of the covariance matrix if the true distribution is not precisely Poisson (Cameron and Trivedi, 1986). Therefore, the robust variance-covariance matrix was used in the ESP estimation. This is the Poisson pseudo-maximum likelihood estimator with statistical inference based on the correct specification of the mean but relaxing the equidispersion assumption (Cameron and Trivedi, 1998: 63).

Seven independent variables plus the constant were selected. Economic theory predictions and statistical performance are the criteria applied when deciding which variables should be kept in the final models. The first criterion explains why some non-significant variables remain in the final specification. The intention is to avoid omitted variable bias. The second justifies the choice between alternative measures for the same variable, namely for income and travel cost variables.

Models are coherent as coefficients' signs do not change. Absolute values of the constant term and of V_w are those with highest differences, but signs do not change across models. Coefficients in NB models are very close such as fit values. The statistical fit differs only modestly among NB models and is quite better than in ESP. However, given that the dispersion parameter in ESNB is not statistically significant, the coefficients of the three models are next analysed.

The five alternative specifications for travel cost were tested. The final choice for the travel cost *proxy* is TC_2. This variable is significant at the 0.01 level in all the models and its coefficient is negative. So demand is downward slopping, as expected. The substitutes' price, TC_S, is also significant in all the models, confirming that as the cost of substitutes

---

[48] The overall F statistic of 39.20 has a p-value of 0.0000. In order to compute the test, a variable $(\hat{\mu}_i)$ is constructed using estimated coefficients, $\hat{\mu}_i = exp(x_i' \hat{\beta})$. Then, an auxiliary ordinary least squares (OLS) regression without constant $\frac{(t_i-\hat{\mu}_i)^2-t_i}{\hat{\mu}_i} = \frac{\alpha g(\hat{\mu}_i)}{\hat{\mu}_i} + u_i$ is run. If α is significant, the null hypothesis is rejected and there is overdispersion.

rises the number of visits to Bussaco increases. While the statistical significance differs from model to model, the income coefficient has a positive sign in all the specifications, confirming visits to Bussaco forest as a normal good. However, this result must be carefully read because it is not statistically significant in NB models.

Bussaco National Forest is an exceptional semi-natural space with ancient trees and a wide variety of plants. Because of this people greatly appreciate walking and/or relaxing while in contact with nature. People preferring natural and semi-natural spaces and stating that contact with nature is a very important motive therefore visit Bussaco more often. On-site time is also positively related to the number of trips, and so the number and duration of visits are complements. The number of visits is positively related to the proportion of visits made in autumn and winter, since the most frequent visitors go there all year round. The less frequent visitors, however, come mainly in spring and summer. It is surprising to find a negative relation between opinion of the Bussaco forest's conditions and the number of visits. Indeed, the D_{Bc} negative coefficient means that individuals with a less favourable opinion are the most frequent visitors. A possible explanation for this unexpected result is that the more frequent visitors are more demanding as to the conditions. This may indicate that if conditions were improved these frequent but less satisfied visitors might visit it even more often. Demographic variables were found not to be statistically significant and were removed. As observed by Martínez-Espiñeira and Amoako-Tuffour (2008: 1328), income and education are perhaps to collinear to allow for independent estimation of the effect of education.

Price and income elasticities of demand were computed based on AME. Their values are also reported in Table 6.3. Demand is inelastic concerning price and income. Taking the ESP results, a change of 1% in TC changes trips the opposite way in about 0.67% in, whereas a variation of 1% in income changes trips in 0.3%. A practical implication of this low price elasticity of demand is that reductions/increases in travel costs, e.g. due to changes in fuels price, are not expected to have much effect on visit levels.

6.5. WELFARE ESTIMATION

Welfare measures derived from recreation demand functions are intended to quantify the site's actual recreation use value. In ITCM it is assumed that people maximize utility subject to an income constraint. Hence, the welfare measure directly derived is the MCS, which can be seen as the net benefit of the trip. The expected MCS is found by integrating the demand curve over the relevant price range. For the linear exponential demand function this integration yields the expression (Hellerstein and Mendelsohn, 1993; Englin *et al.*, 2003: 345):

$$E[MCS] = \int_{TC_0}^{TC_C} \lambda_i \, dTC = \frac{\lambda_i\left(x_i^{TC_C}\right) - \lambda_i\left(x_i^{TC_0}\right)}{-\beta_{TC}}, \tag{6.32}$$

where β_{TC} is the absolute value of the travel cost coefficient; $x_i^{TC_C}$ and $x_i^{TC_0}$ are vectors of explanatory variables for two different travel cost levels, TC_C and TC_0. If TC_C denotes the choke price, the corresponding number of trips would be zero, then $\lambda_i\left(x_i^{TC_C}\right) = 0$. Hence, the MCS *per* individual associated with the actual use of the site is given by:

$$MCS = \frac{\lambda_i\left(x_i^{TC_0}\right)}{\beta_{TC}}. \tag{6.33}$$

The corresponding MCS *per* trip is given by $1/\beta_{TC}$, computed by dividing the equation above by the predicted/observed number of trips. The predicted number of trips *per* person is the correct measure if it is assumed that the measurement error is dominant in the analysis. The sample average should be used instead if specifications errors are assumed to be the dominant (Bockstael and Strand, 1987; Whitehead *et al.*, 2000: 348; Lienhoop and Ansmann, 2011: 1256). Since it is not obvious which to choose, both measures are reported. MCS *per* trip, year and person, for the three models, are presented in Table 6.4.

Table 6.4: Marshallian consumer surplus

	ESP	ESNB	ESNBF
β_{TC}	-0.021016	-0.019552	-0,019902
CS/trip	€47.58	€51.15	€50.25
Observed average		2.31	
CS/person	€109.92	€118.15	€116.10
CS/year	€36.64	€39.38	€38.70
$E(t_i\|x_i, t_i>0)$	$\lambda+1=2.31$	$\lambda(1+\alpha)+1=2.52$	$2.5\lambda+1=2.62$
CS/person	€109.92	€124.77	€131.80
CS/year	€36.64	€41.59	€43.93

Although the TC coefficient is similar across models, its absolute value is slightly higher in the ESP. Consequently, this model produces the lowest MCS *per* trip and the most conservative estimates for the other welfare measures. The ESP's weaker distributional assumptions also make it a defensible choice. Table 6.4 also shows that for the Poisson model the predicted and observed average match[49], but predicted average values in NB models rise above the observed value.

While in the ordinary demand curve income is assumed constant, in the Hicksian compensated demand curve it is the utility that remains unchanged. Thus, Hicksian welfare measures include substitution and income effects and quantify economic benefits more accurately. There are two Hicksian measures that can be applied when the change in welfare is due to a price change. These measures are the CV and EV. They are both computed based on the absolute values of income and travel cost estimated coefficients (β_M and β_{TC}, respectively). The choice between them depends on the property rights assignment (Flores, 2003: 38). The mathematical expressions are (Englin and Shonkwiler, 1995):

$$CV = \frac{1}{\beta_M}\ln\left(1+\frac{\lambda\beta_M}{\beta_{TC}}\right); \quad (6.34)$$

$$EV = -\frac{1}{\beta_M}\ln\left(1-\frac{\lambda\beta_M}{\beta_{TC}}\right). \quad (6.35)$$

---

[49] In the Poisson regression model, predicted and observed average always match because residuals sum zero whenever a constant term is included.

The error incurred by using the MCS depends on the magnitude of the income effect and the relation between consumer surplus and income (Willig, 1976). Given the very slight income effect the three values are expected to be very similar. Furthermore, as the income coefficient is positive, the EV was expected to be higher than the MCS, and this higher than the CV, as explained before in Chapter 2. Table 6.5 reports the values for the three welfare measures based on the ESP coefficients.

Table 6.5: Hicksian welfare measures

	CV		**EV**	
β_{TC}	0.021016			
β_{M}	0.000461			
Per person	€107.25	(29.96)	€112.83	(32.95)
Per trip	€46.42	(12.96)	€48.83	(14.27)
Per person/year	€35.75	(9.99)	€37.61	(10.99)

Standard errors are reported in parentheses.

Using the CV welfare measure, i.e., considering again the most conservative measure, the average individual annual recreational use value estimated is of €35.75 and the average CV *per* trip is €46.42. For an average travel cost of €31.60, the average willingness to pay *per* trip is about €79. Aggregated recreational benefits computed using the CV are displayed in Table 6.6.

Table 6.6: Aggregated recreational benefits

Period	**Respondents** (n=311)	**Respondents' group** (Average=3.2)	**Visitors** (n=25 000)
Annual	11 118	35 579	893 750
3 years	33 355	106 737	2 681 250
10 years	111 185	3 557 912	8 893 500

Respondents' annual CV is approximately €11 120. Assuming that each person in the respondents' group has a similar benefit, the welfare value amounts to €35 580. If the sample is representative of the 25 000 annual visitors[50], annual use value is around €893 750. These figures[51] are a measure of the loss in the individuals' welfare that would

---

[50] Assuming one fourth of the national visitors make visits alike to the ones of individuals surveyed.

[51] Future values are not corrected from expected inflation or discounted.

be experienced if the recreation site happens to be destroyed or public access is forbidden. The total economic value is likely to be an important multiple of this value as passive use values are not included.

6.6. CONCLUDING REMARKS

In this chapter demand for outdoor recreation in Bussaco National Forest was estimated applying the ITCM. We came to the conclusion that besides travel cost, six other variables are meaningful in demand explanation. These variables are: travel cost beard to visit a substitute site, income *per capita*, visit motivations, opinion about Bussaco forest conditions, on-site time and visits distribution throughout the year. Accordingly, the expected substitution relation was confirmed by the positive coefficient sign and visits are normal goods as income has a positive effect on demand. We were surprised by the negative relation found between the number of visits and opinion of Bussaco forest conditions. This result suggests that the most frequent visitors are at the same time more exigent. Probably, if site conditions could be improved, increasing the attractiveness of the space, these frequent but less satisfied visitors could visit it even more often.

Furthermore, according to the price and income elasticities of demand, the number of visits is expected to be modestly affected by variations in price or in income. The low reaction of demand to changes in income and price, allows us to expect that the current economic crises in Portugal and the expected increase in fuels price will not have an extensive impact on visits to Bussaco.

Absolutes values always provide limited information and are of more difficult interpretation than the relative ones. In this sense, the comparison of our results with the ones obtained in similar studies is important. As we mentioned before, for Portugal there are no non-market valuation studies on forest recreation. Hence, the study of Mendes and Proença (2011) is probably the most comparable because their study area, the PGNP, was indicated by respondents as being similar to Bussaco. Our welfare estimates can be further compared with results reported in two meta-analyses on value of recreational forest services using data from studies conducted in European countries.

Zandersen (2005) conducted a meta-analysis on forest recreation valuation studies in Europe, using the TCM. Her survey showed that in preceding studies, at 2000 prices, the MCS range from €0.66 to €112 *per* trip, with an average value of €17.30 and a standard deviation of €28.14. The value we obtained is within that range, but above the average. Two findings of the meta-analysis help to justify this difference. First, the ZTCM tend to produce lower welfare estimates. Second, latitude has a negative effect on estimations and the majority of the observations she included came from studies conducted in north Europe countries. The gap of ten years in the reference period also helps to justify the difference.

Giergiczny *et al.* (2008) provides an estimation of the mean value of recreational forest services in European countries. Values for each country were computed using of a transfer function obtained through a meta-data analysis. The dataset includes 82 forest sites, from 49 studies conducted in 8 European countries using RP and SP methods. The mean estimated WTP/hectare/year for Portugal, at 2005 prices, varies between €6.69 and €13.72, depending on the variables included in the regression. Although these figures are not directly comparable with the ones we estimated, the values we obtain are much higher. Values estimated by Giergiczny *et al.* (2008) are a mean for all the Portuguese forests and it is quite acceptable that Bussaco forest has a high value per hectare due to its distinctive characteristics. Indeed, results of the meta-analysis regression confirm that WTP/hectare/year if forest has a protection status is higher by 106%.

Mendes and Proença (2011) estimated an average individual MCS *per* day of €194, at 2005 prices, very much above our assessment. Three main factors are likely to explain the disparity. First, they used one third of the visitors' *per capita*, *per* hour available recreation income to measure the opportunity cost of time and included on-site and travel time. Our *proxy* for time cost is the MMS and only travel time is accounted. Second, we do not include on-site out-of-pocket costs, while they do. Third, as PGNP is the only national park in Portugal and occupies a greater area, it probably has a larger zone of influence and visitors support higher average travel costs.

CHAPTER 7 – FROM COUNT DATA TO ORDERED CATEGORIES, DATA REINTERPRETATION

7.1. INTRODUCTION

In the previous chapter, demand of Bussaco forest for recreation was analysed making use of count data models. Nevertheless, conclusions achieved through the application of those models may be reinforced (or putted in to question) if a complementary analysis can be conducted. Ordered models have been suggested as an alternative to count data models (Cameron and Trivedi, 1998: 86; 2005: 682; Winkelmann and Boes, 2005: 66), although as far as we know, seldom used.

McKelvey and Zavoina (1975 *apud* Greene and Hensher, 2010) are generally credit with the introduction of ordered models within social sciences[52]. These models have been applied in several fields of the social sciences, namely in sociology, psychology and in many areas of economics (Cunha *et al.*, 2007). Health economics (Rivera, 2001), labour economics (D'Addio *et al.*, 2007) and monetary policy (Kesselring and Bremmer, 2011) are among the matter of application in economics. Although, we are not aware of any application of ordered models in outdoor recreation demand, they might offer some advantages over count data models. First, they are likely to provide a better fit because the threshold parameters allow the flexibility to align predicted and actual frequencies (Winkelmann, 2003: 66). Second, while models from the Poisson family do not follow directly from a utility maximization framework, ordered logit and probit ones do (Greene and Hensher, 2010: 103). For these reasons, these models are used in this research as a complementary tool in data analysis.

---

[52] For a detailed overview of ordered models earlier developments see Greene and Hensher (2010: Chapter 4).

This chapter is structured in the following way. In Section 7.2 the main theoretical features of the standard ordered models are presented, following essentially Cameron and Trivedi (2005) and Winkelmann and Boes (2005). Section 7.3 presents estimation and discusses the results. Section 7.4 sets out the concluding remarks.

7.2. ORDERED MODELS

In ordered category models the dependent variable is discrete and ordinal. Observations are coded and assigned to a limited number of ordered categories which are intended to express a graduation/ranking of the outcomes. The main econometric aspects can be summarized as illustrated below (Cameron and Trivedi, 2005; Winkelmann and Boes, 2005).

Let y_i denote the observation for the i^{th} of n individuals, of a categorical ordered random variable, taking J possible outcomes, coded in a rank preserving mode. Observed ordinal variable values are generated from an unobserved, continuous, latent variable (y_i^*). The two variables are related in the following way:

$$y_i = j \quad if \quad and \quad only \quad if \quad \mu_{j-1} < y_i^* \leq \mu_j, \quad j = 1,...,J . \tag{7.1}$$

Parameters μ_j represent the cut-off points between successive alternatives and are also named threshold levels. They are unknown parameters to be estimated along with the variables' coefficients. In order to ensure well-defined intervals, cut-off points are assumed to fulfil an order constraint $(\forall j \in J : \mu_j > \mu_{j-1})$. It is implicit that μ_j cover the entire real line, hence $\mu_0 = -\infty$ and $\mu_J = +\infty$. As y^* crosses a threshold, y moves to other adjacent category. These categories are taken as continuous intervals on a continuous scale.

The latent variable is generated by a linear regression structure given by:

$$y_i^* = x_i^{'}\beta + \varepsilon_i, \quad i = 1,...,n , \tag{7.2}$$

where x defines a vector of K covariates not including the constant term[53]; β is a column vector of parameters to be estimated and ε_i are the error terms. Error terms are assumed to be iid. When modelling levels of visits the underlying value of the latent variable is not the number of visits but some unknown index determining visit intensity.

The latent dependent variable is not observed and consequently the parameters are computed based on (7.1). The probability of a particular outcome is determined by the area under the density function between the relevant thresholds. From (7.1) it follows that this relevant area is given by:

$$\begin{aligned}\Pr\left[y_i = j \mid x_i\right] &= \Pr\left[\mu_{j-1} < y_i^* \le \mu_j \mid x_i\right] = \Pr\left[\mu_{j-1} < x_i^{'}\beta + \varepsilon_i \le \mu_j \mid x_i\right] \\ &= \Pr\left[\mu_{j-1} - x_i^{'}\beta < \varepsilon_i \le \mu_j - x_i^{'}\beta \mid x_i\right] = F\left(\mu_j - x_i^{'}\beta\right) - F\left(\mu_{j-1} - x_i^{'}\beta\right),\end{aligned} \qquad (7.3)$$

where $F(\cdot)$ is the cumulative distribution function of ε_i. The logistic and the standard normal distribution are the possibilities usually considered, which respectively produce the ordered logit model[54] and to the ordered probit model. The shape of these theoretical distributions is quite similar, except at the tails which are heavier in the logistic distribution (Maddala, 1983: 23). Usually there are no substantial differences in results and it is difficult to justify choosing either of them on theoretical grounds[55]. Coefficients and cut-off points' values are numerically different, but signs match and probabilities and statistical significance are similar. Consequently, economic findings are not affected by the choice of the model.

If P_{ij} is the probability that the i^{th} individual is in j category, the conditional probability function is given by:

---

[53] Different model parameterizations have been adopted to assure model identification. Some authors (e.g., Long, 1997; Greene, 2000) include the constant term in vector x and set $\mu_1=0$. Here, the alternative normalization is adopted, following Cameron and Trivedi (2005) and Winkelmann and Boes (2005).

[54] The ordered logit model can also be derived from the proportional odd ratios instead of the latent variable approach see (see, e.g., Long, 1997).

[55] Unless the sample is so large that there are enough observations at the tails to make the difference.

$$f(y_i|x_i)=(P_{i1})^{d_{i1}}\ldots(P_{iJ})^{d_{iJ}}=\prod_{j=1}^{J}(P_{ij})^{d_{ij}}, \quad d_{ij}=\begin{cases}1 & if \quad y_i=j\\ 0 & if \quad y_i\neq j\end{cases} \quad j=1,\ldots,J. \qquad (7.4)$$

Applying the maximum likelihood technique, $K+J-1$ parameters are obtained (K covariate coefficients and $J-1$ cut-off points). For a sample of K independent variables and n observations, the likelihood is given by:

$$L(\beta,\mu_j|y,x)=\prod_{j=1}^{J}\prod_{i=1}^{n}\left[F(\mu_j-x_i'\beta)-F(\mu_{j-1}-x_i'\beta)\right]^{d_{ij}}. \qquad (7.5)$$

when, as above, cut-off points are assumed to be constant, the parallel regression assumption is implicitly imposed and the standard ordered probit/logit models are in view. This means that β_k coefficients are not allowed to vary between categories and that probability curves generated in the model for each category are assumed parallel. Furthermore, this implies that relative marginal probability effects (MPEs) are constant across categories and that MPEs can change sign only once when moving from the lowest to the highest category[56].

7.3. ESTIMATION AND RESULTS[57]

Some data aggregation may be necessary before applying ordered models to the number of counts. Three use categories were thus defined – *low, medium and high use* – grouping what is considered a homogeneous series. This structure makes y an ordinal variable in which a higher order means a higher use level. The ordered responses are mutually exclusive and exhaustive and the distance between the values has no quantity meaning as different values indicate qualitative differentiation. The categorization is displayed in Table 7.1.

---

[56] Alternatively, generalized ordered models could be considered. Two important differences are that the parallel regression assumption would be relaxed and cut-off points defined as a function of explanatory variables. For a discussion of the main features see, e.g., Greene and Hensher (2010: Chapter 6).

[57] Stata commands and outputs are available in Appendix D.

Table 7.1: Categories definition and distribution

Category	Use	Visits	Percentage
1	Low	< 3	68.48
2	Medium	≥ 3 and < 6	19.57
3	High	≥ 6	11.96

Classification of visits into discrete categories may be a better expression of differences between visitors than the number of counts if the categories better characterize demand than the number of visits *per se*. However, category definition is not free from potential shortcomings. When an ordered dependent variable is specified there is no rigorous guarantee that the categories, although mutually exclusive, are characterized by an authentic intrinsic order. The choice of an ordered variable heralds two types of potential errors. The first is the possibility that a category does not correspond to a likely differentiation, i.e. that j and $j-1$ are not really two different categories. The second is that the categorization may cause inconsistencies or incoherencies in people's allocation (Nunes, 2004: 448). We believe that if any of these errors exists it is not likely to be strong enough to affect the main conclusion(s). Moreover, after estimation, it was tested whether cut-off points can be distinguished from one another, i.e. if ordered categories are truly different. It was examined by checking if their confidence intervals overlap and by testing equality of cut-off points. The acceptance of the null hypothesis ($H_0 : \mu_1 = \mu_2$) would mean that categories are not really different and indistinguishable categories should be collapsed. Our results show that confidence intervals do not overlap and that the null hypothesis is strongly rejected, since $\chi^2_{(1)} = 67.65$ and $\chi^2_{(1)} = 61.15$ for the probit and logit models respectively. Therefore, estimation proceeds with three categories and two cut-off points are estimated. Applying (7.3) to this specific context comes:

$$\begin{aligned}
\Pr[y_i = 1] &= \Pr[y_i^* \leq \mu_1] = F(\mu_1 - x_i^{'}\beta) \\
\Pr[y_i = 2] &= \Pr[\mu_1 < y_i^* \leq \mu_2] = F(\mu_2 - x_i^{'}\beta) - F(\mu_1 - x_i^{'}\beta). \\
\Pr[y_i = 3] &= \Pr[\mu_2 < y_i^* \leq \mu_3] = 1 - F(\mu_2 - x_i^{'}\beta)
\end{aligned} \tag{7.6}$$

Individual characteristics, preferences and opinions, travel cost, income and substitutes' price are among the set of candidate explanatory variables, already listed in Chapter 6 (see Table 6.1). The criteria applied in deciding which variables to keep in the model are the ones used in the preceding chapter. In the final model, coefficients are all statistically significant at the 0.1 level or higher. Values obtained using Stata software package, version 11.2, are reported in Table 7.2 along with standard deviations and fit statistics.

Table 7.2: Estimation results using ordered models

	Ordered probit		**Ordered logit**	
TC	- 0.012818	(0.0041)	- 0.026499	(0.0083)
TC_s	$0.005437^{b)}$	(0.0033)	$0.009526^{b)}$	(0.0058)
M_{pc}	$0.000533^{a)}$	(0.0002)	$0.000862^{a)}$	(0.0004)
Tipau	$0.219696^{b)}$	(0.1237)	$0.383874^{b)}$	(0.2176)
E_p	- 0.437524	(0.1525)	- 0.802513	(0.2713)
DD	0.622529	(0.2395)	1.107536	(0.4157)
hrs	0.157581	(0.0581)	0.296611	(0.1026)
V_w	2.747219	(0.3258)	5.193785	(0.6628)
μ_1	0.999548	(0.5249)	1.574645	(0.9201)
μ_2	2.102851	(0.5392)	3.585103	(0.9514)
Log-likelihood	- 159.5672		- 158.0520	
LR	140.37		143.40	
Pseudo R^2	0.3055		0.3121	
McKelvey & Zavoina R^2	0.557		0.536	
Count R^2	0.775		0.761	
P-value parallel regression	0.4806		0.3322	
AIC	339.134		336.104	
BIC	375.338		372.308	

n=276; Significance: $^{a)}$ $0.01<p\leq0.05$; $^{b)}$ $0.05<p\leq0.1$; $^{c)}$ $p>0.1$; in blank if significant at 0.01 or lower p level.
Count R^2=number of correct predictions/n.

The parallel regression assumption was tested using a likelihood ratio test. The null hypothesis of equality of coefficients across response categories was accepted for both models (p-values are reported in Table 7.2). A Wald test (the Brant test) was also performed for the ordered logit model. The null hypothesis cannot be rejected given that $(\chi^2_{(8)} = 11.28)$ with a corresponding p-value of 0.186. The parallel regression assumption is then accepted, meaning that cut-off points can be assumed to be invariable among individuals and that the standard ordered model can be correctly applied.

If thresholds are fixed, probabilities change with modifications in $x\beta$. Changes in explanatory variables shift the distribution around. In our model, the latent variable value decreases with TC and E_P. Hence, higher travel costs and a better opinion of environmental protection in Portugal increase the probability of belonging to the *low use* group and decrease the probability of belonging to the *high use* group. On the other hand, the latent variable value increases with substitute price, income, percentage of visits in autumn and winter and time on-site. The latent variable is also positively influenced by the dummy identifying those for whom walking and contact with nature are very important motives for visits and whose most-often visited sites are natural and semi-natural spaces. Those living in predominantly rural areas also have a higher probability of being more frequent visitors. Socio-demographic characteristics such as gender and education were not found significant explanatory variables. As the two models typically produce similar results and as fit statistics of the ordered logit model are all slightly better, further analysis focuses on this model.

Considering simultaneous changes in two covariates, such that y^* does not change and probabilities are unaffected, we get:

$$\Delta y^* = 0 \Leftrightarrow \beta_m \Delta x_m + \beta_\ell \Delta x_\ell = 0 \Leftrightarrow \frac{\Delta x_m}{\Delta x_\ell} = -\frac{\beta_\ell}{\beta_m}, \tag{7.7}$$

where ℓ and m denote two elements in x. Applying this result to simultaneous changes in travel cost and income, both measured in monetary units, we obtain $\Delta x_{M_{pc}} / \Delta x_{TC} = 30.76$. This ratio means that, *ceteris paribus*, the latent variable value would remain unchanged if a unit increase in travel cost was offset by an increase of €30.76 in *per capita* monthly income. This is the variation in income required to compensate a unit change in travel cost.

It is also possible to estimate the extent of a change in a covariate needed to move from one response category to another. It can be computed based on the ratio of the threshold interval length to the parameter:

$$\Delta x_m^{\min} = \frac{(\mu_j - \mu_{j-1})}{\beta_m} \tag{7.8}$$

The smaller this ratio, the smaller the minimum change required in x_m to move between categories. Applying it to travel cost, we obtain $\Delta x_{TC}^{\min} = -76$. This means that travel cost *per* person must decrease at least €76 to move the latent variable from the *medium* to the *high* use level. This is to be expected since the price elasticity of demand computed for count data models is low. Thus the reduction in travel cost must be fairly large to move the latent variable to the higher level.

Furthermore, it is possible to estimate the probability that a visitor borning a particular travel cost is in a specific category. Our data shows that if the value for the other variables is set at their mean the probability of being in the first category is always the highest one, no matter how low the travel cost. This result again corroborates the lack of sensitivity of trips to travel costs, in spite of it being significant.

Another interesting possibility is to compute the difference between probabilities concerning two levels of an explanatory variable, *ceteris paribus*. This was analyzed in order to examine how the probability of being in the *high use* category increases with changes in income. However, for people verifying specific characteristics alone the probability of being in the *high use* is greater than of being in the others. We thus fixed $V_w = 0.5$ and the mean for the remaining variables. The conclusion is that if income rises from €2 000 to €2 500, for instance, the probability of being in the *high use* category increases by nearly 10 percentage points.

The impact on probabilities can be computed when an explanatory variable changes. Varying x_k and keeping everything else constant is equivalent to shifting the distribution. For a variable with a positive coefficient the effect of an increase in k is unambiguously to shift some mass out of the leftmost or rightmost cells because the sign of the change in lateral levels depends on the sign of β_k. What happens in the middle cells is *a priori* ambiguous (Greene, 2000: 877). If the independent variable is continuous the MPE is given by:

$$\frac{\partial \Pr(y_i = j|x)}{\partial x_{ki}} = \left[f(\mu_{j-1} - x_i'\beta) - f(\mu_j - x_i'\beta)\right]\beta_k, \tag{7.9}$$

where $f(\cdot)$ denotes the density function of ε_i. As we are working with a non-linear model any marginal effects are functions of x_i and vary among individuals. Therefore there is no single way of computing marginal change in probabilities. The AME, also implicit in the elasticities estimated in Chapter 6, is usually favoured because it best preserves the respondents' characteristics (Cameron and Trivedi, 2009: 334). The average marginal probability effects (AMPE) were therefore computed. The values reported in Table 7.3 show that the effect on *medium use* is in line with *high use*.

Table 7.3: AMPE in ordered logit model

	Low use	***Medium use***	***High use***
TC	0.0030	- 0.0013	- 0.0017
TC_S	- 0.0012	0.0005	0.0007
M_{pc}	- 0.0011	0.0000	0.0000
Tipau	- 0.0468	0.0199	0.0269
E_P	0.0981	- 0.0417	- 0.0563
DD*	- 0.1352	0.0575	0.0777
h_{rs}	- 0.0361	0.0154	0.0208
V_w	- 0.6341	0.2697	0.3644

*$\partial P/\partial x_k$ for factor levels is the discrete change from the base level.

To evaluate the model performance the category corresponding to the higher predicted probability was compared with the category in which each observation lies. Categories match for 77.5% of the sample as already shown by the Count R^2 reported in Table 7.2. Furthermore, the probability of a y_i fit in each of the three categories was computed. The statistics are summarized in Table 7.4, where the observed values are also given. Although the probability of belonging to the *low use* category is underestimated and the probability of belonging to the *high use* one is overestimated, differences are small. In view of these results the model is considered to be quite satisfactory.

Table 7.4: Descriptive statistics for predicted probabilities and observed frequencies

Variable	Mean	Std. Dev.	Min.	Max.
p1	0.6797613	0.3098034	0.0094976	0.9965835
p2	0.1935055	0.1542710	0.0029576	0.4641597
p3	0.1267332	0.2016941	0.0004589	0.9331877
Low	0.6847826	0.4654464	0	1
Medium	0.1956522	0.3974225	0	1
High	0.1195652	0.3250418	0	1

7.4. CONCLUDING REMARKS

In this chapter a reinterpretation of the data used in count data models was presented. Whilst the nature of the dependent variable invites the direct application of count data models, which are standard in an ITCM framework, some well-known convergence problems prevent the straightforward use of all available data. The greater flexibility of ordered models overcomes this drawback and enables advantage to be taken of the entire sample, although at the expense of an assumed categorization. This is important in this research in three ways. Firstly, the most intensive users are kept in the sample. Secondly, given that the small sample size is considered to be one of the main limitations of our analysis, it is desirable to keep all the observations. Finally, ordered models obviate the need to guess a value for visitors stating that they made *many visits*[58], without giving a specific number as requested. Furthermore, different monetary and non-monetary measures can be derived from the models.

Estimations corroborate our hypothesis that it is meaningful to separate visitors into three sub-groups as the categories defined are statistically different. Taking this division for granted, we could assess whether a given change in one explanatory variable is likely to modify the user category. Ordered models findings confirm the importance of travel cost and preferences in the demand level. However, despite significant, the analysis shows that the travel cost is not, *per se*, critical in demand level definition. We also conclude that preferences for nature based activities are related to a higher level of visitation, the most

---

[58] In our questionnaire, respondents were meant to indicate with an *x* if the number of visits was between one and six. For more than six visits, they were asked to write down the number. Those answering many visits were not included in the analysis using count data.

frequent visitors make longer visits and that their visits are less concentrated in spring and summer periods. Furthermore, according to the model results, this semi-natural area is a normal good.

Results obtained here may be compared with those obtained using count data models, although the comparison is not direct. On the whole, the results of the two analyses are fairly coherent strengthening our confidence in the robustness of the conclusions. For the explanatory variables kept in the two final models, statistical significance is reinforced in the ordered model and signs are confirmed. From the count data models, we found an inverse relation between the number of visits and the opinion about Bussaco conditions. The values estimated from the ordered models show an inverse relation between the level of visitation and the level of satisfaction with environmental protection in Portugal. Accordingly, both suggest that the more frequent visitors are more demanding regarding environmental conditions.

CHAPTER 8 – COMBINING OBSERVED AND CONTINGENT TRAVEL BEHAVIOUR[59]

8.1. INTRODUCTION

The TCM is the oldest and most often applied method in the estimation of use value linked to outdoor recreation in public and semi-public natural spaces (Whitehead *et al.*, 2008). The TCM belongs to the group of RP techniques as it is based on observed behaviour. RP techniques have been considered more reliable[60] than SP techniques because valuation always refers to the use of some resource in past and/or present observed conditions. At the same time, RP techniques are limited in scope because they can neither estimate passive use value, nor cope with valuation outside the range of historically observed values. Nevertheless, the analysis typically involves both the observed conditions and the impact of changes on quality/price which are relevant for policy purposes but have not been observed (Rosenberger and Loomis, 1999). These impacts must be assessed using more flexible methods based on stated behaviour.

Combining RP and SP is an efficient manner of taking advantage of both techniques. Several benefits have been ascribed to a combined RP-SP analysis. First, more complete information improves the efficiency of parameter estimation and hence the precision of the estimated preferences pattern (Azevedo *et al.*, 2003; Jeon and Herriges, 2010). Second, it

---

[59] A slightly different version of this Chapter was presented at the International Society for Ecological Economics (ISEE) Conference in Rio de Janeiro, 17th June 2012. This Chapter was presented at the 19th Annual Conference of the European Association of Environmental and Resource Economists (EAERE) in Prague, 28th June 2012. We are grateful to the participants of both conferences for their helpful comments and suggestions. This Chapter was submitted to a scientific journal and is currently under review.

[60] Sugden (2005: 1) quotes the United Kingdom Treasury´s Green Book, where after the description of RP and SP techniques, which are considered alternatives, it is stated that “the technique chosen depend on the individual circumstances, and should be judged on a case-by-case basis. As a general rule, revealed preference methods are fairly reliable, and should be used where the relevant information can be inferred. However, they cannot estimate the value placed on an asset by people who make no direct use of it. In these circumstances, stated preference methods may be useful”. In the same line, US decision makers approached by List (2005) affirmed that they trusted more in empirical estimates from RP than from SP methods but understood that CVM provides the only analytical approach available for estimating total values.

enables the evaluation of a proposed policy that would modify site attributes or the recreational activity cost, but which is not currently or historically observable. Third, a suitable experimental design which introduces hypothetical quality and/or price levels is likely to break down the collinearity among characteristics. Fourth, convergent validity can be tested (Hanley *et al.*, 2003; Jeon and Herriges, 2010; Whitehead *et al.*, 2010). Checking convergent validity is of particular interest because there are always many doubts on the equivalence among SP and RP data. There are, however, some difficulties inherent to this approach. The underlying utility function must be the same in both methods for valid welfare estimates to be computed. Moreover, as additional questions must be introduced into the questionnaire to collect supplementary data, the combination of techniques requires a more complex survey instrument than the one required by the autonomous RP or SP framework.

RP data used in this study were collected using a survey that complied with the requirements for the application of the individual TCM. Questions regarding hypothetical scenarios were introduced in order to collect data on intended behaviour. One scenario is aimed at scrutinizing visitors' reaction to changes in the Bussaco forest entrance fee. The other scenario is meant to ascertain how visitors' would react to deterioration in current conservation conditions caused by a forest fire damaging ¼ of the forest.

In brief, the basic research questions this chapter aims to answer are whether visitors are sensitive to entrance fee changes and how they would react to environmental degradation caused by fire. Three models and three specifications for each model are computed and compared. Models differ with respect to observations used for each respondent and specifications differ in the way pseudo-panel information is modelled.

From the management perspective, the analysis of visitor sensitivity to changes in the entrance fee is always relevant. In periods of severe public budget constraints such as Portugal has been experiencing in recent years, the assessment of alternative financial schemes is of particular importance. In addition, many Portuguese forests and woods, including those lying within protected areas (e.g. in the PGNP and in the *Serra da Estrela* Natural Park), have recently been damaged or threatened by large forest fires. The annual average area affected by fire in protected areas in the past ten years is about 12 800 hectares. In 2010 about 46 000 hectares of woodland in Portugal were damaged by fire, of which 18 400 hectares are in protected areas (Aparício, 2011: 10). Hence, a scenario of destruction caused by a forest fire is not an unrealistic picture and it is important to know how this would affect TEV and specifically the use value. Consequently, these analyses are complementary and should contribute to a better understanding of visitors' demand behaviour.

Some research efforts have been undertaken mainly in North American countries to assess the economic effects of wild and prescribed fire on forest recreation (Vaux *et al.*, 1984; Englin *et al.*, 2001; Hesseln *et al.*, 2004). In general, the results indicate that effects differ among activities and geographic regions and that the outcome depends on the fire intensity. Recent work suggests that recreation trips decrease in initial years following the fire and after about ten years recover to about their previous levels (Rausch *et al.*, 2010). In Europe there are several studies analysing forest recreation demand (see, e.g., the survey conducted by Zandersen and Tol, 2009) and estimating the WTP for afforestation programs (Mogas *et al.*, 2006), but less attention has been paid to the effects of fire on visitation.

This analysis will add to the limited SP-RP literature on European environmental resources in three main aspects. First, while the majority of the studies have focused on water related activities, the recreation site considered here is woodland. Second, contrary to the most usual analysis, the change involves degradation of present conditions instead of a quality improvement. Accordingly, it is the length of the potential loss in use value that is under evaluation. Third, to the best of our knowledge, it is the first research in Portugal to combine TCM with CB data to assess the loss in recreational use value.

The remainder of this chapter is organized as follows. Section 8.2 provides a comprehensive literature review of the joint application of the TCM with contingent methods. Section 8.3 explains how contingent behaviour data was obtained and presents some descriptive statistics. Section 8.4 is devoted to the methods. It describes the econometric approach, reports the estimation results and presents the predicted welfare changes. Section 8.5 sets out the concluding remarks.

8.2. COMBINING OBSERVED AND CONTINGENT TRAVEL DATA: A REVIEW

A number of studies have applied two non-market valuation techniques to the same research analysis. TCM data has been jointly analyzed mainly with data obtained through the application of an SP method. In the context of outdoor nature-based recreation, CVM and CB have been the methods most often chosen. For this reason, the following literature review is restricted to these methods[61].

The CB method addresses directly how the number of trips would be affected by the hypothetical change. In the CB framework respondents are asked to reassess past behaviour or to state their intended behaviour in hypothetical circumstances. These question designs correspond to the RCB and ICB formats, respectively. In the CVM new circumstances are presented and respondents are questioned direct or indirectly about their maximum WTP or their minimum WTA in order to assure, or avoid, the proposed change. When the contingent valuation method is to be combined with TCM, a dichotomous choice question is usually used in the survey to assess whether the number of visits would change in response to the hypothetical change (Cameron, 1992; González *et al.*, 2008). The elicitation question can be presented using any of the formats. Open ended (OE), dichotomous choice (DC), payment cards (PC) and the double-bounded (DB) formats have all been selected to TCM-CVM analysis.

---

[61] For example, Adamowicz *et al.* (1994) compared and combined RP data with SP data obtained by means of CM approach, while Mogas *et al.* (2006) compared CVM and CM results.

While CVM has a longer tradition, if one of the methods is to be combined with TCM, CB encompasses an important advantage. CB data is more directly compatible with TCM data because it also reflects use value[62] only. Nevertheless, both methods are likely to suffer from hypothetical bias (Whitehead *et al.*, 2008). The seminal works of Bishop and Heberlein (1979) and Ribaudo and Epp (1984) are among the pioneer in TCM-CVM and TCM-CB analysis, respectively. Developments can be summarized in phases.

In the earlier phase, the joint application of TCM and an SP technique was designed to evaluate the environmental good in current conditions using different techniques. CVM was the dominant SP method, but was looked with high suspicion because of its hypothetical nature. TCM results were considered more reliable, in spite of method limitations being recognized (Bishop and Heberlein, 1979). One of the main objectives was to address the convergent validity of estimates by evaluating environmental resources using two competing methods (Seller *et al.*, 1985; Cameron, 1992). RP data was assumed to be the most reliable one and was used to validate the SP method.

Next, research moved ahead to a more open vision in which both methods were applied in autonomous estimations without any preconception about which method was the most reliable one (Park *et al.*, 2002; González *et al.*, 2008). In this frame, convergent validity was tested by comparing coefficients' signs, statistical significance and welfare values (see, e.g., Fix and Loomis, 1998). The main objective was to evaluate the environmental good/service using different competing methods, instead of supplement data from observed behaviour with data from stated behaviour. These are the RP-SP *comparison studies* (Whitehead *et al.*, 2008).

The third phase was pioneered by the work of Cameron (1992), who required the *cooperation* among CVM and TCM methods. SP data was meant to enlarge datasets by providing additional observations referring to hypothetical circumstances which go beyond

---

[62] As observed by Gillig *et al.* (2003: 216) the yes/no response of CVM-DC format is likely to include use and non-use values and accordingly the WTP corresponds to the TEV.

the historical data. Consequently, the focus shifted to the analysis of welfare effects resulting from hypothetical changes in quality/price. The assessments of convergent validity between RP and SP data and of the consistency between revealed and stated preferences were also aspects of main importance. In a recent work, Lienhoop and Ansmann (2011) introduced a new element in the research agenda, as they conducted a *comparison study* between TCM-CB and CVM data.

Concerning econometric treatment, three main routes have been followed. One alternative is to use observations from the two methods in autonomous estimations and to compare the results. In Jeon and Herriges (2010) perspective this is the ideal way to test the consistency of the preferences revealed in the two data sets. Accordingly, this econometric treatment was dominant in the first and second phases. A second option is to stack all the observations in a pooled model. The use of pseudo-panel data models with random effects (RE) or fixed effects (FE) is the third option. The application of two of these alternatives is frequent as well.

In the standard pooled models, all the responses are analysed together and each observation is treated as independent. Hence, correlation between the responses of each respondent is ignored. RP and SP data is aggregated assuming that the systematic variation across demand equations is captured by the independent variables and assuming that errors are iid. Therefore, the parameters are constrained to be equal for RP and SP data (e.g. Grijalva *et al.*, 2002; Bergstrom *et al.*, 2004; Grossmann, 2011). Convergent validity is often assessed through the statistical significance of dummy variables distinguishing RP from SP observations. Price elasticity tends to be higher for SP data. The difference has been captured by interaction variables and attributed to hypothetical bias and to differences in errors originated by differences in data (Whitehead *et al.*, 2008: 888). Pooled models are now commonly presented as an initial less sophisticated specification preceding the pseudo-panel analysis.

When combining data regarding actual and intended behaviour, the data set contains multiple observations from the same individual, having a pseudo-panel nature. The recognition of this feature lead to the use of panel data models in a few empirical analysis, following the pioneer work of Englin and Cameron (1996). When data has a panel nature, answers are probably correlated for each respondent due to observable and non-observable specific effects (Rosenberger and Loomis, 1999). Therefore, besides the explanatory variables, the persistence of individual differences may influence the recreation demand level. This means that respondents might share identical observed characteristics but data reveals systematically different choices (Bhat, 2003). Statistical models are inefficient if correlation across the multiple responses from the same individual is not accounted. FE and RE specifications can account efficiently for that possible correlation.

A variety of other econometric approaches have been applied. Joint estimations with a common error structure or allowing for error correlation are among the possibilities (Cameron, 1992; González *et al.*, 2008). In the context of multiple site recreation analysis, the appliance of RE-Poisson and Kuhn-Tucker demand system models have also been experimented (Whitehead *et al.*, 2010).

Table 8.1 presents a summary of the relevant literature focused on the combination of TCM with contingent methods in analysing outdoor recreation. Five main aspects are highlighted for each study: i) the good or service under evaluation; ii) the proposed change(s), when applicable; iii) the SP method(s) chosen and the questionnaire format; iv) econometric model(s) applied to the empirical analysis; and v) whether convergent validity between RP-SP data or consistency between preferences was accepted. Concerning SP data, when more than one observation is available for each individual, the number is displayed in parentheses.

Table 8.1: TCM-CB/CVM literature review

Study	Good	Change	Method	Econometric Model	RP-SP Consistency/ Conv. Validity
Bishop and Heberlein (1979)	Goose hunting	n.a.	TCM CVM-OE	n.a.	Not accepted
Ribaudo and Epp (1984)	Water recreation	Water quality improvement	TCM ICB	OLS OLS	n.a.
Seller *et al.* (1985)	Recreational boating	n.a.	TCM CVM-DC; CVM-OE	SU linear regression system Logit (DC); OLS (OE)	Mixed results
Ward (1987)	Water recreation	Quality	TCM RCB	Restricted OLS	n.a.
Cameron (1992)	Recreational fishing	Travel cost increase	TCM CVM-DC	Joint quadratic utility function-Probit ML estimation	n.a.
Loomis (1993)	Lake visitation	Water level	TCM ICB	n.a.	Accepted
Englin and Cameron (1996)	Recreational fishing	Travel cost increase	TCM RCB (3)	Pooled Poisson FE Poisson	Not accepted
Layman *et al.* (1996)	Recreational fishing	Fishing management	TCM RCB (3)	Pooled OLS Tobit	Not accepted
Loomis (1997)	River recreational activities	Travel cost increase; Quality	TCM CVM-DC; ICB (2)	Pooled Probit; RE Probit	Accepted
Huang *et al.* (1997)	Water-based recreation in estuaries	Quality improvement	TCM CVM	NB Probit	n.a.
Fix and Loomis (1998)	Mountain biking	n.a.	TCM CVM	Truncated Poisson Logit	Accepted
Chase *et al.* (1998)	National parks	Entrance fee	TCM RCB (3)	RE Probit Tobit	n.a.
Rosenberger and Loomis (1999)	Ranch open space	Characteristics	TCM RCB	RE- Poisson	Accepted
Whitehead *et al.* (2000)	Water-based recreation in estuaries	Quality improvement	TCM ICB (2)	RE-Poisson	Not accepted
Park *et al.* (2002)	Snorkelling	n.a.	TCM CVM-DC	TNB Tobit	n.a.
Grijalva *et al.* (2002)	Rock climbing	Access conditions	TCM ICB (2)	Pooled Poisson Pooled NB	Accepted a)

Study	**Good**	**Change**	**Method**	**Econometric Model**	**RP-SP Consistency/ Conv. Validity**
Hanley *et al.* (2003)	Beach recreation	Improvement to costal water quality	TCM ICB	RE-NB	n.a.
Azevedo *et al.* (2003)	Wetland recreation	Travel cost increase	TCM RCB	System of demand equations using ML	Not accepted
Bhat (2003)	Marine reserve	Reef quality improvement	TCM RCB (3)	RE Poisson	n.a.
Gillig *et al.* (2003)	Recreational fishing	n.a.	TCM CVM-DC	ML Probit	n.a.
Hesseln *et al.*(2004)	Hiking trails	Trip cost increase Fire effect	TCM ICB (4)	Poisson	n.a.
Bergstrom *et al.* (2004)	Estuary recreational fishing	Freshwater flows and fish catch	TCM ICB (3)	Pooled GLS	n.a.
Eom and Larson (2006)	River basin	Water quality improvement	TCM CVM-DB	Two equation system	n.a.
Egan and Herriges (2006)	Lake	Travel cost increase	TCM ICB (3)	Multivariate Poisson-log normal SUNB	n.a.
Alberini *et al.* (2007)	Lagoon sports fishing	Travel cost increase; Quality improvement	TCM ICB (3)	OLS RE-GLS	Accepted
González *et al.* (2008)	River	n.a.	TCM CVM-DC	ESNB Probit Joint NB-Probit	n.a.
Landry and Liu (2009)	Beach recreation	Quality and access improvements	TCM CB (2)	Discrete Factors Method	Mixed results
Jeon and Herriges (2010)	Lakes recreation	Quality improvement	TCM ICB (2)	Repeated mixed Logit	Not accepted
Whitehead *et al.* (2010)	Beaches recreation	Improved access; Beach width increase	TCM ICB (3)	NB RE-Poisson RE Poisson; RE Poisson and Kuhn-Tucker Dem. Syst.	Mixed results
Lienhoop and Ansmann (2011)	Water reservoirs recreation	Quality decline	TCM – ICB CVM-PC	Pooled NB ML regression model	Not accepted
Grossmann (2011)	Boating recreation	Water level	ZTCM RCB (2)	Pooled Poisson	Accepted

Seemingly unrelated (SU); a) Authors performed a construct validity test as the stated CB was compared with posterior actual behaviour. n.a – not applicable.

8.3. CONTINGENT BEHAVIOUR DATA

Contingent behaviour was scrutinized considering two sources of variation and four questions referring to the number of visits in the following year. Thus, each respondent was asked to provide data relating to five situations (for simplicity, revealed behaviour is also called a scenario). The scenarios are:

- The number of trips made in the preceding year (scenario 1).
- The number of trips anticipated in the following year, assuming that actual conditions remain unchanged (scenario 2).
- The anticipated number of trips in the following year, considering two possible changes in the entrance fee. The values were chosen from three alternative variations (a decrease of 50%, an increase of 50% or an increase of 100%) randomly distributed among questionnaires[63] (scenarios 3 and 4).
- The anticipated number of trips in the following year, if a forest fire damages ¼ of the forest (scenario 5).

The two requirements mentioned by Ward (1987: 385), so that the change in environmental values resulting from a quality change can be measured using TCM are fulfilled. The recreational site is not uncongested[64] and the quality change is exogenous to the individual.

As explained before (Chapter 5), in order to avoid econometric problems related to many observations of “one”, the questionnaire asked for the number of visits made in the three preceding years. On the other hand, concerning intended behaviour, the longer the period considered, the hard it is to make accurate predictions because of uncertainty factors. For this reason, the period of reference for questions concerning intended behaviour in actual and hypothetical circumstances was the following year. In order to make data compatible,

---

[63] A question similar to the one used by Englin and Cameron (1996) would provide a greater variability in travel costs. However, we believe that respondents would be more reluctant to accept the scenario.

[64] This aspect was checked in question 10 of the questionnaire and respondents classified the forest as uncongested.

the observed number of trips had to be adjusted for the same time period. Hence, when the stated number of visits was equal or less that three, one was the value considered. For larger counts, the observed value was divided by three and non-integer values mathematically rounded off to the nearest integer[65].

Data structure corresponds to a short pseudo-panel. The panel is unbalanced, however. Respondents who entered the park as pedestrians or cyclists did not pay the entrance fee and so the question about changing its value was formulated differently. These people were instead asked if they would park their cars in the forest precincts if the entrance fee was half its present value. Furthermore, there is not complete information for some respondents who answered "*I do not know*" to some hypothetical questions.

Of the 311 returned questionnaires, 9% were excluded because some crucial questions were unanswered (namely, the number of trips or the home zip code). Those reporting more than 52 visits *per* year, i.e., more than one visit *per* week (three respondents) or stating unfeasibly high travel costs for a single site visit (eight respondents) were likewise excluded from the analysis. For contingent behaviour responses, when the number of intended trips and the entrance fee varied in the same way, answers were classified as protest and excluded. The analysis in this chapter is based on the remaining 272 questionnaires.

Table 8.2 reports some informative comparisons taking scenario 2 as reference. As the number of observations varies between scenarios, values for scenario 2 are also computed for each sub-sample in order to enable meaningful interpretation.

---

[65] For example, if t_i=7, the annual average would be 2.(3) hence the value considered is 2. If values were not rounded off, count data models could no longer be used. Given that four out of the five observations are integer numbers, this was the preferred solution. We assume this as a limitation and recommend the inclusion of two questions to assess visit counts in future questionnaires: one question respecting to the preceding year and another covering a larger time period.

Table 8.2: Stated trips in different scenarios

Scenarios	Min	Max	Std. Dev.	Mean	Δ% Mean	Obs.
1	1	30	3.27	1.73	---	272
1	1	30	3.30	1.74	39%	266
2	0	40	4.15	2.41		
2	0	30	2.98	2.07	53%	179
3	0	30	3.23	3.16		
2	0	40	4.48	2.66	-41%	160
4$_{a)}$	0	20	2.57	1.57		
2	0	40	3.95	2.28	-57%	149
4$_{b)}$	0	20	2.20	0.97		
2	0	40	4.25	2.48	-47%	251
5	0	30	2.35	1.32		

Scenario 3 respects to a price reduction to half of the current price; 4$_{a)}$ respects to an increase of 50% in the entrance fee and 4$_{b)}$ to an entrance fee doubling.

The average number of trips during the year prior to the survey was 1.73. The answers to the *status quo* CB question show that if conditions are unchanged, the average number of intended trips in the following year would be 2.41, which is an increase of 39%. It is not new to find statistically significant differences between the observed and the intended number of trips when present conditions are sustained (Huang *et al.*, 1997; Jeon and Herriges, 2010). The main reason indicated for this difference is hypothetical bias. The number of visits in the future is probably inflated by the respondent's good intentions (Whitehead *et al.*, 2000).

A 50% reduction of the entrance fee would lead to an increase in the average intended number of visits of 53%, while an increase of the same magnitude, would lead to a decrease of 41%. A doubling of the entrance fee would imply a reduction in intended visits of 57%. Accordingly, visitors seem more responsive to price reductions than to increases. Damage to part of the woodlands by a forest fire would lead to a decrease in the average number of intended visits to 1.32. Comparison with the *status quo*, indicates a reduction of 47%, while comparison with the observed average gives a reduction of 24%. In general, respondents seem to be sensitive to changes in price and in conservation conditions.

8.4. METHODS

8.4.1. ECONOMETRIC MODELS

Two econometric approaches were applied in this chapter to combine observed and contingent behaviour. One is based on pooled data and the other on panel data models with RE. Since the number of trips is a non-negative integer, count data models from the Poisson family were used in both specifications.

The probability function of the Poisson regression model is:

$$f\left(t_{ij}\middle|x_{ij}\right)=\frac{\exp\left(-\mu_{ij}\right)\mu_{ij}^{t_{ij}}}{t_{ij}!},\quad t_i=0,1,2,...; \tag{8.1}$$

where $i=1,2,...,n$ denotes the respondent; $j=1,...,5$ denotes the scenario; t_{ij} is the number of observed or intended trips *per* year; x_{ij} is the vector of explanatory variables.

The pooled Poisson estimator assumes that t_{ij} is Poisson distributed:

$$t_{ij}\left|x_{it},\beta\sim Poisson\ \left[\exp\left(x_{ij}^{'}\beta\right)\right]\right.; \tag{8.2}$$

$$E\left[t_{ij}\middle|x_{ij}^{'}\right]=\exp\left(x_{ij}^{'}\beta\right)\ . \tag{8.3}$$

Observations are independent across individuals but not necessarily within the same individual. Hence, regression disturbances may be clustered at the individual level, correcting for the possible non-independence of repeated observations for the same individual. This makes the measures of statistical significance robust and control for overdispersion and correlation over *j* for a given *i*. The cluster of disturbances affects the variance co-variance matrix of the estimators, but not the coefficients (StataCorp, 2009: 20).

The Poisson model for panel data, considering the individual specific term multiplicative, as defined by Hausman *et al.* (1984), is expressed as:

$$t_{ij} \left| x_{ij}, \beta, \alpha_i \sim Poisson \left[\mu_{ij} = \alpha_i \lambda_{ij} \right] \right., \tag{8.4}$$

where, $\lambda_{ij} = \exp\left(x_{ij}^{'}\beta\right) > 0$ and α_i is an unobserved individual specific effect not correlated with x_{ij}, otherwise estimations would be inconsistent.

Alternatively, if the individual specific term is assumed additive, the Poisson model for panel data is expressed as:

$$t_{ij} \left| x_{ij}, \beta, \alpha_i \sim Poisson \left[\mu_{ij} = \exp\left(\ln(\alpha_i) + x_{ij}^{'}\beta\right)\right] \right.. \tag{8.5}$$

In the RE specification, α_i are iid random variables.

The model most often used in empirical work has been the RE Poisson-Gamma (RE-Pois-G) resulting from the assumption that the α_i parameter is iid Gamma (δ, δ) (Whitehead *et al.*, 2000). Alternatively, the $\ln(\alpha_i)$ may be assumed to be iid Normal (1, σ_α^2), giving rise to the RE Poisson-Normal (RE-Pois-N)[66]. In both cases:

$$E\left[t_{ij} \left| x_{ij}^{'}, \beta \right.\right] = \exp\left(x_{ij}^{'}\beta\right). \tag{8.6}$$

The Poisson model for panel data has the same properties, in terms of robustness, as when applied to cross-section data. It is consistent as long as the conditional mean is correctly specified, even though data does not exactly follow a Poisson distribution (Cameron and Trivedi, 2009: 620).

The RE-NB model is an alternative to the RE Poisson as the NB estimator is designed to explicitly handle overdispersion. Assuming that $t_{ij} \left| \gamma_{ij} \sim Poisson \left(\gamma_{ij}\right)\right.$ and $\gamma_{ij} \left| \delta_i \sim Gamma \left(\lambda_{ij}, \delta_i\right)\right.$, with $\lambda_{ij} = \exp\left(x_{ij}^{'}\beta\right)$, the corresponding probability function is given by[67]:

---

[66] For the RE-Pois-G a closed-form solution exists, this is the NB distribution. For the RE-Pois-N, a closed-form solution does not exist for this and an option is the use of numerical integration.

[67] Demonstration is available in Appendix E.

$$f(t_{ij}|x_{ij},\delta_i)=\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}\left(\frac{\delta_i}{1+\delta_i}\right)^{\lambda_{ij}}\left(\frac{1}{1+\delta_i}\right)^{t_{ij}}, \qquad (8.7)$$

$t_{ij}=0,1,2,...;\quad i=1,2,...,n;\quad j=1,2,..,J$.

In the RE-NB, the dispersion is assumed to vary randomly across individuals and $\left(\frac{\delta_i}{1+\delta_i}\right)\sim Beta(r,s)$. Integrating (8.7) using the Beta density, and after some algebraic manipulation, the probability function is expressed by:

$$f(t_i|X_i,\beta,r,s)=\frac{\Gamma(r+s)\Gamma\left(r+\sum_{j=1}^{n_i}\lambda_{ij}\right)\Gamma\left(s+\sum_{j=1}^{n_i}t_{ij}\right)}{\Gamma(r)\Gamma(s)\Gamma\left(r+s+\sum_{j=1}^{n_i}\lambda_{ij}+\sum_{j=1}^{n_i}t_{ij}\right)}\prod_{j=1}^{n_i}\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}\,. \qquad (8.8)$$

In our sample, observations for the dependent variable concerning observed behaviour take only strictly positive values due to the on-site data collection. If those observations were to be analyzed alone, a model correcting for truncation and endogenous stratification should be applied. However, since observations concerning the CB framework can take null values, the standard models are applied. In these circumstances inferences cannot be extended to the overall population (Englin *et al.*, 2001: 1838; Lienhoop and Ansmann, 2011: 1255).

If contingent behaviour data is not affected by incidental truncation and endogenous stratification and the Poisson model is used, the correction for the endogenous stratification proposed by Shaw (1988) could be applied to the RP data. This consists in subtracting one to the observed number of visits. However, in our sample is not possible to ensure that those who visited the forest in the past have the same or greater probability of visiting it in the future.

8.4.2. PARAMETER ESTIMATES[68]

As explained before, the basic idea behind the individual version of the TCM is that the number of recreational trips falls as the travel cost to the recreational site rises, *ceteris paribus*. Travel cost is the *proxy* used for the price. The price for each individual is constant, but observing individuals living different distances way and facing different costs makes it possible to draw the demand curve. Quantity demanded is the result of a utility maximization problem subject to budget and time constraints. Hence, other standard variables influencing demand, such as the price of related goods and individual tastes, are included in the set of candidate explanatory variables.

The set of candidate explanatory variables common to the three models is displayed in Table 8.3. A summary of statistics are reported with reference to the 272 observations. The design of these variables was presented before in chapters 5 and 6. Models also include specific slope and intercept shift parameters that distinguish between scenarios.

Table 8.3: Descriptive statistics

Variable	Description	Mean	Std D.	Min	Max
TC	Travel cost *per* person	30.99	27.10	0.22	118.42
TC_S	Travel cost to the substitute site	35.32	26.23	0.51	168
M	Household monthly disposable income	1 813	805	475	3 000
Age	Age	36.53	12.22	16	71
Gen	Gender (1=female)	0.50	0.57	0	1
Educ	Years of formal education	13.59	3.58	4	18
DD	=1 if walking & contact with nature main visit motives	0.14	0.35	0	1
E_p	Importance of environmental protection	4.76	0.34	3.25	5
h_{rs}	On-site time	3.02	1.50	1	7

Econometric computations were performed using Stata software package version 11.2. Three complementary models (A, B and C), differing in scenarios included, were estimated. Model A evaluates the change in the use value if a forest fire damages the woodland. It deals with observations from scenarios 2 and 5. Model B is intended to ascertain visitors' price sensitivity. It works with observations from scenarios 2 to 4. Hence, the three contingent scenarios differ only in travel cost since environmental

---

[68] Stata outputs are available in Appendix F.

conditions remain unchanged. Model C is the most general as it deals with the full set of observations. It takes in the RP observation, not included in the other models. Consequently, for this version pooled and pseudo-panel models require that the two data sources have the same structure in terms of dependent and explanatory variables.

Three specifications were computed for each model: the pooled NB[69] with cluster-robust errors, the RE-Poisson[70] and the RE-NB. All the models were estimated using balanced panels. Panel data specifications accommodate unobserved heterogeneity that here refers to the possibility that unmeasured differences among observationally equivalent individuals affect the number of visits.

---

[69] Poisson and an NB specification were computed, but the statistical fit favoured the NB model.

[70] For the Poisson model two RE specifications were computed, assuming Gamma and Normal distributions.

Table 8.4: Demand models

	Model A (quality change)			Model B (price change)			Model C (RP-SP)		
	Pooled NB	**RE-Pois-G**	**RE NB**	**Pooled NB**	**RE-Pois-G**	**RE NB**	**Pooled NB**	**RE-Pois-G**	**RE NB**
TC (Travel cost *per* person)	- 0.0094 (0.0008)	- 0.0097 (0.0033)	- 0.0083 (0.0027)	- 0.0050 (0.0015)	- 0.0067 $^{a)}$ (0.0026)	- 0.0059 (0.0022)	- 0.0062 (0.0019)	- 0.0074 (0.0027)	- 0.0062 (0.0021)
TC$_s$ (Travel cost to substitute)	0.0075 (0.0013)	0.0074 (0.0027)	0.0054 $^{a)}$ (0.0023)	0.0099 (0.0006)	0.0091 (0.0029)	0.0078 (0.0022)	0.0081 (0.0015)	0.0074 (0.0028)	0.0059 (0.0020)
DD (walking/contacting with nature)	0.5341 (0.0806)	0.5464 (0.2087)	0.4467 (0.1557)	0.4926 (0.0508)	0.4884 $^{a)}$ (0.1914)	0.4827 (0.1597)	0.4440 (0.1046)	0.4320 $^{a)}$ (0.1871)	0.4220 (0.1415)
h$_{rs}$ (on-site time)	0.1901 (0.0097)	0.1851 (0.0443)	0.1694 (0.0381)	0.1936 (0.3061)	0.1947 (0.0375)	0.1793 (0.0385)	0.1979 (0.0232)	0.1959 (0.0382)	0.1567 (0.0341)
Ep (environmental protection)	0.4411 (0.0456)	0.3998 $^{a)}$ (0.1915)	0.2963 $^{b)}$ (0.1792)	---	---	---	---	---	---
Age2 (Age squared)	---	---	---	0.0002 (0.0000)	0.0001 $^{a)}$ (0.0001)	0.0002 (0.0001)	---	---	---
D$_{fire}$	- 0.6790 (0.0077)	-0.7127 (0.1653)	-0.6437 (0.1130)	---	---	---	-0.4148 (0.1483)	-0.4325 (0.1444)	-0.4125 (0.0975)
TC_D$_{fire}$	0.0022 (0.0002)	0.0034 $^{c)}$ (0.0038)	0.0021$^{c)}$ (0.0033)	---	---	---	- 0.0013 $^{c)}$ (0.0038)	-0.0007 $^{c)}$ (0.0034)	-0.0009 $^{c)}$ (0.0028)
D$_{rp}$	---	---	---	---	---	---	-0.3053 $^{b)}$ (0.1735)	-0.3160 $^{a)}$ (0.1564)	-0.3260 (0.0912)
TC_D$_{rp}$	---	---	---	---	---	---	0.0022 $^{c)}$ (0.0037)	0.0021 (0.0033)	0.0030 $^{c)}$ (0.0024)
Const	- 1.9975 (0.1608)	- 1.7726 $^{b)}$ (0.9329)	0.4726 $^{c)}$ (0.9029)	- 0.4846 $^{c)}$ (0.3715)	- 0.3902 $^{a)}$ (0.1769)	1.7218 (0.4558)	-0.1461 $^{c)}$ (0.1160)	-0.0678 $^{c)}$ (0.1658)	2.1986 (0.3013)
α	0.5230 (0.1092)	0.5041 (0.0987)	---	0.6592 (0.2074)	0.5752 (0.1027)	---	0.5314 (0.0612)	0.4597 (0.0666)	---
Log (pseudo)likelihood	- 856.6129	- 842.2450	- 835.1928	- 1 342.5557	- 1 263.0758	- 1 258.5889	- 2 032.4211	- 1 881.9303	- 1 872.7159
Observations	502			732			1 170		
Price elasticity of demand	- 0.2907 (0.0232)	- 0.2997 (0.1008)	- 0.2555 (0.0844)	- 0.1556 (0.0458)	-0.2101 $^{b)}$ (0.0829)	- 0.1852 (0.0695)	- 0.1915 (0.0596)	- 0.2316 (0.0829)	- 0.1921 (0.0647)

Standard errors are reported in parentheses. For pooled models they are cluster-robust standard errors and for the Pois-RE models were estimated using bootstrap robust methods. Significance: $^{a)}$ $0.01<p\leq0.05$; $^{b)}$ $0.05<p\leq0.1$; $^{c)}$ $p>0.1$; in blank if significant at the 0.01 or lower p level.

Table 8.4 reports estimated coefficients and the log (pseudo)likelihood values for the alternative specifications of each model. In addition to the indicator variables for the RP data and fire scenario, models differ in two other explanatory variables. Opinion on environmental protection was found statistically significant only in Model A, while age squared was found statistically significant in Model B. Other socio-demographic variables such as gender, formal education and income were not significant in any of the models and were left out because the literature does not identify a standard effect.

Comparing the results for the econometric models, we come to three main conclusions. First, the coefficients' signs of the significant variables never change, with the single exception of the constant. It is negative in the pooled NB and RE-Pois, while in the RE-NB it is always positive. Second, the RE-Pois produce the highest absolute values for the *TC* coefficient. Consequently, it generates the lowest consumer surplus. Third, the coefficient of the substitute price is always statistically significant at the 0.05 level and has the expected positive sign. As the travel cost to a substitute site increases, *ceteris paribus*, visits to the forest also increase.

The coefficient for the travel cost measure has the expected negative sign and is statistically significant at the 0.01 level in eight out of nine specifications. Hence, demand is downward sloping, according to the law of demand. On the other hand, the interaction variable among TC and the dummy indicator of the forest fire scenario included in Models A and C, is not significant, meaning that demand slopes remain unchanged across the scenarios. Consequently, in the forest fire scenario the MCS *per* trip is not statistically different from the *status quo* and price elasticity does not change either. Furthermore, in Model C, the slope interaction variable among TC and the dummy indicator of the RP data, while having a negative sign, is not statistically significant in either of the specifications.

The dummy DD has a positive sign and is also significant in the three models. Accordingly, respondents who visit Bussaco mainly for walking and be in contact with nature and who are frequent visitors of natural spaces will visit the forest more often in any scenario. This is possible related to the positive and statistically significant coefficient of on-site time. If so, visits and time on-site are complementary.

Respondents' opinion about the importance of environmental protection was included only in Model A, which includes the forest fire scenario. It is significant, with a positive effect in all three specifications. Accordingly, those who are more concerned with environmental protection intend to visit the forest more often.

The dummy variable for the forest fire scenario is statistically significant and has a negative sign in both models, A and C. Hence, it acts as a demand shifter and the effect is that expected. If part of the woodland were to be damaged by a fire, and keeping other variables constant, the number of visits will decrease and a part of the use value will be lost[71].

Model C includes another dummy variable (D_{RP}) which distinguishes contingent from observed behaviour. The high statistical significance of this variable in the pseudo-panel specifications indicates that the two data sources are not statistically equivalent. If so, there is no consistency between the revealed and stated preferences and RP and SP data should not be combined.

Fit statistics indicate that the RE models are a more efficient choice than the pooled models as likelihood values are lower. The likelihood-ratio test also rejects the pooled model. Accordingly, the RE parameter is significant in all the models meaning that there is common variance in individual responses across scenarios. Cameron and Trivedi (2009: 627) argue that the RE-Poisson estimator with cluster-robust standard errors is likely to be more robust and a better choice than the NB estimator because it is based on weaker distributional assumptions. For that reason, predicted values are computed using the

---

[71] When administering the pilot survey some respondents stated that if the woodland was damaged by a fire they would be likely to maintain, or even increase, the number of visits in order to help the recovery through the payment of the entrance fee. Hence, this coefficient may be less meaningful because of this potential effect, not related directly to use value.

results from the RE-Pois. Table 8.5 reports the estimates of the conditional mean, after integrating out the RE, and the stated values.

Table 8.5: Stated and predicted number of trips

Model	**A**		**B**			
Scenario	**2**	**5**	**2**	**3**	**$4_{a)}$**	**$4_{b)}$**
Stated trips	2.48 (4.25)	1.32 (2.35)	2.32 (3.83)	3.16 (3.23)	1.57 (2.57)	0.97 (2.20)
Predicted trips	2.47 (1.47)	1.32 (0.74)	2.09 (1.17)	2.13 (1.15)	2.24 (1.21)	1.95 (1.12)
Observations[a]	251		244	179	160	149

[a] Concerning Model B note that (179+160+149)/2=244 because each respondent was asked about the number of future expected trips considering two price changes.

The average number of trips predicted by Model A is very close to the number of stated trips, for both scenarios. In accordance with both stated and predicted values the planned number of trips would be reduced by about 47% in the following year if a fire damages 25% of the forest. In Model B, however, the difference between stated and predicted values is quite telling. The difference is particularly high for the entrance fee doubling scenario. The mean for model predictions is about twice that of the stated. The difference is also meaningful in the entrance fee-reduction scenario, where predicted mean is 32% lower than the stated.

Paradoxical results concerning the reaction of current users to price changes were thus observed. On the one hand, the descriptive statistics (reported in Table 8.2) show a major reduction in the number of intended trips if the entrance fee increases by 50% or if it doubles. The reduction would be of 41% and 56%, respectively. On the other hand, econometric model predictions point a much weaker reaction of visitors to entrance fee changes. Furthermore, price elasticity of demand is quite low (-0.21, in Model B), showing an inelastic demand. We suggest two possible explanations for these contradictory results.

The first is related to an old issue discussed in TCM analysis, and that is the difference between the price of the trip perceived by visitors and the *proxy* constructed by the researcher for that price (Layman *et al.*, 1996). In the present case, it is possible that respondents do not consider all the implicit costs and pay disproportionate attention to a single explicit cost related to the visit, the entrance fee. The second possibility we envisage is that the stated number of intended trips expresses agreement or disagreement with the proposed change. Hence, the increase (reduction) in the number of visits would be exaggerated in the entrance fee reduction (increase) scenario. If so, and their aim is to influence the decision makers, we are dealing with strategic bias resulting from the hypothetical nature of the question. Respondents may be opposed to this unpopular payment vehicle (Hanley, 1989).

It is difficult to say whether these paradoxical results are due to the specificities of our research. There are very few analysis in the literature similar to the one that yielded these results, especially when price changes are at issue. Lienhoop and Ansmann (2011) is one of the rare examples comparing predicted with the stated number of trips. Their results, based on hypothetical quality changes, are in line with ours as the model predictions are very close to the observed number of trips. Our results are also in line with the conclusions of Whitehead *et al.* (2010: 107): *"*trip overstatement tends to occur in baseline forecasts of behaviour and not in changes in forecast behaviour as quality/conditions change*".* The strickiest question refers to contingent behaviour in response to price changes. Inconsistencies between RP and SP data seem to be more frequent when price changes are considered than when quality changes are at issue (Englin and Cameron, 1996; Azevedo *et al.*, 2003; Whitehead *et al.*, 2010).

In spite of the lack of consistency found between RP and SP data, we admit that respondents are likely to be acting rationally[72], trying to influence decision makers' action in a way that would enhance their well-being. If so, their apparent inconsistent behaviour is not justified by the non-existence of active institutions that reward reliable choice and punish the unreliable choices (Hanley and Shogren, 2005: 14).

---

[72] Assuming that a rational agent: "is one who draws conclusions logically from given premises, who premises are defensible by reasonable argument, who uses evidence dispassionately in evaluating factual assertions, and more technically, who optimizes subject to constraints under conditions of limited information and costly decision making" (Gintis, 2000).

A possible interpretation of our results is that if there is a quality change which escapes from the decision makers control (as in this particular case), stated and revealed behaviour will look consistent because there is no incentive to give a strategic response. Conversely, stated and revealed behaviour will probably seem inconsistent when decision makers are responsible for the proposed change because respondents try to influence the decision. This is a question that certainly deserves future research.

8.4.3. WELFARE CHANGE IN THE FOREST FIRE SCENARIO

Given the doubts as to the validity of Model B, welfare measures are computed for Model A only. The MCS is one of the welfare measures which can be used to compute the net benefit derived from recreational visits. As explained in Section 6.5, total MCS for a season is found by integrating the demand curve over the relevant price range. It can be represented by:

$$MCS = \frac{E\left(t_{ij}\middle|x\right)}{\beta_{TC}} = \frac{\lambda_{ij}}{\beta_{TC}}. \tag{8.10}$$

The change in MCS, resulting from the damage caused by the forest fire, is given by:

$$\Delta MCS = \frac{\lambda_5}{\beta_{TC} + \beta_{TC} * D_{fire}} - \frac{\lambda_2}{\beta_{TC}}, \tag{8.11}$$

where λ_5 is the number of trips in the fire scenario and λ_2 is the number of trips if current conditions do not change.

If the coefficient on the interaction between the scenario indicator variable and the own-price were found significantly different from zero, the relevant price coefficient would differ in the equation above. However, in addition to its unexpected sign in Model A, it is not significant in the RE specification. Accordingly, the equation above simplifies to:

$$\Delta MCS = \frac{\lambda_5 - \lambda_2}{\beta_{TC}}. \tag{8.12}$$

Another consequence of the non-statistical significance of interaction variables is that demand curves have similar slopes. However, they have different positions as the indicator variables of the RP and SP scenarios are statistically significant.

Assuming that measurement errors prevail over specification errors, the estimated number of trips must be considered in the welfare computation (Bockstael and Strand, 1987). Table 8.6 presents some estimates of welfare change in the fire scenario.

Table 8.6: Consumer surplus

	Estimation	95% Conf. Interval
β_{TC}	-0.0097	[-0.0161; -0.0033]
MCS/person, trip	€102.75	[35.02; 170.47]
MCS/person, year – scenario 2	€251.73	[85.80; 417.66]
MCS/person, year – scenario 5	€133.57	[45.53; 221.62]
∆ annual n.º of trips	-1.15	---
∆ MCS/person, year	€-118.16	[-196.05; -40.27]

The figures above show that in the year following the fire damage to the woodland, the average number of trips would be reduced in 1.15. Accordingly, the average loss would be of €102.75 *per* trip and €118.16 *per* year. Considering the 311 respondents, the loss sums up €36 748, while accounting for the average group size (3.2 persons), the loss increases to €117 593. Assuming, further, that the sample is representative of visitors, and considering 25 000 annual visitors with similar characteristics, the total expected loss related to use value would be near three million Euros. It is, however, reasonable to expect that fire would negatively affect visits to the forest for a longer period, resulting in additional use losses. These figures illustrate the huge loss citizens suffer when forest fires damage national woodlands and show how important it is to protect this natural heritage.

When the hypothetical change refers to an improvement in current conditions and the survey is administered on-site, there is always the problem of excluding potential visitors who would visit the site in enhanced conditions. Welfare gains are therefore probably underestimated. Conversely, when deterioration is considered, as in the present study, the probability of non-users becoming users is low. Hence, the estimated welfare change related to use value is likely to involve a low deviation from the true value.

8.5. CONCLUDING REMARKS

This chapter has analyzed the effects of two distinct and independent changes on visiting woodlands. One relates to price variations due to changes in the entrance fee. The other embraces a novel aspect because, instead of site improvements to forest conditions, deterioration due to a forest fire is considered. Model results are theoretically valid as the price of the good, price of substitutes, and preferences are statistically significant in demand explanation.

The econometric model and descriptive statistics show that if ¼ of the woodland was damaged by a fire the number of trips would be reduced by 47% and the MCS would drop to the same extent. For the group of respondents, the total loss is about €117 600. As deterioration is considered, the probability that non-users become users is low. Hence, the estimated welfare change related to use value is likely to involve a low deviation from the true value. On the other hand, this is likely to represent a small part of the TEV since these figures refer to only one year and do not include users' and non-users' passive use values.

The policy implications are very important not just for the Bussaco forest managers, but also for the national authorities responsible for the management of other natural areas. In the past decade hundred thousand hectares of woodland in Portuguese protected areas have been damaged by fire and it is fairly reasonable to assume that recreational visits to those areas have been negatively affected. This research provides some figures indicative of that effect. This is an important cost which must be included in the management decisions, along with the loss of timber.

Results concerning price changes are less clear. While descriptive statistics reveal a high impact of entrance fee changes on the number of visits, the price elasticity of demand is low and model predictions suggest a weaker reaction in demand. These contradictory results are probably due to strategic bias in responses to price changes. Responses are likely to be more indicative of agreement or disagreement than the expression of the true intended behaviour. A deeper analysis on this is left for future research which should also include the study of consistency between RP and SP data. Furthermore, in this research some signs of inconsistency were found as the indicator variable for RP was statistically significant.

Finally, it is worth to note that on-site sampling has the advantage of ensuring that respondents are familiar with the recreation site, in spite of being asked to deal with hypothetical scenarios. However, it imposes endogenous stratification and truncation which are likely limitations of this research.

CHAPTER 9 – CONCLUSION

The main ambition of this study was to contribute to improving knowledge on the value that Portuguese citizens attach to forest/woodland recreation. More specifically, we aimed to trace a demand curve for this type of recreation. This enabled the establishment of a functional relation between quantity and price, and also made it possible to identify other variables explaining the demand, i.e., the demand shifters. To accomplish this aim we carried out an empirical analysis, mainly using data collected through the administration of a survey designed specifically for this research. Compared with other surveys on non-market valuation conducted in Portugal, ours had the novelty of collecting data simultaneously on revealed and contingent behaviour. Moreover, the count data models traditionally used in the analysis of revealed travel behaviour were complemented with ordered models, which enabled a deeper understanding of demand. Two additional original features of the analysis are the use of a correction factor, which enabled us to keep the multi-purpose trips in the estimation, and the use of the minimum monthly salary in the computation of the opportunity cost of time. Furthermore, instead of the usual improvement scenario, a hypothetical deterioration in the forest conditions due to a fire was considered. These can be considered the main contributions of this study to the non-market environmental valuation literature.

In Part I we showed that the first theoretical ideas which sowed the seeds for the development of the main non-market valuation techniques date back to the first half of the 20th century. Since then, the literature on environmental valuation has been proliferating and significant developments have been achieved. Two main factors appear to be responsible for this dynamism. First, there is a growing awareness that natural resources offer a wide range of benefits to society and that many of these resources can be damaged or irrevocably lost if active measures are not adopted. At the same time, it is recognized that conservation and protection programmes are costly and that public budgets are not sufficient to achieve all the goals. Consequently, economic valuation is perceived as essential to inform policy-makers about the benefits of alternative measures as their benefits are not properly reflected in the market prices. Second, there has been a continuous attempt to refine the methodological framework of non-market valuation. This includes the enhancement of theoretical models, the improvement of analytical and econometric treatment of the data, the enrichment of datasets and the refinement of the conceptual framework. Significantly, as methods are enhanced, the reliability of the research results is reinforced.

As was made clear from the literature review, environmental valuation research in Portugal began in the early 1990s and has mainly focused on the evaluation of natural parks and traditional landscapes. The CVM has proved to be the preponderant method. Typically, price, income and the resource used for recreational purposes are among the most important explanatory variables. In general, the results confirm the method's theoretical validity.

The main purpose of the empirical analysis developed in Part II of this study was to contribute to knowledge of the demand for forest recreation in Portugal. The Bussaco National Forest was chosen as the case study. In addition to the other aspects which enable the application of the TCM, this forest can be classified as a well defined entity and, as such, it is likely to ensure that what we wanted to measure tallied with what was in fact measured.

In Chapter 5 we described the survey procedures followed in the data collecting process. The design of the questionnaire was explained and descriptive statistics were presented. Contrary to other countries, in Portugal there are no national recreation surveys and no data on forest recreation was available, hence this was an essential stage in the research. Of the 1055 questionnaires distributed, 311 were returned, resulting in a rate of return of 29.5%. This modest rate confirms the difficult balance between the extension of the questionnaire and the response rate. However, the questionnaire was able to collect essential data and was also successful in the tricky task of distinguishing and characterizing multiple destination trips and in identifying Bussaco substitutes.

From the data analysis we concluded that environmental concerns and preferences for open air recreational activities are perceptible. Visits to the Bussaco National Forest are motivated mainly by wanting to be in contact with nature and to socialize with friends and family. The average annual visit frequency is 1.9 and the average distance travelled is 80 km. Bussaco is regarded as being a good recreational space and not facing serious conservation threatens at present.

In Chapter 6 we focused on the observed trip behaviour. The ITCM was used to estimate the demand function. The dependent variable was the number of trips made by the respondent in the three years prior to the administration of the questionnaire. Accordingly, count data models were used in the econometric analysis. Statistical distributions were corrected for endogenous stratification of data and zero truncation of counts because of the questionnaire being administered on site. The standard explanatory variables were used in the estimation. However, instead of excluding multiple destination trips, a correction factor accounting for the influence of the visit to Bussaco in the trip decision was introduced in the travel cost variable.

Travel cost, substitute price and household income are among the relevant variables in demand explanation. The most conservative figures, based on the compensating variation and ignoring potential substitution effects, show that the destruction or the interdiction of this national forest to visitors would impose a loss in welfare of about €106 700/year. However, in the case of destruction, higher values are likely to be involved because passive use values have not been included in this analysis. Furthermore, our estimates show a low demand reaction to changes in income and in price.

In Chapter 7 we employed ordered models to analyse the observed trip behaviour, supplementing the analysis by count data models. While count data models enabled the estimation of monetary welfare measures, ordered models utilized the entire sample because observations associated with the higher counts could be kept in the estimation.

The ordered models confirmed our hypothesis that it was meaningful to separate visitors into three sub-groups of users – low, medium and high – as the categories were found to be statistically different. Overall, ordered models confirmed that travel cost and substitute prices are important explanatory variables, that visits are a normal good and that the number of visits and time on-site are complements. Furthermore, these models suggest that despite its statistical significance, travel cost is not, *per se*, critical to demand level definition. This conclusion is in line with the low price elasticity of demand computed based on count models. Thus, the results of this chapter reinforced the confidence in previous results.

In Chapter 8 we combined contingent and observed behaviour to estimate the change in recreational use value of the Bussaco National Forest due to quality and price changes. A rich dataset containing five observations for each respondent allowed for the estimation of three models, differing in the number of scenarios. Instead of the usual improvement scenario, a hypothetical deterioration in forest conditions due to a forest fire was considered. A RE Poisson model was ultimately chosen. Overall, the results showed that visitors are sensitive to price and quality changes. The model confirms that in the forest fire scenario, the intended number of trips would be seriously reduced and that respondents would suffer an important welfare loss.

Although no payments have been directly asked for, this research also provides some indications of bias resulting from the contingent scenarios. Answers to the hypothetical behaviour question about if current conditions remain unchanged seem to be affected by hypothetical bias, probably due to respondents having “good intentions” regarding future behaviour. This is a sign of inconsistency between revealed and intended behaviour and may also be taken as a symptom of bounded rationality. Additionally, reactions to price changes seem to be affected by strategic bias. In light of the results, we hypothesize that when there is a quality change which is beyond the decision makers’ control (e.g., in the case of a forest fire), observed and stated behaviour will be consistent because there is no incentive to give a strategic answer. Conversely, when decision makers are responsible for the proposed change, behaviour will seem inconsistent because respondents try to influence the decision. But, this is not a definitive conclusion and this topic certainly requires further research.

As noted by Martínez-Espiñeira and Amoako-Tuffour (2008: 1330) substantial research judgement is necessary in every TCM application. The design of travel cost components is the most important of them. This involves, in particular, the definition of the opportunity cost of time and the definition of the appropriate cost *per* kilometre travelled. In this respect, our proxies were the minimum monthly salary and the reimbursement rate *per* kilometre paid by the Portuguese government. Furthermore, although a considerable effort was made to accurately identify the substitutes, we must recognize that some subjectivity remains. While recognizing that some degree of imprecision is involved, we are in agreement with Hanemann (1994: 38), who, quoting Douglass North, remembers that “the price you pay for precision is an inability to deal with real-world issues”.

Finally, we should identify four main limitations in our study. The first is the absence of non-users in the sample, which was motivated by budget constraints. The same reason is behind the second limitation, which is the sample size, which may be considered as somewhat smaller than desirable. A larger sample would have strengthened confidence in the results. The third limitation is the inability to estimate the economic impact of visits to Bussaco on the local economy. A question was included in the questionnaire with this aim but the response to this particular question was low. Lastly, the analysis would be more complete if data on the cost side was available.

To the best of our knowledge, this dissertation is the first study in Portugal to analyse the demand for forest recreation, combining stated and revealed preferences to examine the effects of forest fires on visits to forests and woodlands. Hence, it partially fills an information gap and is useful to policy-makers. But more knowledge about forest recreation behaviour and preferences is needed. Extra knowledge can be gained using different channels. Three distinct possibilities are: the combination of individual and zonal models[73]; the application of a similar questionnaire to other national forests, and the development of a model covering a set of substitute sites which should be analysed using an RUM able to explain participation decision and site choice. There are three main issues that we became particularly interested in during the research conducted for this dissertation, but they could not be addressed here and have been left for future research. One, of methodological interest, is the analysis of the impact that the use of the internet for survey administration can have on results. This survey mode is recognized to be advantageous with respect to time and budget constraints, but the effect on results is not clear. Another topic of interest is the assessment of the impact that the CB question format can have on results. Two formats have been used indiscriminately without questioning and analysing if they can have distinct effects on the results. The third and final issue is the analysis of survey results to explicitly account for the contributions of the behavioural economics field.

---

[73] Data is being collected with this aim.

APPENDICES

APPENDIX A: QUESTIONNAIRE

I. PERCEPTIONS E PREFERENCES

1. Please express your opinion about environmental protection and preservation of natural spaces in Portugal, indicating your level of agreement concerning the following sentences.

Consider a scale from 1 to 5, with the following interpretation:
1. Totally Disagree, 2. Disagree, 3. Do not disagree or agree, 4. Agree, 5. Totally agree.

	Totally disagree	Disagree	Do not agree or disagree	Agree	Totally agree
1. Environmental protection is very important	1	2	3	4	5
2. As an individual, you can play a role in protecting the environment	1	2	3	4	5
3. A good quality of life requires high environmental quality	1	2	3	4	5
4. Environmental problems have a direct effect on your daily life	1	2	3	4	5
5. Among the main issues in Portugal there are the environmental ones	1	2	3	4	5
6. Environmental protection measures adopted in Portugal are adequate	1	2	3	4	5
7. Natural spaces in Portugal are sufficient	1	2	3	4	5
8. Natural spaces in Portugal are well preserved	1	2	3	4	5
9. Wild life in Portugal is well preserved	1	2	3	4	5
10. Green spaces in Portuguese cities are sufficient	1	2	3	4	5
11. Green spaces in Portuguese cities offer good quality	1	2	3	4	5

2. In order to know your preferences concerning free time occupation, please indicate the recreational activities in which you participate.

Consider a scale from 1 to 5, with the following interpretation:
1. Never, 2. Occasionally, 3. Sometimes, 4. Often, 5. Very often.

	Never	Occasionally	Sometimes	Often	Very often
1. Sports requiring specific infrastructures (ex. soccer, tennis, golf, swimming, aerobic, yoga)	1	2	3	4	5
2. Open air sports not requiring specific infrastructures (e.g. walking, bicycle, jogging)	1	2	3	4	5
3. Nature sports (e.g. pedestrianism, mountaineering, surf, canoeing, BTT, rappel)	1	2	3	4	5
4. Hunting, fishing	1	2	3	4	5
5. Attending cinema or theatre and visiting expositions and museums	1	2	3	4	5
6. Attending sport events	1	2	3	4	5
7. Attending musical concerts	1	2	3	4	5
8. Visits to historic places	1	2	3	4	5
9. Social activities (e.g. picnicking, non-credit educational courses)	1	2	3	4	5
10. Visiting beaches, rivers or lakes	1	2	3	4	5
11. Other. Please specify: ________________________________	1	2	3	4	5

3. From activities above, indicate the 3 favourite.

1st______________________________ 2nd______________________________ 3th______________________________

4. Are you member of any environmentalist group? ☐ No ☐ Yes

5. Are you member of any sport nature group? ☐ No ☐ Yes, Specify the sport:__________________

II. BUSSACO FOREST

6. Is this your first visit to Bussaco forest? ☐ Yes (go to question 9) ☐ No

7. Approximately how many times have you visited Bussaco forest in the past **3 years**, including today?

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ ______

8. How many of those visits were made in Autumn/Winter seasons? _____

9. What is the importance of the following motives in your visits to Bussaco forest?
Consider a scale from 1 to 5, where 1 means "Not Important" e 5 means "Very important".

		Not Important				Very Important
1.	To know the heritage	1	2	3	4	5
2.	To contact with nature	1	2	3	4	5
3.	Socializing with friends/family	1	2	3	4	5
4.	Walking	1	2	3	4	5
5.	Bicycle	1	2	3	4	5
6.	Participating in entertaining activities organized in the forest	1	2	3	4	5
7.	To take photos	1	2	3	4	5
8.	Participating in religious activities (except weddings ceremonies)	1	2	3	4	5
9.	Other. Please specify: ____________________________________	1	2	3	4	5

10. Please, select your degree of agreement with the following statements:

	10.1. Bussaco forest	**Totally disagree**	**Disagree**	**Do not agree or disagree**	**Agree**	**Totally agree**
1.	Is an excellent recreational space	1	2	3	4	5
2.	Is a unique space in the Centro Region	1	2	3	4	5
3.	Is a unique space in the country	1	2	3	4	5
4.	Has remarkable environmental conditions	1	2	3	4	5
5.	Has remarkable architectonic heritage	1	2	3	4	5
6.	Has remarkable religious heritage	1	2	3	4	5
7.	Has trees and plants in good conditions	1	2	3	4	5
8.	Has poorly maintained infrastructures	1	2	3	4	5
9.	Has adequate information and signalling	1	2	3	4	5
10.	Provides the necessary information to the visitors	1	2	3	4	5
11.	Is overcrowded	1	2	3	4	5
12.	Has good parking conditions	1	2	3	4	5
13.	Is a safe place	1	2	3	4	5
14.	Is currently threatened by forest fires	1	2	3	4	5
15.	Is currently threatened by biotic agents	1	2	3	4	5
16.	Is currently threatened by human action	1	2	3	4	5

	10.2. Would you visit Bussaco forest more often if:					
1.	More entertaining activities were organized in the forest (e.g. workshops and ateliers)	1	2	3	4	5
2.	There were better infrastructures (e.g. picnic areas, wc's)	1	2	3	4	5
3.	Walking trails were correctly signalled	1	2	3	4	5
4.	Had a maintenance circuit	1	2	3	4	5
5.	There were bicycle trails available	1	2	3	4	5
6.	The forest area surrounding the forest offers better environmental conditions	1	2	3	4	5
7.	It was a safer place	1	2	3	4	5
8.	Other. Please specify: ____	1	2	3	4	5

11. Do you know some place with characteristics similar to Bussaco forest?

☐ No ☐ Yes. Please, indicate the local: ____

12. What recreational place do you visit more often? ____

III. TRIP AND RELATED EXPENSES

13. What was your transportation mode when travelling to Bussaco forest?

☐ Car ☐ Moto ☐ Bicycle ☐ Bus ☐ Other. Please specify: ________________________

14. Including yourself, what was the number of people in the vehicle when travelling to Bussaco forest? _____

15. How long was your visit to the Bussaco forest?

☐ ≤ 1h ☐ ≤ 2h ☐ ≤ 3h ☐ ≤ 4h ☐ ≤ 5h ☐ > 5h ☐ ___ days

16. Did you visited or will visit other places during this trip?

☐ No (go to question 17)

☐ Yes. **16.1.** The visit to Bussaco forest is the primary objective of the trip? ☐ Yes ☐ No

16.2. Specify the other(s) place(s) you visited or will visit during this trip:

____

16.3. Considering the time spent in the various recreational places (ignore travelling time), specify the fraction of time devoted to Bussaco forest: _______%

16.4. What was the weight of the visit to Bussaco forest in the trip decision? ☐ ______ % ☐ Do not know

17. If current conditions keep up, how many visits would you make to Bussaco forest next year?

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ ______

18. Have you parked inside the forest? ☐ Yes (go to question 20) ☐ No

19. If access fee was half of the current

19.1 Would you park inside the forest? ☐ No (go to question 21) ☐ Yes

19.2 How many visits would you make next year? ________ (go to question 21)

20. How many visits to Bussaco forest would you make next year, if:

20.1. Access fee *per* vehicle is €2.5 ☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ ______

20.2. Access fee *per* vehicle is €10 ☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ ______

21. If a forest fire destroys 25% of Bussaco forest, how many visits would you make next year?

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ ______

22. In order to estimate the contribution of tourism to the local economy, please specify, for each of the following categories, the amount you spent (or predict to spend) in this trip, and the locality where the expense occurred (or is predictable to occur). If you are not travelling alone, specify the group expenses.

Expense	Amount	Locality
1. Expenses made at home related to the trip	€	
2. Gas and oil	€	
3. Toll highway	€	
4. Other transport costs	€	
5. Entrance fee in the forest	€	
6. Entrance fee in the museum/monastery	€	
7. Restaurants	€	
8. Grocery and convenience stores	€	
9. Supermarket (food/beverage)	€	
10. Road maps/ guidebooks	€	
11. Accommodation (hotel, motel, etc.)	€	
12. Camping fees	€	
13. Local products	€	
14. Souvenirs	€	
15. Other. Please specify?______________________________	€	

IV. PERSONAL DATA

23. Gender: ☐ Female ☐ Male

24. Age: _______ years

25. What is the highest level of formal education that you have completed?

☐ Primary/Elementary (4 years or less completed)
☐ 2nd Cycle (6 years completed)
☐ Lower secondary level (9 years completed)
☐ Upper secondary level (12 years completed
☐ Bachelor or university degree
☐ Post graduation degree or a higher degree

26. Please provide the zip code of your home (*or holidays home, if you travelled from there)*

_________-________ ______________________________________

27. How many people are in your household? _____________

28. How many people from your household received income during the past 12 months? _____________

29. What is your household monthly income after taxes?

☐ ≤€499 ☐ €500 to €999 ☐ €1 000 to €1 499 ☐ €1 500 to €1 999 ☐ €2 000 to €2 499 ☐ ≥€2 500

30. Occupation:

☐ Student ☐ Retired ☐ Housekeeper

Profession: ____ ☐ Employed ☐ Unemployed

31. How would you occupy your time had not you visit Bussaco forest?

☐ Would have visited another place. Please specify ____
☐ With another activity. Please specify ____
☐ Do not know.

32. If you want to, could you increase the number of working hours?

☐ No (go to question 33)
☐ Yes **32.1.** Which category best describes the hourly rate you would have received?

☐ €0 ☐ up to €4.99 ☐ €5 to €9.99 ☐ €10 to €14.99
☐ €15 to €19.99 ☐ €20 to €24.99 ☐ €25 a €29.99 ☐ €30 or more

33. Would you like to receive a copy of the results upon completion of the survey?

☐ No ☐ Yes. Please, indicate your e-mail address:____

You can use the space bellow to comments concerning Bussaco forest and/or this questionnaire?

____

____

____

____

Thank you very much for your assistance.

APPENDIX B: NB DISTRIBUTION DENSITY FUNCTION

The function $f(\lambda_i)$ is a mixing distribution for the Poisson. The Gamma distribution is usually chosen because it satisfies some basic requirements:

i) The condition $\lambda > 0$ must be assured since the Poisson distribution is a model of counts and λ is the mean of that distribution;

ii) A flexible function that can describe a wide range of possibilities for λ must be chosen since there is no specific knowledge about λ;

iii) A function complementary to the Poisson distribution is advisable because it allows for the integral computation without the use of very sophisticated integration procedures.

The marginal density of a non-homogeneous Poisson process, when the Gamma distribution is used as the mixing distribution, is the NB density function. A NB constructed in this way is usually called a continuous mixture distribution or a compound Poisson distribution. The developments presented in this appendix are based on a few authors (Greene, 2000: 887; Cameron and Trivedi, 2005: 675; Hilbe, 2010). However, the demonstration presented here is more detailed.

The Gamma distribution is a two-parameter continuous distribution taking the form (Murteira *et al.*, 2002: 242):

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{x^{\alpha-1}\beta^{\alpha}\exp(-\beta x)}{\Gamma(\alpha)} & ,x>0\\ 0 & ,x\le 0\end{cases};\quad \alpha,\beta>0\,. \tag{B.1}$$

If $x \sim Gamma\ (\alpha,\beta)$, $E(x)=\dfrac{\alpha}{\beta}$ and $Var(x)=\dfrac{\alpha}{\beta^2}$. (B.2)

It is assumed that $t|x,\mu \sim Poisson(\mu)$, $\mu=\lambda\upsilon,\ \upsilon>0$, $\lambda=\exp(x'\beta)$ and $\upsilon \sim Gamma\ (\delta,\delta)$ then, $E(\upsilon)=1$ and $Var(\upsilon)=1/\delta$. The marginal distribution of a Poisson-Gamma mixture is:

$$f(t) = \int_0^\infty \frac{\exp(-\lambda\upsilon)(\lambda\upsilon)^t}{t!} \frac{\upsilon^{\delta-1}\delta^\delta \exp(-\delta\upsilon)}{\Gamma(\delta)} d\upsilon = \frac{\delta^\delta \lambda^t}{t!\,\Gamma(\delta)} \int_0^\infty \upsilon^{t+\delta-1} \exp(-(\lambda+\delta)\upsilon) d\upsilon . \qquad \text{(B.3)}$$

Multiplying and dividing (B.3) by the expression $(\lambda+\delta)^{(t+\delta)}\Gamma(t+\delta)$, the equality is preserved:

$$\frac{\delta^\delta \lambda^t}{t!\,\Gamma(\delta)} \int_0^\infty \upsilon^{t+\delta-1} \exp(-(\lambda+\delta)\upsilon) \frac{(\lambda+\delta)^{(t+\delta)}\Gamma(t+\delta)}{(\lambda+\delta)^{(t+\delta)}\Gamma(t+\delta)} d\upsilon . \qquad \text{(B.4)}$$

After rearranging the terms, the equation above becomes:

$$\frac{\delta^\delta \lambda^t\, \Gamma(t+\delta)}{t!\,\Gamma(\delta)(\lambda+\delta)^{(t+\delta)}} \int_0^\infty \frac{\upsilon^{t+\delta-1} + (\lambda+\delta)^{(t+\delta)} \exp(-(\lambda+\delta)\upsilon)}{\Gamma(t+\delta)} d\upsilon . \qquad \text{(B.5)}$$

Since the integral in (B.5) is the integral over the domain of a Gamma distribution with parameters $(t+\delta)$ and $(\lambda+\delta)$, it is equal to 1. The equation (B.5) can be rewritten as:

$$\frac{\Gamma(t+\delta)}{\Gamma(t+1)\Gamma(\delta)} \frac{\delta^\delta}{(\lambda+\delta)^\delta} \frac{\lambda^t}{(\lambda+\delta)^t} = \frac{\Gamma(t+\delta)}{\Gamma(t+1)\Gamma(\delta)} \left(\frac{\delta}{\delta+\lambda}\right)^\delta \left(\frac{\lambda}{\lambda+\delta}\right)^t . \qquad \text{(B.6)}$$

This is the probability function of a variable following a NB distribution, which in statistical text books is usually written as (Murteira *et al.*, 2002: 175):

$$f(t) = \binom{t+r-1}{t} p^r (1-p)^t ; \quad r = \delta, \quad p = \frac{\delta}{\delta+\lambda} . \qquad \text{(B.7)}$$

The standard way of writing is: $t \sim NB(r,p)$, where r is the number of successes and p is the probability of success.

The mean and variance are given by:

$$E(t) = \frac{r(1-p)}{p} = \frac{\delta\left(1-\frac{\delta}{\delta+\lambda}\right)}{\left(\frac{\delta}{\delta+\lambda}\right)} = \lambda , \qquad \text{(B.8)}$$

$$Var(t) = \frac{r(1-p)}{p^2} = \lambda + \frac{1}{\delta}\lambda^2 . \qquad \text{(B.9)}$$

If *t* follows a NB distribution, then $Var(t)=\frac{1}{p}E(t)$. Since $\delta>0$ and $\lambda>0$, the variance exceeds the mean and the model allows for overdispersion.

Different NB regression models can be generated by linking parameters δ and λ to the explanatory variables. A wide range of variance-mean relationships can be obtained by defining (Cameron and Trivedi, 1986):

$$\delta=\frac{1}{\alpha}(\lambda)^k,\quad \alpha>0 \text{ and } k \text{ constant.} \tag{B.10}$$

In accordance:

$$Var(t)=\lambda+\frac{1}{\frac{1}{\alpha}\lambda^k}\lambda^2=\lambda+\alpha(\lambda)^{2-k}. \tag{B.11}$$

Two main parameterizations implying different assumptions about the functional form of heteroscedasticity have been used. The NB2 is obtained when $k=0$ and the NB1 is obtained when $k=1$.

To assure the non-negativity of the mean the usual specification is:

$$E(t)=\lambda=\exp(x'\beta). \tag{B.12}$$

Setting the parameterization $k=0 \Rightarrow \delta=\frac{1}{\alpha}$, we are choosing the NB2, which can be rewritten as:

$$f(t_i)=\frac{\Gamma\left(t_i+\frac{1}{\alpha}\right)}{\Gamma(t_i+1)\Gamma\left(\frac{1}{\alpha}\right)}\left(\frac{\frac{1}{\alpha}}{\frac{1}{\alpha}+\lambda}\right)^{1/\alpha}\left(\frac{\lambda}{\lambda+\frac{1}{\alpha}}\right)^{t_i}. \tag{B.13}$$

This is the NB probability function applied in Chapter 6 developments.

APPENDIX C: CHAPTER 6 STATA COMMANDS AND OUTPUTS

C.1. STATA COMMANDS

- Definition of the log-likelihood of the ESP and instruction for coefficients estimation

```
* Endogenous Stratified Poisson

capture program drop TPoisson
program TPoisson
version 11.2
args lnf xb
local y = "$ML_y1"
tempvar m
quietly generate double `m'=exp(`xb')
quietly replace `lnf'=(`y'-1)*`xb'-`m'-lngamma(`y')
end

************************
*Likelihood maximization
cd "C:\Users\Paula\Desktop\PhD\Stata"
use DadosCh6e7, clear

ml model lf TPoisson (eq1: Ti = TC2C TCs Mpc DD hrs Bc Vw) if (D==1), robust
title(Endogenous Stratified Poisson)
ml check
ml maximize
estat ic
```

- Definition of the log-likelihood of the ESNB and instruction for coefficients estimation

```
* Endogenous Stratified Negative Binomial 2 (on-site sampling)

capture program drop MESNB
program MESNB
version 11.2
args lnf xb alfa
local y = "$ML_y1"
tempvar m
quietly generate double `m'=exp(`xb')
quietly replace `lnf'=lngamma(`y'+(1/`alfa'))-lngamma(`y'+1)-lngamma(1/`alfa')-
(`y'+(1/`alfa'))*ln(1+`alfa'*`m')+`y'*ln(`alfa')+(`y'-1)*ln(`m')+ln(`y')
end

************************
*Likelihood maximization
cd "C:\Users\Paula\Desktop\PhD\Stata"
use DadosCh6e7, clear

ml model lf MESNB (eq1: Ti = TC2C TCs Mpc DD hrs Bc Vw) (alpha:) if (D==1),
robust title (Endogenous Stratified Negative Binomial)
ml check                                       // basta verificar a primeira vez
se usa
ml maximize
estat ic
```

- Definition of the log-likelihood of the ESNBF and instruction for coefficients estimation

```
*Endogenous Stratified Negative Binomial alfa=1.25(on-site sampling)

capture program drop ESNBAfixo
program ESNBAfixo
version 11.2
args lnf xb
local y = "$ML_y1"
tempvar m
quietly generate double `m'=exp(`xb')
quietly replace `lnf'=lngamma(`y'+(1/1.25))-lngamma(`y'+1)-lngamma(1/1.25)-
(`y'+(1/1.25))*ln(1+1.25*`m')+`y'*ln(1.25)+(`y'-1)*ln(`m')+ln(`y') if $ML_samp ==
1
end

************************
*Likelihood maximization
cd "C:\Users\Paula\Desktop\PhD\Stata"
use DadosCh6e7, clear

ml model lf ESNBAfixo (eq1: Ti = TC2C TCs Mpc DD hrs Bc Vw) if (D==1), robust
title (Endogenous Stratified Negative Binomial)
ml check                                     // basta verificar a primeira vez
se usa
ml maximize
estat ic
```

C.2. STATA OUTPUTS

- **POISSON CORRECTED FROM ENDOGENOUS STRATIFICATION**

```
. poisson Tspois TC2C TCs Mpc DD hrs Bc Vw if D==1, vce(robust)
Iteration 0:   log pseudolikelihood = -385.41915
Iteration 1:   log pseudolikelihood = -384.13596
Iteration 2:   log pseudolikelihood = -384.13243
Iteration 3:   log pseudolikelihood = -384.13243

Poisson regression                                Number of obs   =        264
                                                  Wald chi2(7)    =     116.27
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -384.13243                 Pseudo R2       =     0.3142
------------------------------------------------------------------------------
             |               Robust
      Tspois |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0210157   .0059945    -3.51   0.000    -.0327648   -.0092667
         TCs |   .0066616   .0028532     2.33   0.020     .0010694    .0122538
         Mpc |   .0004608   .0002154     2.14   0.032     .0000386    .0008829
          DD |   .6666698   .2054116     3.25   0.001     .2640704    1.069269
         hrs |   .2334645   .0620673     3.76   0.000     .1118147    .3551142
          Bc |  -.4799914   .1807988    -2.65   0.008    -.8343505   -.1256323
          Vw |   1.930858   .3292559     5.86   0.000     1.285528    2.576187
       _cons |  -.7815849   .3509088    -2.23   0.026    -1.469353   -.0938163

. estat ic

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    264   -560.1334   -384.1324      8     784.2649    812.8725
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note

. mfx, eyex
Elasticities after poisson
      y  = Predicted number of events (predict)
         =  .76748554
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    TC2C |  -.6719395      .19166   -3.51   0.000  -1.04759 -.296284   31.9732
     TCs |   .2315049      .09916    2.33   0.020   .037165  .425845   34.7521
     Mpc |   .3004266      .14044    2.14   0.032   .025167  .575686    652.03
      DD |   .0757579      .02334    3.25   0.001   .030008  .121508   .113636
     hrs |   .6942031      .18456    3.76   0.000   .332479  1.05593   2.97348
      Bc |  -.3472665      .13081   -2.65   0.008   -.60364 -.090893   .723485
      Vw |   .2342628      .03995    5.86   0.000   .155968  .312558   .121326
------------------------------------------------------------------------------


. nlcom ((1/_b[Mpc])*ln(1+2.310606*(_b[Mpc])/-_b[TC2C]))

       _nl_1:  (1/_b[Mpc])*ln(1+2.310606*(_b[Mpc])/-_b[TC2C])

------------------------------------------------------------------------------
      Tspois |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   107.2523   29.95646     3.58   0.000     48.53872    165.9659
------------------------------------------------------------------------------
```

```
. nlcom ((1/_b[Mpc])*ln(1+2.310606*(_b[Mpc])/-_b[TC2C]))/3

       _nl_1:  ((1/_b[Mpc])*ln(1+2.310606*(_b[Mpc])/-_b[TC2C]))/3

------------------------------------------------------------------------------
      Tspois |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   35.75077   9.985486     3.58   0.000     16.17957    55.32196
------------------------------------------------------------------------------

. nlcom ((1/_b[Mpc])*ln(1+2.310606*(_b[Mpc])/-_b[TC2C]))/2.310606

       _nl_1:  ((1/_b[Mpc])*ln(1+2.310606*(_b[Mpc])/-_b[TC2C]))/2.310606

------------------------------------------------------------------------------
      Tspois |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   46.41739   12.96476     3.58   0.000     21.00692    71.82786
. nlcom (-(1/_b[Mpc])*ln(1-2.310606*(_b[Mpc])/-_b[TC2C]))

       _nl_1:  -(1/_b[Mpc])*ln(1-2.310606*(_b[Mpc])/-_b[TC2C])

------------------------------------------------------------------------------
      Tspois |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   112.8292   32.96941     3.42   0.001     48.21033     177.448
------------------------------------------------------------------------------

. nlcom (-(1/_b[Mpc])*ln(1-2.310606*(_b[Mpc])/-_b[TC2C]))/3

       _nl_1:  (-(1/_b[Mpc])*ln(1-2.310606*(_b[Mpc])/-_b[TC2C]))/3

------------------------------------------------------------------------------
      Tspois |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   37.60973    10.9898     3.42   0.001     16.07011    59.14935
------------------------------------------------------------------------------

. nlcom (-(1/_b[Mpc])*ln(1-2.310606*(_b[Mpc])/-_b[TC2C]))/2.310606

       _nl_1:  (-(1/_b[Mpc])*ln(1-2.310606*(_b[Mpc])/-_b[TC2C]))/2.310606

------------------------------------------------------------------------------
      Tspois |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   48.83099   14.26873     3.42   0.001      20.8648    76.79719
------------------------------------------------------------------------------
```

- **NB CORRECTED FROM ENDOGENOUS STRATIFICATION**

```
. nbstrat Ti TC2C TCs Mpc DD hrs Bc Vw if D==1

initial:       log likelihood = -466.43877
alternative:   log likelihood = -470.77977
rescale:       log likelihood = -454.10284
rescale eq:    log likelihood = -441.02072
Iteration 0:   log likelihood = -441.02072
Iteration 1:   log likelihood = -348.20149
Iteration 2:   log likelihood = -339.74052
Iteration 3:   log likelihood =   -338.566
Iteration 4:   log likelihood = -338.30128
Iteration 5:   log likelihood = -338.27824
Iteration 6:   log likelihood = -338.27767
Iteration 7:   log likelihood = -338.27767

Negative Binomial with Endogenous Stratification  Number of obs   =        264
                                                  Wald chi2(7)    =     123.22
Log likelihood = -338.27767                       Prob > chi2     =     0.0000
------------------------------------------------------------------------------
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0195517   .0043606    -4.48   0.000    -.0280984   -.0110051
         TCs |   .0065137   .0031944     2.04   0.041     .0002529    .0127746
         Mpc |   .0003164   .0002411     1.31   0.190    -.0001562     .000789
          DD |   .6644273   .2518552     2.64   0.008     .1708002    1.158054
         hrs |   .1677798   .0617781     2.72   0.007     .0466969    .2888628
          Bc |  -.3574178   .1941833    -1.84   0.066      -.73801    .0231745
          Vw |   3.021694   .4038819     7.48   0.000       2.2301    3.813287
       _cons |  -2.762005   1.364006    -2.02   0.043    -5.435408   -.0886019
-------------+----------------------------------------------------------------
    /lnalpha |   1.794417   1.525938     1.18   0.240    -1.196367    4.785201
-------------+----------------------------------------------------------------
       alpha |   6.015968   9.179995                      .3022905    119.7254
------------------------------------------------------------------------------
AIC Statistic   =        2.623                      BIC Statistic =  -1427.443
Deviance        =        0.000                      Dispersion    =      0.000

. estat ic

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    264           .   -338.2777      9     694.5553    726.7389
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note


. generate Tesnb=exp(-0.0195517*TC2C+0.0065137*TCs+0.0003164*Mpc+0.6644273*DD+
0.1677798*hrs -0.3574178*Bc+3.021694*Vw-2.762005) if D==1

Predictnl
Tesnb=exp(_b[TC2C]*TC2C+_b[TCs]*TCs+_b[Mpc]*Mpc+_b[DD]*DD+_b[hrs]*hrs+_b[Bc]*Bc+_
b[Vw]*Vw+_b[_cons]) if D==1

. sum Tesnb

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       Tesnb |       264    .2166073    .3924095   .0059163   3.042902
```

- **NB CORRECTED FROM ENDOGENOUS STRATIFICATION, α=1.25**

```
C:\Users\Paula\Desktop\PhD\Stata

. do DefESNBF

. ml maximize

initial:       log pseudolikelihood =  -712.7001
rescale:       log pseudolikelihood = -472.06088
Iteration 0:   log pseudolikelihood = -472.06088
Iteration 1:   log pseudolikelihood = -341.40191
Iteration 2:   log pseudolikelihood = -340.09005
Iteration 3:   log pseudolikelihood = -340.08809
Iteration 4:   log pseudolikelihood = -340.08809

Endogenous Stratified Negative Binomial          Number of obs   =        264
                                                 Wald chi2(7)    =     125.54
Log pseudolikelihood = -340.08809                Prob > chi2     =     0.0000

------------------------------------------------------------------------------
             |               Robust
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |   -.019902   .0057558    -3.46   0.001    -.0311831   -.0086208
         TCs |   .0065713     .00287     2.29   0.022     .0009462    .0121964
         Mpc |   .0003291   .0001924     1.71   0.087     -.000048    .0007061
          DD |   .6662763   .2061276     3.23   0.001     .2622738    1.070279
         hrs |   .1772471   .0548483     3.23   0.001     .0697464    .2847477
          Bc |  -.3671107   .1848573    -1.99   0.047    -.7294244    -.004797
          Vw |   2.828316   .4218394     6.70   0.000     2.001526    3.655106
       _cons |  -1.610564     .32966    -4.89   0.000    -2.256685   -.9644419
------------------------------------------------------------------------------

. estat ic

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    264           .   -340.0881      8     696.1762    724.7838
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

APPENDIX D: CHAPTER 7 STATA COMMANDS AND OUTPUTS

D.1. STATA COMMANDS

- Definition of the log-likelihood of Probit model and instructions for coefficients estimation

```
* Ordered Probit with 3 categories

capture program drop OPROB
program OPROB
version 11.2
args lnf xb miu1 miu2 miu3
quietly replace `lnf' = ln(normal(`miu1' - `xb')) if $ML_y1==1
quietly replace `lnf' = ln(1 - normal(`miu2' - `xb')) if $ML_y1==3
quietly replace `lnf' = ln(normal(`miu2' - `xb') - normal(`miu1' - `xb')) if
$ML_y1==2
end

*************************************

* Ordered Probit Maximization
cd "C:\Users\Paula\Desktop\PhD\Stata"
use DadosCh6e7, clear
ml model lf OPROB (Eq1: OTib = TC2C TCs Mpc Tipau Ep DD hrs Vw, nocons) (cut1: )
(cut2: ) if (TC1B<150 & Age>15), title (Ordered Probit regression)
*ml check
ml maximize
```

- Definition of the log-likelihood of Logit model and instructions for coefficients estimation

```
* Ordered Logit with 3 categories

capture program drop OLOG
program OLOG
version 11.2
args lnf xb miu1 miu2 miu3
quietly replace `lnf' = ln(invlogit(`miu1' - `xb')) if $ML_y1==1
quietly replace `lnf' = ln(1 - invlogit(`miu2' - `xb')) if $ML_y1==3
quietly replace `lnf' = ln(invlogit(`miu2' - `xb') - invlogit(`miu1' - `xb')) if
$ML_y1==2
end

************************************

* Ordered Logit Maximization
cd "C:\Users\Paula\Desktop\PhD\Stata"
use DadosCh6e7, clear
ml model lf OLOG (Eq1: OTib = TC2C TCs Mpc Tipau Ep DD hrs Vw, nocons) (cut1: )
(cut2: ) if (TC1B<150 & Age>15), title (Ordered Logit regression)
ml check
ml maximize
```

D.2. STATA OUTPUTS

▪ PROBIT

```
. omodel probit OTib TC2C TCs Mpc Tipau Ep DD hrs Vw if (D==1)

Iteration 0:   log likelihood = -229.75057
Iteration 1:   log likelihood = -162.88043
Iteration 2:   log likelihood = -159.63966
Iteration 3:   log likelihood = -159.56723
Iteration 4:   log likelihood = -159.56717

Ordered probit estimates                          Number of obs   =        276
                                                  LR chi2(8)      =     140.37
                                                  Prob > chi2     =     0.0000
Log likelihood = -159.56717                       Pseudo R2       =     0.3055

------------------------------------------------------------------------------
       OTib |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0128181   .0041519    -3.09   0.002    -.0209557   -.0046806
         TCs |   .0054374   .0032739     1.66   0.097    -.0009794    .0118542
         Mpc |   .0005334    .000245     2.18   0.029     .0000532    .0010136
       Tipau |   .2196961   .1237354     1.78   0.076    -.0228209     .462213
          Ep |  -.4375238   .1525043    -2.87   0.004    -.7364268   -.1386208
          DD |    .622529   .2395104     2.60   0.009     .1530972    1.091961
         hrs |    .157581   .0580933     2.71   0.007     .0437203    .2714418
          Vw |   2.746329   .3258277     8.43   0.000     2.107719     3.38494
-------------+----------------------------------------------------------------
       _cut1 |    .999548   .5249144          (Ancillary parameters)
       _cut2 |   2.102851   .5392383
Approximate likelihood-ratio test of equality of coefficients across response
categories:
            chi2(8) =      7.53
          Prob > chi2 =    0.4806

. fitstat

Measures of Fit for oprobit of OTib

Log-Lik Intercept Only:       -229.751   Log-Lik Full Model:           -159.567
D(266):                        319.134   LR(8):                         140.367
                                         Prob > LR:                       0.000
McFadden's R2:                   0.305   McFadden's Adj R2:               0.262
ML (Cox-Snell) R2:               0.399   Cragg-Uhler(Nagelkerke) R2:      0.492
McKelvey & Zavoina's R2:         0.536
Variance of y*:                  2.157   Variance of error:               1.000
Count R2:                        0.761   Adj Count R2:                    0.241
AIC:                             1.229   AIC*n:                         339.134
BIC:                          -1175.892  BIC':                          -95.404
BIC used by Stata:             375.338   AIC used by Stata:             339.134


. test _b[/cut1] = _b[/cut2]

 ( 1)  [cut1]_cons - [cut2]_cons = 0

           chi2(  1) =   67.65
         Prob > chi2 =    0.0000
```

- **LOGIT**

```
. ologit OTib TC2C TCs Mpc Tipau Ep DD hrs Vw if (D==1), nolog

Ordered logistic regression                       Number of obs   =        276
                                                  LR chi2(8)      =     143.40
                                                  Prob > chi2     =     0.0000
Log likelihood = -158.05204                       Pseudo R2       =     0.3121

------------------------------------------------------------------------------
        OTib |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0264995   .0082555    -3.21   0.001    -.0426799    -.010319
         TCs |    .009526   .0057716     1.65   0.099    -.0017861    .0208381
         Mpc |   .0008615   .0004256     2.02   0.043     .0000272    .0016957
       Tipau |   .3838741     .21761     1.76   0.078    -.0426338    .8103819
          Ep |  -.8025128   .2713358    -2.96   0.003    -1.334321   -.2707043
          DD |   1.107536   .4156823     2.66   0.008     .2928134    1.922258
         hrs |   .2966112   .1025982     2.89   0.004     .0955225    .4976999
          Vw |   5.193785   .6627648     7.84   0.000      3.89479     6.49278
-------------+----------------------------------------------------------------
       /cut1 |   1.574645   .9201437                     -.2288034    3.378094
       /cut2 |   3.585103   .9514359                      1.720323    5.449883
------------------------------------------------------------------------------
Approximate likelihood-ratio test of proportionality of odds across response
categories:
         chi2(8) =      9.12
       Prob > chi2 =    0.3322


. fitstat
Measures of Fit for ologit of OTib
------------------------------------------------------------------------------
Log-Lik Intercept Only:       -229.751   Log-Lik Full Model:           -158.052
D(266):                        316.104   LR(8):                         143.397
                                         Prob > LR:                       0.000
McFadden's R2:                   0.312   McFadden's Adj R2:               0.269
ML (Cox-Snell) R2:               0.405   Cragg-Uhler(Nagelkerke) R2:      0.500
McKelvey & Zavoina's R2:         0.557
Variance of y*:                  7.422   Variance of error:               3.290
Count R2:                        0.775   Adj Count R2:                    0.287
AIC:                             1.218   AIC*n:                         336.104
BIC:                         -1178.923   BIC':                          -98.434
BIC used by Stata:             372.308   AIC used by Stata:             336.104
------------------------------------------------------------------------------

. test _b[/cut1] = _b[/cut2]

 ( 1)  [cut1]_cons - [cut2]_cons = 0

           chi2(  1) =   61.15
         Prob > chi2 =    0.0000

. brant
Brant Test of Parallel Regression Assumption
    Variable |      chi2   p>chi2    df
-------------+--------------------------
         All |     11.25    0.188     8
-------------+--------------------------
        TC2C |      0.02    0.878     1
         TCs |      3.25    0.072     1
         Mpc |      0.17    0.684     1
       Tipau |      0.63    0.427     1
          Ep |      2.41    0.120     1
          DD |      0.11    0.735     1
         hrs |      4.62    0.032     1
          Vw |      0.13    0.720     1
----------------------------------------
A significant test statistic provides evidence that the parallel regression
assumption has been violated.
```

```
. margins, dydx(*) predict(outcome(1))

Average marginal effects                          Number of obs   =        276
Model VCE    : OIM
Expression   : Pr(OTib==1), predict(outcome(1))
dy/dx w.r.t. : TC2C TCs Mpc Tipau Ep DD hrs Vw

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |   .0032344   .0009541     3.39   0.001     .0013643    .0051045
         TCs |  -.0011627     .00069    -1.69   0.092    -.0025151    .0001897
         Mpc |  -.0001051   .0000513    -2.05   0.041    -.0002058   -4.53e-06
       Tipau |  -.0468539   .0263274    -1.78   0.075    -.0984547    .0047468
          Ep |   .0979511   .0319716     3.06   0.002     .0352878    .1606143
          DD |  -.1351808   .0493998    -2.74   0.006    -.2320025   -.0383591
         hrs |   -.036203     .01196    -3.03   0.002    -.0596442   -.0127619
          Vw |  -.6339298   .0520481   -12.18   0.000    -.7359423   -.5319173
------------------------------------------------------------------------------

. margins, dydx(*) predict(outcome(2))

Average marginal effects                          Number of obs   =        276
Model VCE    : OIM
Expression   : Pr(OTib==2), predict(outcome(2))
dy/dx w.r.t. : TC2C TCs Mpc Tipau Ep DD hrs Vw

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0013758   .0004115    -3.34   0.001    -.0021824   -.0005692
         TCs |   .0004946   .0002907     1.70   0.089    -.0000752    .0010643
         Mpc |   .0000447   .0000226     1.98   0.048     4.68e-07     .000089
       Tipau |   .0199295   .0114927     1.73   0.083    -.0025958    .0424547
          Ep |  -.0416638   .0144982    -2.87   0.004    -.0700797   -.0132479
          DD |   .0574996   .0225576     2.55   0.011     .0132875    .1017117
         hrs |   .0153991   .0051475     2.99   0.003     .0053102     .025488
          Vw |   .2696443   .0394837     6.83   0.000     .1922576     .347031
------------------------------------------------------------------------------


. margins, dydx(*) predict(outcome(3))

Average marginal effects                          Number of obs   =        276
Model VCE    : OIM

Expression   : Pr(OTib==3), predict(outcome(3))
dy/dx w.r.t. : TC2C TCs Mpc Tipau Ep DD hrs Vw

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0018586   .0005932    -3.13   0.002    -.0030213   -.0006959
         TCs |   .0006681   .0004085     1.64   0.102    -.0001325    .0014688
         Mpc |   .0000604   .0000298     2.03   0.042     2.07e-06    .0001188
       Tipau |   .0269244   .0152299     1.77   0.077    -.0029257    .0567746
          Ep |  -.0562872   .0188739    -2.98   0.003    -.0932795    -.019295
          DD |   .0776812   .0285569     2.72   0.007     .0217107    .1336517
         hrs |   .0208039   .0073224     2.84   0.004     .0064523    .0351555
          Vw |   .3642856    .040778     8.93   0.000     .2843622    .4442089
```

```
. margins, at(Vw=0.5 Mpc=2000) predict(outcome(3)) atmeans

Adjusted predictions                              Number of obs   =        276
Model VCE    : OIM

Expression   : Pr(OTib==3), predict(outcome(3))
at           : TC2C            =    31.60047 (mean)
               TCs             =    35.69735 (mean)
               Mpc             =        2000
               Tipau           =    1.478261 (mean)
               Ep              =     2.46971 (mean)
               DD              =    .1268116 (mean)
               hrs             =    3.014493 (mean)
               Vw              =          .5

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .4644278     .15389     3.02   0.003     .1628089    .7660468
------------------------------------------------------------------------------

. margins, at(Vw=0.5 Mpc=2500) predict(outcome(3)) atmeans

Adjusted predictions                              Number of obs   =        276
Model VCE    : OIM

Expression   : Pr(OTib==3), predict(outcome(3))
at           : TC2C            =    31.60047 (mean)
               TCs             =    35.69735 (mean)
               Mpc             =        2500
               Tipau           =    1.478261 (mean)
               Ep              =     2.46971 (mean)
               DD              =    .1268116 (mean)
               hrs             =    3.014493 (mean)
               Vw              =          .5

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .5715574   .1993954     2.87   0.004     .1807497    .9623651
```

```
. quietly generate Low=1 if OTib==1
. quietly generate Medium=1 if OTib==2
. quietly generate High=1 if OTib==3

. sum p1 p2 p3 Low Medium High if D==1, separator (3)

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
          p1 |       276    .6797613    .3098034   .0094976   .9965835
          p2 |       276    .1935055     .154271   .0029576   .4641597
          p3 |       276    .1267332    .2016941   .0004589   .9331877
-------------+--------------------------------------------------------
         Low |       276    .6847826    .4654464          0          1
      Medium |       276    .1956522    .3974225          0          1
        High |       276    .1195652    .3250418          0          1
```

APPENDIX E: RE-NB DISTRIBUTION FUNCTION

Let t_{ij} be the count for the j^{th} observation in the i^{th} group. Following Hausman *et al.* (1984), it is assumed that $t_{ij}|\gamma_{ij} \sim Poisson(\gamma_{ij})$ and $\gamma_{ij}|\delta_i \sim Gamma\ (\gamma_{ij}, \delta_i)$, with $\lambda_{ij} = \exp(x_{ij}\beta)$, where x_{ij} is the vector of explanatory variables and β is a vector of coefficients. This specification produces a NB distribution for t_{ij}.

The steps which must be followed in order to obtain the probability density function of the NB are the ones presented in Appendix B. For simplicity in the demonstration subscript are omitted.

$$f(t) = \int_0^\infty \frac{\exp(-\gamma)\gamma^t}{t!} \frac{\gamma^{\lambda-1}\delta^\lambda \exp(-\delta\gamma)}{\Gamma(\lambda)} d\gamma = \frac{\delta^\lambda}{t!\Gamma(\lambda)} \int_0^\infty \exp(-\gamma(\delta+1))\gamma^{t+\lambda-1} d\gamma$$

$$f(t) = \frac{\delta^\lambda}{t!\Gamma(\lambda)} \int_0^\infty \exp(-\gamma(\delta+1))\gamma^{t+\lambda-1} \frac{(\delta+1)^{t+\lambda}\Gamma(t+\lambda)}{(\delta+1)^{t+\lambda}\Gamma(t+\lambda)} d\gamma$$

$$f(t) = \frac{\delta^\lambda \Gamma(t+\lambda)}{t!\Gamma(\lambda)(\delta+1)^{t+\lambda}} \int_0^\infty \exp(-\gamma(\delta+1))\gamma^{t+\lambda-1} \frac{(\delta+1)^{t+\lambda}}{\Gamma(t+\lambda)} d\gamma$$

$$f(t) = \frac{\Gamma(\lambda+t)}{\Gamma(\lambda)\Gamma(t+1)} \left(\frac{\delta}{\delta+1}\right)^\lambda \left(\frac{1}{1+\delta}\right)^t .$$

Introducing the subscript in the probability density function of the NB, the equation above becomes:

$$\Pr(T_{ij} = t_{ij}|x_{ij}, \delta_i) = \frac{\Gamma(\lambda_{ij} + t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)} \left(\frac{\delta_i}{1+\delta_i}\right)^{\lambda_{ij}} \left(\frac{1}{1+\delta_i}\right)^{t_{ij}} . \qquad (E.1)$$

In the RE specification, δ_i is assumed to vary randomly across individuals and $\left(\frac{\delta_i}{1+\delta_i}\right)$ is assumed to follow a $Beta(r,s)$ distribution.

Generically, the density function of the $Beta(\alpha,\beta)$ distribution is given by (Murteira *et al.*, 2002: 250):

$f\left(x|\alpha,\beta\right)=\frac{1}{B(\alpha,\beta)}(x)^{\alpha-1}(1-x)^{\beta-1},\quad 0<x<1,\quad \alpha>0,\quad \beta>0$, where B is the Beta function,

$E(x)=\frac{\alpha}{\alpha+\beta}$ and $Var\,(x)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

Replacing the variable x by $\left(\frac{\delta_i}{1+\delta_i}\right)$ and parameters α and β by r and s, respectively, the Beta density function is written as:

$$f\left(\frac{1}{1+\delta_i}\right)=\frac{1}{B(r,s)}\left(\frac{\delta_i}{1+\delta_i}\right)^{r-1}\left(1-\frac{\delta_i}{1+\delta_i}\right)^{s-1}. \tag{E.2}$$

Integrating (E.1) using the Beta density defined in (E.2), as follows:

$$\Pr\left(T_{i1}=t_{i1},\ldots,T_{in_i}=t_{in_i}\middle|X_i\right)=\int_0^\infty\prod_{j=1}^{n_i}\Pr\left(T_{ij}=t_{ij}\middle|x_{ij},\delta_i\right)f(\delta_i)\,d\delta_i$$

$$=\int_0^\infty\prod_{j=1}^{n_i}\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}\left(\frac{\delta_i}{1+\delta_i}\right)^{\lambda_{ij}}\left(\frac{1}{1+\delta_i}\right)^{t_{ij}}\frac{1}{B(r,s)}\left(\frac{\delta_i}{1+\delta_i}\right)^{r-1}\left(1-\frac{\delta_i}{1+\delta_i}\right)^{s-1}d\delta_i$$

$$=\frac{1}{B(r,s)}\prod_{j=1}^{n_i}\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}\int_0^\infty\prod_{j=1}^{n_i}\left(\frac{\delta_i}{1+\delta_i}\right)^{\lambda_{ij}+r-1}\left(\frac{1}{1+\delta_i}\right)^{t_{ij}+s-1}d\delta_i.$$

Taking into consideration that $B(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$,

$$\Pr\left(T_{i1}=t_{i1},\ldots,T_{in_i}=t_{in_i}\middle|X_i\right)=$$

$$=\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}\prod_{j=1}^{n_i}\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}\int_0^\infty\prod_{j=1}^{n_i}\left(\frac{\delta_i}{1+\delta_i}\right)^{\lambda_{ij}+r-1}\left(\frac{1}{1+\delta_i}\right)^{t_{ij}+s-1}d\delta_i. \tag{E.3}$$

Multiplying and dividing (E.3) by $B(\lambda_{ij}+r,t_{ij}+s)$, the equation can be rewritten as:

$$\Pr\left(T_{i1}=t_{i1},\ldots,T_{in_i}=t_{in_i}\middle|X_i\right)=$$

$$=\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}\prod_{j=1}^{n_i}\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}\int_0^\infty\prod_{j=1}^{n_i}\frac{1}{B(\lambda_{ij}+r,t_{ij}+s)}\left(\frac{\delta_i}{1+\delta_i}\right)^{\lambda_{ij}+r-1}\left(\frac{1}{1+\delta_i}\right)^{t_{ij}+s-1}B(\lambda_{ij}+r,t_{ij}+s)d\delta_i$$

$$=\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}\prod_{j=1}^{n_i}\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}B(\lambda_{ij}+r,t_{ij}+s)\int_0^\infty\prod_{j=1}^{n_i}\frac{1}{B(\lambda_{ij}+r,t_{ij}+s)}\left(\frac{\delta_i}{1+\delta_i}\right)^{\lambda_{ij}+r-1}\left(\frac{1}{1+\delta_i}\right)^{t_{ij}+s-1}d\delta_i$$

Since the integral above is the integral over the domain of a Beta distribution with parameters $(\lambda_{ij}+r)$ and $(t_{ij}+s)$, and it is equal to 1. The equation can be rewritten as:

$$\Pr\left(T_{i1}=t_{i1},\ldots,T_{in_i}=t_{in_i}\middle|X_i\right)=$$

$$=\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}\prod_{j=1}^{n_i}\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}\int_0^{\infty}\prod_{j=1}^{n_i}\frac{1}{B(\lambda_{ij}+r,t_{ij}+s)}\left(\frac{\delta_i}{1+\delta_i}\right)^{\lambda_{ij}+r-1}\left(\frac{1}{1+\delta_i}\right)^{t_{ij}+s-1}B(\lambda_{ij}+r,t_{ij}+s)d\delta_i$$

$$=\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}\prod_{j=1}^{n_i}B(\lambda_{ij}+r,t_{ij}+s)\times\prod_{j=1}^{n_i}\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}$$

$$=\frac{\Gamma(r+s)\Gamma\left(r+\sum_{j=1}^{n_i}\lambda_{ij}\right)\Gamma\left(s+\sum_{j=1}^{n_i}t_{ij}\right)}{\Gamma(r)\Gamma(s)\Gamma\left(r+s+\sum_{j=1}^{n_i}\lambda_{ij}+\sum_{j=1}^{n_i}t_{ij}\right)}\prod_{j=1}^{n_i}\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}.$$

The previous equation is the joint probability of the counts for the i^{th} group, hence:

$$\Pr\left(T_{i1}=t_{i1},\ldots,T_{in_i}=t_{in_i}\middle|X_i\right)=\frac{\Gamma(r+s)\Gamma\left(r+\sum_{j=1}^{n_i}\lambda_{ij}\right)\Gamma\left(s+\sum_{j=1}^{n_i}t_{ij}\right)}{\Gamma(r)\Gamma(s)\Gamma\left(r+s+\sum_{j=1}^{n_i}\lambda_{ij}+\sum_{j=1}^{n_i}t_{ij}\right)}\prod_{j=1}^{n_i}\frac{\Gamma(\lambda_{ij}+t_{ij})}{\Gamma(\lambda_{ij})\Gamma(t_{ij}+1)}\ . \quad \text{(E.4)}$$

APPENDIX F: CHAPTER 8 STATA OUTPUTS

▪ MODEL A

Pooled NB

```
. nbreg Ti TC2C TCs DD hrs Ep Dfire TC2C_Dfire if (D==1 & D_25==1 & Obs5om==0 &
Obs2om==0), vce (cluster Obs)

Fitting Poisson model:
Iteration 0:   log pseudolikelihood = -1010.5862
Iteration 1:   log pseudolikelihood = -1010.4906
Iteration 2:   log pseudolikelihood = -1010.4905

Fitting constant-only model:
Iteration 0:   log pseudolikelihood = -937.89022
Iteration 1:   log pseudolikelihood = -937.59479
Iteration 2:   log pseudolikelihood = -937.59473

Fitting full model:
Iteration 0:   log pseudolikelihood = -877.91718
Iteration 1:   log pseudolikelihood = -861.71122
Iteration 2:   log pseudolikelihood = -856.66452
Iteration 3:   log pseudolikelihood = -856.61294
Iteration 4:   log pseudolikelihood = -856.61293

Negative binomial regression                      Number of obs   =        502
Dispersion           = mean                       Wald chi2(0)    =          .
Log pseudolikelihood = -856.61293                 Prob > chi2     =          .

                                    (Std. Err. adjusted for 2 clusters in Obs)
------------------------------------------------------------------------------
             |               Robust
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |   -.009441    .000753   -12.54   0.000    -.0109169   -.0079651
         TCs |   .0075389   .0013058     5.77   0.000     .0049796    .0100982
          DD |   .5341135     .08063     6.62   0.000     .3760816    .6921453
         hrs |   .1901353   .0097304    19.54   0.000      .171064    .2092065
          Ep |   .4411186   .0456014     9.67   0.000     .3517414    .5304957
       Dfire |  -.6789874   .0076674   -88.56   0.000    -.6940152   -.6639597
  TC2C_Dfire |   .0022171   .0002003    11.07   0.000     .0018245    .0026098
       _cons |  -1.997548   .1607997   -12.42   0.000    -2.312709   -1.682386
-------------+----------------------------------------------------------------
    /lnalpha |  -.6480923   .2088661                     -1.057462   -.2387222
-------------+----------------------------------------------------------------
       alpha |   .5230427   .1092459                      .3473361    .7876337
------------------------------------------------------------------------------

. mfx, eyex

Elasticities after nbreg
      y  = Predicted number of events (predict)
         =  1.5402604
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    TC2C |  -.2907148      .02319  -12.54   0.000  -.336161 -.245268   30.7928
     TCs |   .2718658      .04709    5.77   0.000   .179572   .36416   36.0617
      DD |   .0787339      .01189    6.62   0.000   .055438  .102029    .14741
     hrs |   .5741934      .02939   19.54   0.000     .5166  .631787   3.01992
      Ep |   2.100779      .21717    9.67   0.000   1.67513  2.52643   4.76239
   Dfire |  -.3394937      .00383  -88.56   0.000  -.347008  -.33198        .5
TC2C_D~e |   .0341359      .00308   11.07   0.000    .02809  .040181   15.3964
------------------------------------------------------------------------------
```

RE Poisson, Gamma mixture

```
. xtpoisson Ti TC2C TCs DD hrs Ep Dfire TC2C_Df if (D==1 & D_25==1 & Obs5om==0 &
Obs2om==0), re vce(boot, reps(500) seed (10101) nodots)

Random-effects Poisson regression               Number of obs      =       502
Group variable: Inq                             Number of groups   =       251

Random effects u_i ~ Gamma                      Obs per group: min =         2
                                                               avg =       2.0
                                                               max =         2

                                                Wald chi2(7)       =    130.88
Log likelihood  = -842.24495                    Prob > chi2        =    0.0000

                                   (Replications based on 251 clusters in Inq)
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0097326   .0032732    -2.97   0.003    -.0161478   -.0033173
         TCs |    .007442   .0027255     2.73   0.006     .0021002    .0127838
          DD |   .5463572   .2086819     2.62   0.009     .1373483    .9553662
         hrs |   .1851275    .044251     4.18   0.000     .0983971     .271858
          Ep |   .3998224   .1914927     2.09   0.037     .0245037    .7751411
       Dfire |  -.7127152   .1652717    -4.31   0.000    -1.036642   -.3887887
     TC2C_Df |   .0033824   .0038072     0.89   0.374    -.0040797    .0108444
       _cons |  -1.772595   .9328565    -1.90   0.057     -3.60096    .0557704
-------------+----------------------------------------------------------------
    /lnalpha |  -.6849391    .195848                      -1.068794   -.3010841
-------------+----------------------------------------------------------------
       alpha |   .5041209   .0987311                       .3434224    .7400155
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0: chibar2(01) =   336.49 Prob>=chibar2 = 0.000


. mfx, eyex predict (nu0)

Elasticities after bootstrap:xtpoisson
      y  = Predicted number of events (assuming u_i=0) (predict, nu0)
         =  1.5457709
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    TC2C |  -.2996926      .10079   -2.97   0.003  -.497237 -.102148   30.7928
     TCs |   .2683718      .09828    2.73   0.006   .075738  .461006   36.0617
      DD |   .0805387      .03076    2.62   0.009   .020247  .140831    .14741
     hrs |   .5590704      .13363    4.18   0.000   .297151  .820989   3.01992
      Ep |    1.90411      .91196    2.09   0.037   .116696  3.69152   4.76239
   Dfire |  -.3563576      .08264   -4.31   0.000  -.518321 -.194394        .5
 TC2C_Df |   .0520764      .05862    0.89   0.374  -.062813  .166965   15.3964
```

```
. predict T_MA if (D==1 & D_25==1 & Obs5om==0 & Obs2om==0), nu0
(1055 missing values generated)

. sum T_MA if (D==1 & Dfire==1)

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        T_MA |       251    1.301752    .7395878   .3136086    4.73173

. sum T_MA if (D==1 & Dfire==0)

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        T_MA |       251    2.445041    1.475823   .4451776   9.032353
```

```
. nlcom (-1/_b[TC2C])

       _nl_1:  -1/_b[TC2C]

------------------------------------------------------------------------------
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   102.7479   34.55523     2.97   0.003     35.02088    170.4749
------------------------------------------------------------------------------

. nlcom (-2.45/_b[TC2C])

       _nl_1:  -2.45/_b[TC2C]

------------------------------------------------------------------------------
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   251.7324   84.66032     2.97   0.003     85.80117    417.6635
------------------------------------------------------------------------------

. nlcom (-1.30/_b[TC2C])

       _nl_1:  -1.30/_b[TC2C]

------------------------------------------------------------------------------
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   133.5723    44.9218     2.97   0.003     45.52715    221.6174
------------------------------------------------------------------------------

. nlcom (1.15/_b[TC2C])

       _nl_1:  1.15/_b[TC2C]

------------------------------------------------------------------------------
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |  -118.1601   39.73852    -2.97   0.003    -196.0462   -40.27402
```

RE NB

```
. xtnbreg Ti TC2C TCs DD hrs Ep Dfire TC2C_Df if (D==1 & D_25==1 & Obs5om==0 &
Obs2om==0), re

Fitting negative binomial (constant dispersion) model:
Iteration 0:   log likelihood = -1010.5862
Iteration 1:   log likelihood = -1010.4906
Iteration 2:   log likelihood = -1010.4905

Iteration 0:   log likelihood = -952.86292
Iteration 1:   log likelihood = -938.49427
Iteration 2:   log likelihood = -937.59518
Iteration 3:   log likelihood = -937.59473
Iteration 4:   log likelihood = -937.59473

Iteration 0:   log likelihood = -937.59473
Iteration 1:   log likelihood =  -912.6093
Iteration 2:   log likelihood = -896.10215
Iteration 3:   log likelihood = -895.85075
Iteration 4:   log likelihood = -895.85066

Fitting full model:
Iteration 0:   log likelihood = -930.81462
Iteration 1:   log likelihood = -848.76169
Iteration 2:   log likelihood = -836.62222
Iteration 3:   log likelihood = -835.21727
Iteration 4:   log likelihood = -835.19285
Iteration 5:   log likelihood = -835.19284

Random-effects negative binomial regression     Number of obs      =       502
Group variable: Inq                             Number of groups   =       251

Random effects u_i ~ Beta                       Obs per group: min =         2
                                                               avg =       2.0
                                                               max =         2

                                                Wald chi2(7)       =    110.26
Log likelihood  = -835.19284                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0082973   .0027392    -3.03   0.002     -.013666   -.0029287
         TCs |   .0053613   .0022717     2.36   0.018     .0009088    .0098137
          DD |   .4466609   .1557329     2.87   0.004     .1414301    .7518917
         hrs |   .1694195   .0381212     4.44   0.000     .0947032    .2441358
          Ep |   .2963003   .1791843     1.65   0.098    -.0548946    .6474951
       Dfire |  -.6436863   .1129529    -5.70   0.000    -.8650699   -.4223027
     TC2C_Df |   .0021104   .0032681     0.65   0.518    -.0042948    .0085157
       _cons |     .47262   .9029013     0.52   0.601    -1.297034    2.242274
-------------+----------------------------------------------------------------
       /ln_r |   2.711108   .2892871                      2.144116    3.278101
       /ln_s |   .9840582   .1797482                      .6317582    1.336358
-------------+----------------------------------------------------------------
           r |   15.04594   4.352597                      8.534493    26.52534
           s |   2.675291   .4808788                      1.880915    3.805161
------------------------------------------------------------------------------
Likelihood-ratio test vs. pooled: chibar2(01) =   121.32 Prob>=chibar2 = 0.000
```

```
. mfx, eyex predict (nu0)

Elasticities after xtnbreg
      y  = Predicted number of events (assuming u_i=0) (predict, nu0)
         =  1.5705181
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    TC2C |  -.2554981      .08435   -3.03   0.002  -.420813 -.090183   30.7928
     TCs |   .1933356      .08192    2.36   0.018   .032772    .3539   36.0617
      DD |   .0658424      .02296    2.87   0.004   .020848  .110837    .14741
     hrs |   .5116334      .11512    4.44   0.000   .285996  .737271   3.01992
      Ep |   1.411098      .85335    1.65   0.098  -.261429  3.08362   4.76239
   Dfire |  -.3218432      .05648   -5.70   0.000  -.432535 -.211151        .5
 TC2C_Df |   .0324931      .05032    0.65   0.518  -.066125  .131111   15.3964
```

- **MODEL B**

Pooled NB

```
. nbreg Ti TC2C TCs DD hrs Age2 if (D==1 & D_234==1 & Obs2om==0 & Obs3om==0 &
Obs4om==0), vce (cluster Obs)

Fitting Poisson model:

Iteration 0:   log pseudolikelihood =  -1587.283
Iteration 1:   log pseudolikelihood = -1587.1283
Iteration 2:   log pseudolikelihood = -1587.1282

Fitting constant-only model:

Iteration 0:   log pseudolikelihood = -1422.1044
Iteration 1:   log pseudolikelihood =  -1422.104
Iteration 2:   log pseudolikelihood =  -1422.104

Fitting full model:

Iteration 0:   log pseudolikelihood = -1357.0899
Iteration 1:   log pseudolikelihood = -1343.2271
Iteration 2:   log pseudolikelihood = -1342.5564
Iteration 3:   log pseudolikelihood = -1342.5557
Iteration 4:   log pseudolikelihood = -1342.5557

Negative binomial regression                      Number of obs   =        732
Dispersion           = mean                       Wald chi2(0)    =          .
Log pseudolikelihood = -1342.5557                 Prob > chi2     =          .

                                    (Std. Err. adjusted for 3 clusters in Obs)
------------------------------------------------------------------------------
             |               Robust
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0049656   .0014604    -3.40   0.001    -.0078279   -.0021033
         TCs |   .0098979    .000551    17.96   0.000      .008818    .0109778
          DD |   .4925892   .0507985     9.70   0.000     .3930259    .5921525
         hrs |    .193649     .03609     5.37   0.000      .122914    .2643841
        Age2 |   .0001518   .0000285     5.32   0.000     .0000959    .0002078
       _cons |  -.4846242   .3714956    -1.30   0.192    -1.212742    .2434937
-------------+----------------------------------------------------------------
    /lnalpha |  -.4167508   .3146552                     -1.033464    .1999621
-------------+----------------------------------------------------------------
       alpha |   .6591852   .2074161                      .3557725    1.221356
------------------------------------------------------------------------------

. mfx, eyex
Elasticities after nbreg
      y  = Predicted number of events (predict)
         =  1.8049769
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    TC2C |  -.1555747      .04575   -3.40   0.001  -.245253 -.065897   31.3305
     TCs |   .3493423      .01945   17.96   0.000   .311227  .387457   35.2947
      DD |   .0666207      .00687    9.70   0.000   .053155  .080086   .135246
     hrs |   .5976136      .11138    5.37   0.000   .379321  .815907   3.08607
    Age2 |   .2171701      .04083    5.32   0.000    .13715   .29719   1430.45
------------------------------------------------------------------------------
```

RE Poisson, Gamma mixture

```
. xtpoisson Ti TC2C TCs DD hrs Age2 if (D==1 & D_234==1 & Obs2om==0 & Obs3om==0 &
Obs4om==0), re vce(boot, reps(500) seed (10101) nodots)
Random-effects Poisson regression               Number of obs      =       732
Group variable: Inq                             Number of groups   =       244
Random effects u_i ~ Gamma                      Obs per group: min =         3
                                                               avg =       3.0
                                                               max =         3
                                                Wald chi2(5)       =     70.22
Log likelihood  = -1263.0758                    Prob > chi2        =    0.0000
                                   (Replications based on 244 clusters in Inq)
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0067048    .002647    -2.53   0.011    -.0118928   -.0015168
         TCs |   .0090562   .0029439     3.08   0.002     .0032862    .0148262
          DD |   .4883549   .1913569     2.55   0.011     .1133024    .8634075
         hrs |   .1947256   .0375084     5.19   0.000     .1212106    .2682407
        Age2 |   .0001444   .0000576     2.51   0.012     .0000315    .0002572
       _cons |  -.3902431   .1768534    -2.21   0.027    -.7368695   -.0436167
-------------+----------------------------------------------------------------
    /lnalpha |  -.5529884   .1785102                     -.9028621   -.2031148
-------------+----------------------------------------------------------------
       alpha |   .5752282   .1026841                      .4054077    .8161846
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0: chibar2(01) =   648.10 Prob>=chibar2 = 0.000

. mfx, eyex predict (nu0)
Elasticities after bootstrap:xtpoisson
      y  = Predicted number of events (assuming u_i=0) (predict, nu0)
         =  1.8090965
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    TC2C |  -.2100651      .08293   -2.53   0.011  -.372608 -.047522   31.3305
     TCs |    .319636       .1039    3.08   0.002   .115986  .523286   35.2947
      DD |    .066048      .02588    2.55   0.011   .015324  .116772   .135246
     hrs |   .6009361      .11575    5.19   0.000   .374064  .827808   3.08607
    Age2 |   .2065157      .08235    2.51   0.012   .045103  .367928   1430.45
```

```
. predict T_MB if (D==1 & D_234==1 & Obs2om==0 & Obs3om==0 & Obs4om==0), nu0
(825 missing values generated)

. sum T_MB if (D==1 & Obs==2 & Obs2om==0)
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        T_MB |       244    2.097875    1.366846   .4337409   12.61561

. sum T_MB if (D==1 & Obs==3 & Obs3om==0 & Redpreco==1)
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        T_MB |       177    2.162051    1.408002   .4373914   12.82885

. sum T_MB if (D==1 & D_34==1 & Obs3om==0 & Obs4om==0 &  medpreco==1)
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        T_MB |       160     2.20656    1.505392   .4301209   12.40591

. sum T_MB if (D==1 & Obs==4 & Obs4om==0 & duppreco==1)
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        T_MB |       149    1.889092    1.114244    .549911   9.526676
```

RE NB

```
. xtnbreg Ti TC2C TCs DD hrs Age2 if (D==1 & D_234==1 & Obs2om==0 & Obs3om==0 &
Obs4om==0), re

Fitting negative binomial (constant dispersion) model:
Iteration 0:   log likelihood =  -1587.283
Iteration 1:   log likelihood = -1587.1283
Iteration 2:   log likelihood = -1587.1282

Iteration 0:   log likelihood = -1456.6233
Iteration 1:   log likelihood = -1427.7208
Iteration 2:   log likelihood = -1422.1118
Iteration 3:   log likelihood =  -1422.104
Iteration 4:   log likelihood =  -1422.104

Iteration 0:   log likelihood =  -1422.104
Iteration 1:   log likelihood = -1373.3683
Iteration 2:   log likelihood = -1368.5498
Iteration 3:   log likelihood = -1368.5125
Iteration 4:   log likelihood = -1368.5125

Fitting full model:
Iteration 0:   log likelihood = -1364.1448
Iteration 1:   log likelihood = -1276.6376
Iteration 2:   log likelihood = -1260.1096
Iteration 3:   log likelihood = -1258.6514
Iteration 4:   log likelihood = -1258.5893
Iteration 5:   log likelihood = -1258.5889
Iteration 6:   log likelihood = -1258.5889

Random-effects negative binomial regression     Number of obs      =       732
Group variable: Inq                             Number of groups   =       244
Random effects u_i ~ Beta                       Obs per group: min =         3
                                                               avg =       3.0
                                                               max =         3
                                                Wald chi2(5)       =     73.46
Log likelihood  = -1258.5889                    Prob > chi2        =    0.0000
------------------------------------------------------------------------------
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0059124   .0022181    -2.67   0.008    -.0102599    -.001565
         TCs |   .0077802   .0022892     3.40   0.001     .0032934     .012267
          DD |   .4827275   .1597312     3.02   0.003     .1696601    .7957949
         hrs |   .1792824   .0384942     4.66   0.000     .1038351    .2547297
        Age2 |   .0001676   .0000532     3.15   0.002     .0000633    .0002718
       _cons |   1.721803   .4558026     3.78   0.000      .828446    2.615159
-------------+----------------------------------------------------------------
       /ln_r |   2.883915   .3621985                      2.174019    3.593811
       /ln_s |    .746585   .1469763                      .4585167    1.034653
-------------+----------------------------------------------------------------
           r |   17.88415   6.477613                      8.793556    36.37243
           s |   2.109783   .3100882                      1.581726    2.814131
------------------------------------------------------------------------------
Likelihood-ratio test vs. pooled: chibar2(01) =   219.85 Prob>=chibar2 = 0.000


. mfx, eyex predict (nu0)
Elasticities after xtnbreg
      y  = Predicted number of events (assuming u_i=0) (predict, nu0)
         =  1.8032908
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    TC2C |  -.1852398       .0695   -2.67   0.008  -.321449 -.049031   31.3305
     TCs |   .2745996       .0808    3.40   0.001    .11624  .432959   35.2947
      DD |   .0652869       .0216    3.02   0.003   .022946  .107628   .135246
     hrs |   .5532773       .1188    4.66   0.000   .320442  .786113   3.08607
    Age2 |   .2396771      .07611    3.15   0.002   .090506  .388849   1430.45
```

- ## MODEL C

Pooled NB

```
. nbreg Ti TC2C TCs DD hrs Dfire TC2C_Df  Drp TC2C_Drp if (D==1 & Obs2om==0 &
Obs3om==0 & Obs4om==0 & Obs5om==0), vce(robust)

Fitting Poisson model:
Iteration 0:   log pseudolikelihood = -2357.7508
Iteration 1:   log pseudolikelihood = -2357.6039
Iteration 2:   log pseudolikelihood = -2357.6039

Fitting constant-only model:
Iteration 0:   log pseudolikelihood = -2172.5387
Iteration 1:   log pseudolikelihood = -2167.2844
Iteration 2:   log pseudolikelihood = -2167.2738
Iteration 3:   log pseudolikelihood = -2167.2738

Fitting full model:
Iteration 0:   log pseudolikelihood = -2055.3755
Iteration 1:   log pseudolikelihood = -2033.2686
Iteration 2:   log pseudolikelihood = -2032.4218
Iteration 3:   log pseudolikelihood = -2032.4211
Iteration 4:   log pseudolikelihood = -2032.4211

Negative binomial regression                      Number of obs   =       1170
Dispersion           = mean                       Wald chi2(8)    =     210.96
Log pseudolikelihood = -2032.4211                 Prob > chi2     =     0.0000


------------------------------------------------------------------------------
             |               Robust
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0061513   .0019131    -3.22   0.001     -.009901   -.0024016
         TCs |   .0080801   .0015282     5.29   0.000     .0050849    .0110752
          DD |   .4439881   .1046195     4.24   0.000     .2389377    .6490386
         hrs |   .1979064   .0231754     8.54   0.000     .1524835    .2433293
       Dfire |  -.4147856   .1483398    -2.80   0.005    -.7055262   -.1240449
     TC2C_Df |  -.0013071   .0037853    -0.35   0.730     -.008726    .0061119
         Drp |  -.3052891    .173458    -1.76   0.078    -.6452606    .0346823
    TC2C_Drp |    .002205   .0036964     0.60   0.551    -.0050399    .0094499
       _cons |  -.1460985    .115971    -1.26   0.208    -.3733975    .0812005
-------------+----------------------------------------------------------------
    /lnalpha |  -.6322788   .1151908                     -.8580487   -.4065089
-------------+----------------------------------------------------------------
       alpha |   .5313795     .06121                      .4239886    .6659711
------------------------------------------------------------------------------


. mfx, eyex

Elasticities after nbreg
      y  = Predicted number of events (predict)
         =   1.626805
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    TC2C |  -.1915216      .05957   -3.22   0.001  -.308269 -.074775   31.1351
     TCs |   .2869163      .05426    5.29   0.000   .180561  .393272   35.5092
      DD |   .0626137      .01475    4.24   0.000   .033696  .091531   .141026
     hrs |   .6131715       .0718    8.54   0.000   .472438  .753905   3.09829
   Dfire |  -.0829571      .02967   -2.80   0.005  -.141105 -.024809        .2
 TC2C_Df |  -.0080816       .0234   -0.35   0.730  -.053953  .037789   6.18293
     Drp |  -.0610578      .03469   -1.76   0.078  -.129052  .006936        .2
TC2C_Drp |   .0136332      .02285    0.60   0.551  -.031162  .058428   6.18293
```

RE Poisson, Gamma mixture

```
. xtpoisson Ti TC2C TCs DD hrs Dfire TC2C_Df  Drp TC2C_Drp if (D==1 & Obs2om==0 &
Obs3om==0 & Obs4om==0 & Obs5om==0), re vce(boot, reps(
> 500) seed (10101) nodots)

Random-effects Poisson regression               Number of obs      =      1170
Group variable: Inq                             Number of groups   =       234

Random effects u_i ~ Gamma                      Obs per group: min =         5
                                                               avg =       5.0
                                                               max =         5

                                                Wald chi2(8)       =    117.00
Log likelihood  = -1881.9303                    Prob > chi2        =    0.0000

                                   (Replications based on 234 clusters in Inq)
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0074373   .0026619    -2.79   0.005    -.0126546     -.00222
         TCs |   .0073722   .0027929     2.64   0.008     .0018983     .012846
          DD |    .432005   .1871164     2.31   0.021     .0652636    .7987464
         hrs |    .195861   .0381739     5.13   0.000     .1210415    .2706804
       Dfire |  -.4324556   .1444035    -2.99   0.003    -.7154814   -.1494299
     TC2C_Df |   -.000659   .0034155    -0.19   0.847    -.0073532    .0060353
         Drp |  -.3160147   .1564731    -2.02   0.043    -.6226963   -.0093332
    TC2C_Drp |   .0021494   .0032561     0.66   0.509    -.0042324    .0085312
       _cons |  -.0677664   .1658102    -0.41   0.683    -.3927484    .2572156
-------------+----------------------------------------------------------------
    /lnalpha |  -.7772598    .144848                      -1.061157   -.4933629
-------------+----------------------------------------------------------------
       alpha |   .4596639   .0665814                       .3460553    .6105696
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0: chibar2(01) =   951.35 Prob>=chibar2 = 0.000


. mfx, eyex predict (nu0)

Elasticities after bootstrap:xtpoisson
      y  = Predicted number of events (assuming u_i=0) (predict, nu0)
         =  1.6318782
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    TC2C |  -.2315602      .08288   -2.79   0.005  -.394001 -.069119   31.1351
     TCs |   .2617791      .09917    2.64   0.008   .067406  .456152   35.5092
      DD |   .0609238      .02639    2.31   0.021   .009204  .112644   .141026
     hrs |   .6068342      .11827    5.13   0.000   .375022  .838647   3.09829
   Dfire |  -.0864911      .02888   -2.99   0.003  -.143096 -.029886        .2
 TC2C_Df |  -.0040743      .02112   -0.19   0.847  -.045464  .037316   6.18293
     Drp |  -.0632029      .03129   -2.02   0.043  -.124539 -.001867        .2
TC2C_Drp |   .0132895      .02013    0.66   0.509  -.026169  .052748   6.18293
```

RE NB

```
. xtnbreg Ti TC2C TCs DD hrs Dfire TC2C_Df  Drp TC2C_Drp if (D==1 & Obs2om==0 &
Obs3om==0 & Obs4om==0 & Obs5om==0), re

Fitting negative binomial (constant dispersion) model:

Iteration 0:   log likelihood = -2357.7508
Iteration 1:   log likelihood = -2357.6039
Iteration 2:   log likelihood = -2357.6039

Iteration 0:   log likelihood = -2184.2468
Iteration 1:   log likelihood = -2167.8632
Iteration 2:   log likelihood =  -2167.274
Iteration 3:   log likelihood = -2167.2738

Iteration 0:   log likelihood = -2167.2738
Iteration 1:   log likelihood = -2124.5904
Iteration 2:   log likelihood = -2098.4941
Iteration 3:   log likelihood = -2098.3352
Iteration 4:   log likelihood = -2098.3351

Fitting full model:

Iteration 0:   log likelihood = -2046.6592
Iteration 1:   log likelihood = -1904.4806
Iteration 2:   log likelihood =  -1874.313
Iteration 3:   log likelihood = -1872.7602
Iteration 4:   log likelihood = -1872.7161
Iteration 5:   log likelihood = -1872.7159

Random-effects negative binomial regression     Number of obs      =      1170
Group variable: Inq                             Number of groups   =       234

Random effects u_i ~ Beta                       Obs per group: min =         5
                                                               avg =       5.0
                                                               max =         5

                                                Wald chi2(8)       =    111.06
Log likelihood  = -1872.7159                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
          Ti |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        TC2C |  -.0061707   .0020789    -2.97   0.003    -.0102454   -.0020961
         TCs |   .0059145   .0020094     2.94   0.003     .0019761    .0098529
          DD |   .4219776   .1414829     2.98   0.003     .1446762    .6992789
         hrs |   .1566934    .034066     4.60   0.000     .0899252    .2234616
       Dfire |  -.4124927   .0975368    -4.23   0.000    -.6036613   -.2213241
     TC2C_Df |  -.0009088    .002785    -0.33   0.744    -.0063673    .0045498
         Drp |  -.3259551   .0912448    -3.57   0.000    -.5047916   -.1471186
    TC2C_Drp |   .0030124   .0024293     1.24   0.215    -.0017489    .0077738
       _cons |   2.198553   .3013407     7.30   0.000     1.607936     2.78917
-------------+----------------------------------------------------------------
       /ln_r |   3.120895   .2648445                      2.601809     3.63998
       /ln_s |   .9403177   .1227998                      .6996344    1.181001
-------------+----------------------------------------------------------------
           r |   22.66665   6.003137                      13.48812    38.09109
           s |   2.560795   .3144652                      2.013017    3.257633
------------------------------------------------------------------------------
Likelihood-ratio test vs. pooled: chibar2(01) =   451.24 Prob>=chibar2 = 0.000
```

```
. mfx, eyex predict (nu0)

Elasticities after xtnbreg
      y  = Predicted number of events (assuming u_i=0) (predict, nu0)
         =  1.6344342
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    TC2C |  -.1921258      .06473   -2.97   0.003   -.31899 -.065261   31.1351
     TCs |   .2100175      .07135    2.94   0.003   .070168  .349867   35.5092
      DD |   .0595097      .01995    2.98   0.003   .020403  .098616   .141026
     hrs |   .4854817      .10555    4.60   0.000   .278614  .692349   3.09829
   Dfire |  -.0824985      .01951   -4.23   0.000  -.120732 -.044265        .2
 TC2C_Df |   -.005619      .01722   -0.33   0.744  -.039369  .028131   6.18293
     Drp |   -.065191      .01825   -3.57   0.000  -.100958 -.029424        .2
TC2C_Drp |   .0186256      .01502    1.24   0.215  -.010814  .048065   6.18293
```

APPENDIX G: THE LR TEST AND GOODNESS-OF-FIT STATISTICS

- The likelihood ratio (LR) test is one of the classical statistical techniques for testing hypotheses. Let $L(\beta) = f(y|X,\beta)$ denote the likelihood function and the null and alternative hipoteses are:

$$H_0 : g(\beta) = 0$$
$$H_a : g(\beta) \neq 0$$

$LR = -2[\ln(\beta_r) - \ln(\beta_u)] \sim \chi^2(h)$ under H_0

where $\ln(\beta_r)$ is the maximized log-likelihood of the constrained model, the, $\ln(\beta_u)$ is the maximized log-likelihood of the unconstrained models and h is the number of constraints. H_0 is rejected at a significance level α if the computed test statistics exceeds $\chi^2(h;\alpha)$. In the LR test if H_0 is accepted, the constrained and unconstrained maximum of likelihood function should be the same (Cameron and Trivedi, 1998: 44).

- Goodness-of-fit is interpreted as closeness of fitted values to sample value of the dependent variable. This can be measure using the pseudo-R^2 that is an extension of the R^2 to the linear regression model. A higher value is usually preferred. One possible way of calculus is given by:

$$Pseudo\ R^2 = 1 - \frac{\ln L_{fit}}{\ln L_0}$$

- Information criteria are log-likelihood criteria with degrees of freedom adjustment. The model with the smallest information criteria is preferred. Among the most used there are the AIC and the BIC (Cameron and Trivedi, 2005: 278).

– Akaike Information Criterion (AIC)

$AIC = -2\ln L_{fit} + 2p$, where p is the number of unknown parameters.

– Bayes Information Criterion (BIC)

$BIC = -2\ln L_{fit} + p\ln(n)$, where p is the number of unknown parameters and n is the sample dimension

REFERENCES

Adamowicz, W., Boxall, P., Louviere, J., Swait, J., and Williams, M. (1999). Stated-Preference Methods for Valuing Environmental Amenities. In I. Bateman and K. Willis (Eds.), *Valuing Environmental Preferences, Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*. Oxford: Oxford University Press

Adamowicz, W., Boxall, P., Williams, M., and Louviere, J. (1998). Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation. *American Journal of Agricultural Economics,* 80(1), 64-75.

Adamowicz, W., Louviere, J., and Williams, M. (1994). Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities. *Journal of Environmental Economics and Management,* 26(3), 271-292.

Agnello, R., and Han, Y. (1993). Substitute Site Measures in Varying Parameter Model with Application to Recreational Fishing. *Marine Resource Economics,* 8(1), 65-77.

Alberini, A., Zanatta, V., and Rosato, P. (2007). Combining actual and contingent behaviour to estimate the value of sports fishing in the Lagoon of Venice. *Ecological Economics,* 61(2-3), 530-541.

Allen, R. (1990). The Concise Oxford Dictionary of Current English. In R. Allen (Ed.) (Eighth ed.). Oxford: Oxford University Press.

Alriksson, S., and Öberg, T. (2008). Conjoint Analysis for Environmental Evaluation - A review of methods and applications. *Environmental Science and Pollution Research,* 15(3), 244-257.

Álvarez-Farizo, B., Hanley, N., and Barberán, R. (2001). The Value of Leisure Time: A Contingent Rating Approach. *Journal of Environmental Planning and Management,* 44(5), 681-699.

Anderson, D. (2010). Estimating the economic value of ice climbing in Hyalite Canyon: An application of travel cost count data models that account for excess zeros. *Journal of Environmental Management,* 91(4), 1012-1020.

Aparício, A. (2011). *Incêndios Rurais/Florestais na Rede Nacional de Áreas Protegidas 2010*. Lisboa: Instituto de Conservação da Natureza e da Biodiversidade.

Arrow, K., Solow, R., Portney, P., Leamer, E., Radner, R., and Schuman, W. (1993). Report of the National Oceanic and Atmospheric Administration Panel on Contingent Valuation. *Federal Register,* 58(10), 4602-4614.

Azevedo, C., Herriges, J., and Kling, C. (2003). Combining Revealed and Stated Preferences: Consistency Tests and Their Interpretations. *American Journal of Agricultural Economics,* 85(3), 525-537.

Bateman, I., and Langford, I. (1997). Non-users' Willingness to Pay for a National Park: An Application and Critique of the Contingent Valuation Method. *Regional Studies,* 31(6), 571-582.

Beal, D. (1995). A Travel Cost Analysis of the Value of Carnarvon Gorge National Park for Recreational Use. *Review of Marketing and Agricultural Economics,* 63(2), 292-303.

Bell, S., Tyrväinen, L., Sievänen, T., Pröbstl, U., and Simpson, M. (2007). Outdoor Recreation and Nature Tourism: A European Perspective, *Living Rev. Landscape Res.* (Vol. 1). Müncheberg.

Bergstrom, J., Cordell, H., Ashley, G., and Watson, A. (1990). Economic Impacts of Recreational Spending on Rural Areas: A Case Study. *Economic Development Quarterly,* 4(1), 29-39.

Bergstrom, J., Dorfman, J., and Loomis, J. (2004). Estuary Management and Recreational Fishing Benefits. *Coastal Management,* 32(4), 417-432.

Bestard, A., and Font, A. (2010). Estimating the aggregate value of forest recreation in a regional context. *Journal of Forest Economics,* 16(3), 205-216.

Betz, C., Bergstrom, J., and Bowker, J. (2003). A Contingent Trip Model for Estimating Rail-trail Demand. *Journal of Environmental Planning and Management,* 46(1), 79-96.

Bhat, M. (2003). Application of non-market valuation to the Florida Keys marine reserve management. *Journal of Environmental Management,* 67(4), 315-325.

Bishop, K. (1992). Assessing the Benefits of Community Forests: An Evaluation of the Recreational Use Benefits of Two Urban Fringe Woodlands. *Journal of Environmental Planning and Management,* 35(1), 63-76.

Bishop, R. (1982). Option Value: An Exposition and Extension. *Land Economics,* 58(1), 1-15.

Bishop, R. (2003). Where to from here? In P. Champ, K. Boyle and T. Brown (Eds.), *A Primer on Nonmarket Valuation* (Vol. 3, pp. 537-566). Dordrecht: Kluwer Academic Publishers.

Bishop, R., and Heberlein, T. (1979). Measuring Values of Extramarket Goods: Are Indirect Measures Biased? *American Journal of Agricultural Economics,* 61(5), 926-930.

Bockstael, N., and Freeman, A. (2003). Welfare Theory and Valuation. In K.-G. Mäler and J. Vincent (Eds.), *Handbook of Environmental Economics* (Vol. II, pp. 517-570): Elsevier.

Bockstael, N., Hanemann, W., and Kling, C. (1985). Modeling Recreational Demand in a Multiple Site Framework, *AERE Workshop on Recreation Demand Modelling*. Boulder, Colorado.

Bockstael, N., Hanemann, W., and Kling, C. (1987a). Estimating the value of water quality improvements in a recreational demand framework. *Water Resources Research,* 23(5), 951-960.

Bockstael, N., Hanemann, W., and Strand, I. (1989a). *Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models, Vol. I*. Washington, D.C.: United Stated Environmental Protection Agency.

Bockstael, N., Hanemann, W., and Strand, I. (1989b). *Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models, Vol. II*. Washington, D.C.: United Stated Environmental Protection Agency.

Bockstael, N., and McConnell, K. (1980). Calculating Equivalent and Compensating Variation for Natural Resource Facilities. *Land Economics,* 56(1), 56-63.

Bockstael, N., and McConnell, K. (1999). Revisiting the Hedonic Travel Cost Model. College Park: University of Maryland.

Bockstael, N., McConnell, K., and Strand, I. (1989c). Measuring the Benefits of Improvements in Water Quality: The Chesapeake Bay. *Marine Resource Economics,* 6(1), 1-18.

Bockstael, N., and Strand, I. (1987). The Effect of Common Sources of Regression Error on Benefit Estimates. *Land Economics,* 63(1), 11-20.

Bockstael, N., Strand, I., and Hanemann, W. (1987b). Time and the Recreational Demand Model. *American Journal of Agricultural Economics,* 69(2), 293-302.

Bowes, M., and Loomis, J. (1980). A Note on the Use of Travel Cost Models with Unequal Zonal Populations. *Land Economics,* 56(4), 465-470.

Bowker, J., English, D., and Donovan, J. (1996). Toward a value for Guided Rafting on Southern Rivers. *Journal of Agricultural and Applied Economics,* 28(2), 423-432.

Boyce, R., Brown, T., McClelland, G., Peterson, G., and Schulze, W. (1992). An Experimental Examination of Intrinsic Values as a Source of the WTA-WTP Disparity. *The American Economic Review,* 82(5), 1366-1373.

Brainard, J., Lovett, A., and Bateman, I. (1997). Using isochrone surfaces in travel-cost models. *Journal of Transport Geography,* 5(2), 117-126.

Brown, G., and Mendelsohn, R. (1984). The Hedonic Travel Cost Method. *The Review of Economics and Statistics,* 66(3), 427-433.

Brown, R., and Hansen, W. (1974). *A Generalized Recreation Day Use Planning Model*. Sacramento, California: U.S. Army Engineer District, Sacramento.

Brown, T., and Gregory, R. (1999). Why WTA-WTP disparity matters. *Ecological Economics,* 28(3), 323-335.

Brown, W., and Nawas, F. (1973). Impact of Aggregation on the Estimation of Outdoor Recreation Demand Functions. *American Journal of Agricultural Economics,* 53(2), 246-249.

Brown, W., Sorhus, C., and Gibbs, K. (1980). Estimated expenditure by sport anglers and net economic value of salmon and steelhead for specified fisheries in the pacific northwest. Oregon State University: Dep. Agr. and Res. Econ. .

Brox, J., and Kumar, R. (1997). Valuing Campsite Characteristics: A Generalized Travel-Cost Model of Demand for Recreational Camping. *Environmetrics,* 8(2), 87-106.

Buchli, L. (2004). *Valuing quality improvements in environmental goods - A case study for the Ticino River.* Unpublished Dissertation, University of Lugano, Lugano.

Buhyoff, G., Williams, S., and Klemperer, W. (1981). Gravity Model Formulation for an Extensive National Parkway Site. *Environmental Management,* 5(3), 253-262.

Burt, O., and Brewer, D. (1971). Estimation of net social benefits from outdoor recreation. *Econometrica,* 39(5), 813-827.

Cameron, A., and Trivedi, P. (1986). Econometric Models Based on Count Data: Comparisons and Applications of some Estimators and Tests. *Journal of Applied Econometrics,* 1(1), 29-53.

Cameron, A., and Trivedi, P. (1998). *Regression analysis of count data* (Vol. 30). New York: Cambridge University Press.

Cameron, A., and Trivedi, P. (2005). *Microeconometrics, Methods and Applications*. Cambridge Cambridge University Press.

Cameron, A., and Trivedi, P. (2009). *Microeconometrics Using Stata*. College Station,Texas: Stata Press.

Cameron, T. (1992). Combining Contingent Valuation and Travel Cost Data for the Valuation of Nonmarket Goods. *Land Economics,* 68(3), 302-317.

Cameron, T., Shaw, W., Ragland, S., Callaway, J., and Keefe, S. (1996). Using Actual and Contingent Behavior Data with Differing Levels of Time Aggregation to Model Recreation Demand. *Journal of Agricultural and Resource Economics,* 21(1), 130-149.

Carlsson, F. (2010). Design of Stated Preference Surveys: Is There More to Learn from Behavioral Economics? *Environmental and Resource Economics,* 46(2), 167-177.

Carpio, C., Wohlgenant, M., and Boonsaeng, T. (2008). The Demand for Agritourism in the United States. *Journal of Agricultural and Resource Economics,* 33(2), 254-269.

Carson, R., Flores, N., Martin, K., and Wright, J. (1996). Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-Public Goods. *Land Economics,* 72(1), 80-99.

Carson, R., Flores, N., and Meade, N. (2001). Contingent Valuation: Controversies and Evidence. *Environmental and Resource Economics,* 19(2), 173-210.

Carson, R., Hanemann, M., and Mitchell, R. (1986). Determining the demand for public goods by simulating referendums at different tax prices, *Working Paper, Dep Econ.* San Diego: University of California.

Caulkins, P., Bishop, R., and Bouwes, N. (1986). The Travel Cost Model for Lake Recreation: A Comparison of Two Methods for Incorpoating Site Quality and Substitution Effects. *American Journal of Agricultural Economics,* 68(2), 291-297.

Caussade, S., Ortúzar, J., Rizzi, L., and Hensher, D. (2005). Assessing the influence of design dimensions on stated choice experiments estimates. *Transport Research Part B,* 39(7), 621-640.

Cesario, F. (1976). Value of Time in Recreation Benefit Studies. *Land Economics,* 52(1), 32-41.

Cesario, F., and Knetsch, J. (1976). A Recreation Site Demand and Benefit Estimation Model. *Regional Studies,* 10(1), 97-104.

Chakraborty, K., and Keith, J. (2000). Estimating the Recreation Demand and Economic Value of Mountain Biking in Moab, Utah: An Application of Count Data Models. *Journal of Environmental Planning and Management,* 43(4), 461-469.

Champ, P. (2003). Collecting survey data for nonmarket valuation. In P. Champ, K. Boyle and T. Brown (Eds.), *A Primer on Nonmarket Valuation* (Vol. 3, pp. 59-98). Dordrecht: Kluwer Academic Publishers.

Champ, P., Bishop, R., Brown, T., and McCollum, D. (1997). Using Donation Mechanisms to Value Nonuse Benefits from the Public Goods. *Journal of Environmental Economics and Management,* 33(2), 151-162.

Chase, L., Lee, D., Schulze, W., and Anderson, D. (1998). Ecotourism Demand and Differential Pricing of National Park Access in Costa Rica. *Land Economics,* 74(4), 466-482.

Christie, M., Hanley, N., and Hynes, S. (2007). Valuing enhancements to forest recreation using choice experiment and behaviour methods. *Journal of Forest Economics,* 13(2-3), 75-102.

Cicchetti, C., Fisher, A., and Smith, V. (1976). An econometric evaluation of a generalized consumer surplus measure: the mineral king controversy. *Econometrica,* 44(6), 1259-1276.

Clawson, M. (1959). *Methods of measuring the demand for and value of outdoor recreation*. Washington, D.C.: Resources for the Future.

Cooper, J., Hanemann, M., and Signorello, G. (2002). One-and-one-half-bound dichotomous choice contingent valuation. *The Review of Economics and Statistics,* 84(4), 742-750.

Creel, M., and Loomis, J. (1990). Theoretical and Emprirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California. *American Journal of Agricultural Economics,* 71(2), 434-441.

Cruz, A., and Royuela, J. (2009). *Assessing Socio-Economic Benefits of Natura 2000 – a Case Study on the ecosystem service provided by SPA Pico da Vara/Ribeira do Guilherme. Output of the project Financing Natura 2000: Cost estimate and benefits of Natura 2000 (Contract No.: 070307/2007/484403/MAR/B2)*.

Culas, R. (2007). Deforestion and the environmental Kuznets curve: An institutional perspective. *Ecological Economics,* 61(2-3), 429-437.

Cummings, R., and Harrison, G. (1995). The Measurement and Decomposition of Nonuse Values: A Critical Review. *Environmental and Resource Economics,* 5(3), 225-247.

Cunha-e-Sá, M., Madureira, L., Nunes, L., and Otrachshenko, V. (2012). Protesting and Justifying: A Latent Class Model for Contingent Valuation with Attitudinal Data. *Environmental and Resource Economics,* 52(4), 531-548.

Cunha, F., Heckman, J., and Navarro, S. (2007). The Identification and Economics Content of Ordered Choice Models with Stocastic Thresholds. *International Economic Review,* 48(4), 1273-1309.

Curtis, J. (2002). Estimating the Demand for Salmon Angling in Ireland. *The Economic and Social Review,* 33(3), 319-332.

D'Addio, A., Eriksson, T., and Frijters, P. (2007). An analysis of the determinants of job satisfaction when individuals' baseline satisfaction levels may differ. *Applied Economics,* 39(19), 2413 - 2423.

Davis, R. (1964). The Value of Big Game Hunting in a Private Forest, *Transactions of the 29th North American Wildlife and Natural Resources Conference*. Washington, D.C.: Wildlife Management Institute.

Desvousges, W., Smith, V., and McGivney, M. (1983). *A Comparison of Alternative Approaches for Estimating Recreation and Related Benefits of Water Quality Impromments*. Washington, D.C.: U.S Environmental Protection Agency.

Dickie, M. (2003). Defensive behavior and damage cost methods. In P. Champ, K. Boyle and T. Brown (Eds.), *A Primer on Nonmarket Valuation* (Vol. 3, pp. 395-444). Dordrecht: Kluwer Academic Publishers.

Donovan, G., and Champ, P. (2009). The Economic Benefits of Elk Viewing at the Jewell Meadows Wildlife Area in Oregon. *Humand Dimensions of Wildlife,* 14(1), 51-60.

Dwyer, J., Kelly, J., and Bowes, M. (1977). *Improved procedures for valuation of the contribution of recreation to national economic development*. Urbana: Water Resources Center, University of Illinois, Research Report n.º 128

Egan, K., and Herriges, J. (2006). Multivariate count data regression models with individual panel data from an on-site sample. *Journal of Environmental Economics and Management,* 52(2), 567-581.

Englin, J., and Cameron, T. (1996). Augmenting Travel Cost Models with Contingent Behavior Data, Poisson Regression Analysis with Individual Panel Data. *Environmental and Resource Economics,* 7(2), 133-147.

Englin, J., Holmes, T., and Sills, E. (2003). Estimating Forest Recreation Demand Using Count Data Models. In S. E and A. K (Eds.), *Forest in a Market Economy* (pp. 314-359). Dordrecht: Kluwer Academic Publishers.

Englin, J., Loomis, J., and González-Cabán, A. (2001). The dynamic path of recreational values following a forest fire: a comparative analysis of states in the Intermountain West. *Canadian Journal of Forest Research,* 31(10), 1837-1844.

Englin, J., and Shonkwiler, J. (1995). Estimating Social Welfare Using Count Data Models: An Application to Long-Run Recreation Demand Under Conditions of Endogenous Stratification and Truncation. *The Review of Economics and Statistics,* 77(1), 104-112.

English, D., and Bowker, J. (1996). Economic impacts of guided whitewater rafting: a study of five rivers. *Water Resource Bulletin,* 32(6), 1319-1328.

Eom, Y.-S., and Larson, D. (2006). Improving environmental valuation estimates through consistent use of revealed and stated preference information. *Journal of Environmental Economics and Management,* 52(1), 501-516.

Eurobarometer. (2008). *Attitudes of European citizens towards the environment*: European Comission.

Faber, M., Petersen, T., and Schiller, J. (2002). Homo oeconomicus and homo piliticus in Ecological Economics. *Ecological Economics,* 40(3), 323-333.

Feather, P., and Shaw, W. (1999). Estimating the Cost of Leisure Time for Recreation Demand Models. *Journal of Environmental Economics and Management,* 38(1), 49-65.

Figueira, D. (1994). *Método de Avaliação Contingencial - Aplicação à qualidade da água de uma rede de abastecimento público: o caso da Vila do Sardoal.* Unpublished Dissertation, Universidade Técnica de Lisboa, Instituto Superior de Economia e Gestão, Lisboa.

Fix, P., and Loomis, J. (1998). Comparing the Economic Value of Mountain Biking Estimated Using Revealed and Stated Preference. *Journal of Environmental Planning and Management,* 41(2), 227-236.

Fleming, C., and Bowden, M. (2009). Web-based surveys as an alternative to traditional mail methods. *Journal of Environmental Management,* 90(1), 284-292.

Fleming, C., and Cook, A. (2008). The recreational value of Lake McKenzie Island: An application of the travel cost method. *Tourism Management,* 29(6), 1197-1205.

Flores, N. (2003). Conceptual Framework for Nonmarket Valuation. In P. Champ, K. Boyle and T. Brown (Eds.), *A Primer on Nonmarket Valuation* (Vol. 3, pp. 27-58). Dordrecht: Kluwer Academic Publishers.

Freeman, A. (2003a). Economic valuation: what and why. In P. Champ, K. Boyle and T. Brown (Eds.), *A Primer on Nonmarket Valuation* (Vol. 3, pp. 1-25). Dordrecht: Kluwer Academic Publishers.

Freeman, A. (2003b). *The Measurement of Environmental and Resource Values* (Second ed.). Washington, D.C.: Resources for the Future.

Garrod, G., and Willis, K. (1992). The Amenity Value of Woodland in Great Britain: A Comparison of Economic Estimates. *Environmental and Resource Economics,* 2(4), 415-434.

Garrod, G., and Willis, K. (2001). *Economic Valuation of the Environment*. Northampton, MA, USA: Edward Elgar.

Giergiczny, M., Mavsar, R., and Wenchao, Z. (2008). *A new environmental accounting framework using externality data for policy analysis*: European Comission, EXIOPOL project.

Gillig, D., Woodward, R., Ozuna, T., and Griffin, W. (2003). Joint Estimation of Revealed and Stated Preference Data: An Application to Recreational Red Snapper Valuation. *Agricultural and Resource Economics Review,* 32(2), 209-221.

Gintis, H. (2000). Beyond Homo economicus: evidence from experimental economics. *Ecological Economics,* 35(3), 311-322.

González, J., Loomis, J., and González-Caban, A. (2008). A joint estimation method to combine dichotomous choice CVM models with count data models corrected for truncation and endogenous stratification. *Journal of Agricultural and Applied Economics,* 40(2), 681-695.

Greene, W. (2000). *Econometric Analysis* (Fourth ed.). New Jersey: Prince Hall International, Inc.

Greene, W., and Hensher, D. (2010). *Modeling Ordered Choices, A primer*. Cambridge: Cambridge University Press.

Grijalva, T. (2000). *Valuing the Loss in Access: An Institutional and Welfare Analysis of Rock Climbing on U.S. Public Lands.* Unpublished Dissertation, University of New Mexico, Albuquerque, New Mexico.

Grijalva, T., Berrens, R., Bohara, A., and Shaw, W. (2002). Testing the Validity of Contingent Behavior Trip Responses. *American Journal of Agricultural Economics,* 84(2), 401-414.

Grogger, J., and Carson, R. (1991). Models for truncated counts. *Journal of Applied Econometrics* 6(3), 225-238.

Grossmann, M. (2011). Impacts of boating trip limitations on the recreational value of the Spreewald wetland: a pooled revealed/contingent behaviour application of the travel cost method. *Journal of Environmental Planning and Management,* 54(2), 211-226.

Guimarães, P. (2002). The state of Portuguese research in economics: an analysis based on publications in international journals. *Portuguese Economic Journal,* 1(1), 3-25.

Gum, R., and Martin, W. (1975). Problems and Solutions in Estimating the Demand for Value of Rural Outdoor Recreation. *American Journal of Agricultural Economics,* 57(4), 558-566.

Haab, T., and McConnell, K. (1996). Count Data Models and the Problem of Zeros in Recreation Demand Analysis. *American Journal of Agricultural Economics,* 78(1), 89-102.

Hanemann, W. (1984). Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses. *American Journal of Agricultural Economics,* 66(3), 332-341.

Hanemann, W. (1991). Willingness To Pay and Willingness To Accept: How Much Can They Differ? *The American Economic Review,* 81(3), 635-647.

Hanemann, W. (1994). Valuing the Environment Through Contingent Valuation. *Journal of Economic Perspectives,* 8(4), 19-43.

Hanley, N. (1989). Valuing Non-Market Goods Using Contingent Valuation. *Journal of Economic Surveys,* 3(3), 235-252.

Hanley, N., Bell, D., and Alvarez-Farizo, B. (2003). Valuing the Benefits of Coastal Water Quality Improvements Using Contingent and Real Behaviour. *Environmental and Resource Economics,* 24(3), 273-285.

Hanley, N., Koop, G., Álvarez-Farizo, B., Wright, R., and Nevin, C. (2001a). Go climb a mountain: an application of recreation demand modelling to rock climbing in Scotland. *Journal of Agricultural Economics,* 52(1), 36-52.

Hanley, N., Kriström, B., and Shogren, J. (2009). Coherent Arbitrariness: On Value Uncertainty for Environmental Goods. *Land Economics,* 85(1), 41-50.

Hanley, N., Mourato, S., and Wright, R. (2001b). Choice Modelling approaches: a superior alternative for environmental valuations? *Journal of Economic Surveys,* 15(3), 435-462.

Hanley, N., and Shogren, J. (2005). Is Cost–Benefit Analysis Anomaly-Proof? *Environmental and Resource Economics,* 32(1), 13-24.

Hanley, N., and Spash, C. (1993). *Cost-Benefit Analysis and the Environment*. Cheltenham, UK and Brookfield, US: Edward Elgar.

Haspel, A., and Johnson, F. (1982). Multiple Destination Trip Bias in Recreation Benefit Estimation. *Land Economics,* 58(3), 364-372.

Hausman, J., Hall, B., and Griliches, Z. (1984). Econometric models for count data with an application to the patents - R&D relationship. *Econometrica,* 52(4), 909-938.

Heberling, M., and Templeton, J. (2009). Estimating the Economic Value of National Parks with Count Data Models Using On-site, Secondary Data: The Case of the Great Sand Dunes National Park and Preserve. *Environmental Management,* 43(4), 619-627.

Hellerstein, D. (1991). Using Count Data Models in Travel Cost Analysis with Aggregate Data. *American Journal of Agricultural Economics,* 73(3), 860-866.

Hellerstein, D. (1993). Intertemporal Data and Travel Cost Analysis. *Environmental and Resource Economics,* 3(2), 193-207.

Hellerstein, D., and Mendelsohn, R. (1993). A Theoretical Foundation for Count Data Models. *American Journal of Agricultural Economics,* 75(3), 604-611.

Henderson, J., and Allen, D. (1994). The Regional Recreation Demand Model - A Tool for Decisionmaking. *Recnotes,* R-94-1.

Hensher, D., Louviere, J., and Swait, J. (1999). Combining sources of preference data. *Journal of Econometrics,* 89(1-2), 197-221.

Hesseln, H., Loomis, J., and González-Cabán, A. (2004). Comparing the economic effects of fire on hiking demand in Montana and Colorado. *Journal of Forest Economics,* 10(1), 21-35.

Hilbe, J. M. (2010). *Negative Binomial Regression* (Second ed.). Cambridge: Cambridge University Press.

Horne, P., Boxall, P., and Adamowicz, W. (2005). Multiple-use management of forest recreation sites: a spatially explicit choice experiment. *Forest Ecology and Management,* 207(1-2), 189-199.

Horowitz, J., and McConnell, K. (2002). A Review of WTA/WTP studies. *Journal of Environmental Economics and Management,* 44(3), 426-447.

Hotelling, H. (1947). *The economics of public recreation*. Washington, D.C.: National Park Service. Original proposal for TCM.

Huang, J.-C., Haab, T., and Whitehead, J. (1997). Willingness to Pay for Quality Improvements: Should Revealed and Stated Preference Data Be Combined? *Journal of Environmental Economics and Management,* 34(3), 240-255.

Hynes, S., and Hanley, N. (2006). Preservation versus development on Irish rivers: whitewater kayaking and hydro-power in Ireland *Land Use Policy,* 23(2), 170-180.

Hynes, S., Hanley, N., and O'Donoghue, C. (2009). Alternative treatments of the cost of time in recreational demand models: an application to whitewater kayaking in Ireland. *Journal of Environmental Management,* 90(2), 1014-1021.

INE. (2009). Tipologia das Áreas Urbanas. Lisboa: Instituto Nacional de Estatística.

INE. (2011a). Estimativas Anuais da População Residente. Lisboa: Instituto Nacional de Estatística.

INE. (2011b). Inquérito ao Emprego. Lisboa: Instituto Nacional de Estatística.

INE. (2011c). Inquérito às Organizações não Governamentais de Ambiente. Lisboa: Instituto Nacional de Estatísitica.

Jeon, Y., and Herriges, J. (2010). Convergent Validity of Contingent Behavior Responses in Models of Recreation Demand. *Environmental and Resource Economics,* 45(2), 223-250.

Kaoru, Y. (1993). Differentiating use and nonuse values for coastal pond water quality improvements *Environmental and Resource Economics,* 3(5), 487-494.

Kesselring, R., and Bremmer, D. (2011). Setting the target for the federal funds rate: the determinants of Fed behaviour. *Applied Economics,* 43(11), 1341 - 1349.

King, G. (1989). Event Count Models for International Relations: Generalizations and Applications. *International Studies Quarterly,* 33(2), 123-147.

Kling, C. (1988). Comparing Welfare Estimates of Environmental Quality Changes from Recreation Demand Models. *Journal of Environmental Economics and Management,* 15(3), 331-340.

Kling, C., and Crooker, J. (1999). Recreation demand models for environmental valuation. In V. Bergh (Ed.), *Handbook of Environmental and Resource Economics*. Cheltenham: Edward Elgar.

Krutilla, J. (1967). Conservation Reconsidered. *The American Economic Review,* 57(4), 777-786.

Kumar, S., and Kant, S. (2007). Exploded logit modeling of stakeholders' preferences for multiple forest values. *Forest Policy and Economics,* 9(5), 516-526.

Landry, C., and Liu, H. (2009). A semi-parametric estimator for revealed and stated preference data—An application to recreational beach visitation. *Journal of Environmental Economics and Management,* 57(2), 205-218.

Layman, R., Boyce, J., and Criddle, K. (1996). Economic Valuation of the Chinook Salmon Sport Fishery of the Gulkana River, Alaska, under Current and Alternate Management Plans. *Land Economics,* 72(1), 113-128.

León, C. (1996). Double Bounded Survival Values for Preserving the Landscape of Natural Parks. *Journal of Environmental Management,* 46(2), 103-118.

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.

Lienhoop, N., and Ansmann, T. (2011). Valuing water level changes in reservoirs using two stated preference approaches: An exploration of validity. *Ecological Economics,* 70(7), 1250-1258.

Lindhjen, H., and Navrud, S. (2011). Are Internet surveys an alternative to face-to-face interviews in contingent valuation? *Ecological Economics,* 70(9), 1628-1637.

List, J. (2005). Scientific Numerology, Preference Anomalies, and Environmental Policymaking. *Environmental and Resource Economics,* 32(1), 35-53.

Liston-Heyes, C., and Heyes, A. (1999). Recreational benefits from the Dartmoor National Park. *Journal of Environmental Management,* 55(2), 69-80.

Long, J. (1997). *Regression Models for Categorical and Dependent Variables*. Thousand Oaks: Sage Publications.

Loomis, J. (1993). An Investigation into the Reliability of Intended Visitation Behavior. *Environmental and Resource Economics,* 3(1), 183-191.

Loomis, J. (1997). Panel Estimators to Combine Revealed and Stated Preference Dichotomous Choice Data. *Journal of Agricultural and Resource Economics,* 22(2), 233-245.

Loomis, J., and Ekstrand, E. (1998). Alternative approaches for incorporating respondent uncertainty when estimating willingnes to pay: the case of the Mexican spotted owl. *Ecological Economics,* 27(1), 29-41.

Loomis, J., Yorizane, S., and Larson, D. (2000). Testing Significance of Multi-Destination and Multi-PurposeTrip Effects in a Travel Cost Method Demand Model for Whale Watching Trips. *Agricultural and Resource Economics Review,* 29(2), 183-191.

Loureiro, M., and Lotade, J. (2005). Interviewer Effects on the Valuation of Goods with Ethical and Environmental Attributes. *Environmental and Resource Economics,* 30(1), 49-72.

Lourenço, O. (2007). *Unveiling health care consumption groups, a latent class approach in the Portuguese health data context.* Unpublished Dissertation, Universidade de Coimbra, Coimbra.

Machado, F., and Mourato, S. (1998). Improving the assessment of water related health impacts: evidence from costal waters in Portugal, *First World Congress on Environmental and Resource Economics*. Venice.

Mackenzie, J. (1990). Conjoint Analysis of Deer Hunting. *Northeastern Journal of Agricultural and Resource Economics,* 19(2), 109-117.

Mackenzie, J. (1993). A Comparison of Contingent Preferences Models. *American Journal of Agricultural Economics,* 75(3), 593-603.

Maddala, G. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Madureira, L. (2001). *Valorização Económica dos Atributos Ambientais e Paisagísticos através de Escolhas Contingentes: O amendoal tradicional do Douro Superior.* Unpublished Dissertation, Universidade de Trás-os-Montes e Alto Douro, Vila Real.

Madureira, L., Nunes, L., Borges, J., and Falcão, A. (2011). Assessing forest management strategies using a contingent valuation approach and advanced visualisation techniques: A Portuguese case study. *Journal of Forest Economics,* 17(4), 399-414.

Marta-Pedroso, C., Freitas, H., and Domingos, T. (2007). Testing for survey mode effect on contingent valuation data quality: A case study of web based versus in-person interviews. *Ecological Economics,* 62(3-4), 388-398.

Martinez-Alier, J., Munda, G., and O´Neill, J. (1998). Weak comparability of values as a foundation for ecological economics. *Ecological Economics,* 26(3), 277-286.

Martínez-Espiñeira, R., and Amoako-Tuffour, J. (2008). Recreation demand analysis under truncation, overdispersion and endogenous stratification: An application to Gros Morne National Park. *Journal of Environmental Management,* 88(4), 1320-1332.

Martínez-Espiñeira, R., and Amoako-Tuffour, J. (2009). Multi-Destination and Multi-Purpose Trip Effects in the Analysis of the Demand for Trips to a Remote Recreational Site. *Environmental Management,* 43(6), 1146-1161.

McCollum, D., Hellerstein, D., and Peterson, G. (1990). A Reverse Gravity Specification for Travel Cost Model, *Meetings of Western Regional Research Project W-133: Benefits and Cost in Natural Resources Planning* (pp. 45-59). Molokai, Hawaii.

McConnell, K. (1975). Some Problems in Estimating the Demand for Outdoor Recreation. *American Journal of Agricultural Economics,* 57(2), 330-334.

McConnell, K. (1992). On-Site Time in the Demand for Recreation. *Agricultural and Resource Economics Review,* 74(4), 918-925.

McConnell, K., and Strand, I. (1981). Measuring the Cost of Time in Recreation Demand Analysis: An Application to Sportfishing. *American Journal of Agricultural Economics,* 63(1), 153-156.

McKean, J., Johnson, D., and Walsh, R. (1995). Valuing Time in Travel Cost Demand Analysis: An Empirical Investigation. *Land Economics,* 71(1), 96-105.

McKean, J., Walsh, R., and Johnson, D. (1996). Closely Related Good Prices in the Travel Cost Model. *American Journal of Agricultural Economics,* 78(3), 640-646.

McKelvey, R., and Zavoina, W. (1975). A statistical model for the analysis of ordinal dependent variables. *Journal of Mathematical Sociology,* 4(1), 103-120.

Mendelsohn, R. (1985). Modelling the Demand for Outdoor Recreation, *AERE Workshop on Recreation Demand Modelling*. Boulder, Colorado.

Mendes, I. (1996). *O Valor de Recreio das Áreas Protegidas - Uma Aplicação ao Caso do Parque Nacional da Peneda-Gerês.* Unpublished Dissertation, Universidade Técnica de Lisboa, Instituto Superior de Economia e Gestão, Lisboa.

Mendes, I., and Proença, I. (2011). Measuring the Social Recreation Per-Day Net Benefit of the Wildlife Amenities of a National Park: A Count-Data Travel-Cost Approach. *Environmental Management,* 48(5), 920-932.

Mill, G., Rensburg, T., Hynes, S., and Dooley, C. (2007). Preferences for multiple use forest management in Ireland: citizen and consumer perspectives. *Ecological Economics,* 60(3), 642-653.

Mitchell, R., and Carson, R. (1981). *An Experiment in Determining Willingness to Pay for National Water Improvements*. Washington, D.C.: Draft report for U.S. Environmental Protection Agency.

Mitchell, R., and Carson, R. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washignton, D.C.: Resources for the Future.

Mogas, J., Riera, P., and Bennett, J. (2006). A comparison of contingent valuation and choice modelling with second-order interactions. *Journal of Forest Economics,* 12(1), 5-30.

Montgomery, M., and Needelman, M. (1997). The Welfare Effects of Toxic Contamination in Freshwater Fish. *Land Economics,* 73(2), 211-223.

Morais, A. (1994). Dicionário de Inglês-Português, *Dicionários Editora*. Porto: Porto Editora.

Morey, E., Rowe, R., and Watson, M. (1993). A Repeated Nested-Logit Model of Atlantic Salmon Fishing. *American Journal of Agricultural Economics,* 75(3), 578-592.

Murteira, B., Ribeiro, C., Silva, J., and Pimenta, C. (2002). *Introdução à Estatística*. Lisboa: McGraw-Hill.

Nakatani, T., and Sato, K. (2010). Truncation and endogenous stratification in various count data models for recreation demand analysis. *Journal of Development and Agricultural Economics,* 2(1), 293-303.

Norton, B., and Noonan, D. (2007). Ecology and valuation: Big changes needed. *Ecological Economics,* 63(4), 664-675.

Nunes, F. (2004). *Dinâmicas de Pobreza e Eficácia do Sistema de Solidariedade e Segurança Social - Uma Aplicação a Portugal.* Unpublished Dissertation, Instituto Superior de Economia e Gestão, Universidade Técnica de Lisboa, Lisboa.

Nunes, P. (2002a). Measuring the economic benefits of protecting the Parque Natural do Sudoeste Alentejano e Costa Vicentina from commercial tourism development: results from a contingent valuation survey. *Portuguese Economic Journal,* 1(1), 71-87.

Nunes, P. (2002b). Using factor analysis to identify consumer preferences for the protection of a natural area in Portugal. *European Journal of Operational Research,* 140(2), 499-516.

Ojea, E., and Loureiro, M. (2010). Valuing the recovery of overexploited fish stocks in the context of existence and option values. *Marine Policy,* 34(3), 514-521.

Oster, J., Taylor, D., Jacobs, J., and Bradley, E. (1987). *Reservoir Eutrophication and the Value of Recreational Acivities: a Case Study of Flaming Gorge Reservoir*. Laramie, Wyoming: Agricultural Economics Department, University of Wyoming.

Palmquist, R., Phaneuf, D., and Smith, V. (2010). Short Run Constraints and the Increasing Marginal Value of Time in Recreation. *Environmental and Resource Economics,* 46(1), 19-41.

Park, T., Bowker, J., and Leeworthy, V. (2002). Valuing snorkeling to the Florida Keys with stated and revealed preference models. *Journal of Environmental Management,* 65(3), 301-312.

Parsons, G. (2003). The Travel Cost Model. In P. Champ, K. Boyle and T. Brown (Eds.), *A Primer on Nonmarket Valuation* (Vol. 3, pp. 269-329). Dordrecht: Kluwer Academic Publishers.

Parsons, G., Kang, A., Leggett, C., and Boyle, K. (2009). Valuing Beach Closures on the Padre Island National Seashore. *Marine Resource Economics,* 24(2), 213-235.

Parsons, G., and Wilson, A. (1997). Incidental and Joint Consumption in Recreation Demand. *Agricultural and Resource Economics Review,* 26(1), 1-6.

Pascual, U., and Muradian, R. (2010). The economics of valuing ecosystem services and biodiversity. In P. Kumar (Ed.), *The Economics of Ecosystems and Biodiversity - Ecological and Economic Foundations*. London: Earthscan.

Pearce, D., Atkinson, G., and Mourato, S. (2006). *Cost-Benefit Analysis and the Environment: Recent Developments*. Paris: OCDE Publishing.

Pendleton, L. (1999). Reconsidering the hedonic vs. RUM debate in the valuation of recreational environmental amenities. *Resource and Energy Economics,* 21(2), 167-189.

Pereira, H. (2004). *Contribuição para a valorização, geoconservação e gestão da jazida fossilífera de Cacela (Parque Natural da Ria Formosa, Algarve, Portugal).* Unpublished Dissertation, Universidade do Algarve, Faculdade de Ciências do Mar e Ambiente, Faro.

Perman, R., Man, Y., McGilvray, J., and Common, M. (2003). *Natural Resource and Environmental Economics* (Third ed.): Pearson Education.

Perna, F. (1994). *Avaliação de benefícios em economia do ambiente: o caso da Ilha da Culatra.* Unpublished Dissertation, Universidade Técnica de Lisboa, Instituto Superior de Economia e Gestão, Lisboa.

Perna, F. (2001). *Técnicas de Valorização Contingencial Aplicadas ao Ambiente: Estudo de Cenários para a Ilha da Culatra.* Unpublished Dissertation, Universidade do Algarve, Faculdade de Economia, Faro.

Phaneuf, D., Kling, C., and Herriges, J. (2000). Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand. *Review of Economics & Statistics,* 82(1), 83-92.

Phaneuf, D., and Smith, V. (2005). Recreation Demand Models. In K.-G. Mäler and J. Vincent (Eds.), *Handbook of Environmental Economics* (Vol. II): Elsevier.

Plottu, E., and Plottu, B. (2007). The concept of Total Economic Value of environment: A reconsideration within a hierarchical rationality. *Ecological Economics,* 61(1), 52-61.

Portney, P. (1994). The Contingent Valuation Debate: Why Economists Should Care. *Journal of Economic Perspectives,* 8(4), 3-17.

Rae, D. (1981). *Visibility Impairment at Mesa Verde National Park: An Analysis of Benefits and Costs of Controlling Emissions in the Four Corners Area, Research Project 1742.* Palo Alto, California: Electric Power Research Institute.

Randall, A. (1994). A difficulty with Travel Cost Method. *Land Economics,* 70(1), 88-96.

Randall, A., and Stoll, J. (1980). Consumer's Surplus in Commodity Space. *The American Economic Review,* 70(3), 449-455.

Rausch, M., Boxall, P., and Verbyla, A. (2010). The development of fire-induced damage functions for forest recreation activity in Alberta, Canada. *International Journal of Wildland Fire,* 19(1), 63-74.

Ready, R., Navrud, S., Day, B., Dubourg, R., Machado, F., Mourato, S., Spanninks, F., and Rodriquez, M. (2004). Contingent valuation of ill health caused by pollution: testing for context and ordering effects. *Portuguese Economic Journal,* 3(2), 145-156.

Ready, R., Whitehead, H., and Blomquist, G. (1995). Contingent Valuation When Respondents Are Ambivalent. *Journal of Environmental Economics and Management,* 29(2), 181-196.

Ribaudo, M., and Epp, D. (1984). The Importance of Sample Discrimination in Using Travel Cost Method to Estimate the Benefits of Improved Water Quality. *Land Economics,* 60(1), 398-403.

Ribeiro, P. (2002). *Determinação do valor económico da pesca desportiva nas albufeiras do Alentejo: uma aplicação dos métodos do custo de viagem e da avaliação contingente.* Unpublished Dissertation, Universidade Técnica de Lisboa, Instituto Superior de Agronomia, Lisboa.

Rivera, B. (2001). The effects of public health spending on self-assessed health status: an ordered probit model. *Applied Economics,* 33(10), 1313-1319.

Rolfe, J., and Bennett, J. (2009). The impact of offering two versus three alternatives in choice modelling experiments. *Ecological Economics,* 68(4), 1140-1148.

Rosenberger, R., and Loomis, J. (1999). The Value of Ranch Open Space to Tourists: Combining Observed and Contingent Behavior Data. *Growth and Change,* 30(3), 366-383.

Rosenthal, D. (1987). The Necessity for Substitute Prices in Recreation Demand Analyses. *American Journal of Agricultural Economics,* 69(4), 828-836.

Samnaliev, M., Stevens, T., and More, T. (2006). A comparison of alternative certainty calibration techniques in contingent valuation. *Ecological Economics,* 57(3), 507-519.

Samples, K., and Bishop, R. (1985). Estimating the Value of Variation in Anglers' Success Rates: Multiple-Site Travel Cost Method. *Marine Resource Economics,* 2(1), 55-74.

Santos, J. (1998). *The Economic Valuation of Landscape Change: Theory and Policies for Land Use and Conservation*. Cheltenham: Edward Elgar Publishing Limited.

Santos, J. (2002). *Novo Guia Histórico do Buçaco*. Coimbra: Edições Minerva.

Santos, J., Ribeiro, P., Gonçalves, A., and Oliveira, I. (2001). *Estudo estratégico para a gestão das pescas continentais - Relatório Final*. Lisboa: Instituto Superior de Agronomia.

Sattout, E., Talhouk, S., and Caligari, P. (2007). Economic value of cedar in Lebanon: An application of contingent valuation method for conservation. *Ecological Economics,* 61(2-3), 315-322.

Seller, C., Stoll, J., and Chavas, J. (1985). Validation of Empirical Measures of Welfare Change: A Comparison of Nonmarket Techniques. *Land Economics,* 61(2), 156-175.

Shaikh, S., Sun, L., and Kooten, G. (2007). Treating respondent uncertainty in contingent valuation: A comparison of empirical treatments. *Ecological Economics,* 62(1), 115-125.

Shaw, D. (1988). On-site samples' regression, Problems of Non-negative Integers, Truncation, and Endogenous Stratification. *Journal of Econometrics,* 37(2), 211-223.

Shaw, W., and Feather, P. (1999). Possibilities for Including the Opportunity Cost of Time in Recreation Demand Systems. *Land Economics,* 75(4), 592-602.

Shogren, J., Shin, S., Hayes, D., and Kliebenstein, J. (1994). Resolving Differences in Willigness to Pay and Willingness To Accept. *The American Economic Review,* 84(1), 255-270.

Shogren, J., and Taylor, L. (2008). On Behavioral-Environmental Economics. *Review of Environmental Economics and Policy* (1), 26-44.

Shrestha, R., Seidl, A., and Moraes, A. (2002). Value of recreational fishing in Brazilian Pantanal: a travel cost analysis using count data models. *Ecological Economics,* 42(1-2), 289-299.

Shrestha, R., Stein, T., and Clark, J. (2007). Valuing nature-based recreation in public natural areas of the Apalachicola River region, Florida. *Journal of Environmental Management,* 85(4), 977-985.

Simões, P., Cruz, L., and Barata, E. (2012). Non-market recreational value of a national forest: survey design and results, *Discussion Paper 09-2012*: GEMF, FEUC.

Sinden, J. (1974). A Utility Approach to the Valuation of Recreational and Aesthetic Experiences. *American Journal of Agricultural Economics,* 56(1), 61-72.

Smailes, P., and Smith, D. (2001). The growing recreational use of state forest lands in Adelaide hills. *Land Use Policy,* 18(2), 137-152.

Smith, V. (1989). Taking Stock of Progress with Travel Cost Recreation Demand Methods: Theory and Implementation. *Marine Resource Economics,* 6(4), 279-310.

Smith, V., and Desvousges, W. (1985). The Generalized Travel Cost Model and Water Quality Benefits: A Reconsideration. *Southern Economic Journal,* 52(2), 371-381.

Smith, V., Desvousges, W., and McGivney, M. (1983). Estimating Water Quality Benefits: An Econometric Analysis. *Southern Economic Journal,* 50(4), 422-437.

Smith, V., and Kaoru, Y. (1987). The Hedonic Travel Cost Model: A View From the Trenches. *Land Economics,* 63(2), 179-186.

Smith, V., and Kaoru, Y. (1990). Signals or Noise? Explaining the Variation in Recreation Benefit Estimates. *American Journal of Agricultural Economics,* 72(2), 419-433.

Sorg, C., Loomis, J., Donnelly, D., Peterson, G., and Nelson, L. (1985). *Net Economic Value of Cold and Warm Recreational Fishing in Idaho*. Fort Collins, Colorado: United Stated Department of Agriculture.

StataCorp. (2009). *Stata Longitudinal-Data/Panel-Data Reference Manual* (Vol. Release 11). College Station, Texas: Stata Corp LP.

Sugden, R. (2005). Anomalies and Stated Preference Techniques: A Framework for a Discussion of Coping Strategies. *Environmental and Resource Economics,* 32(1), 1-12.

Sutherland, R. (1982a). A Regional Approach to Estimating Recreation Benefits of Improved Water Quality. *Journal of Environmental Economics and Management,* 9(3), 229-247.

Sutherland, R. (1982b). The Sensitivity of Travel Cost Estimates of Recreation Demand to the Functional Form and Definition of Origin Zones. *Western Journal of Agricultural Economics,* 7(1), 87-98.

Sutherland, R. (1983). *A Regional Recreation Demand and Benefits Model*. Los Alamos, New Mexico: US Environmental Protection Agency, Freshwater division, Environmental Research Laboratory.

Tacconi, L. (1995). Rethinking the economic analysis of forests: theory and practice. *Forest Ecology and Management,* 73(1-3), 229-238.

Taylor, L. (2003). The Hedonic Method. In P. Champ, K. Boyle and T. Brown (Eds.), *A Primer on Nonmarket Valuation* (Vol. 3, pp. 331-393). Dordrecht: Kluwer Academic Publishers.

Tietenberg, T. (2003). *Environmental and Natural Resource Economics* (Sixth ed.). Boston: Addison-Wesley, Pearson International Edition.

Torras, M. (2000). The total economic value of Amazonian deforestation, 1978-1993. *Ecological Economics,* 33(2), 283-297.

Trice, A., and Wood, S. (1958). Measurement of Recreational Benefits. *Land Economics,* 34(3), 195-207.

Turner, R., Paavola, J., Cooper, P., Farber, S., Jessamy, V., and Georgiou, S. (2003). Valuing nature: lessons learned and future research directions. *Ecological Economics,* 46(3), 493-510.

Tversky, A., and Kahneman, D. (1991). Loss aversion in riskless choice: a reference-dependent model. *The Quarterly Journal of Economics,* 106(4), 1039-1061.

Varian, H. (2006). Revealed Preference. In M. Szenberg (Ed.), *Samuelsonian Economics and the 21th Century*.

Vaughan, W., and Russell, C. (1982). Valuing a Fishing Day: An Application of a Systematic Varying Parameter Model. *Land Economics,* 58(4), 450-463.

Vaux, H., Gardner, P., and Mills, T. (1984). *Methods for Assessing the Impact of Fire on Forest Recreation*. Berkeley: United States Department of Agriculture.

Venkatachalam, L. (2008). Behavioral economics for environmental policy. *Ecological Economics,* 67(4), 640-645.

Walsh, R., Loomis, J., and Gillman, R. (1984). Valuing Option, Existence and Bequest Demands for Wilderness. *Land Economics,* 60(1), 14-29.

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.

Ward, F. (1987). Economics of Water Allocation to Instream Uses in a Fully Appropriated River Basin: Evidence from a New Mexico Wild River. *Water Resources Research,* 23(3), 381-392.

Ward, F., and Beal, D. (2000). *Valuing Nature With Travel Cost Models, A Manual*. Cheltenham, UK and Northampton, Massachusetts, USA: Edward Elgar.

Ward, F., and Loomis, J. (1986). The Travel Cost Demand Model as an Environmental Policy Assessment Tool: A Review of Literature. *Western Journal of Agricultural Economics,* 11(2), 164-178.

Ward, F., Roach, B., Loomis, J., Ready, R., and Hendersen, J. (1996). *Regional Recreation Demand Models for Large Reservoirs: Database, Development, Modle Estimation and Management Applications* (No. Technical Report R-96-2). Washington, D.C.: U.S. Army Corps of Engineers, Waterways Experiment Station.

Wegge, T., Hanemann, W., and Stand, I. (1986). *An Economic Assessment of Marine Recreational Fishing in Southern California*: Terminal Island CA:US Commerce Department NOAA Tech. Memo.

Weisbrod, B. (1964). Collective-consumption services of individual-consumption goods. *The Quarterly Journal of Economics,* 78(3), 471-477.

Welsh, M., and Poe, G. (1998). Elicitation Effects in Contigent Valuation: Comparisons to a Multiple Bounded Discrete Choice Approach. *Journal of Environmental Economics and Management,* 36(2), 170-185.

Whitehead, J., Haab, T., and Huang, J.-C. (2000). Measuring recreation benefits of quality improvements with revealed and stated behavior data. *Resource and Energy Economics,* 22(4), 339-354.

Whitehead, J., Pattanayak, S., Houtven, G., and Gelso, B. (2008). Combining Revealed and Stated Preference Data to Estimate the Nonmarket Value of Ecological Services: an Assessement of the State of the Science. *Journal of Economic Surveys,* 22(5), 872-908.

Whitehead, J., Phaneuf, D., Dumas, C., Herstine, J., Hill, J., and Buerger, B. (2010). Convergent Validity of Revealed and Stated Recreation Behavior with Quality Change: A Comparison of Multiple and Single Site Demands. *Environmental and Resource Economics,* 45(1), 91-112.

Willig, R. (1976). Consumer's Surplus Without Apology. *The American Economic Review,* 66(4), 589-597.

Willis, K., and Garrod, G. (1991). Valuing Open Access Recreation on Inland Waterways: On-site Recreation Surveys and Selection Effects. *Regional Studies,* 25(6), 511-524.

Winkelmann, R. (2003). *Econometric Analysis of Count Data* (Fourth ed.). Berlin: Springer-Verlag.

Winkelmann, R., and Boes, S. (2005). *Analysis of Microdata*. Berlin: Springer.

Winkelmann, R., and Zimmermann, K. (1995). Recent Developments in Count Data Modelling: Theory and Application. *Journal of Economic Surveys,* 9(1), 1-24.

Yeh, C.-Y., Haab, T., and Sohngen, B. (2006). Modeling multiple-objective recreation trips with choices over trip duration and alternative sites *Environmental and Resource Economics,* 34(2), 189-218.

Zandersen, M. (2005). *Valuing Forest Recreation in Europe: Time and Spatial Considerations.* Unpublished Dissertation, Hamburg University and International Max Planck Research School on Earth System Modelling, Hamburg.

Zandersen, M., and Tol, R. (2009). A meta-analysis of forest recreational values in Europe. *Journal of Forest Economics,* 15(1-2), 109-130.

Zhao, J., and Kling, C. (2001). A new explanation for the WTP/WTA disparity. *Economic Letters,* 73(3), 293-300.