Universidade do Minho
Escola de Economia e Gestão

Tânia Cristina do Cima Gonçalves

Consumer preferences for wine: an application of discrete choice experiments

October 2019

Consumer preferences for wine: an application of discrete choice experiments
Tânia Cristina do Cima Gonçalves
UMinho | 2019

Universidade do Minho
Escola de Economia e Gestão

Tânia Cristina do Cima Gonçalves

Consumer preferences for wine: an application of discrete choice experiments

Doctoral thesis
Ph. D. in Economics

Work conducted under the guidance of
Professor Doctor Lígia Maria Costa Pinto
and
Professor Doctor Lina Sofia de Matos Lourenço Gomes

October 2019

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.



Atribuição-NãoComercial-SemDerivações
CC BY-NC-ND

Acknowledgements

I would like to extend my sincere gratitude to people who generously offered me advice, encouragement, and inspiration during these four years of hard work.

To my advisors, Professor Lígia Pinto and Professor Lina Lourenço-Gomes, I thank them for sharing their knowledge, for the opportunities they have provided me, for their patience, guidance and friendship, and for encouraging me to deepen my research, by helping me find solutions and sharing new ideas.

I thank all Professors, from EEG and FEUC, who taught me during my PhD and who enriched my experience and knowledge.

I owe a deep sense of gratitude to Professor João Rebelo, from the University of Trás-os-Montes and Alto Douro, for his constant motivation and supporting my academic pursuits, and particularly allowing me to attend renowned courses in discrete choice modelling. For this reason, I have the opportunity to thank all professors who taught me in these courses. To Professor Moshe Ben-Akiva and Michel Bierlaire, and Professor Stephane Hess and Dr Thijs Dekker, thank you for listening and advising me on conducting my research. An especial thanks to Dr. David Palma Araneda, from the Choice Modelling Centre – University of Leeds, for sharing his knowledge and with whom I started a collaborative research stage.

To Professor William Greene, who provided assistance with my estimation software and the empirical studies, as well as continuous guidance for my research and clarification of my doubts.

I want to express my highest appreciation for the professors and staff in the Department of Economics, Sociology and Management - UTAD for their words of motivation and friendship during these years.

I would like to thank the entities that supported this work financially, namely the project INNOVINE & WINE – Innovation Platform of Vine & Wine, Portugal (NORTE-01-0145-FEDER-000038) and the Portuguese Foundation for Science and Technology under the projects UID/SOC/04011/2019 and UID/ECO/03182/2019.

To my friends and colleagues who have made this journey more enjoyable. To those who were my classmates and shared the same anxiety, thank you for making this phase less difficult.

Finally, thanks to my family, who provide me continued support, trust and encouragement. To my parents, who instilled a strong work ethic, and for teaching me countless philosophies of life, which during this process proved to be invaluable. To Daniel, for his love, patience, care and dedication.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Abstract

Consumer preferences for wine: an application of discrete choice experiments

The wine industry faces an increasingly competitive environment, and the emergence of new producers and consumers reinforced the global dimension of this market. Despite its traditional nature, this industry is rapidly changing, while wine becomes a particular product linking tradition with innovation. Producers are increasingly searching for new opportunities to differentiate their products, hence understanding consumer preferences is crucial.

The present dissertation represents an insightful contribution to understand wine choice and the decision-making process, in the international market. Moreover, it addresses important methodological issues related to discrete choice experiments and the underlying assumptions that individuals are rational utility maximizers and they choose their preferred option taking into account all available information. However, increasing evidence shows that people use simplifying heuristics to make choices, supporting the alternative of bounded rational behavior which is likely to increase when the good is complex and unfamiliar. The four articles that constitute this dissertation cover some of the challenges that researchers face when valuing a complex good, such as the one under study - wine, by identifying two critical properties of discrete choice models: (i) consumers heterogeneity, both in the scale and in preference, and (ii) the use of heuristics, namely attribute non-attendance, in two different ways, using stated and inferred approaches.

On one hand, results show the importance of accounting for both sources of heterogeneity, while the use of latent class model allows the identification of distinct market segments in terms of preference heterogeneity. On the other hand, it provides evidence of attribute non-attendance both in the stated and inferred approaches, showing biased willingness to pay measures when attribute non-attendance is neglected. However, there is little concordance between the rates of non-attendance, which suggests that respondents stated a higher level of attendance than inferred by the model. Finally, we develop an inferred method in a stepwise approach finding all possible combination patterns between attendance and non-attendance. Results show the potential benefits of using this approach combined with other patterns assuming extreme behavior. The improvements considered in each study lead to a better understanding of consumers' preferences and the choice process in general, showing the adaptability of choice models and their potential as tools for modelling preferences and conduct welfare analysis.

Keywords: attribute non-attendance; consumer preferences; decision making; discrete choice experiments; wine choice

Resumo

Preferências do consumidor por vinho: aplicação da técnica de escolhas discretas

A indústria do vinho enfrenta um ambiente cada vez mais competitivo, pelo que a emergência de novos produtores e consumidores veio reforçar a dimensão global deste mercado. Apesar da sua natureza tradicional, esta indústria tem vindo a mudar rapidamente, enquanto o vinho se torna um produto singular unindo a tradição à inovação. Cada vez mais os produtores procuram novas oportunidades para diferenciar os seus produtos sendo, por isso, crucial conhecer as preferências do consumidor.

A presente tese contribui para a compreensão do processo de decisão e escolha de vinho, no âmbito internacional. Metodologicamente, aborda importantes questões associadas à aplicação da técnica de escolhas discretas e, sobretudo, ao pressuposto de que os indivíduos escolhem a sua alternativa preferida tendo em conta toda a informação disponível. Contudo, evidências crescentes mostram que as pessoas usam estratégias simplificadoras quando escolhem, o que indicia um comportamento racional limitado, suscetível de aumentar quando o bem é complexo e desconhecido. Os quatro artigos que constituem esta tese abordam alguns dos desafios quando se estuda um bem complexo, como o vinho, identificando duas propriedades críticas da técnica de escolhas discretas: (i) tratar a heterogeneidade do consumidor, na consistência (escala) e nas preferências e (ii) incluir o uso de heurísticas, nomeadamente o ignorar certos atributos, através das abordagens declarada e inferida.

Por um lado, os resultados mostram a importância de considerar ambas as fontes de heterogeneidade, enquanto o modelo de classes latentes permite definir segmentos de consumidores de vinho distintos em termos de diferenças preferenciais. Por outro lado, ambas as abordagens, declarada e inferida, evidenciam a omissão de atributos na escolha, gerando medidas de disposição a pagar enviesadas. No entanto, a concordância entre a intensidade deste comportamento é baixa, o que sugere que os respondentes declararam um nível de atenção aos atributos maior do que o inferido pelo modelo. Por fim, desenvolvemos a técnica inferida que identifica, de forma gradual, todas as combinações possíveis entre ignorar/atender atributos. Os resultados mostram os potenciais benefícios desta metodologia combinada com outros padrões de modelos que assumem comportamentos extremos. Os avanços considerados em cada estudo permitem uma melhor compreensão das preferências dos consumidores e do processo de escolha em geral, revelando a capacidade de adaptação dos modelos de escolha e o seu potencial como ferramentas de modelização de preferências e análise de bem-estar.

Palavras-chave: escolha de vinho; heurísticas; preferências do consumidor; técnica de escolhas discretas; tomada de decisão

Contents

Acknowledgements ..... **iii**

Abstract ..... **v**

Resumo ..... **vi**

Contents ..... **vii**

List of tables ..... **x**

List of figures ..... **xii**

Abbreviations ..... **xiii**

Chapter 1 ..... **1**

Introduction 2

1. Theoretical Foundations of Quantitative Models of Discrete Choice Data 4
2. Survey design and data collection 9
3. Aims and thesis outline 11

Chapter 2 ..... **17**

Exploring distinct sources of heterogeneity in discrete choice experiment: An application to wine choice across European consumers 18

1. Introduction 19
2. Sample and methodology 20
   - 2.1. Sample 20
   - 2.2. Methodology 20
3. Empirical results 22
4. Conclusion 27

References 28

Chapter 3 ..... **30**

The Role of Attribute Non-Attendance on Consumer Decision Making: Theoretical Insights and Empirical Evidence 31

1. Introduction 32
2. Heuristics in decision making process 34
3. Dealing with Attribute nonattendance 37
   - 3.1. Identifying ANA 37
   - 3.2. Econometric framework 38

3.2.1. Mixed logit model using stated ANA 38
3.2.2. Model inference 38
3.2.3. Latent class framework 39
3.2.3.4 Models of preference homogeneity but processing heterogeneity 39
3.2.3.5 Models of both preference and processing heterogeneity 41
4. Overview of application 43
4.1. Methods for review 43
4.2. Description of studies 44
4.3. Synthesis of findings 44
4.3.1. Transportation 44
4.3.2. Environmental economics 46
4.3.3. Health economics 49
4.3.4. Food choice 50
5. Concluding remarks 51
References 55
Chapter 4 62
Attribute non-attendance in wine choice: contrasts between stated and inferred approaches 63
1. Introduction 64
2. Survey, sample, and data collection 66
3. Modelling techniques to account for attribute non attendance 67
3.1. Mixed logit model using stated non-attendance 68
3.2. Latent Class model to infer non-attendance 68
4. Results 69
4.1. Stated ANA 69
4.2. Inferred ANA and concordance with the stated approach 74
4.3. Impact on welfare estimates 77
5. Conclusion 79
References 80
Chapter 5 84
Dealing with ignored attributes through an inferred approach in wine choice experiments 85
1. Introduction 86
2. Survey, sample, and data collection 88

3. Methods 89
3.1. Latent Class Model to account for attribute non-attendance 89
3.2. Exploring ANA patterns through different class models 90
4. Results 91
4.1. ANA patterns 91
4.2. Impact of ANA on model estimates and WTP 94
5. Conclusion 96
References 97
Chapter 6 101
Discussion and conclusion 102
Appendices 110

List of tables

Chapter 1

Table 1 - Attributes and levels used in the study 10

Table 2 - Overview of the studies performed in this thesis 12

Chapter 2

Table 1 - Attributes and levels 20

Table 2 - MNL, SMNL, MIXL, GMXL model results 24

Table 3 - LCM results 25

Table 4 - WTP for wine attributes 26

Chapter 3

Table 1 - Classification scheme to deal with ANA 43

Chapter 4

Table 1 - Attributes and levels 66

Table 2 - An overview of self-reported attribute non-attendance 70

Table 3 - MNL and MIXL model results with all attributes assumed to be attended 72

Table 4 - MNL and MIXL model results with attribute non-attendance 73

Table 5 - ECLC model results with attribute non-attendance 74

Table 6 - Attribute non-attendance patterns based on ECLC model results and contrasts with stated ANA 76

Table 7 - WTP estimates, in US$ 78

Chapter 5

Table 1 - Attributes and levels used in the choice sets 88

Table 2 - Preliminary results of LCM 92

Table 3 - Class memberships from the different ECLC models 93

Table 4 - Estimates for the MNL and ECLC models 95

Table 5 - Willingness to pay estimates for wine, for each model, in euros 96

Chapter 6

Table 1 - Overview of the main findings of the thesis 108

Appendices

Appendix B

Table 1 - Summary statistics of European sample descriptive variables 130

Appendix C

Table 1 - DCE studies on ANA 131

Appendix D

Table 1 - Summary statistics of descriptive variables 135
Table 2 - Attitudinal and perceptual variables towards wine 137
Table 3 - Self-reported attribute non-attendance for all ANA combinations 139

Appendix E

Table 1 - ANA patterns tested in each model and average class membership probabilities 142
Table 2 - Nlogit output for predicted choices 145

List of figures

Chapter 4

Figure 1 - Choice set example 67

Chapter 5

Figure 1 - Choice set example 89

Abbreviations

AA	Attribute Attendance
AIC	Akaike Information Criterion
ANA	Attribute Non-Attendance
ASC	Alternative Specific Constant
BIC	Bayes Information Criterion
COO	Country of Origin
CV	Coefficient of Variation
DCE	Discrete Choice Experiments
DM	Discrete Mixture
EAA	Endogenous Attribute Attendance
ECLC	Equality Constrained Latent Class
ECLM	Error Component Logit Model
GMXL	Generalized Mixed Logit
HCM	Hybrid Choice Model
HH	Hess and Hensher (2010) method
LC	Latent Class
LCM	Latent Class Model
LL	Log Likelihood
MEAA	Mixed Endogenous Attribute Attendance
MIXL	Mixed Logit
MMNL	Mixed Multinomial Logit
MNL	Multinomial Logit

RP	Random Parameter
RPANA	Random Parameters Attribute Non-Attendance
RPL	Random Parameters Logit
RPLC	Random Parameters Latent Class
RUT	Random Utility Theory
SC	Stated Choice
SMNL	Scaled Multinomial Logit Model
SP	Stated Preference
SQ	Status Quo
VTTS	Value of Travel Time Savings
WTP	Willingness to Pay

Chapter 1

Introduction

Discrete choice experiments (DCE) is a stated preference method to elicit decision-maker preferences, by modelling its choice among a set of mutually exclusive alternatives. These methods have strong foundations in economic theory, and assume that individuals are rational utility maximizers. Nevertheless, some behaviors appear to violate the microeconomic foundations of discrete choice analysis. In an attempt to address these irregularities, much work has been developed to improve discrete choice models, adding extensions that relax restrictive assumptions, and provide more realistic representations of the choice behavior process. One of the major deviations, the bounded rationality (Simon, 1955), routed in psychology, defines decision making processes based on the utility maximizing compensatory framework. The concept of "rationality" is the core difference. While in economics rationality refers to the choice outcome, in other social sciences it refers to the processes employed (Simon, 1982). The last definition leads to a different assumption in terms of information use in DCE, by recognizing the use of decision heuristics and simplifying strategies by respondents.

Given its versatility, choice models have been applied in several research contexts, including transportation, economics, public policy, among others, with contributions to a better understanding of behavioral hypotheses and also to the definition of policy and marketing strategies. In the wine economics field, the use of DCE to analyze wine consumer preferences and behavior has been increasing in the last years (Lockshin, Corsi, Cohen, Lee, & Williamson, 2017; Mueller, Lockshin, Saltman, & Blanford, 2010; Perrouty, D´Hauteville, & Lockshin, 2006).

The consumption of wine was established over many years, with the world wine production being dominated by European countries, such as France, Spain, Italy and Portugal, known as Old World. This pattern has recently been changing due to the emergence of new producing countries from the New World, such as Australia, Chile, USA (California), South Africa, among others. The new dynamics of global economic arena have dictated changing international trade patterns, with the economic power shifting from the traditional wine markets to emerging ones. This growing presence of emerging producers, coupled with a globalized context of persistent decrease in the volume of the overall consumption, have encouraged the Old World to seek out its share, and consequently demanded the understanding of the drivers of wine choice for wine companies and marketing practitioners.

Wine is considered a difficult and confusing product to choose, as the number of cues become vast when compared to other products (Lockshin, Jarvis, D´Hauteville, & Perrouty, 2006). There are, not only

different wine types, such as red, white, sparkling, among others, but also the number of cues remain large within these categories. Intrinsic cues are related to the physical–chemical characteristics of wine, such as grape variety, wine style, and alcohol content, while the extrinsic cues are those that are not physically part of the product, such as brand, price, packaging, among others. Thus, the perception of quality of wine depends on many distinct attributes, which makes the taste of wine different, even when the extrinsic characteristics remains the same (Lockshin, Jarvis, D´Hauteville, & Perrouty, 2006). Moreover, as wine is an experience good, its quality, or taste, cannot be assessed prior to consumption. Therefore, wine consumers will make their choices based on the extrinsic cues such as those presented on the bottle label, or derived from previous experiences.

However, not all information provided is used by consumers to make a choice. As reported by many studies, price is repeatedly one of the main key drivers of wine choice and it is used as an indicator of quality (Corsi, Mueller, & Lockshin, 2012; Lockshin et al., 2006; Mtimet & Albisu, 2006; Mueller, Lockshin, et al., 2010; Mueller, Osidacz, Francis, & Lockshin, 2010). Medals awarded are also perceived as signals of quality, and several studies show their importance for consumers' choice (Corsi et al., 2012; Lockshin, Cohen, & Goodman, 2009). The origin of wine is also a major cue for wine choice, being well documented, either in the case of a region or country (Kallas, Escobar, & Gil, 2013; Lockshin et al., 2017; Tang, Tchetchik, & Cohen, 2015). A nation's image, composed of contemporary and historical associations, is indeed a choice factor considered when taking a wine buying decision. Given the changes in the international wine environment due to the emergence of new producing countries, there is a clear division between the Old World and New World wines. Presently in any given market, there are wines coming from countries with a long-documented history of wine production (primarily wine made in Europe), and wines from the new wine regions which have no tradition in wine production. Other studies found the grape varietals as key choice drivers, particularly for the New World wines who construct their reputation upon this attribute (Corsi et al., 2012; Kallas et al., 2013; Mtimet & Albisu, 2006). On the contrary, the Old World wines appear to be more related to the *terroir* and tradition. The work of Tempesta et al. (2010) shows that the link between wine production cultural heritage and regional viticulture significantly influences consumers' preferences in Italy. For Lockshin and Corsi (2012) wine tells us about the history and culture of a country, which is a mirror of the environment where it has been produced. Tempesta et al. (2010) point out that the beauty of a landscape can positively affect the wine quality perception, since landscape can reflect the environment and the culture in which the vineyard is cultivated, and wine produced. As such, we explore the role of "Landscape" as a wine attribute reflecting

the environment in which the vines are cultivated, which can potential be important for wine valuation studies.

Additionally, literature shows that the choice of wine is related to the situation and context defining the occasion on which the wine will be consumed. Some wine attributes may be more or less important depending on whether the wine is to consume in a social situation or during a personal relaxation private moment (Corduas, Cinquanta, & Ievoli, 2013).

Given its complexity, different alternatives have been used to understand how consumers choose wine. This thesis aims to apply and show that discrete choice models are a useful instrument to understand and measure consumers' preferences for complex beverage products, by focusing on a bottle of red wine selected for a formal dinner with guests at home. Additionally, this thesis focuses on two main methodological issues: 1) analyzing sources of consumers heterogeneity; and 2) studying information processing heterogeneity. These two issues have generated four articles, one related to issue 1), while the last three are related to the issue 2).

1. Theoretical Foundations of Quantitative Models of Discrete Choice Data

DCE is routed on Lancaster's model of consumer behavior (Lancaster, 1966), as well as on the random utility theory (RUT) from McFadden (1974). The first assumes that the value of a good is derived from its specific attributes (or characteristics) and their levels, and not a function of the good itself. The second states that the overall utility can be expressed as the sum of a deterministic (observable) component and also a random (unobservable) component. Formally, the basic axiom of RUT is:

$$U_{nj} = V_{nj} + \varepsilon_{nj} \quad (1)$$

where U_{nj} is the latent, unobservable utility of the respondent *n* when choosing the alternative *j*, $V_{nj} = \beta_{0j} + \sum_{k=1}^{K} \beta_{jk} X_{nk} + \sum \alpha Z_n$ is the deterministic component depending on the product attributes (X_{nk}) and respondents' characteristics (Z_n), and ε_{nj} is the random component. The weighting coefficients β_{jk} and α_w reflect the relative importance of each of the explanatory variables. Individual *n* chooses the alternative *j* rather than alternative *i* if $U_{nj} > U_{ni}$, that is, consumer chooses the alternative that gives the highest utility. On probabilistic terms, we have:

$$P_{nj} = prob\left(V_{nj} + \varepsilon_{nj} > V_{ni} + \varepsilon_{ni}\right) = prob\left(V_{nj} - V_{ni} > +\varepsilon_{nj} - \varepsilon_{ni}\right) \forall\, j \neq i \in C \quad (2)$$

where C is the choice set of J alternatives, j=1, ..., J.

Different assumptions about the utility function and the distribution of random components may result in different models. The conditional logit model (McFadden, 1974), widely known as multinomial logit (MNL)[1] model, is "the workhorse" for analyzing discrete choice data. If the error terms are independent and identically distributed, the probability given by this model is:

$$P_{nj} = \frac{\exp\left(V_{nj}\right)}{\sum_{j=1}^{J} \exp\left(V_{nj}\right)} \quad (3)$$

such that $0 < P_{nj} < 1$ and $\sum_{j=1}^{J} P_{nj} = 1$.

The utility parameters are to be estimated by maximum likelihood, where the log likelihood function of the model can be written as:

$$LL = \sum_{n}^{N} \sum_{j=1}^{J} I\ ln\ P_{nj} = \sum_{n}^{N} \left(\sum_{j=1}^{J} I(y_n = j)\, V_{nj} - ln\left(\sum_{j=1}^{J} \exp\left(V_{nj}\right) \right) \right) \quad (4)$$

$I(\cdot)$ is an indicator function with value 1 if the alternative j was chosen ($y_n = j$) and zero otherwise.

However, this model specification assumes preference homogeneity among respondents, such that the deterministic part of the utility does not vary. Furthermore, it assumes that choices are independent from each other, such that the ratio of probabilities between two alternatives remains constant, independently of the other alternatives in the choice set. These assumptions may lead to behaviorally implausible outcomes, and have been relaxed over time, giving rise to advanced developments in discrete choice analysis (some of them are covered in the following chapters of this thesis).

In practice, DCE consists of simulating a choice situation by presenting a sequence of finite choice sets with mutually exclusive alternatives, differing in terms of attributes' levels, and asking respondents to

[1] These models are commonly assumed to be the same. However, differences arise depending on the information included on X. While the MNL model includes data on the characteristics of the individual, the conditional logit also includes attributes of the choices that differ across alternatives.

select their preferred alternative. A baseline alternative corresponding to "not buying" or to choosing the current situation (status quo) is usually included to add more realism to the experiment. Respondents make a trade-off in attribute levels against the cost of making these changes (Hanley, Mourato, & Wright, 2001), and implicitly, the technique assumes that the individual is able to make the required choice and that there are no limitations on its ability to make it. DCE intends to determine whether preferences are significantly influenced by the attributes and their relative importance.

This method provides a diverse range of outputs, such as choice share predictions, elasticities, and marginal rates of substitution, which are commonly denoted as willingness to pay (WTP) for changes in specific attributes of alternatives.

Preference and scale heterogeneity

A key improvement to discrete choice models is related to the role of heterogeneity of preferences among respondents. The introduction of socio-economic and demographic variables in the deterministic component of the utility function (i.e., in V_{nj}) is the most straightforward way to capture the "observed heterogeneity". On the other hand, there are several methods to capture "unobserved heterogeneity". Random parameters logit (RPL) or the so-called Mixed logit (MIXL) models allow the capture of stochastic variation in preferences, and have been widespread accepted. In the same way, Latent class (LC) models capture preference heterogeneity by identifying different preference segments of respondents.

The MIXL model is a powerful tool, where the variation of preferences or taste is described by continuous distributions (Revelt & Train, 1998; Train, 2009). Under this model, the choice specific constants, β_{0j} and the elements of β_{jk} may be randomly distributed across individuals, such that $\beta_n = \beta + \Delta z_n + \Gamma w_n$. $\beta + \Delta z_n$ is the heterogeneous means and w_n the random unobserved taste variation, characterized by $E[w_n] = 0$ and $Var[w_n] = diag[\sigma_1, \dots, \sigma_k]$, and σ_k is a known constant. Γ is the unknown lower triangular matrix (Greene & Hensher, 2003). The presence of a standard deviation of the parameter accommodates preference variation, often referred to as unobserved heterogeneity. Nevertheless, the researcher has the challenge of making distributional assumptions for preferences, being the most common the normal, log normal, uniform and triangular (Train, 2009). These distributions have advantages and disadvantages, and may complicate when calculating WTP given that its distribution becomes a ratio of two independent distributions. To avoid this, the most common procedure is to set the price/cost parameter as fixed. However, it implies the same marginal utility of money for everybody, which can be behaviorally restrictive. Choosing a distribution for the monetary attribute is advised, but it

may lead to undefined moments (see Daly, Hess and Train (2012) for proofs of defined moments in WTP calculation). To form the unconditional likelihood, w_n needs to be integrated according to:

$$P_{nj}|w_n = \int_{w_n} P_{nj}|\, w_n h(w_n) \mathrm{d}w_n \tag{5}$$

where $h(w_n)$ is the density of the standardized random vector w_n (Greene & Hensher, 2003). Given its intractability, the only practical method is to use simulations, i.e. w_n may be random draw from the assumed distribution of β.

The LCM is similar to the MIXL model. The heterogeneity across individuals is modeled with a discrete distribution, or set of classes. The LCM postulates a pre-defined number of latent classes, C, that distinguish types of individuals in the sample according to their preferences, without the requirement of any distributional assumption. Individuals are then sorted into different groups, and the probability of choosing the alternative j in choice situation t depends on the class membership:

$$P_{njt|c} = Prob(y_n = j|class = c) \tag{6}$$

Preference structures are not observed, and the choice of the number of classes is usually determined by information criteria. It is assumed heterogeneity across classes but homogeneity within class (Greene & Hensher, 2003). Recent developments, which we cover on Chapter 3, allow to relax the former assumption by using MIXL models to describe preference structures within class (Greene & Hensher, 2013).

While most of the studies has focused on capturing preference heterogeneity, scale heterogeneity has been ignored by researchers, who assume that scale is constant and equal to one, i.e. individuals are homogeneous in their error variances. This may lead to confounding between the model parameter estimates and scale (Louviere & Eagle, 2006) since that parameter estimates may vary with error variance. However, considering scale heterogeneity brings behavioral implications, since when the variance of the error term tends to infinity, the scale parameter limits to zero which implies equal probability of choosing an alternative across alternatives, characterizing a random choice process (Louviere & Eagle, 2006; Train & Weeks, 2005).

Thus, modelling scale heterogeneity allows to account for more or less consistency in the respondents' choices. That is, some respondents may behave more consistently with the predicted by the RUT than others. If we neglect it, estimates of the utility parameters may be biased. Hence, to identify scale

heterogeneity and preference heterogeneity, the generalized multinomial logit model (GMXL) appears to be a useful method (e.g. Fiebig, Keane, Louviere, & Wasi, 2010; Greene & Hensher, 2010). The specification and application of this model is presented in Chapter 2.

Information or attribute processing heterogeneity

Some research in related literature has focused on respondents' information processing routine to make a choice, bringing to the fore deviations from the fully compensatory choice process. It is recently accepted that there are limits to how much information respondents can process, known as cognitive burdens, leading to the use of particular processing rules to simplify their decisions. It seems reasonable that this type of behavior appears more likely when the good under consideration is complex and/or unfamiliar to individuals.

The literature suggests that individuals may use several strategies to make a choice, such as ignore some attributes below or above a given threshold, or simply choose based on a single attribute, exhibiting what is known as lexicographic behavior (e.g. Campbell, Hensher, & Scarpa, 2012; Caputo, Van Loo, Scarpa, Nayga, & Verbeke, 2018; Hensher & Rose, 2009; Ortega & Ward, 2016; Scarpa, Thiene, & Hensher, 2010). This literature became relevant in choice modelling, particularly the stream of research that recognizes that any individual might only attend to a subset of attributes that defines an alternative, known as Attribute Non-Attendance (ANA). For individuals who ignore some attributes, it implies that such attributes have no bearing on the choice and then no influence on utility. ANA impacts on DCE outputs have been studied in transportation, health an environmental economics fields, and more recently on food choice.

To identify ANA, researchers have two alternative approaches. First, ANA can be self-reported by respondents, by asking them their attendance to the attributes in the choice experiment. Second, ANA can be inferred through the use of probabilistic models, without using any additional information. In particular, the second alternative assumes that each individual has a given probability of belonging to a particular latent class described by a specific decision rule (ignoring a determined attribute or combination of attributes).

There is mixed evidence of which method to undertake, but evidence clearly shows that ignoring this behavior may lead to biased DCE outputs, specially WTP (Campbell, Hensher, & Scarpa, 2011; Kragt, 2013; Lagarde, 2013; Scarpa, Gilbride, Campbell, & Hensher, 2009).

2. Survey design and data collection

This thesis was developed within a research project, INNOVINE & WINE - Vineyard and Wine Innovation Platform, Research Line Competitiveness, such that we always try to match the thesis objectives within the project goals.

We employ a DCE on wine choice, using a web panel of wine consumers from different countries: Portugal, France, Germany, UK, Switzerland, Denmark, China, Russia, Brazil, USA, and Canada. The thesis covers the European market, comprising Portugal, France and Germany, and the North American market, including USA and Canada.

To survey wine consumers, we designed a questionnaire composed of three general sections, and six subsections (see Appendix A). The first section includes a wide range of questions covering wine consumption habits, consumer attitudes, perceptions and knowledge as well as wine purchasing behavior. The second section includes the discrete choice scenarios generated by a D-efficient design, using the Ngene software. Respondents were asked to complete ten choices between three wines, plus a non-purchase alternative. We designed the bottles of wine to be as realistic as possible, we therefore fixed five attributes in all wines - a fictional brand (RUMO), an arbitrary vintage (2016), bottle volume (750ml), bottle style (Bordeaux bottle), and the appellation of origin. All wines have an appellation classification according to the country of origin system, being equivalent to the "Protected Designation of Origin". We defined a choice context as the choice of wine for a formal consumption occasion, such as a dinner with guests at home, according to the following sentence: "Imagine you are standing in front of a shelf of wine. You have 3 different bottles of PDO wine available. Which of the following wines would you buy to enjoy over a dinner with guests at home?"

Based on the literature review, focus group, advice from experts in the wine market, and a pilot study, six wine attributes were selected: price, country of origin, landscape, alcohol content, grape variety and medals awarded (Table 1). After the choice exercise, we included a set of debriefing questions for ANA and importance ranking question for each attribute.

The third and last section consisted of questions regarding the respondent's sociodemographic characteristics. All questionnaires were translated to the native language of the sample, including English, Portuguese, German, French, Danish, Chinese, and Russian. Data were collected in August 2018 through an online survey conducted by a market research company.

Table 1 - Attributes and levels used in the study

Attribute	Reason	Levels	
Country of origin	Wine origin is an important cue and it is well documented, either in the case of a region or at the level of a country (e.g. Kallas, Escobar, & Gil, 2013; Lockshin, Corsi, Cohen, Lee, & Williamson, 2017). Taking into account the changes in the international wine market due to the emergence of new producing countries, we choose wines that came from countries with a long-documented history of wine production, and wines from the new players.	Portugal, France, Italy, USA, Australia, Chile	
Landscape	The interest from the wine industry, specially from Old Wine countries, argue that landscape certification may be a way to differentiate their wine. In our research, we analyze the value that consumers associate to the historical and natural qualities of the landscape in which the vines are cultivated	1= certified landscape 0 = otherwise	
Medals	Consumers perceive this attribute as an indication of quality, and several studies show their importance for wine choice (Corsi, Mueller, & Lockshin, 2012; Lockshin et al., 2006).	1= gold medal 0 = none	
Alcohol level	There is growing concern about the effects of overconsumption of alcohol, and some results of the focus groups revealed alcohol level as relevant for consumers' choices.	12% vol. 13.5% vol. 15% vol.	
Grape variety	Grape varietals are key choice drivers, particularly for New World wines (Corsi et al., 2012; Kallas et al., 2013). We focused on red varieties, with Cabernet Sauvignon and Syrah as very well-known varieties around the world, the Red blend, or then that information is not available in the bottle label.	Cabernet Sauvignon, Syrah, Red blend, none	
Price	As one of the main key drivers of wine choice, price is used as an indicator of quality and as a possible consumers' risk reduction strategy (e.g. Corsi et al., 2012; Lockshin et al., 2006; Mueller, Lockshin, Saltman, et al., 2010). Four price levels were selected to cover a range that approximates the wine sales across the countries under analysis.	European countries	North America
		3,99€	8.99$
		7,5€	12.99$
		12,5€	17.99$
		18€	24.99$

3. Aims and thesis outline

This thesis compiles different studies regarding theoretical and empirical questions concerning individual economic decision-making (as explained in Table 2). The overall aim is to contribute to a better understanding of decision-making process, with focus on the role of consumers heterogeneity and on the potential that the recognition of the use of heuristics, specifically attribute non-attendance, plays in explaining individual choice behavior. It is also our goal to understand wine valuation, to deepen the knowledge about consumers' choice and preferences towards international wines and their attributes, providing important insights to wineries and marketing practitioners, and decision makers, who want to position their wines in the international sphere.

Specifically, in Chapter 2 we allow for preference and scale heterogeneity, and we evaluate different models to check whether heterogeneity may be caused by unobserved factors, i.e. scale, rather than differences in the underlying preference structure. Moving to the issue of processing information heterogeneity, in Chapter 3 we provide a detailed literature review on attribute non-attendance in four distinct fields using DCE, presenting a brief comparison between studies to explore how to identify and deal with this heuristic, and whether it takes place, thus determining its relevance. This chapter integrates findings from many studies, and synthesizes the literature in a useful manner, harmonizing different notation and language, by suggesting a general econometric framework to account for non-attendance, which is a lack in the literature.

While the literature has not resolved whether to use the stated or the inferred approach to identify and to deal with ANA, the inference of ANA using an analytical approach has strong appeal. In Chapter 4 we analyze the performance of these two different methods, examining whether in fact respondents do what they state to do. To assess whether self-reported attendance maps onto inferred attendance, we compare mixed logit models that use supplementary statements on ANA, to equality constrained latent class models, which estimate the probability of ANA.

In Chapter 5, we focus on inferred ANA to accommodate the full set of possible patterns of ignored and attended attributes that wine consumers might use. To infer and focus only on ANA we need to include an equality constraint for parameters across classes, assuming equal preference structure, but allowing for different non-attendance patterns. We follow a stepwise approach, where 2^K classes or decision rules were estimated in our study, being K the number of attributes.

Finally, the chapter Discussion and Conclusion discusses the main findings of the present thesis and reflect on methodological considerations, practical implications and outline directions for further research.

Table 2 - Overview of the studies performed in this thesis

	Aim(s) of the study	Methodological approach	Status
Studying consumers heterogeneity			
Ch2	To model preferences underlying wine, a highly differentiated product, accounting for distinct behavioral patterns and/or sources of heterogeneity, by accommodating mixtures of preference heterogeneity and scale heterogeneity.	Scaled multinomial logit model; Mixed logit model; Generalized mixed logit model	Published
Studying heuristics			
Ch3	Review on ANA effects on DCE outputs, reflecting on the theoretical insights, econometric modelling, and developing a typology or organizing framework to classify choice models incorporating attribute non-attendance.	Literature review	Submitted for publication
Ch4	To contrast the two methodologies to identify and deal with ANA, the stated and inferred ANA, to explore the possible biases of ignoring ANA on WTP measures and to identify the most important attributes for wine consumers.	Mixed logit model accounting for stated ANA; Equality constrained latent class model accounting for inferred ANA	Submitted for publication Oral presentation
Ch5	To identify all ANA possible strategies to detect the relevant ones used by wine consumers, to find the welfare values for wine attributes, and to find the ANA effects on the outcome of the different models.	Stepwise inferred approach	Submitted for publication

Descriptive statistics of the sample for each chapter and the attitudinal and perceptual variables values are available in the Appendices section, namely in the Appendix B and Appendix D.

References

Campbell, D., Hensher, D. A., & Scarpa, R. (2011). Non-attendance to attributes in environmental choice analysis: a latent class specification. *Journal of Environmental Planning and Management*, *54*(8), 1061–1076. https://doi.org/10.1080/09640568.2010.549367

Campbell, D., Hensher, D. A., & Scarpa, R. (2012). Cost thresholds, cut-offs and sensitivities in stated choice analysis: Identification and implications. *Resource and Energy Economics*, *34*(3), 396–411. https://doi.org/10.1016/j.reseneeco.2012.04.001

Caputo, V., Van Loo, E. J., Scarpa, R., Nayga, R. M., & Verbeke, W. (2018). Comparing Serial, and Choice Task Stated and Inferred Attribute Non-Attendance Methods in Food Choice Experiments. *Journal of Agricultural Economics*, *69*(1), 35–57. https://doi.org/10.1111/1477-9552.12246

Corduas, M., Cinquanta, L., & Ievoli, C. (2013). The importance of wine attributes for purchase decisions: A study of Italian consumers' perception. *Food Quality and Preference*, *28*(2), 407–418. https://doi.org/10.1016/j.foodqual.2012.11.007

Corsi, A. M., Mueller, S., & Lockshin, L. (2012). Let's See What They Have . . .: What Consumers Look For in a Restaurant Wine List Armando. *Cornell Hospitality Quarterly*, *XX*(X), 1–12. https://doi.org/10.1177/1938965511428448

Daly, A., Hess, S., & Train, K. (2012). Assuring finite moments for willingness to pay in random coefficient models. *Transportation*, *39*(1), 19–31. https://doi.org/10.1007/s11116-011-9331-3

Fiebig, D. G., Keane, M. P., Louviere, J., & Wasi, N. (2010). The Generalized Multinomial Logit Model : Accounting for Scale and Coefficient Heterogeneity. *Marketing Science*, 1–29. https://doi.org/10.1287/mksc.1090.0508

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transportation Research Part B*, *37*, 681–698. https://doi.org/10.1016/S0191-2615(02)00046-2

Greene, W. H., & Hensher, D. A. (2010). Does scale heterogeneity across individuals matter ? An empirical assessment of alternative logit models. *Transportation*, *37*, 413–428. https://doi.org/10.1007/s11116-010-9259-z

Greene, W. H., & Hensher, D. A. (2013). Revealing additional dimensions of preference heterogeneity in

a latent class mixed multinomial logit model. *Applied Economics*, *45*, 1897–1902. https://doi.org/10.1080/00036846.2011.650325

Hanley, N., Mourato, S., & Wright, R. E. (2001). Choice Modelling Approaches: A Superior Alternative for Environmental Valuatioin? *Journal of Economic Surveys*, *15*(3), 435–462.

Hensher, D. A., & Rose, J. M. (2009). Simplifying choice through attribute preservation or non-attendance: Implications for willingness to pay. *Transportation Research Part E*, *45*(4), 583–590. https://doi.org/10.1016/j.tre.2008.12.001

Kallas, Z., Escobar, C., & Gil, J. M. (2013). Analysis of consumers' preferences for a special-occasion red wine: A dual response choice experiment approach. *Food Quality and Preference*, *30*(2), 156–168. https://doi.org/10.1016/j.foodqual.2013.05.008

Kragt, M. E. (2013). Stated and inferred attribute attendance models: A comparison with environmental choice experiments. *Journal of Agricultural Economics*, *64*(3), 719–736. https://doi.org/10.1111/1477-9552.12032

Lagarde, M. (2013). Investigating attribute non-attendance and its consequences in choice experiments with latent class models. *Health Economics*, *22*, 554–567. https://doi.org/10.1002/hec

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*(2), 132–157. https://doi.org/10.1086/259131

Lockshin, L., Cohen, E., & Goodman, S. (2009). Overcoming Measurement Errors: Segmenting Wine Consumers across 11 Countries. *Australian and New Zealand Wine Industry Journal*, *24*, 46–51.

Lockshin, L., & Corsi, A. M. (2012). Consumer behaviour for wine 2.0: A review since 2003 and future directions. *Wine Economics and Policy*, *1*(1), 2–23. https://doi.org/10.1016/j.wep.2012.11.003

Lockshin, L., Corsi, A. M., Cohen, J., Lee, R., & Williamson, P. (2017). West versus East: Measuring the development of Chinese wine preferences. *Food Quality and Preference*, *56*, 256–265. https://doi.org/10.1016/j.foodqual.2016.02.014

Lockshin, L., Jarvis, W., D´Hauteville, F., & Perrouty, J.-P. (2006). Using simulations from discrete choice experiments to measure consumer sensitivity to brand, region, price, and awards in wine choice. *Food Quality and Preference*, *17*, 166–178. https://doi.org/10.1016/j.foodqual.2005.03.009

Louviere, J., & Eagle, T. (2006). Confound it! That Pesky Little Scale Constant Messes Up Our Convenient

Assumptions. *Proceedings of the Sawtooth Software Conference*, 211–228.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In Z. P (Ed.), *Frontiers in Econometrics* (pp. 105–142). New York: Academic Press.

Mtimet, N., & Albisu, L. M. (2006). Spanish Wine Consumer Behavior: A Choice Experiment Approach. *Agribusiness*, *22*(3), 343–362. https://doi.org/10.1002/agr.20090

Mueller, S., Lockshin, L., Saltman, Y., & Blanford, J. (2010). Message on a bottle: The relative influence of wine back label information on wine choice. *Food Quality and Preference*, *21*(1), 22–32. https://doi.org/10.1016/j.foodqual.2009.07.004

Mueller, S., Osidacz, P., Francis, I. L., & Lockshin, L. (2010). Combining discrete choice and informed sensory testing in a two-stage process: Can it predict wine market share? *Food Quality and Preference*, *21*(7), 741–754. https://doi.org/10.1016/j.foodqual.2010.06.008

Ortega, D. L., & Ward, P. S. (2016). Information processing strategies and framing effects in developing country choice experiments: results from rice farmers in India. *Agricultural Economics (United Kingdom)*, *47*(5), 493–504. https://doi.org/10.1111/agec.12249

Perrouty, J. P., D´Hauteville, F., & Lockshin, L. (2006). The Influence of Wine Attributes on Region of Origin Equity: An Analysis of the Moderating Effect of Consumer's Perceived Expertise. *Agribusiness*, *22*(3), 323–341. https://doi.org/10.1002/agr

Revelt, D., & Train, K. (1998). Mixed Logit with repeated choices: Households' choices of appliance efficiency level. *The Review of Economics and Statistics*, *80*(4), 647–657. https://doi.org/10.1162/003465398557735

Scarpa, R., Gilbride, T. J., Campbell, D., & Hensher, D. A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics*, *36*(2), 151–174. https://doi.org/10.1093/erae/jbp012

Scarpa, R., Thiene, M., & Hensher, D. A. (2010). Monitoring Choice Task Attribute Attendance in Nonmarket Valuation of Multiple Park Management Services : Does It Matter ? *Land Economics*, *86*(4), 817–839.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Simon, H. A. (1982). *Models of bounded rationality*. MIT Press.

Tang, V. C. M., Tchetchik, A., & Cohen, E. (2015). Perception of wine labels by Hong Kong Chinese consumers. *Wine Economics and Policy*, *4*(1), 12–21. https://doi.org/10.1016/j.wep.2015.02.002

Tempesta, T., Arboretti, R., Corain, L., Salmaso, L., Tomasi, D., & Boatto, V. (2010). The importance of landscape in wine quality perception: An integrated approach using choice-based conjoint analysis and combination-based permutation tests. *Food Quality and Preference*, *21*(7), 827–836. https://doi.org/10.1016/j.foodqual.2010.04.007

Train, K. (2009). *Discrete choice methods with simulation* (2nd ed.). https://doi.org/10.1017/CBO9780511805271

Train, K., & Weeks, M. (2005). Discrete Choice Models in Preference Space and Willingness-to-Pay Space. In R. Scarpa & A. Alberini (Eds.), *Applications of Simulation Methods in Environmental and Resource Economics* (pp. 1–16). https://doi.org/10.1007/1-4020-3684-1_1

Chapter 2

Exploring distinct sources of heterogeneity in discrete choice experiment: An application to wine choice across European consumers

Published in
Economics Letters 178 (2019) 28-32
DOI: 10.1016/j.econlet.2019.02.019



Abstract

This paper explores the role of unobserved heterogeneity using different approaches: mixed, scaled and the generalized mixed logit model, and the latent class model, applied to wine choice. Results show that preference heterogeneity appears to be more relevant than scale heterogeneity, and the model accounting for both performs best. This finding suggests that wine choice is complex, demanding a higher cognitive burden, which may lead to the use of simplification strategies when choosing. For the purpose of wine valuation, the latent class model gives useful insights for market segmentation.

Keywords: Discrete choice models; Preference heterogeneity; Wine choice

JEL: C25, C52, D11, Q18

1. Introduction

When making a wine purchase decision, consumers are presented with a large variety of wines, with different brands of domestic and foreign wineries. On many purchase occasions, wine remains an untasted product, such that consumers rely only on the information provided by the label, including the details on both intrinsic and extrinsic quality cues, such as grape variety, origin, producer, brand or medals. Despite this information, the wine has a unique but highly complex nature which could lead to distinctive choice processes compared to other product categories, implying different sources of preference heterogeneity.

Researchers have tested different alternatives to understand how consumers choose wine, namely through stated preference experiments, as it allows the simulation of consumer choices (Lockshin et al., 2017). Using the discrete choice experiment technique (DCE), the present article intends to model unobserved heterogeneity of wine consumers. We estimate and compare different models that accommodate mixtures of preference or taste heterogeneity and scale heterogeneity (i.e., in error variance across consumers), such as the multinomial logit model allowing for scale heterogeneity (SMNL), the mixed or random parameters logit model (MIXL), and the generalized mixed logit model (GMXL). As part of our interest in wine valuation, we estimate a latent class model (LCM) to find market segments of wine consumers. In addition, we also provide the willingness to pay (WTP) for the wine attributes.

Several studies (e.g. Fiebig, Keane, Louviere, & Wasi, 2010; Keane & Wasi, 2010; Thiene, Scarpa, & Louviere, 2015) have compared different models through a number of criteria, such as Akaike Information criterion (AIC) and Bayes information criterion (BIC). In this study, we compare and evaluate different models based on model fit measures, and also using a matrix of actual versus predicted choices to calculate the percentage of correct predictions, which shows the model performance in capturing the actual behavior.

To this end, section 2 describes the sample and methods used to address the issue posed. Section 3 presents the empirical results, section 4 concludes the paper.

2. Sample and methodology

2.1. Sample

We explore consumers preferences for wine with stated preference data from a DCE using a European web panel. Six attributes were selected through literature review, focus group and advice from experts (Table 1). Ten choice sets formed by 3 distinct wines and a none-option were generated by a D-efficient design, using the Ngene software. A total of 600 wine consumers completed the discrete choices, under a specific wine-drinking context, a formal dinner with guests at home.

Table 1 - Attributes and levels

Attribute	Levels
Country of origin (COO)	Portugal[+], France, Italy, USA, Australia, Chile
Landscape	1=certified; 0=not certified
Medals	1=gold medal; 0=none
Alcohol level	12, 13.5, 15
Grape variety	Cabernet Sauvignon, Syrah, Red blend, none[+]
Price	3.99€; 7.5€; 12.5€; 18€

[+]base level for dummy coding

2.2. Methodology

The GMXL extends the MIXL to include scale heterogeneity across respondents, that is, a variation on the variance condition associated with the random component (Fiebig et al., 2010; Greene & Hensher, 2010). Accounting for scale heterogeneity across choices implies that:

$$\beta_{\mathrm{n}} = \sigma_{\mathrm{n}}\,\beta + [\gamma + \sigma_{\mathrm{n}}\,(1-\gamma)]\; v_{\mathrm{n}} \qquad (1)$$

where $\sigma_n = exp(\bar{\sigma} + \tau w_n)$ is the standard deviation of the individual specific scale factor, where τ captures the unobserved scale heterogeneity, and $w_{\mathrm{n}} \sim \mathrm{N}\ (0,1)$. Particularly, $\overline{\sigma}$ is not identified independently from τ, such that σ_n is normalized to 1, by setting $\overline{\sigma} = -\tau^2/2$. γ is a weighting parameter, bounded between 0 and 1, and reflects the extent to which σ_{n} influences the person-specific component of β_{n}. v_n is a random term with zero mean and follows a specific distribution. Variants of

GMXL arise, such that, when $\sigma_n=1$, GMXL reverts to MIXL, when $\mathrm{var}(v_n) = 0$ it reverts to SMNL, and when both conditions hold, it reverts to MNL.

Given the interest in establishing welfare measures, the WTP is estimated as the ratio of the marginal utility of one attribute and the marginal disutility of price. Under MIXL and GMXL models, the price may be specified as a random or non-random parameter. In this paper we define all attributes as random parameters, being the WTP the ratio of two randomly distributed terms. In this case, it may happen that the resulting distribution of WTP estimates may not be defined or may produce unrealistic results (Daly et al., 2012). Alternatively, the model can be reparametrized to estimate taste parameters in WTP space (Greene & Hensher, 2010), rather than the parameters of the usual coefficients (Train & Weeks, 2005). Hence, for the GMXL model, the utility expression becomes:

$$U_{nj} = \sigma_n\left(\beta_{np}p_{nj} + \beta_{nx}x_{nj}\right) + \varepsilon_{nj} \tag{2}$$

$$U_{nj} = \sigma_n\beta_{n,p}\left[p_{nj} + \left(1/\beta_{np}\right)\beta_{nx}x_{nj}\right] + \varepsilon_{nj} = \sigma_n\left(p_{nj} + \beta^*_{nx}x_{nj}\right) + \varepsilon_{nj} \tag{3}$$

Equation (3) represents the WTP space form of (2), where β_{np} becomes the normalizing price constant in the WTP space representation, and β^*_n directly gives the individual-specific WTP.

The LCM is a variant of the MIXL model, replacing the parameters' continuous distribution with a discrete distribution. This model provides an alternative approach to accommodate unobserved preference heterogeneity by assuming that the population can be sorted into a discrete number C of separate classes, where individuals in each segment share homogeneous preferences. Heterogeneity exists across classes, but not within each class, such that:

$$\beta_n = \beta_c \text{ with probability } P_{nc} \text{ for c=1, ..., C and } \sum P_{nc} = 1, P_{nc} > 0 \;\forall c \tag{4}$$

P_{nc} is the probability person *n* belongs to a particular class *c*, and may depend on individuals' characteristics or may be assumed identical across individuals. The probability of respondent *n* selecting the alternative *j*, given membership of class *c* becomes:

$$P_{nj|c} = \frac{\exp\left(\beta_c{}'x_{nj}\right)}{\sum_{j=1}^{Jn}\exp\left(\beta_c{}'x_{nj}\right)} \tag{5}$$

3. Empirical results

Table 2 reports the results of the estimated models, using maximum simulated likelihood methods with 500 Halton draws in NLOGIT 6.0. As a conventional procedure, Alternative Specific Constant is not a random parameter (Fiebig et al., 2010) and it is included for the None option.

In terms of model fit, GMXL provides a significant improvement compared to MNL, with a higher McFadden's pseudo-R^2, lower BIC and AIC, and higher percentage of correctly predicted choices. Results show that the introduction of random parameters substantially improves the model fit, and when both sources of heterogeneity are captured (in the GMXL), it provides a better fit than either SMNL, MIXL and LC models (in line with Keane and Wasi, 2013). The MIXL outperformed the SMNL indicating that the scale heterogeneity is less important than preference heterogeneity towards the choice attributes. For instance, the significant standard deviations for the random parameters show either taste differences across wine consumers, or the presence of individual preference heterogeneity.

In both SMNL and GMXL models, the scale heterogeneity parameter (τ) is statistically significant, suggesting the presence of substantial scale heterogeneity, such that respondents varied in terms of certainty/consistency in their choices. The estimate of γ in GMXL is 0.53, and it is statistically significantly different from zero, which suggests the adoption of GMXL instead SMNL.

In contrast to MNL, the coefficient on Cabernet Sauvignon becomes insignificant when introducing random coefficients and scale heterogeneity, suggesting that this grape variety does not affect wine choice. Alcohol content is not significant in the SMNL, while other models show the opposite. In general, the estimates are higher in the GMXL than in the other models, and when scale effect is taken into account, some of the taste heterogeneity from the MIXL becomes relatively higher in magnitude. However, after accounting for scale heterogeneity, individuals are more similar in their valuation of Australian and American wines, as the estimated standard deviation becomes small relative to the magnitude of the mean in GMXL compared to MIXL. Despite the standard deviation coefficients for red blend are not consistent across models, results from the GMXL suggest preference homogeneity for a red blend wine. The estimates for the mean of medals and landscape are positive in all models, and the consistency of country of origin shows a general predilection for Portuguese wines, especially compared to wines from the *new world*.

Regarding the LCM (Table 3), the number of classes was selected considering the lowest value of AIC, the magnitude of class probabilities and also huge standard errors (Greene & Hensher, 2013). Therefore,

the model with four classes was found to have the best trade-off between fit and plausibility. Each class considers the main effect of each attribute and the probability of belonging to a class is assumed to be constant for every respondent. In this model, the performance of the adjustment is not the best, however it correctly predicted the actual choice outcome for 83 percent of the total number of observations. The first class (15.1%) involves consumers who use price as a cue for quality. They prefer Portuguese wines over wines from other countries, wines produced in a certified place, and they are not influenced by the alcoholic content. The second class (43.2%) is constituted by consumers who value awarded wines and are positively influenced by grape variety, such as Cabernet Sauvignon and red blend. The third class (21.4%) seems to represent more price-sensitive consumers, whose choice depends most on the wine´s medals and prefer French wines, with high alcohol concentration. The fourth class (20.3%) includes consumers who value medals awarded, and they are not influenced by the grape cultivar. In classes 2-4, price has a negative sign, showing the disutility for higher prices in these segments.

Table 2 - MNL, SMNL, MIXL, GMXL model results

	MNL	SMNL	MIXL		GMXL	
Variables			**Mean**	**SD**	**Mean**	**SD**
Medals	0.3832***	0.6146***	0.8474***	0.6828***	1.5387***	1.0902***
	(0.0460)	(0.0688)	(0.0721)	(0.0918)	(0.1554)	(0.1055)
Alcohol	0.0376***	0.0199	0.0411**	0.2355***	0.091***	0.3153***
	(0.0136)	(0.0151)	(0.0206)	(0.0127)	(0.0223)	(0.0174)
COO (base: Portugal)						
France	-0.1972***	0.3091***	-0.4604***	1.3436***	-1.0532***	1.7931***
	(0.0539)	(0.0400)	(0.0929)	(0.0916)	(0.1169)	(0.1153)
Italy	-0.7492***	-0.7474***	-1.3536***	0.8615***	-2.3275***	0.9323***
	(0.0704)	(0.0937)	(0.1044)	(0.1227)	(0.1864)	(0.1401)
USA	-1.1164***	-1.2339***	-1.7883***	1.1423***	-3.2783***	0.9859***
	(0.0612)	(0.1093)	(0.0986)	(0.1099)	(0.2452)	(0.1499)
Chile	-0.6953	-1.0801***	-1.2933***	1.5155***	-2.9341***	1.8186***
	(0.0776)	(0.1177)	(0.1363)	(0.1732)	(0.3087)	(0.2194)
Australia	-1.1435***	-1.6523***	-2.1684***	1.1215***	-3.5316***	0.7425*
	(0.0896)	(0.1734)	(0.1588)	(0.2310)	(0.3679)	(0.4058)
Landscape	0.1606***	0.2498***	0.3120***	0.2334**	0.6728***	0.3988***
	(0.0353)	(0.0400)	(0.0484)	(0.1056)	(0.0937)	(0.1024)
Grape variety (base: None)						
Cabernet Sauvignon	0.2343***	0.1607***	0.1304**	0.5382***	0.0145	0.6130***
	(0.0477)	(0.0473)	(0.0649)	(0.0861)	(0.0979)	(0.1000)
Syrah	0.0714	-0.1918***	-0.2343***	0.3504**	-0.6717***	0.5388***
	(0.0634)	(0.0678)	(0.0861)	(0.1436)	(0.1451)	(0.1218)
Red blend	0.4235***	0.5282***	0.3813***	0.6281***	0.4740***	0.2903
	(0.0854)	(0.1263)	(0.1199)	(0.1638)	(0.1685)	(0.1819)
Price	-0.0670***	-0.0777***	-0.1797***	0.2365***	-0.3043***	0.2860***
	(0.0039)	(0.0089)	(0.0123)	(0.0116)	(0.0250)	(0.0148)
Constant	0.0536	-0.9905***	-3.3686***		-3.8325***	
	(0.1963)	(0.2220)	(0.2938)		(0.2063)	
τ (scale)		0.9820***			0.9822***	
		(0.0884)			(0.0620)	
γ					0.5272***	
					(0.0452)	
LL	-7557.6	-7398.8	-6108.0		-6015.9	
AIC	15141.2	14825.7	12266.1		12085.9	
BIC	15228.3	14919.5	12433.6		12266.7	
Pseudo-R^2	0.076	0.110	0.266		0.277	
Correctly predicted	42%	40%	42%		43%	

Standard errors in parenthesis; ***, **, * significance at 1%, 5%, 10% level respectively

Table 3 - LCM results

	LCM			
Variables	**Class1**	**Class2**	**Class3**	**Class4**
Medals	0.3940*	0.6349***	0.8832***	0.6691***
	(0.2043)	(0.0840)	(0.3016)	(0.1535)
Alcohol	-0.0216	0.0066	0.1848**	0.1231***
	(0.0569)	(0.0212)	(0.0793)	(0.0447)
COO (base: Portugal)				
France	-2.4419***	-0.2223*	0.5765*	0.1757
	(0.4543)	(0.1149)	(0.3362)	(0.1650)
Italy	-2.7258***	-0.7479***	-0.4417	-1.0851***
	(0.5436)	(0.1332)	(0.3608)	(0.2059)
USA	-3.2535***	-0.7340***	-1.1264***	-1.8691***
	(0.3915)	(0.1149)	(0.2919)	(0.2550)
Chile	-1.8173***	-0.2117	-0.8663	-1.0773***
	(0.2911)	(0.1423)	(0.5681)	(0.2881)
Australia	-3.2369***	-1.4505***	-1.8622**	-1.3797***
	(0.4129)	(0.1532)	(0.8335)	(0.3502)
Landscape	0.7314**	0.0596	0.3386*	0.2654*
	(0.2581)	(0.0575)	(0.1738)	(0.1374)
Grape variety (base: None)				
Cabernet Sauvignon	-0.6566**	0.2337***	0.4126**	0.2455
	(0.2554)	(0.0755)	(0.1873)	(0.1715)
Syrah	-0.8293**	0.1065	0.1064	-0.0305
	(0.3356)	(0.1180)	(0.2842)	(0.1794)
Red blend	-0.7310**	0.5469***	1.7100*	0.0258
	(0.3458)	(0.1738)	(0.8917)	(0.2101)
Price	0.1139***	-0.0574***	-0.4146***	-0.1075***
	(0.0211)	(0.0091)	(0.0766)	(0.0176)
Constant	-2.2611***	-3.2371***	-1.2893	2.3471***
	(0.8504)	(0.3441)	(1.0294)	(0.6207)
LC probability (%)	15.1	43.2	21.4	20.3
LL	-6233.9			
AIC	12577.7			
BIC	12946.2			
Pseudo-R^2	0.250			
Correctly predicted	83%			

Standard errors in parenthesis; ***, **, * significance at 1%, 5%, 10% level respectively

Table 4 presents the WTP estimates for the GMXL derived from the WTP-Space and WTP for each class from the LCM. In both models, the highest WTP estimates are for medals. From the GMXL, consumers are willing to pay more than 9€ for an awarded wine, followed by 4.86€ for a red blend, and 1.61€ for a wine produced in a certified landscape. As expected, there is a range of WTP estimates across the latent classes. The WTP for medals is variable over classes, being low for class 1 and 3 (3.46€ and 2.13€ respectively), whereas for classes 2 and 4 are about 11€ and 6.23€, respectively. Regarding grape variety, the WTP between classes 1 and 2 was totally opposite, which suggests that individuals in these two segments have different tastes when choosing a type of grape. The findings from both models also indicate that consumers are willing to pay more for a wine produced in a certified landscape (with a high value in class one). The negative WTP estimates for wine origin, in particular for class 1, confirm the preference for Portuguese wines. WTP for French wines are positive in classes 3 and 4, however not statistically significant.

Table 4 - WTP for wine attributes

Variables	GMXL[2]	LCM[3]			
		Class 1	Class 2	Class 3	Class 4
Medals	9.45***	3.46*	11.07***	2.13***	6.23***
Alcohol	1.22***	-0.19	0.11	0.45**	1.15***
COO (base: Portugal)					
France	-3.10***	-21.44***	-3.87*	1.39	1.64
Italy	-13.37***	-23.94***	-13.04***	-1.07	-10.10***
USA	-15.12***	-28.43***	-12.79***	-2.72***	-17.40***
Chile	-10.49***	-28.58***	-3.69	-2.09*	-10.03***
Australia	-18.66***	-15.96***	-25.28***	-4.49**	-12.84***
Landscape	1.61***	6.42***	1.04	0.82**	2.47**
Grape variety (base: None)					
Cabernet Sauvignon	1.32**	-5.77***	4.07***	0.99*	2.28
Syrah	1.12***	-7.28***	1.86	0.26	-0.28
Red blend	4.86***	-6.42**	9.53***	4.12*	0.24

***, **, * significance at 1%, 5%, 10% level respectively

[2] Results from the GMXL estimated in WTP-Space, despite the model fit be inferior to the preference space specification (in line with Hole & Kolstad, 2012 and Train & Weeks, 2005).

[3] WTP of an attribute k is given by the ratio of the attribute coefficient, β_k, and the price coefficient, β_{price}, for each class.

4. Conclusion

We have compared alternative discrete choice models to understand consumers heterogeneity. Overall, the GMXL model performs best, showing that accommodating both sources of unobserved heterogeneity is important. These results are consistent with previous literature (e.g. Greene & Hensher, 2010; Hossain, Saqib, & Masudul, 2018; Scarpa, Thiene, & Hensher, 2012). Our study shows that the complexity of wine may lead to a higher cognitive burden, with some consumers choosing wines on the basis of one or two attributes or having a more random decision-making process.

Although the LCM does not achieve the best fit, this model provides better prediction of respondents' choice, and it is useful for gaining an intuitive understanding of the nature of heterogeneity and of consumers segmentation. Our findings provide guidance to policy makers and wineries in developing adequate strategies. Having in mind the specific wine-drinking context, consumers who value higher priced wines, use price as a quality cue to reduce the risk of buying a bad wine. In addition to medals, a group of consumers place more importance in the origin, and also are willing to pay for the certification of landscape, which may be linked to traditional wine producing places classified as world heritage, as the case of some regions in the countries considered. This result suggests that this might be a relevant information to include in the label, especially for consumers of more expensive wines from traditional origins.

Acknowledgements

This work is financed by the project NORTE-01-0145-FEDER-000038 (INNOVINE & WINE – Innovation Platform of Vine & Wine, Portugal) and by European and Structural and Investment Funds in the FEDER, Spain component, and supported through the Operational Competitiveness and Internationalization Program, Portugal (COMPETE 2020) [Project No. 006971 UID/SOC/04011)]; and national funds, through the FCT – Portuguese Foundation for Science and Technology under the projects UID/SOC/04011/2019 and UID/ECO/03182/2019.

References

Daly, A., Hess, S., & Train, K. (2012). Assuring finite moments for willingness to pay in random coefficient models. Transportation, 39(1), 19–31. https://doi.org/10.1007/s11116-011-9331-3

Fiebig, D. G., Keane, M. P., Louviere, J., & Wasi, N. (2010). The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity. Marketing Science, 1–29. https://doi.org/10.1287/mksc.1090.0508

Greene, W. H., & Hensher, D. A. (2010). Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models. Transportation, 37, 413–428. https://doi.org/10.1007/s11116-010-9259-z

Greene, W. H., & Hensher, D. A. (2013). Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. Applied Economics, 45, 1897–1902. https://doi.org/10.1080/00036846.2011.650325

Hole, A. R., & Kolstad, J. R. (2012). Mixed logit estimation of willingness to pay distributions: a comparison of models in preference and WTP space using data from a health-related choice experiment. Empirical Economics, 42(2), 445–469. https://doi.org/10.1007/s00181-011-0500-1

Hossain, I., Saqib, N. U., & Masudul, M. (2018). Scale heterogeneity in discrete choice experiment: An application of generalized mixed logit model in air travel choice. Economics Letters, 172, 85–88. https://doi.org/10.1016/j.econlet.2018.08.037

Keane, M. P., & Wasi, N. (2010). The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity, 29(3), 393–421. https://doi.org/10.1287/mksc.1090.0508

Keane, M., & Wasi, N. (2013). Comparing Alternative Models of Heterogeneity in Consumer Choice Behavior. Journal of Applied Econometrics, 28, 1018–1045. https://doi.org/10.1002/jae

Lockshin, L., Corsi, A. M., Cohen, J., Lee, R., & Williamson, P. (2017). West versus East: Measuring the development of Chinese wine preferences. Food Quality and Preference, 56, 256–265. https://doi.org/10.1016/j.foodqual.2016.02.014

Scarpa, R., Thiene, M., & Hensher, D. A. (2012). Preferences for tap water attributes within couples: An exploration of alternative mixed logit parameterizations. Water Resources Research, 48, 1–11. https://doi.org/10.1029/2010WR010148

Thiene, M., Scarpa, R., & Louviere, J. J. (2015). Addressing Preference Heterogeneity, Multiple Scales and Attribute Attendance with a Correlated Finite Mixing Model of Tap Water Choice. Environmental and Resource Economics, 62, 637–656. https://doi.org/10.1007/s10640-014-9838-0

Train, K., & Weeks, M. (2005). Discrete Choice Models in Preference Space and Willingness-to-Pay Space. In R. Scarpa & A. Alberini (Eds.), Applications of Simulation Methods in Environmental and Resource Economics (pp. 1–16). Dordrecht: Springer Netherlands. https://doi.org/10.1007/1-4020-3684-1_1

Chapter 3

The Role of Attribute Non-Attendance on Consumer Decision Making: Theoretical Insights and Empirical Evidence

Submitted for publication

Abstract

This paper analyses the recent research on choice modelling and the issue of attribute nonattendance in discrete choice experiments, from 2005 to 2018, providing valuable insights into how this specific research area has evolved over time and within distinct fields. The aim of this study is threefold. First, to present the theoretical background on the decision-making process, especially related to the use of heuristics, with focus on attribute nonattendance. Second, to examine the most frequent econometric techniques for dealing with this issue, and to develop a classification scheme by using a common language and notation. Third, to provide a structured review that provides a guide to earlier research on the subject of attribute nonattendance behavior, concerning the impacts on willingness to pay measures. This literature review shows that the applications on environmental and health economics fields are likely to present more problematic non-attendance rates, especially for the monetary attribute. The majority of the reviewed studies show a decrease in the willingness to pay measures, and only four papers find similar values when accounting for attribute nonattendance. This study reinforces the growing recognition that when faced with a typical stated choice scenario, different respondents will process the information in different ways, highlighting the importance of accounting for the decision-making process in stated experiments.

Keywords: Consumer behavior; Heuristics; Attribute non-attendance; Discrete choice models; Willingness to pay

1. Introduction

The application of stated preference methods has increased in the last decades in several research contexts, such as agricultural and food economics, marketing, environmental, ecological and cultural economics, health economics and transportation. Discrete Choice Experiments (DCE), a quantitative technique in nature, being the most popular method to elicit stated preferences. DCE builds hypothetical scenarios through repeated choice questions, asking individuals to state their preference when valuing different attribute combinations for each product or service. Grounded in Lancaster (1966) framework, this technique assumes that respondents make trade-offs among all the product's attributes included in the choice sets. In addition, it considers that individuals are able to process all information in a fully rational manner, in light of the continuity axiom, a standard hypothesis of the neoclassic model of consumer behavior. Violations to this principle have been gaining increasing attention in the field of choice modelling, specially the situations where respondents only attend a subset of attributes, ignoring others, commonly referred to as attribute non-attendance (ANA), or conversely, selective attendance. The non-attendance behavior can be seen as one mechanism prior to choice that respondents might employ when processing information, also called heuristics, information processing strategies, decision rules, and/or attribute processing strategies, and can be viewed as a building block upon which many of the heuristics depend. These rules are commonly applied to simplify the information in order to make a choice. Therefore, it is not realistic to continue assuming that all attributes and levels are relevant and consequently invoke a fully compensatory choice behavior.

Despite the different types of heuristics, within choice modelling field the evidence for non-compensatory behavior has focused on ANA. This heuristic has been particularly prominent as an element of the broader literature on choice modelling attribute processing (see Leong & Hensher, 2012 for a review). ANA has been studied in different fields, such as transportation (e.g., Hensher, Collins, & Greene, 2013; Hensher, Rose, & Greene, 2005), environmental economics (e.g., Campbell, Hensher, & Scarpa, 2011; Scarpa, Gilbride, Campbell, & Hensher, 2009), health economics (e.g., Hole, 2016; Lagarde, 2013), and more recently, on food/agricultural choice (Caputo, Van Loo, Scarpa, Nayga, & Verbeke, 2018; Chalak, Abiad, & Balcombe, 2016; Scarpa et al., 2013). The main focus are the implications either at the level of not considering ANA at all, or of ignoring the price/cost attribute. Over time, while the dominant themes were established, some important questions remained unsolved.

To detect and deal with this ANA, the literature proposes two main methods – one consists in making additional questions on whether specific attributes are ignored (stated method), and the other requires a

model that can reveal the extent to which each attribute is preserved across a sample (inferred method). Nevertheless, there is evidence that not all respondents who stated to have not considered an attribute really did it (Hess & Hensher, 2013), which raise problems about the reliability of responses to supplementary questions (e.g. Hensher & Rose, 2012; Hess & Hensher, 2010). This concern has increased the interest in detecting and understanding the impacts of ANA through model inference, accommodating a latent ANA behavior (e.g. Caputo, Van Loo, Scarpa, Nayga, & Verbeke, 2018; Hess & Hensher, 2010; Hole, 2011; Scarpa, Gilbride, Campbell, & Hensher, 2009; Scarpa, Zanoli, Bruschi, & Naspetti, 2012).

The impact of accounting for ANA depends on the assumed source of this behavior. When ANA results from a choice simplification strategy, it could lead to erroneous inferences. Inversely, if the evidence on ANA is interpreted as an expression of preference heterogeneity, this heuristic may not be a problem, but a valid behavioral phenomenon. In this context, the econometric specifications and the robustness of the methods play an important role, especially to avoid suboptimal policy and pricing decisions. For the same reason, there is a greater concern related to nonattending to the monetary attribute by decision makers, as this precludes the identification of the willingness to pay (WTP) values. In this special case, respondents who are indifferent to cost have no upper limit in their marginal WTP (Campbell, Hensher, & Scarpa, 2012). So, ANA has serious consequences on the calculation of predicted probabilities and welfare estimates. Empirical studies show that not accounting for attribute processing strategies results in misspecification bias (e.g. Hensher, 2006), such that the rationality assumption in choice models could lead to doubts about the accuracy of the behavioral outputs.

Motivated by the issues that may arise from not accounting for ANA and the still unknown of their extent, this paper provides a review offering a unified presentation of previous literature on this issue, by synthesizing findings, integrating research streams from different fields, and identifying research priorities and future developments. Specifically, this research investigates the impact of ANA, with focus in the four main areas of non-attendance - transportation, food/agricultural economics, health economics and environmental economics. We also offer a general classification scheme for the econometric framework to account for ANA, and recognizing the commonalities between the different studies, we transform them into common language and notation. Further, the paper summarizes the empirical findings, and provides a systematization of the evidence for each research area, to gain a more complete understanding of the consequences of this behavior.

The reminder of this paper is organized as follows. Section 2 presents the role of heuristics in decision making process, with special attention to ANA patterns. Section 3 provides a unified framework to deal with ANA. Section 4 includes the criteria of the review process, the description of studies and the main findings in the four most relevant areas of non-attendance. Finally, section 5 presents the discussion and directions for future research.

2. Heuristics in decision making process

DCE assume that individuals behave in a fully compensatory fashion, such that there is always an additional amount of an attribute that compensates the reduction in another and keeps the respondent on a given indifference curve. That curve should be continuous, otherwise the WTP for some changes in attributes is not defined. Two assumptions of consumer theory underpin the DCE: i) individuals make compensatory decisions, and they are willing to make trade-offs between the attributes (gains or losses); ii) preferences are continuous, which implies that if the individuals carry out the trade-off process, two similar combinations of goods to each other will occupy a near location in the ordering of consumer preferences. However, psychology literature provides empirical evidence showing that there are limits to how much information respondents can process, leading to the use of particular processing rules to simplify their decisions and to minimize cognitive effort. This literature contrasts with the basic assumption of the neoclassic model of consumer choice, which stipulates a utility maximizing agent with complete information and full knowledge of preferences. Within the DCE this assumption implies that all attributes are considered by individuals and are traded off in a compensatory manner. Nevertheless, efforts of the choice decision increase with choice complexity (Caussade, Ortúzar, Rizzi, & Hensher, 2005) and the use of heuristics appears to overcome cognitive needs to answer to DCE requirements (Heiner, 1983).

Given the recognition in behavioral research since the 1950s (see Simon, 1955) that consumers have a rule-driven behavior, based on a bounded rationality framework considering only a part of the available information, it is accepted that cognitive processes have a key role in preference elicitation exercises. Over the years, the recognition of heuristics in the choice modelling literature has made important advances in terms of modelling these decision rules. So, understanding the individuals/respondents' choice process is an important concern in the recent literature, given its possible nonconformity with the relevant assumptions of consumer theory on which DCE relies.

Among the several existing strategies that individuals might adopt, ANA became the focus of choice modelers as the most relevant heuristic to elicit individual's preferences. This non-compensatory behavior precludes the computation of the marginal rates of substitution and the estimation of the WTP measures (Campbell, Hutchinson, & Scarpa, 2008; Lancsar & Louviere, 2006). An exception may be ignoring an attribute by influence of the levels of other attributes, thus the respondent behaves in a compensatory or non-compensatory manner as attribute levels change in the choice experiment, as can be seen in Puckett and Hensher (2008).

Reasons for ANA

Besides the importance of considering ANA, an issue of relevance for welfare analysis is also to be able to distinguish the reasons for ignoring attributes. It is still an incipient strand of the literature concerning the extent to which respondents fall back on heuristics. ANA may be the result of several factors, we propose to organize into three groups: behavioral reasons, DCE design problems, and psychological reasons.

In the first group we include three main reasons for non-attendance: (1) task simplification to deal with perceived complexity of the choice task (Hensher et al., 2005; Carlsson et al., 2010; Lagarde, 2013; Alemu et al., 2013), in addition to the unfamiliarity with the issue under valuation which requires a higher cognitive effort (Carlsson et al., 2010; Lagarde, 2013); (2) attribute irrelevance for the individual (Hensher et al., 2005; Lagarde, 2013; Alemu et al., 2013); and (3) refusing to consider the attribute against others, specially the trade-off between the attribute money and the environmental attribute (Alemu et al., 2013; Carlsson et al., 2010; Lagarde, 2013). Scarpa et al. (2009) found a high rate of ANA to cost attribute and the reason behind may be associated with the difficulty in trading off aspects of rural landscape with money. This conduct may result from a strategic behavior, or from individuals' beliefs about certain attributes trade-offs. This may be particularly prominent in non-market valuation studies, corroborating the findings of Carlsson et al. (2010) and Lagarde (2013). In addition, Alemu et al. (2013) distinguished real zero preferences (when the attribute does not affect respondent's utility) from a task simplification strategy and found that the statement "It made it easier to choose between alternatives" best described respondents' motivation.

In the second group we include the reasons for ANA as a consequence of the experimental design. On the one hand, the degree to which individuals ignore attributes is influenced by the dimensionality and complexity of the DCE and also by the choice of attributes and levels. For example, as the number of alternatives in each choice set increase, the cognitive burden is higher and it has a large influence on

error variance (Caussade et al., 2005). The deviation of attribute levels from a reference alternative (Hensher, 2006), the presence of a dominant attribute, or the inappropriate ranges of attribute levels may also not justify the trade-offs in the choice process (Alemu et al., 2013; Carlsson et al., 2010; Sælensminde, 2006). In fact, to ensure the trade-off process, the attribute levels should be sufficiently differentiated. Nevertheless, the paper of Weller, Oehlmann, Mariel, and Meyerhoff (2014) whose main focus is the influence of experimental design on stated and inferred ANA, varying the number of choice sets, alternatives, or attributes provides evidence that the connection between design dimensions and ANA is not strong. On the other hand, the way the information is presented can also contribute to the use of simplifying decision rules. Hoehn, Lupi, and Kaplowitz (2010) explore how the format in which information is presented affects stated choice outcomes and show that alternatives and attributes as text format makes comparison more difficult, leading to the use of heuristics, particularly the attribute elimination. Day et al. (2012) investigate a variety of different forms of position-dependent and precedent-dependent ordering effect in preferences and argue that when the respondent does not know the choice tasks before the exercise, this could lead to more problems compared to when respondents are made aware of the full set of choice tasks.

Finally, the third group includes other reasons, as a result from the psychology research contributions. The presence of very strong preferences and beliefs about the importance of an attribute can eliminate the use of any exchange process (Scott, 2002). The possible absence of *a priori* preferences (Payne et al., 1992) and the resulting construction as the choices are made (Kjær, Bech, Gyrd-Hansen, & Hart-Hansen, 2006; Scott, 2002), or throughout the choice task, means that preferences will be context specific. This imply that preferences in a choice set might be influenced by other factors than the attributes of the alternative under analysis.

In order to reduce the probability of using heuristics, Vanniyasingam et al. (2018) present a very recent systematic review, with the aim of defining the best approaches to structure a DCE, to reduce cognitive burden and to get better statistical efficiency. They conclude that this depends on the number of choice tasks or alternatives, arguing for fewer attributes, and levels. It also improves with model-based approaches including covariates for response behavior, considering previous knowledge, specifying good Bayesian priors, accounting for heterogeneity, minimizing parameter prior variances, and using the appropriate method for the experimental design.

3. Dealing with Attribute nonattendance

In this section we review the main approaches that have been used in the literature to identify ANA, as well as, the related modelling issues and econometric framework to deal with ANA. This constitutes a piece of uniformed information, in which we identify the commonalities between models and provide a common notation and modelling framework.

3.1. Identifying ANA

The stated ANA approach defines two distinct ways to identify ANA: i) the commonly referred as "Choice task ANA", by asking respondents to identify the ignored attributes immediately after each choice set; and ii) "Serial ANA", where respondents are asked the same question only once, at the end of all choice tasks. When an attribute has been ignored, its marginal utility is conditioned to zero. The main concern between these variants relates to the fact that, in the serial ANA, a respondent who specified to have ignored an attribute, may instead have a lower intensity of preferences for that attribute (Hess, Stathopoulos, Campbell, O'Neill, & Caussade, 2013). So, the choice task ANA approach may be susceptible to give deeper insights about attribute attendance. While some studies adopted this variant (e.g. Ortega and Ward, 2016; Scarpa et al., 2010), others (e.g. Balcombe 2015) highlight the additional costs and survey length of collecting ANA after each choice task, also suggesting its theoretical justification weakness. Several studies, including the first stated choice studies on ANA, based their identification of such behavior through self-statements (e.g. Alemu et al., 2013; Carlsson, Kataria, & Lampi, 2010; Hensher et al., 2005; Rose, Hensher, Greene, & Washington, 2012), but concerns related to what respondents say and what they really do when choosing arise. In this case, it may happen that a respondent states that he ignored an attribute but the statistical evidence of their choices suggests otherwise. This leads to questioning whether these respondents' statements are adequate to measure such behavior (Hensher & Rose, 2009; Hess & Hensher, 2010). By this reason, we face a lively debate in the current literature about the reliability of self-statement on ANA.

On the other hand, rather than asking respondents to state their attribute attendance, model inference (e.g. Hess, Stathopoulos, Campbell, O'Neill, & Cau ssade, 2013; Hole, 2011; Kravchenko, 2014; Scarpa, Zanoli, Bruschi, & Naspetti, 2012) appears to be a valid alternative approach, and do not require further data. This method relies on the estimation of probabilistic models to infer non-attendance. Respondents are grouped into classes, and each class corresponds to a specific attribute non-attendance pattern, with the parameters of ignored attributes in a particular class being restricted to zero.

3.2. Econometric framework

3.2.1. Mixed logit model using stated ANA

Within the discrete choice models, the conditional or commonly called multinomial logit (MNL) model from McFadden (1974) is the most widely used and constitutes a reference point when compared to other models. One of the main limitations of this model is the strong assumption of preference homogeneity between the sampled respondents. The Mixed logit (MIXL), also known as random parameter logit model, generalizes the standard logit by allowing for preference heterogeneity. This is one of the most used models to account for stated ANA (e.g. Colombo et al., 2013; Rose et al., 2012). The basis for this model is the following utility function for the individual n = 1, ..., N when choosing the alternative j = 1..., J, in choice situation t = 1, ..., T:

$$U_{njt} = \beta_n' x_{njt} + \varepsilon_{njt} \quad (1)$$

where $\beta_n = \beta + \Delta z_n + w_n$, and varies over respondents introducing individual heterogeneity. The first two terms denote heterogeneity in the mean of the distribution of the random parameter, representing z_n the individuals´ characteristics when included, and w_n captures non-observable effects. x_{njt} the K attributes of alternative j, and ε_{njt} the error term which follows an iid (independent and identically distributed) extreme value distribution. The choice probabilities of the model, conditional on β_n are:

$$P_{njt} = \frac{\exp\left(\beta_n' x_{njt}\right)}{\sum_{j=1}^{J} \exp\left(\beta_n' x_{njt}\right)} \quad (2)$$

To account for ANA, β_n may be conditioned depending on whether a respondent ignored the corresponding attribute or not. The common procedure is to restrict the parameter on the ignored attribute to zero, according to respondents´ statements about ANA in the supplementary questions.

3.2.2. Model inference

The challenge here is how to define and incorporate the ANA phenomenon into models. An inferred method proposed by Hess & Hensher (2010), sometimes called as HH method (e.g. Ortega and Ward, 2016; Scarpa et al., 2013), relies on the analysis of respondent-specific coefficient distributions, by conditioning on observed behavior. Two groups of respondents, attribute attenders and attribute nonattenders, should be identified from the self-statements about ANA. When a respondent state to have ignored a specific attribute, it should be reflected on his choice behavior. When the attribute k was stated to be attended the estimated coefficient β_k should be higher than when the attribute was ignored. The

authors suggest to formulate and indirect utility function to allow the estimation of separate parameters for each attribute. This strategy recommends to work with the coefficient of variation (CV) of estimated conditional mean (μ_{kn}) and variances (σ_{kn}^2) for each respondent *n*, calculated as $CV_{kn} = \sigma_{kn}/\mu_{kn}$. So, the allocation between the two groups should be based on CV rather than on respondents' self-statements, by comparing the CV value to a defined threshold λ[4]. If the $CV_{kn} > \lambda$, this respondent should be allocated to the group of nonattenders for *k*, and vice-versa.

3.2.3. Latent class framework

Without the need of additional data, the most common approach and the one that gained more relevance in the last years, is the latent class framework to identify probabilistic decision rules. In this approach, different alternative specifications can be performed according to the assumptions made about non-attendance, by restricting parameters in each class, and also about modelling preferences. Respondents are divided into a discrete number of classes, defined by the analyst according to the attribute non-attendance patterns to be analyzed, such that each class represents different ANA strategies used by respondents. Under this approach, we identify two main model classifications: i) the ones assuming preference homogeneity but information processing heterogeneity, and ii) models assuming both preference and processing heterogeneity.

3.2.3.4 Models of preference homogeneity but processing heterogeneity

Within this classification, we include the models assuming the same taste intensities between respondents, but allowing for attribute processing heterogeneity, such as the equality constrained latent class (ECLC) models and the endogenous attribute attendance (EAA) models. The ECLC model, proposed by Scarpa et al. (2009), is one of the most parsimonious ways to modelling ANA, by constraining ignored attributes coefficients to zero, and assuming utility weights to take the same value across classes for attended attributes. In each class, an ignored attribute will have 0-utility weight, while an attended attribute will take the same nonzero coefficient across latent classes, as in multinomial logit models. The model infers class probabilities from observed choices, such that respondents will be assigned to a class that best defines their choice strategies according to the pattern defined in that class. A full attendance strategy is commonly assumed, where some respondents attended all attributes, these are the ones who behave in a fully compensatory manner. On the opposite side, not always considered, we may also group

[4] Hess and Hensher (2010) defined a threshold of 2, which they considered as a conservative value, and leaving space to more work on the impact of threshold choice.

respondents who ignore all attributes, exhibiting random choice based on chance alone. In the intermediate classes, we can define several patterns. One of the most employed strategy is to consider classes for non-attendance to one attribute, and also classes for lexicographic choices, where only one of the K attributes is considered, setting all K–1 coefficients to zero (e.g. Glenk et al., 2015; Scarpa et al., 2012). However, other combinations can also be defined (e.g. Kragt, 2013). In fact, with K attributes we have 2^K possible combinations of attendance and non-attendance patterns employed by respondents. As can be noted, the complexity of this appealing approach increases with the increasing number of attributes. So, the analyst faces an important decision by defining the number of ANA patterns to be explored, being it a combination of two ignored attributes, three, etc., or a mixed mapping at different ANA-levels. Within this context, the probability of respondent n=1, ..., N to select alternative j=1, ..., J, conditional on belonging to a particular class c which includes the choice of attribute subset Q_c, is formally given by:

$$P_{njt|Qc} = \frac{\exp\left(\sum_{k\in Qc} \beta^k x'_{njt}\right)}{\sum_{j=1}^{Jn} \exp\left(\sum_{k\in Qc} \beta^k x'_{njt}\right)} \tag{3}$$

The probability that respondent belongs to a particular class c, is given by equation 4 (Greene & Hensher, 2003):

$$H_{nc(ECLC)} = \frac{\exp\left(\theta_c z_n\right)}{\sum_{c\in C} \exp\left(\theta_c z_n\right)}, \qquad c = 1, \dots, C; \;\; \theta_C = 0 \tag{4}$$

θ_c are class-specific parameters, C^{th} is a normalized vector for model identification purposes, and z_n is a vector of alternative attributes and individual's characteristics, when included. In this class assignment model remains the biggest difference between ECLC and EAA models. As in the ECLC model, in the EAA (Hole, 2011) model it is assumed that respondents who attend to an attribute have identical preferences. However, the EAA approach assumes that ANA probabilities are independent over attributes (also called independent ANA model). Hole (2011) proposes a two-stage structure, where the decision of accounting an attribute is endogenously considered. In the 1st stage, the respondent selects the subset of attributes to consider in the choice process of an alternative (Q_c). Conditional on the attributes' choice in the first stage, in the 2nd stage the preferred alternative is chosen. The utility function is conditioned on the subset of attributes chosen such that:

$$U_{njt} = \sum_{k \epsilon Qc} \beta^k \, x_{njt}^k + \varepsilon_{njt} \tag{5}$$

with x_{nj}^k as the value of attribute *k* relating to alternative *j* and β^k is the preference weight.

But while the probability that respondent belongs to a particular class *c,* choosing the subset Q_c, in the ECLC models follow eq. 4, in the EAA it becomes:

$$H_{nQc(EAA)} = \prod_{k \in Qc} \frac{\exp\left(\theta_k z_n\right)}{1 + \exp\left(\theta_k z_n\right)} \prod_{k \notin Qc} \frac{1}{1 + \exp\left(\theta_k z_n\right)} \tag{6}$$

where the first part of the equation represents the probability of taking the attribute *k* into account in the first stage, being θ_k a vector of parameters, only identified when $\beta_k \neq 0$. This model has the advantage of not having to define the number of classes a priori. We should note that this approach is more parsimonious than the ECLC, having the latter a higher degree of complexity, and being more data demanding.

3.2.3.5 Models of both preference and processing heterogeneity

The accommodation of individual preference heterogeneity is an important issue, since it is not realistic to assume that all individuals have an identical preference and choice system. The homogeneity of taste intensity assumed above can be relaxed when there is interest in identifying preference heterogeneity across classes, allowing β to vary. These models include another layer of heterogeneity, by assuming variation in behavior due to differences in the tastes added to differences in the decision-making process of individuals. Within these models, we also identify two variants related to the underlying preference structure. Those whose preference structure is a LCM or finite mixture, and those where the preference structure is a mixed logit model. This extension appears to be important, given the evidence of confounding attribute non-attendance with weak preferences (Alemu et al., 2013; Hess et al., 2013), and leading to high ANA rates.

i) Preference structure is a LCM

As in the section above, two sub-models are estimated. First, a model for class allocation, where the (multinomial logit) probability of class membership (H_{nc}) is expressed as given by eq. 4. Second, a model for choice within class, whose probability of respondent *n* selecting the alternative *j*, is weighted by the class-allocation choice probabilities for that specific individual, as:

$$P_{njt|c} = \sum_{c=1}^{C} H_{nc} \frac{\exp(\beta_c x'_{njt})}{\sum_{j=1}^{Jn} \exp(\beta_c x'_{njt})} \quad (7)$$

β_c is a vector of all possible ANA strategies, where all non-zero elements (for Q_c) are different across classes. Each class now has a behavioral meaning. This formulation uses fixed parameters within each class, but can allow for random parameters, which we will see next.

ii) Preference structure is a mixed logit model

These models, known as LC_MMNL or RPLCM, which generally stands for LCM with random parameters, combine preference heterogeneity with ANA. These are an extension of the fixed parameter (or standard) latent class model, and allow for continuous variation of the parameter vector within classes, i.e. instead of a discrete mixture of taste coefficients, these can vary over respondents with continuous density (some applications are given in Collins et al., 2013; Greene and Hensher, 2013; Hensher et al., 2013; Hess et al., 2013; Hole et al., 2013). This heterogeneity is similar to mixed logit and is given by eq. 8:

$$\beta_{n|c} = \beta_c + w_{n|c} \quad (8)$$

being $w_{n|c}$ uncorrelated with exogenous data. The conditional probability of respondent *n* selects alternative *j* becomes:

$$P_{njt|Qc} = \frac{\exp\left[\sum_{j=1}^{J} y_{njt} (\beta_c + w_n)' x_{njt}\right]}{\sum_{j=1}^{Jn} \exp\left[\sum_{j=1}^{J} y_{njt} (\beta_c + w_n)' x_{njt}\right]} \quad (9)$$

where $y_{njt} = 1$ when *j* corresponds to the alternative chosen, and 0 otherwise. The class probabilities can be parametrized using an MNL formulation such in the eq. 4 or a discrete mixture like the class structure following eq. 6 assuming full independence of ANA across attributes (e.g. Hess et al., 2013; Hole et al., 2013). The utility function given by eq. 5 comes with the subscript *n* for β, which is now an individual-specific coefficient for the attribute. Table 1 presents our classification scheme to deal with ANA.

Table 1 - Classification scheme to deal with ANA

Heuristic	Identification	Modelling	
		Preference homogeneity and attribute processing heterogeneity	**Preference heterogeneity and attribute processing heterogeneity**
Attribute non-attendance	**Stated approach** Restricting parameters of the ignored attributes to zero	Multinomial logit model (MNL) Parameter: β	Mixed logit model (MIXL) Parameter: β_n
	Inferred approach Restricting parameters of the ignored attributes in a particular class c=1, ..., C to zero	Equality constrained latent class (ECLC) and endogenous attribute attendance (EAA) models where β_1=(0,0) β_2=(β_a,0) ... β_C=(β_a, β_b) Such that parameter: $\beta_{a\|1}$=$\beta_{a\|2}$=...=$\beta_{a\|C}$	Preference structure is a LCM: LCM with parameter: $\beta_{a\|1}\neq\beta_{a\|2}\neq...\neq\beta_{a\|C}$ Preference structure is a Mixed logit: RPLCM with parameter: $\beta_{na\|1}\neq\beta_{na\|2}\neq...\neq\beta_{na\|C}$

4. Overview of application

This section seeks to provide an overview of the DCE studies accounting for ANA included for review. We analyze the work and present the general findings following the proposed classification scheme in the previous section, to facilitate the reading, through the standardization of language used in each study.

4.1. Methods for review

We select different libraries (Econlit, ScienceDirect, ISI Web of Science, EBSCO Host JSTOR) to search for the most relevant literature on attribute non-attendance, with a complementary search using Google Scholar. The search was based on combinations of the following terms: “discrete choice experiment”, “discrete choice experiments”, “discrete choice modeling”, “discrete choice modelling”, “choice experiment”, “stated preferences”, “discrete choice”, “attribute nonattendance”, “attribute non-attendance”, “heuristics”, “decision rules”. A particular study was included if it explicitly uses DCE methodology, has an hypothetical nature, analyses the impacts of attribute processing strategies on WTP

and/or model fit, provides information about the experimental design, namely, the number of attributes, alternatives and choice tasks, clearly belongs to one of the four fields (transportation, health, environmental, food/agricultural economics), and was published as a full-text article in English between January 2005 and December 2018. This review provides a synthesis and comparison of the findings of the identified studies, which were compared in relation to study design, methodology and key results in terms of non-attendance consequences on model fit and WTP.

4.2. Description of studies

Forty-three studies met the inclusion criteria. More than half of the sample belong to the transportation and environmental field, eight to food/agriculture and six to health. Appendix C contains a summary of the studies selected. Pioneering studies belong to the transportation field, and gradually moved towards the fields of health and food/agricultural. Concerning the experimental design, in general, the number of attributes appear to be higher in both transportation and health studies, with five or six attributes, although there are some exceptions. In environmental valuations, despite the majority of studies include five attributes, the most recent studies have included a smaller number. In food studies, the number of attributes ranges from four to seven. In environmental and health valuations respondents face six or less choice sets, with two or three alternatives. Inversely, in transportation most of the papers present sixteen choice tasks, however only comprising two alternatives. Despite the low number of choice tasks, ranging from four to seven, and in some cases low number of attributes, in food applications the choice sets are formed by a higher number of alternatives, presenting an additional cognitive effort. Moreover, most of the DCE applications allow respondent to choose the status quo (none option), which is encouraged to add more realism to the choice experiment.

4.3. Synthesis of findings

4.3.1. Transportation

Until 2012, the majority of transportation studies identified ANA through supplementary questions, and using the MIXL model by restricting the individual-specific coefficient for the ignored attribute to zero. A set of papers (Hensher, 2006; Hensher et al., 2007, 2005; Hensher and Rose, 2009) focused on travel time savings for car commuter trips in Sydney, and found a high share of respondents ignoring at least one attribute. Regarding the impacts on the mean value of travel time savings (VTTS), the evidence suggests that accounting for non-attendance leads to a significant downward trend in VTTS estimates. However, the results from Hensher et al. (2007) and Hensher and Rose (2009) suggest a deflating effect

when ANA behavior is ignored, while Puckett and Hensher (2008) show overestimated VTTS. On the other hand, Puckett and Hensher (2009) found that ignoring decision process heterogeneity in ANA strategies have significant impacts in the estimates of VTTS. In the context of choice of airline carrier for an interstate holiday, Rose et al. (2012) explore stated ANA, and found the flight time variability as the most ignored attribute, by 66% of respondents, while the price was the most attended, by almost 90%. The significant differences in parameter estimates and WTP when accounting for ANA demonstrate the importance of including the processing strategies undertaking stated preferences work.

Other studies employ the LCM and contrast it with the results from the MNL and MIXL model. Hensher and Greene (2010) consider 3 types of heuristics: 1) ANA; 2) a heuristic related to the way common-metric attributes are jointly evaluated; 3) a heuristic as a variant of the previous. They inferred non-attendance through LCM, which increased the goodness-of-fit and for which the mean estimate of VTTS is 39.4% higher than MNL and 25.42% higher than MIXL. This result highlights that ignoring process heterogeneity may underestimate the marginal WTP for travel time savings. Hensher, Rose, and Greene (2012) also infer attribute processing strategies through the ECLC model. This model outperformed the MNL based on full attribute attendance, also suggesting an under-estimation of the mean welfare measures. In terms of VTTS, the impacts depend on the respondents' sample included. Both ANA and VTTS measures depend on the relevance of the range of levels for time and cost to respondents, which highlights the importance of attribute selection and experimental design conception. Hess and Hensher (2010) also infer ANA through an analysis of respondent-specific coefficient distributions from conditioning on observed choices, which we already denoted as HH method, providing an interesting comparison of both ways to identify ANA. Results support the existence of this heuristic, but also reveal differences in the two approaches, such that the inferred approach using the coefficient of variation fits the data well.

Since 2013, the transportation applications included random parameters in the latent class framework. Hensher et al. (2013) show that allowing for taste heterogeneity increases the probability of full attendance, and more importantly it provides evidence that dealing with ANA through latent classes but fixed parameters could lead to an exaggerated ANA rate. However, in terms of VTTS, there is no statistically significant difference when adding random parameters to the model that accounts for full attendance and ANA. Collins et al. (2013) present a generalized RPLCM that allows for an arbitrary degree of correlation of non-attendance across attributes. The stated ANA is introduced as a covariate, allowing ANA rate to vary across decision makers, which improves the model fit. The model specification yields

inferred ANA rates more aligned with stated ANA responses, and lower ANA rates than the other models. The fare attribute was ignored by 6.9%, while the rate is somewhat higher in both correlated ANA specification and independent ANA, 13.63% and 13.26%, respectively.

To study travelers' choices between alternative trips in various contexts, Hess et al. (2013) added random parameters to a latent class structure but assuming independence across class probabilities, which is a variant both of the EAA and of the RPLCM with independent ANA. They found a problematic rate of non-attendance to fare, being ignored by 44% of commuters. The random nature of the employed approach leads to a significant drop in the shares for the non-attendance classes, and in several cases, it drops to zero. In fact, Hess et al. (2013) argue that the latent class approach is misguided, and its results are likely to be affected by confounding with regular taste heterogeneity. So, as in Campbell et al. (2012), Hess et al. (2013) show that high ANA rates may be caused by not adequately deal with taste heterogeneity. Greene and Hensher (2013) also apply the two-stage EAA allowing for preference heterogeneity (such as Hole, 2011). The improvement by adding random parameters is recognized, and results show that the WTP estimates are upward biased in MNL and MIXL.

4.3.2. Environmental economics

As in the transportation field, the first studies accounting for ANA in environmental economics deal with stated ANA in a multinomial logit framework, which was the dominant approach until 2013. In a study of multiple park management services Scarpa, Thiene, and Hensher (2010) compared the *choice task* ANA and *serial* ANA within the MNL framework. Evidence indicates that individuals ignore different attributes, and this intra-panel variation is a source of bias regarding welfare estimates. The direction of the sign in the WTP changes remains unclear, but in general, higher WTP measures were found in serial ANA, but lower at the choice task level, when compared to the full attendance model. Several studies (Campbell et al., 2008; Kosenius, 2013; Meyerhoff & Liebe, 2009) apply the error component logit model incorporating stated non-attendance. Studying onshore wind power generation in Germany, Meyerhoff and Liebe (2009) found no significant differences in marginal WTP values when accounting or not for ANA, and also little difference in the implicit prices according to how ANA is classified. On the other hand, Campbell et al. (2008) evaluated the implications of discontinuous preferences in the context of rural landscape improvements in Ireland, and found a higher rate of non-attendance to cost (70%). More than thirty per cent of respondents ignored at least one of the attributes, and one-fifth of respondents exhibited a lexicographic choice behavior. Introducing ANA decreases WTP measures by approximately 50%. In a DCE on water quality in Finland, Kosenius (2013) show that accounting for information on preference

discontinuity improves the model performance, and the elimination of less important attributes increases the efficiency of WTP measures.

In turn, Carlsson et al. (2010) explored the impacts of stated ANA on WTP in three different environmental DCE, by mean of MIXL models. The cost and cultural assets were the most ignored attributes, while environmental attributes were considered by a higher rate of respondents (90%). The authors found support that not all respondents who stated to have not considered an attribute really did it, which can be an argument against using respondent' stated information about the choice process. Similar to Meyerhoff and Liebe (2009), Carlsson et al. (2010) did not find systematic differences in WTP between models with and without consideration of non-attendance. Instead of using a binary question for ANA, Colombo et al. (2013) considered three cases of attendance (may always, sometimes, or never pay attention) in the valuation of ecosystem services and biodiversity enhancements in UK. The highest rate of non-attendance was found for the wild food attribute, never considered by 15% of the respondents. Different mixed logit models were estimated to account for non-attendance, and results show important impacts in WTP depending on the assumptions in modelling attribute attendance. This highlights the importance of partial attendance, since a dichotomous question will allocate respondents into the always or never considered group, when they actually consider the attribute sometimes. Seeking to find the reasons behind ANA, Alemu et al. (2013) include follow-up questions in their stated choice study. The main findings, from the estimation of ECLM and the HH method, suggest that respondents who state to have ignored attributes to simplify their choices, have really not completely ignored such attributes, and asking the reasons for ANA improve parameter estimation, more than just taking stated non-attendance into account.

The inference of ANA using equality constraints started with Scarpa et al. (2009) who conducted a stated preference survey to value rural landscapes in Ireland. Authors infer that accounting for non-attendance improves model fit in comparison to models that assume full attribute attendance, in addition it produces substantially smaller WTP values. The study reports an important result for non-market valuation by showing a staggering rate, between 80 and 90 per cent, of respondents not attending to cost. Significant rates were also found in Campbell et al. (2011) with 61% of respondents revealing not attending to cost, and the same direction bias in WTP. Making a comparison between both stated and inferred non-attendance for natural resources management in Tasmania, Kragt (2013) estimated a MIXL model to incorporate serial ANA and an ECLC model for the inferred ANA. The main conclusions for serial non-attendance suggest that less than 55% of respondents make trade-offs between all attributes, and

approximately 18% exhibited lexicographic preferences. Comparing the results of both approaches, the supplementary questions exhibit a much lower level of non-attendance, 21.9%, when this rate actually increases to 39% through model inference. However, no significant differences were found in terms of WTP measures when ANA is accounted. Contrariwise, from the ECLC models it is possible to conclude that WTP estimates appear to be overestimated when ANA is ignored. Glenk, Manuel, and Potts (2015) conducted a DCE to elicit preferences for ecological improvements of water bodies in two Spanish river basins. Authors inferred non-attendance through ECLC models which provide empirical evidence of ANA impacts in the WTP measures, with the direction of bias depending on the sample. Jourdain and Vivithkeyoonvong (2017) explore the valuation of ecosystem services associated with irrigated rice agriculture, in Thailand. The main results show that when compared with the MNL, the inferred approach for ANA improved the model fit and shows a high rate of full attendance, 45%. In line with other studies, the findings suggest that failing to consider ANA would overestimate the WTP.

Including preference heterogeneity, Colombo and Glenk (2014) investigated two types of decision rules: the attribute non-attendance and the non-attendance to alternatives due to unacceptable attribute levels in the choice task, defining a LCM preference structure. The authors assume that respondents would employ the "eliminating by aspects rule" to eliminate alternatives, and then the remaining alternatives are evaluated following a compensatory decision rule or through an ANA approach. According to this model, each respondent determines the most important attribute and the alternatives with smaller values for that attribute are then eliminated, such that the number of alternatives is gradually reduced. Different LCM were compared. Results show that non-attendance to alternatives occurred in 13.8% of the choices, and modelling ANA and non-attendance to alternatives improve the model fit and provides an explanation for a larger degree of the choice variability. This paper clearly shows that ANA may occur jointly with other decision rules, and also shows the need for some caution regarding the conclusion of Scarpa et al. (2010). Focused on household preferences for tap water Thiene, Scarpa, and Louviere (2015) employed a latent class model to address simultaneously three sources of inter-personal variation: scale, taste and attribute attendance. Model comparisons were performed and successfully isolate the three behavioral issues. In terms of WTP measures, the model that captures preference, scale and ANA differences gives the most plausible results, and shows that WTP decreases when compared to the MNL model. Finally, in a more recent paper, Sandorf, Campbell, and Hanley (2017) incorporate random parameters in a latent class approach to explore the role of knowledge about cold-water coral on Norwegian respondents' propensity to ignore one or more attributes. In the same vein, results for the RPLCM show that the conditional WTP estimates decrease substantially when ANA is accounted for.

4.3.3. Health economics

In DCE health applications, ANA is potentially highly relevant, as assuming that attendance applies to all attributes would be erroneous and could lead to misguided policy recommendations. The majority of the studies reviewed account for this heuristic by using the inferred approach. Considering an application to patient's choice of general practitioner, Hole (2011) employ the EAA model, which inferred that the cost attribute was ignored by 70% of the sampled population. This approach has better fit and provides lower WTP measures when compared to a basic MNL. Later, Hole et al. (2013) extended the EAA model by incorporating preference heterogeneity, and allowing the ANA probability to depend on the respondents' stated non-attendance, being it an extension of Hess et al. (2013) methodology. Results show that the model fit increases when stated ANA is accounted for, which suggests that the self-reported data contain useful information about the respondents' attribute processing strategy. Allowing for preference heterogeneity also reduces the influence of ANA. The estimated probabilities of non-attendance are lower in the extended model than in the EAA model, showing that it could result in an incorrect classification of respondents with weak preferences as "non-attenders" in the simpler model. Contrary to the previous study, Hole et al. (2013) did not find a substantial difference in WTP across models. In 2016, Hole (2016) contrasts two alternative approaches: a variant of the EAA model and an ECLC model. Again, the results show that allowing for non-attendance improves model fit, but its impacts depend on the source of non-attendance. When it is considered as a form of preference heterogeneity, both models are similar to those from standard models. However, the models perform differently if respondents ignore attributes as a simplification decision rule.

Using an appealing method by estimating ECLC models, Lagarde (2013) investigated all possible ANA patterns (a total of 64) by healthcare providers, concerning adherence to clinical guidelines in Ghana. Results show that under 3% of respondents considered all attributes when making a choice, and the majority only takes into account one or two attributes. A high rate of 46.3% ignored the cost attribute, being the treatment approach the most ignored attribute, by 91.2%. Once ANA strategies were included in the analysis, the model fit improved and the welfare estimates showed an upward bias. This result is consistent with the study of Erdem et al. (2014) about public preferences for health service innovations in UK, finding a problematic rate of 85% of respondents ignoring the cost attribute. The study recognizes non-attendance not only to attributes but also to the attribute's levels (ALNA). The processing strategies are accounted using a modified latent class model (for more details see Erdem et al., 2014). Results show that respondents process each attribute' level differently, suggesting that this may be an important

issue, particularly with qualitative attribute levels. Moreover, allowing for ALNA leads to substantial improvements in the model fit and a significant reduction in marginal WTP estimates. However, only few studies have extended the analysis of non-attendance to the levels of attributes (e.g. Campbell et al. 2012). Finally, the most recent study, Heidenreich (2017), investigated whether respondents' choice and the subsequent ANA is a heuristic or a genuine preference, using the LCM. The paper explores the role of respondents' familiarity with the good, such that familiar DCE respondents are less likely to use a decision heuristic compared to unfamiliar. Almost 50% of familiar respondents ignore the cost, while for unfamiliar this rate decreases to 40%. Two WTP measures were calculated, one for ANA being a heuristic and the other for ANA as a preference (for more details see Heidenreich, 2017). Authors found that having wrong assumptions about ANA results in overestimated welfare effects when ANA is a preference instead of a heuristic.

4.3.4. Food choice

The literature on ANA in the field of food choice is recent. A mixed mapping of modelling techniques is identified, with some studies dealing with stated ANA and others dealing with both stated and inferred ANA. Within the MIXL model, Balcombe, Burton, & Rigby (2011) accounted for ANA in a context of genetically modified food. A rate of 27% of respondents appear to ignore at least one attribute, being the price the least frequently ignored attribute. Results also show that respondents who ignored an attribute in fact assigned a zero weight to this attribute. Later, Balcombe, Bitzios, and Fraser (2014) used a Bayesian econometric specification incorporating respondents' ranking over the importance of bread attributes, after all the choice tasks. Two different ways of dealing with this type of data were explored. First a modified MIXL model with ranking as covariates in marginal utilities was developed, followed by a new alternative designated by "contraction approach", with ranking data to scale the parameters (for more details see Balcombe et al., 2014). Evidence shows that ANA has impacts on WTP measures, though depending on the approach considered the bias has a different sign. But even so, the best performing model (contraction approach) shows a tendency for an upward revision in WTP estimates and improves the model performance when considering ANA. Kehlbacher et al. (2013) investigated the effect of stated attribute non-attendance for meat choice and for animal welfare, in Britain. Two DCE were conducted and respondents were asked which attributes they considered after the completion of the choice tasks. In the first choice experiment, the price/tax was ignored by 47.1% of respondents. Results from different MIXL models show that WTP estimates decrease substantially when ANA is taken into account, and evidence was found that respondents' accuracy to report their ANA behavior improves if

they know a priori of the existence of such a question. Chalak, Abiad, and Balcombe (2016) jointly used non-attendance and ranking data about consumers' preferences for safety certification in food, through a modification of the MIXL. The findings show a high correspondence between ranking data and non-attendance, and the stated non-attendance does not in fact correspond to having zero marginal utility for the ignored attributes.

Moving to the inferred approach, ANA effects are also explored in Ortega and Ward (2016), who include follow up questions after each choice task, and infer serial non-attendance based on the posterior distributions of random coefficient estimates, i.e. using the HH method. Results show that choice task stated non-attendance affects the marginal utility coefficients and the WTP values, but the inferred approach suggests that some attributes are not likely ignored as respondents indicate. Within the latent class framework, Scarpa et al. (2013) provide a comparison between the inferred and stated ANA on meat preferences. Between 28% and 35% of respondents never or rarely attend to price. In general, the estimates from the ECLC model addressing non-attendance increase the model fit, and leads to smaller WTP values when compared to the competing models. The paper of Caputo et al. (2018) explores consumer preferences for chicken breast product, and shows that WTP varies according to the way information on stated ANA is collected. The price was the most attended attribute, ignored by 26% of respondents in serial ANA. In line with Kragt (2013), the findings point out for low concordance between ECLC and serial ANA, being the choice task-level ANA measures more congruent with model estimates. The work of Thiene et al. (2013) on wine preferences is an example of the few applications of ANA based on mixed Revealed and Stated Preferences data. LCM where employed, and findings show that accommodating ANA reduces the differences in the estimates of preferences from revealed and stated choice experiments.

5. Concluding remarks

The scope of this paper provides broad contributions in several areas. First, it presents the theoretical insights regarding ANA, built in the context of heuristics use. Second, it standardizes the notation and language of existing methodologies dealing with ANA issues, making the analyst aware of the possible econometric specifications, its assumptions and commonalities. Third, it selectively reviews related literature providing a deep understanding of the impact of ANA on current choice modelling exercises, in four distinct and relevant areas of research.

All reviewed papers assume violations of the continuity of preferences principle, and account for the use of heuristics by respondents, although in an isolated way, with focus on ANA. In the stated approach, the majority of the reviewed papers employ the MIXL model, though others sought to endogenously account for ANA. The focus has shifted from the serial non-attendance to the choice-task non-attendance, with potential advantages of accounting for intra respondent heterogeneity in attribute attendance. Nevertheless, requesting for attribute attendance after each choice task is known as to be more complex for the respondent, especially with unfamiliar products. On the other hand, most of the findings seem to suggest that respondents who stated they do not attend some attributes do not fully ignore such attributes. It seems that respondents who indicate ANA place lower importance but do not ignore them altogether (Hess and Hensher, 2010; Alemu et al., 2013). Moreover, Colombo et al. (2013) show that relying on a measure of attribute attendance through a dichotomous option may lead to wrong conclusions, allocating respondents who only sometimes consider the attribute into the always or never considered group. Thus, a yes or no question does not allow for the possibility that a specific attribute only has lower instead of null importance.

Regarding the incidence of non-attendance, the most problematic rates were specially found in the environmental field, with cost attribute being ignored between 60% and 90% of respondents (e.g. Campbell et al., 2011, 2008; Scarpa et al., 2009). Higher rates were also found in health studies, with 85% of the respondents ignoring the cost attribute in Erdem et al. (2014) and 70% in Hole (2011). For lower, but still problematic rates, there is a mix of studies where the cost is ignored. Contrary, it is in the transportation field where attendance to the monetary attribute is notably higher, although there is some evidence in food choice experiments (Balcombe et al., 2011; Caputo et al., 2018; Thiene et al., 2013). This seems to indicate that, to some extent even independent of the experimental design, in the fields where there is a greater concern with the product under evaluation, especially in health, there is less price sensitivity, when compared to food and transportation choices. In the case of the environmental field, individuals do not appear to be available for trading-off the environment with money. Thus, if full attendance is assumed, this demonstrates the tremendous impact, not only numerically but also behaviorally, ANA regarding cost, can induce. This is particularly crucial for non-market valuation (e.g. agricultural public goods), where no relative implicit price can be calculated for those who behaved in a non-compensatory manner. On the other hand, the results for health policy recommendations require an additional robustness check, given the focus of health systems to strive for the higher health status of the population, dealing with serious problems, such as morbidity and mortality.

In general, in the studies using self-stated ANA, estimates are much lower and more realistic when the heuristic behavior is embedded in the model. However, in addition to costs in terms of survey time, from the literature review we can conclude that answering to supplementary questions may lead to reliability problems. To overcome this issue, inferred attribute non-attendance is in fact a valuable alternative, and appears to fit the data well. The most prevalent inferred econometric approach to identify ANA is an adaptation of the LCM, with the transportation field applying the most advanced models, by combining random parameters with LCM. Both in the stated or inferred approaches, the degree of ANA and its causal factors are critical to derive the estimates from the choice experiments since this behavior does not result in the truthful revelation of preferences. The empirical evidence analyzed in this literature review consistently shows that there is consensus that failing to account for non-attendance results in biased estimates and poor-quality models. The direction of bias is not clear a priori, and there is mixed evidence as to the direction of change of WTP measures when accounting for non-attendance. Despite the general agreement that taking ANA into account should improve the model fit and affect the related welfare estimates, there is no dominant way to do it. In any case, the inferred ANA using the analytical approach has strong appeal.

An important issue is that non-attendance may be confounded with preference heterogeneity, such that several studies found a reduction in the non-attendance rates when allowing for preference heterogeneity among respondents. A plausible explanation is that models that assume preference homogeneity, respondents may be incorrectly classified as non-attenders for an attribute, when they actually have weak preferences.

Insights for future research

There are many possible focuses for future research. To guarantee that the changes in a given attribute are important to respondents, the inclusion of a wider range of attribute levels should be considered in future studies. Given the plethora of heuristics that individuals may use, it is time to similarly consider multiple heuristics and how to include them simultaneously in choice models. For example, Hess et al. (2012), started to investigate a pair of heuristics in a latent class approach. They tested varying them across different data sets, using combinations of fully compensatory models and the lexicography-based paradigm. Results evidenced significant gains in model fit with behavioral mixing models. The very recent development from Balbontin et al. (2017) which simultaneously accounts for ANA and value learning, shows an improvement in models' goodness to fit, and helps to better understand choice behavior. It is thus of the upmost importance to continue developing combined heuristics approaches, with the

probabilistic models being a proven way. Some other literature (e.g. Balcombe, Fraser, & Mcsorley, 2015; Selivanova & Krabbe, 2018) point to the use of eye-tracking data, or visual inspection, as a future creative way in this research area including indicators of attribute attendance, attention, importance and the like. Accordingly, we point eye tracking as an emerging third approach to identify ANA. However, there still seems to be space to assess inferred versus stated approaches. Nevertheless, the ongoing debate in the literature related to the validity of stated responses, either at choice-task and even serial non-attendance, brings to the fore the issue of endogeneity bias. Dealing with this problem is a future key question. The designated Integrated Choice and Latent Variable Model, or simply Hybrid Choice Model (Ben-Akiva et al., 1999), appears as an important avenue to deal with endogeneity problems and to incorporate other sources of heterogeneity, such as attitudes, perceptions, knowledge, and social influence that may introduce bias when not accounted (Ben-Akiva et al., 2002; Hess & Beharry-Borg, 2012). This approach integrates a choice model combined with a latent variable model and presents an efficient use of data. Additionally, the hybrid structure may be able to provide more information, by capturing more closely the choice process through attitudinal data, being an interesting approach to study respondents' heterogeneity. Furthermore, this review may serve as a basis for a meta-analysis which can have different research purposes.

Acknowledgement

This work is financed by the project NORTE-01-0145-FEDER-000038 (INNOVINE & WINE – Innovation Platform of Vine & Wine, Portugal) and by European and Structural and Investment Funds in the FEDER, Spain component, and supported through the Operational Competitiveness and Internationalization Program, Portugal (COMPETE 2020) [Project No. 006971 UID/SOC/04011)]; and national funds, through the FCT – Portuguese Foundation for Science and Technology under the projects UID/SOC/04011/2019 and UID/ECO/03182/2019.

References

Alemu, M., Mørkbak, M., Olsen, S., & Jensen, C. (2013). Attending to the Reasons for Attribute Non-attendance in Choice Experiments. *Environmental Resource Economics*, *54*, 333–359. https://doi.org/10.1007/s10640-012-9597-8

Balbontin, C., Hensher, D. A., & Collins, A. T. (2017). Integrating attribute non-attendance and value learning with risk attitudes and perceptual conditioning. *Transportation Research Part E: Logistics and Transportation Review*, *97*, 172–191. https://doi.org/10.1016/j.tre.2016.11.002

Balcombe, K., Bitzios, M., & Fraser, I. (2014). Using Attribute Importance Rankings Within Discrete Choice Experiments : An Application to Valuing Bread Attributes. *Journal of Agricultural Economics*, *0*(2), 446–462. https://doi.org/10.1111/1477-9552.12051

Balcombe, K., Burton, M., & Rigby, D. (2011). Skew and attribute non-attendance within the Bayesian mixed logit model. *Journal of Environmental Economics and Management*, *62*(3), 446–461. https://doi.org/10.1016/j.jeem.2011.04.004

Balcombe, K., Fraser, I., & Mcsorley, E. (2015). Visual Attention and Attribute Attendance in Multi-Attribute Choice Experiments. *Journal of Applied Econometrics*, *30*(3), 447–467. https://doi.org/10.1002/jae.2383

Ben-akiva, M., McFadden, D., Garling, T., Gopinath, D., Walker, J., Bolduc, D., … Rao, V. (1999). Extended Framework for Modeling Choice Behavior. *Marketing Letters*, *10*(3), 187–203. https://doi.org/10.1023/A

Ben-Akiva, M., Walker, J., Bernardino, A., Gopinath, D. A., Morikawa, T., & Polydoropoulou, A. (2002).

Intergration of Choice and Latent Variable Models. In *In Perpetual Motion-Travel Behavior Research Opportunities and Application Challenges* (Mahmassani, pp. 431–470). https://doi.org/10.1016/B978-008044044-6/50022-X

Campbell, D., Hensher, D. A., & Scarpa, R. (2011). Non-attendance to attributes in environmental choice analysis: a latent class specification. *Journal of Environmental Planning and Management*, *54*(8), 1061–1076. https://doi.org/10.1080/09640568.2010.549367

Campbell, D., Hensher, D. A., & Scarpa, R. (2012). Cost thresholds, cut-offs and sensitivities in stated choice analysis: Identification and implications. *Resource and Energy Economics*, *34*(3), 396–411. https://doi.org/10.1016/j.reseneeco.2012.04.001

Campbell, D., Hutchinson, W. G., & Scarpa, R. (2008). Incorporating Discontinuous Preferences into the Analysis of Discrete Choice Experiments. *Environmental and Resource Economics*, *41*, 401–417. https://doi.org/10.1007/s10640-008-9198-8

Caputo, V., Van Loo, E. J., Scarpa, R., Nayga, R. M., & Verbeke, W. (2018). Comparing Serial, and Choice Task Stated and Inferred Attribute Non-Attendance Methods in Food Choice Experiments. *Journal of Agricultural Economics*, *69*(1), 35–57. https://doi.org/10.1111/1477-9552.12246

Carlsson, F., Kataria, M., & Lampi, E. (2010). Dealing with ignored attributes in choice experiments on valuation of Sweden's environmental quality objectives. *Environmental and Resource Economics*, *47*(1), 65–89. https://doi.org/10.1007/s10640-010-9365-6

Caussade, S., Ortúzar, J. D. D., Rizzi, L. I., & Hensher, D. A. (2005). Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B*, *39*(2005), 621–640. https://doi.org/10.1016/j.trb.2004.07.006

Chalak, A., Abiad, M., & Balcombe, K. (2016). Joint use of attribute importance rankings and non-attendance data in choice experiments. *European Review of Agricultural Economics*, *43*(5), 1–24. https://doi.org/10.1093/erae/jbw004

Collins, A. T., Rose, J. M., & Hensher, D. A. (2013). Specification issues in a generalised random parameters attribute nonattendance model. *Transportation Research Part B*, *56*, 234–253. https://doi.org/10.1016/j.trb.2013.08.001

Colombo, S., Christie, M., & Hanley, N. (2013). What are the consequences of ignoring attributes in choice experiments? Implications for ecosystem service valuation. *Ecological Economics*, *96*, 25–

35. https://doi.org/10.1016/j.ecolecon.2013.08.016

Colombo, S., & Glenk, K. (2014). Social preferences for agricultural policy instruments: joint consideration of non-attendance to attributes and to alternatives in modelling discrete choice data. *Journal of Environmental Planning and Management*, *57*(2), 215–232. https://doi.org/10.1080/09640568.2012.738190

Day, B., Bateman, I. J., Carson, R. T., Dupont, D., Louviere, J. J., Morimoto, S., ... Wang, P. (2012). Ordering effects and choice set awareness in repeat-response stated preference studies. *Journal of Environmental Economics and Management*, *63*(1), 73–91. https://doi.org/10.1016/j.jeem.2011.09.001

Erdem, S., Campbell, D., & Hole, A. R. (2014). Accounting for Attribute-Level Non-Attendance in a Health Choice Experiment: Does it Matter? *Health Economics*, *24*, 773–789. https://doi.org/10.1002/hec

Glenk, K., Manuel, J. M., & Potts, J. (2015). *Inferring Attribute Non-attendance from Discrete Choice Experiments : Implications for Benefit Transfer*, *60(4)*, 497–520. https://doi.org/10.1007/s10640-014-9777-9

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transportation Research Part B*, *37*, 681–698. https://doi.org/10.1016/S0191-2615(02)00046-2

Greene, W. H., & Hensher, D. A. (2013). Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. *Applied Economics*, *45*, 1897–1902. https://doi.org/10.1080/00036846.2011.650325

Heidenreich, S. (2017). Decision heuristic or preference ? Attribute non - attendance in discrete choice problems. *Health Economics*, *27,* 157–171. https://doi.org/10.1002/hec.3524

Heiner, R. A. (1983). The Origin of Predictable Behavior. *The American Economic Review*, *73*(4), 560–595. https://doi.org/10.1016/0165-4896(83)90055-0

Hensher, D. A. (2006). How Do Respondents Process Stated Choice Experiments? Attribute Consideration under Varying Information Load. *Journal of Applied Econometrics*, *21*, 861–878. https://doi.org/10.1002/jae.877

Hensher, D. A., Collins, A. T., & Greene, W. H. (2013). Accounting for attribute non-attendance and

common-metric aggregation in a probabilistic decision process mixed multinomial logit model: A warning on potential confounding. *Transportation*, *40*(5), 1003–1020. https://doi.org/10.1007/s11116-012-9447-0

Hensher, D. A., & Greene, W. H. (2010). Non-attendance and dual processing of common-metric attributes in choice analysis : a latent class specification. *Empirical Economics*, *39*, 413–426. https://doi.org/10.1007/s00181-009-0310-x

Hensher, D. A., & Rose, J. M. (2009). Simplifying choice through attribute preservation or non-attendance: Implications for willingness to pay. *Transportation Research Part E*, *45*(4), 583–590. https://doi.org/10.1016/j.tre.2008.12.001

Hensher, D. A., & Rose, J. M. (2012). *The Influence of Alternative Acceptability , Attribute Thresholds and Choice Response Certainty on Automobile Purchase Preferences*. *46*(September), 451–468.

Hensher, D. A., Rose, J. M., & Bertoia, T. (2007). The implications on willingness to pay of a stochastic treatment of attribute processing in stated choice studies. *Transportation Research Part E*, *43*, 73–89. https://doi.org/10.1016/j.tre.2005.07.006

Hensher, D. A., Rose, J. M., & Greene, W. H. (2005). The implications on willingness to pay of respondents ignoring specific attributes. *Transportation*, *32*(3), 203–222. https://doi.org/10.1007/s11116-004-7613-8

Hensher, D. A., Rose, J. M., & Greene, W. H. (2012). Inferring attribute non-attendance from stated choice data: implications for willingness to pay estimates and a warning for stated choice experiment design. *Transportation*, *39*, 235–245. https://doi.org/10.1007/s11116-011-9347-8

Hess, S., & Beharry-Borg, N. (2012). Accounting for Latent Attitudes in Willingness-to-Pay Studies: The Case of Coastal Water Quality Improvements in Tobago. *Environmental and Resource Economics*, *52*(1), 109–131. https://doi.org/10.1007/s10640-011-9522-6

Hess, S., & Hensher, D. A. (2010). Using conditioning on observed choices to retrieve individual-specific attribute processing strategies. *Transportation Research Part B: Methodological*, *44*(6), 781–790. https://doi.org/10.1016/j.trb.2009.12.001

Hess, S., & Hensher, D. A. (2013). Making use of respondent reported processing information to understand attribute importance: A latent variable scaling approach. *Transportation*, *40*(2), 397–412. https://doi.org/10.1007/s11116-012-9420-y

Hess, S., Stathopoulos, A., Campbell, D., O'Neill, V., & Caussade, S. (2013). It's not that I don't care, I just don't care very much: Confounding between attribute non-attendance and taste heterogeneity. *Transportation*, *40*(3), 583–607. https://doi.org/10.1007/s11116-012-9438-1

Hoehn, J. P., Lupi, F., & Kaplowitz, M. D. (2010). Stated choice experiments with complex ecosystem changes: The effect of information formats on estimated variances and choice parameters. *Journal of Agricultural and Resource Economics*, *35*(3), 568–590. https://doi.org/10.2307/23243072

Hole, A. R. (2011). A discrete choice model with endogenous attribute attendance. *Economics Letters*, *110*(3), 203–205. https://doi.org/10.1016/j.econlet.2010.11.033

Hole, A. R., Kolstad, J. R., & Gyrd-Hansen, D. (2013). Inferred vs. stated attribute non-attendance in choice experiments: A study of doctors' prescription behaviour. *Journal of Economic Behavior and Organization*, *96*, 21–31. https://doi.org/10.1016/j.jebo.2013.09.009

Hole, A. R., Norman, R., & Viney, R. (2016). Response patterns in health state valuation using endogenous attribute attendance and latent class analysis. *Health Economics*, *25*, 212–224. https://doi.org/10.1002/hec

Jourdain, D., & Vivithkeyoonvong, S. (2017). Valuation of ecosystem services provided by irrigated rice agriculture in Thailand: a choice experiment considering attribute nonattendance. *Agricultural Economics*, *48*(5), 655–667. https://doi.org/10.1111/agec.12364

Kehlbacher, A., Balcombe, K., & Bennett, R. (2013). Stated Attribute Non-attendance in Successive Choice Experiments. *Journal of Agricultural Economics*, *64*(3), 693–706. https://doi.org/10.1111/1477-9552.12021

Kjær, T., Bech, M., Gyrd-Hansen, D., & Hart-Hansen, K. (2006). Ordering effect and price sensitivity in discrete choice experiments: Need we worry? *Health Economics*, *15*(11), 1217–1228. https://doi.org/10.1002/hec.1117

Kosenius, A. K. (2013). Preference discontinuity in choice experiment: Determinants and implications. *Journal of Socio-Economics*, *45*, 138–145. https://doi.org/10.1016/j.socec.2013.05.004

Kragt, M. E. (2013). Stated and inferred attribute attendance models: A comparison with environmental choice experiments. *Journal of Agricultural Economics*, *64*(3), 719–736. https://doi.org/10.1111/1477-9552.12032

Lagarde, M. (2013). Investigating attribute non-attendance and its consequences in choice experiments

with latent class models. *Health Economics*, *22*, 554–567. https://doi.org/10.1002/hec

Lancsar, E., & Louviere, J. (2006). Deleting "irrational" responses from discrete choice experiments: A case of investigating or imposing preferences? *Health Economics*, *15*(8), 797–811. https://doi.org/10.1002/hec.1104

Leong, W., & Hensher, D. A. (2012). *Transport Reviews : A Transnational Embedding Decision Heuristics in Discrete Choice Models : A Review Embedding Decision Heuristics in Discrete Choice Models : A Review*. *32*(August), 37–41.

Meyerhoff, J., & Liebe, U. (2009). Status quo effect in choice experiments: Empirical evidence on attitudes and choice task complexity. *Land Economics*, *85*(3), 515–528. https://doi.org/10.3368/le.85.3.515

Ortega, D. L., & Ward, P. S. (2016). Information processing strategies and framing effects in developing country choice experiments: results from rice farmers in India. *Agricultural Economics (United Kingdom)*, *47*(5), 493–504. https://doi.org/10.1111/agec.12249

Puckett, S. M., & Hensher, D. A. (2008). The role of attribute processing strategies in estimating the preferences of road freight stakeholders. *Transportation Research Part E: Logistics and Transportation Review*, *44*(3), 379–395. https://doi.org/10.1016/j.tre.2007.01.002

Puckett, S. M., & Hensher, D. A. (2009). Revealing the extent of process heterogeneity in choice analysis: An empirical assessment. *Transportation Research Part A: Policy and Practice*, *43*(2), 117–126. https://doi.org/10.1016/j.tra.2008.07.003

Rose, J. M., Hensher, D. A., Greene, W. H., & Washington, S. P. (2012). Attribute exclusion strategies in airline choice: accounting for exogenous information on decision maker processing strategies in models of discrete choice. *Transportmetrica*, *8*(5), 344–360. https://doi.org/10.1080/18128602.2010.506897

Sandorf, E., Campbell, D., & Hanley, N. (2017). Disentangling the influence of knowledge on attribute non-attendance. *Journal of Choice Modelling*, *24*, 36–50. https://doi.org/10.1016/j.jocm.2016.09.003

Scarpa, R., Gilbride, T. J., Campbell, D., & Hensher, D. A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics*, *36*(2), 151–174. https://doi.org/10.1093/erae/jbp012

Scarpa, R., Thiene, M., & Hensher, D. A. (2010). Monitoring Choice Task Attribute Attendance in Nonmarket Valuation of Multiple Park Management Services : Does It Matter ? *Land Economics*, *86*(4), 817–839.

Scarpa, R., Zanoli, R., Bruschi, V., & Naspetti, S. (2012). Inferred and stated attribute non-attendance in food choice experiments. *American Journal of Agricultural Economics*, *95*(1), 165–180. https://doi.org/10.1093/ajae/aas073

Scarpa, R., Zanoli, R., Bruschi, V., & Naspetti, S. (2013). Inferred and stated attribute non-attendance in food choice experiments. *American Journal of Agricultural Economics*, *95*(1), 165–180. https://doi.org/10.1093/ajae/aas073

Scott, A. (2002). Identifying and analysing dominant preferences in discrete choice experiments: An application in health care. *Journal of Economic Psychology*, *23*(3), 383–398. https://doi.org/10.1016/s0167-4870(02)00082-x

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Thiene, M., Scarpa, R., Galletto, L., & Boatto, V. (2013). Sparkling wine choice from supermarket shelves: the impact of certification of origin and production practices. *Agricultural Economics*, *44*, 523–536. https://doi.org/10.1111/agec.12036

Thiene, M., Scarpa, R., & Louviere, J. J. (2015). Addressing Preference Heterogeneity , Multiple Scales and Attribute Attendance with a Correlated Finite Mixing Model of Tap Water Choice. *Environmental and Resource Economics*, *62*, 637–656. https://doi.org/10.1007/s10640-014-9838-0

Vanniyasingam, T., Daly, C., Jin, X., Zhang, Y., Foster, G., Cunningham, C., & Thabane, L. (2018). Investigating the impact of design characteristics on statistical efficiency within discrete choice experiments: A systematic survey. *Contemporary Clinical Trials Communications*, *10*, 17–28. https://doi.org/10.1016/j.conctc.2018.01.002

Weller, P., Oehlmann, M., Mariel, P., & Meyerhoff, J. (2014). Stated and inferred attribute non-attendance in a design of designs approach. *Journal of Choice Modelling*, *11*, 43–56. https://doi.org/10.1016/j.jocm.2014.04.002

Chapter 4

Attribute non-attendance in wine choice: contrasts between stated and inferred approaches

Submitted for publication

Presented in the

13th Annual Conference of the American Association of Wine Economists,
Vienna, Austria (July, 2019)

Abstract

A growing number of studies employing discrete choice experiments provides evidence that respondents ignore some attributes while making a choice, thus violating preference structure assumptions. Using a discrete choice experiment, we explore consumers preferences for wine. A greater understanding of the choice process accounting for attribute non-attendance may contribute to the definition of more accurate policy recommendations and market strategies. The result is checked in three different databases referring to three countries. This paper adds to the literature by contrasting the stated and inferred modelling approaches accounting for attribute non-attendance. The main contributions of the paper are (1) to compare the respondents stated non-attendance with probabilities inferred by the model, and explores the concordance between both methods; (2) examine the differences in model fit measures and willingness to pay values, adding as reference a standard model that assumes all attributes are attended. Results provide evidence of the presence of attribute non-attendance in all markets, the inferred approach reaching the best model fit. Few respondents behave in a fully compensatory manner, but a high proportion considers the price attribute, which is an important tip for marketers and wine firms. In addition, it is a key factor for the robust estimation of willingness to pay by attribute. Moreover, respondents overstate their levels of attendance and we found little concordance between both methods, in all datasets. The willingness to pay measures are lower in the inferred approach, while the estimates from the mixed logit model are found to be more congruent with those from the standard model.

Keywords: attribute non-attendance, stated preference, discrete choice modelling, wine choice

JEL classification: C25, D12

1. Introduction

Discrete choice experiments (DCE) is a quantitative approach widely used to elicit consumer preferences. This technique simulates a purchase situation by presenting a sequence of choice sets with two or more alternatives, differing in terms of attributes and levels. Respondents are asked to select their preferred alternative, with the possibility of choosing to "not buying" or to choose the current situation (status quo). This framework relies on the continuity axiom, which is a standard assumption of neoclassical consumer theory, and considers that individuals make trade-offs among all the product´s attributes on the basis of unlimited substitutability among them. Accordingly, under the principle of utility maximization, respondents will make their best choice exhibiting compensatory preference structures (Hensher, Rose, and Greene 2005).

However, some individuals only attend a subset of attributes, ignoring others, while making a choice. This situation characterizes what is commonly referred to as attribute non-attendance (ANA), as a result of using simplifying decision rules, unfamiliarity with the good, or simply because an attribute or its levels is not relevant in the choice process (Alemu et al. 2013). Hence, the assumptions stated above are commonly violated in light of a non-compensatory behavior. In recent years a growing body of literature from transportation, health and environmental economics (e.g. Campbell, Hensher, & Scarpa, 2011; Hensher, Collins, & Greene, 2013; Scarpa, Gilbride, Campbell, & Hensher, 2009) seeks to understand the implications of assuming a model specification that assumes all attributes as relevant, when that might not be the case and also the role that ANA might play in improving the predictive capability of discrete choice models. Although recently ANA was studied in the context of food choice (Caputo, Van Loo, Scarpa, Nayga, & Verbeke, 2018; Chalak, Abiad, & Balcombe, 2016; Scarpa et al., 2013), when applied to wine, several applications have ignored this possibility. These studies show the importance of extrinsic cues in wine choice, such as wine origin (Mueller and Szolnoki 2010; Platania, Platania, and Santisi 2016), grape variety (Kallas, Escobar, and Gil 2013; Lockshin and Corsi 2012), awards or medals (Combris et al. 2009; Corsi, Mueller, and Lockshin 2012), and price (Mueller et al. 2010; Palma et al. 2016). Still, the number of studies accounting for ANA remains very limited (see Thiene et al., 2013 for example).

ANA may be captured in two different ways, by using the stated and/or the inferred techniques. Stated ANA relies on self-assessment, where respondents are directly asked whether or not they ignored one or more attributes (e.g. Carlsson, Kataria, & Lampi, 2010; Hensher et al., 2005; Rose, Hensher, Greene, & Washington, 2012). The main disadvantage of this approach is being susceptible to measurement errors.

The stated approach can be implemented by questioning respondents at the end of the choice sets (serial level), or asking them after each choice set (choice task level). Alternatively, it is possible to infer non-attendance through analytical modeling techniques, without additional data (e.g. Greene & Hensher, 2003; Hess, Stathopoulos, Campbell, O'Neill, & Caussade, 2013; Scarpa, Zealand, Gilbride, Campbell, & Hensher, 2009). This approach requires the definition, by the researcher, of how to incorporate ANA. The inferred and stated approaches have been compared in the literature and divergences have been found raising the question on which method should be applied in a particular situation (e.g. Kragt, 2013; Scarpa et al., 2013; Weller, Oehlmann, Mariel, & Meyerhoff, 2014). Findings also suggest that respondents may not truthfully reveal their behavior, making wrong statements about attribute attendance.

Given the evidence, the issue of ANA may be critical in understanding agents´ behavior and getting accurate estimates of utility parameters, hence preventing related biases. It is hypothesized that due to its complexity, the choice of wine may be a result of simplifying strategies, such that consumers make a purchase decision based on a subset of attributes. Hence, this paper investigates consumer preferences for wine, exploring the violation of continuity axiom through ANA. We infer ANA using a latent class framework and contrast it with the respondent's self-statements about non-attendance, through a mixed logit model. In particular, we focus on the differences in the implied rates of non-attendance across the two approaches to deal with ANA. Finally, we examine the effects on willingness to pay values and model fit across the different methods applied, and also when compared to a model where all attributes are assumed to be relevant.

This article contributes to the growing literature studying the robustness of conventional methods to elicit willingness to pay estimates to ANA. Also, it adds to that literature by analyzing the concordance between stated and inferred ANA, given that only few studies have addressed that issue. Finally, it contributes to the study of heuristics in the food/agricultural literature, which is a limited literature and where ANA has been generally ignored despite its importance and potential relevance for the decision and policy making.

The paper is structured as follows. Section 2 describes the survey design and the data. Section 3 introduces the modeling techniques, and explains the mixed logit and latent class models. Section 4 presents the results. Section 5 concludes the paper.

2. Survey, sample, and data collection

We constructed a discrete choice experiment study on wine choice in two European markets (France and Portugal) and in North America (comprising USA and Canada together). Our motivation for not using just a single case is to produce more reliable conclusions, not dependent on one particular group of consumers. Following a D-efficient design using the Ngene v1.1.2 software (ChoiceMetrics, 2014), all respondents in each market faced ten different choice tasks, involving the choice between three unlabeled alternatives (wines), plus a non-purchase alternative if they rather wished to opt out (see Figure 1 for an example of a choice set). A total of 4000 observations in North America and 2000 in each European market was obtained. The three alternatives were described by six attributes, namely price, country of origin, landscape, alcohol level, grape variety and medals awarded. Linear effects were specified for the continuous attributes, while dummy terms were included for the country of origin and grape variety. These attributes, as well as their respective levels (see Table 1), were determined through a literature review, focus group, advice from wine market experts, and through a pilot study.

Table 1 - Attributes and levels

Attributes	Levels
Country of origin	Portugal (baseline), France, Italy, USA, Australia, Chile
Landscape	0=not certified (baseline)
	1=certified
Medals	0=none (baseline)
	1=gold medal
Alcohol level	12, 13.5, 15
Grape variety	Cabernet Sauvignon, Syrah, Red blend, none (baseline)
Price (in $)	Europe: 4.50, 8.50, 14.00, 20.50
	North America: 8.99, 12.99, 17.99, 24.99

To give more realism to the experiment, five constant attributes always appear in all bottles of wine: a fictional brand, an arbitrary vintage, volume, bottle style, and appellation of origin. We used a Bordeaux bottle style with the standard size (750ml). All wines have an appellation classification according to the country of origin system and equivalent to the "Protected Designation of Origin"- PDO. We controlled for situation by stating for each choice task, "Imagine you are standing in front of a shelf of wine. You have

3 different bottles of PDO wine available. Which of the following wines would you buy to enjoy over a dinner with guests at home?". So, we delimited our wine selection by focusing on a wine purchased for a dinner with guests. A market research company collected the data through an online survey in August 2018.



Figure 1 - Choice set example

3. Modelling techniques to account for attribute non attendance

A sampled respondent *n* (n=1, ..., N) makes repeated choices among different *J* alternatives in *T* choice situations. Each alternative *j* is represented by a random utility expression given by the following equation:

$$U_{njt} = \beta' x_{njt} + \varepsilon_{njt} \quad (1)$$

where x_{njt} is the vector of explanatory variables, including product attributes and respondents' characteristics, and ε_{njt} is the random component (or stochastic influence). Individual *n* chooses the alternative that gives the highest utility, such that $P_{nj} = prob(\beta' x_{nj} + \varepsilon_{nj} > \beta' x_{nk} + \varepsilon_{nk}) \; \forall \, j \neq k \in C$, where C is the choice set of J alternatives, j=1,...,J.

3.1. Mixed logit model using stated non-attendance

The Mixed (also known as a Random Parameter) Logit model generalizes the standard logit by allowing the coefficients on the attributes to vary randomly across the sampled population, accounting for unobserved individual preference heterogeneity (Revelt & Train, 1998). The Mixed Logit (MIXL) is the most general specification of discrete choice models (Greene and Hensher 2003). This model introduces individual heterogeneity through the vector of coefficient β_n, such that

$$\beta_n = \beta + \Delta z_n + u_n \tag{2}$$

where $\beta + \Delta z_n$ denote heterogeneity in the mean of the distribution of the random parameters and u_n is a vector of random components, which captures non-observable effects. The choice probabilities of the model are:

$$Prob(choice_n = j|'x_{njt}, u_n) = \frac{\exp\,(\beta'_n x_{njt})}{\sum_{j=1}^{Jn} exp\,(\beta'_n x_{njt})} \tag{3}$$

Regarding ANA, the value of β_n can be conditioned depending on whether a respondent ignored a specific attribute or not. ANA is accounted for by restricting the attribute coefficients to zero when a respondent did not attend to the attribute, according to his statement in the supplementary questions.

3.2. Latent Class model to infer non-attendance

Alternatively, we can infer non-attendance based on observed choice behavior, without requiring additional information on ANA. In general, the Latent Class model captures preference heterogeneity through membership of distinct classes of utility description (Hensher, Collins, and Greene 2013). The probability that a sampled respondent selects the alternative *j*, in T choice situations, given membership of class $c = 1, \dots, C$, is:

$$P_{njt|c} = \frac{\exp\left(\beta'_c x_{njt}\right)}{\sum_{j=1}^{Jn} \exp\left(\beta'_c x_{njt}\right)} \tag{4}$$

The probability of belonging to a specific class *c*, is given in the following equation (Greene and Hensher 2003):

$$H_{nc} = \frac{\exp\left(\theta_c z_n\right)}{\sum_{c \in C} \exp\left(\theta_c z_n\right)} \quad (5)$$

where θ_c are class-specific parameters, with C^{th}vector normalized to ensure model identification, and z_n is observed data. In the case of ANA, given that the individual is assumed to choose a subset from the K attributes, each class represents different ANA strategies employed by the respondent. The researcher determines the restrictions on the utility coefficients, by specifying the classes where some attributes have their parameters restricted to zero. The parameters different from zero are generally constrained to the same value across classes, and thus imposing homogeneity of taste intensity. One of the approaches to analyze the forms of ANA is through the inference of membership class probabilities from the observed choices. A higher predicted probability of belonging to a given class will be attached to respondents that behave according to the specified pattern or strategy of non-attendance in that class.

4. Results

4.1. Stated ANA

An analysis of self-attribute attendance at serial level is presented in Table 2. All attributes were considered by almost 25% of respondents in the North American and French samples, while only 15.5% stated full attended in the Portuguese sample. This is a small proportion when compared to the results obtained in other food studies, such as Scarpa et al. (2013) and/or environmental studies such as Kragt (2013). Just over 30% stated they had ignored only one attribute, and less that 4% of all respondents based their choice solely on one attribute exhibiting lexicographic preferences. The alcohol level was the most ignored attribute in the North American (31.5%) and French (29.5%) samples, while landscape was in the Portuguese sample (34.5%). Grape variety was the least ignored attribute by North American and French consumers (22.4% and 19.5%, respectively), while the price was the least ignored by the Portuguese (20.5%). Nevertheless, there are no trade-offs for almost 23% of respondents, both in French and North American samples.[5]

[5] Results from summing up all stated probabilities where each attribute was ignored (see Appendix D).

Table 2 - An overview of self-reported attribute non-attendance

ANA pattern	North America		France		Portugal	
	Number of respondents	Proportion of total	Number of respondents	Proportion of total	Number of respondents	Proportion of total
Full attendance	102	0.255	49	0.245	31	0.155
Ignoring 1 attribute	121	0.303	74	0.37	71	0.355
Country of origin	16	0.040	14	0.07	15	0.075
Medals	21	0.053	4	0.02	8	0.04
Alcohol level	23	0.058	24	0.12	13	0.065
Landscape	27	0.068	12	0.06	15	0.075
Grape variety	9	0.023	4	0.02	12	0.06
Price	25	0.063	16	0.08	8	0.04
Ignoring 2 attributes	81	0.203	44	0.22	54	0.27
Ignoring 3	56	0.140	22	0.11	31	0.155
Ignoring 4	25	0.063	7	0.035	10	0.05
Lexicographic choices	13	0.033	4	0.02	3	0.015
Ignoring all attributes	2	0.005	0	0	0	0
TOTAL	**400**	**1**	**200**	**1**	**200**	**1**

Tables 3 and 4 report the different model results, estimated with Nlogit 6.0. The models presented in Table 3 assume full attribute attendance, i.e., passive bounded rationality, while models in Table 4 account for attribute non-attendance. For those who ignore an attribute, the individual parameter estimate is zero and then not inclusive of the entire sample. The Alternative Specific Constant was included for the "none" option, and has the expected sign, implying that respondents prefer the three alternatives over the none option. For the MIXL models ignoring ANA, the landscape and the Chilean origin parameters

are not significant in the North American sample, such as the Cabernet Sauvignon is not in the French and Syrah in the Portuguese. It would result in wrongly assuming that these attributes are not important for the entire population when, in fact, there is a proportion of the sample for which the attributes are significant and important in the wine choice process. Additionally, the results suggest that the exclusion of specific attributes has impacted in a significant way on the discriminating contribution of the landscape, wine origin and also grape varieties relative to the respective baseline.

Table 3 - MNL and MIXL model results with all attributes assumed to be attended

Variables	North America		France		Portugal	
	MNL	MIXL	MNL	MIXL	MNL	MIXL
		Mean		Mean		Mean
Medals	0.426***	0.318***	0.467***	0.714***	0.564***	1.119***
	(0.054)	(0.087) rp	(0.079)	(0.106) rp	(0.091)	(0.126)
Landscape	0.042	0.075	0.044	0.224***	0.347***	0.770***
	(0.042)	(0.047) rp	(0.065)	(0.077) rp	(0.067)	(0.088)
Alcohol	-0.023	0.192***	0.016	0.272***	0.058**	0.330***
	(0.015)	(0.011) rp	(0.024)	(0.017) rp	(0.028)	(0.024)
Country of origin (Portugal)						
France	0.145**	-0.242***	0.753***	0.685***	-1.126***	-2.133***
	(0.069)	(0.083) rp	(0.096)	(0.136) rp	(0.097)	(0.150)
Italy	0.060	-0.631***	-0.407***	-1.028***	-1.546***	-2.901***
	(0.087)	(0.132) rp	(0.128)	(0.185) rp	(0.127)	(0.196)
USA	0.112	-0.182**	-0.403***	-0.743***	-2.236***	-4.101***
	(0.070)	(0.080) rp	(0.107)	(0.134) rp	(0.121)	(0.275)
Australia	-0.267***	-0.232**	-0.809***	-0.568***	-2.461***	-3.323***
	(0.096)	(0.117) rp	(0.155)	(0.192) rp	(0.214)	(0.382)
Chile	-0.063	-0.127	-0.092	-0.383**	-1.930***	-4.053***
	(0.087)	(0.110) rp	(0.147)	(0.192) rp	(0.157)	(0.331)
Grape variety (None)						
Cab. Sauv.	0.358***	0.270***	0.294***	0.085	0.108	-0.116
	(0.055)	(0.061) rp	(0.091)	(0.102) rp	(0.084)	(0.103)
Syrah	0.076	-0.382***	0.240**	0.277*	-0.285**	-0.611
	(0.077)	(0.111) rp	(0.114)	(0.146) rp	(0.116)	(0.150)
Red blend	0.340***	0.709***	0.634***	0.660***	0.415***	0.659***
	(0.103)	(0.168) rp	(0.154)	(0.216) rp	(0.157)	(0.239)
Price	-0.037***	-0.103***	-0.047***	-0.157***	-0.094***	-0.266***
	(0.004)	(0.007) rp	(0.006)	(0.013) rp	(0.007)	(0.018)
ASC	-1.423***	-0.166	-0.063	1.734***	-1.408***	-0.762***
	(0.219)	(0.186)	(0.325)	(0.236)	(0.308)	(0.280)
Log Likelihood	-5106.3	-4813.6	-2486.1	-2200.6	-2273.4	-1910.2
AIC	10238.5	9653.2	4998.1.4	4427.1	4572.4	3846.5
BIC	10320.3	9735.0	5071.0	4499.9	4645.5	3919.3

Notes: Elasticities are uninformative given that this experiment has unlabeled alternatives. BIC – Bayesian Information Criterion. AIC – Akaike Information Criterion. For the MIXL we used constrained triangular distributions and 500 Halton draws (e.g. Greene and Hensher, 2013). rp denotes random parameter. Standard errors in parenthesis; ***, **, * significance at 1%, 5%, 10% level respectively.

Table 4 - MNL and MIXL model results with attribute non-attendance

Variables	North America		France		Portugal	
	MNL	MIXL	MNL	MIXL	MNL	MIXL
		Mean		Mean		Mean
Medals	0.570***	0.914***	0.425***	0.805***	0.514***	0.770***
	(0.050)	(0.067) ^{rp}	(0.071)	(0.097) ^{rp}	(0.071)	(0.090) ^{rp}
Landscape	0.097**	0.114**	-0.046	-0.008	0.128*	0.208***
	(0.044)	(0.048) ^{rp}	(0.065)	(0.072) ^{rp}	(0.066)	(0.078) ^{rp}
Alcohol	-0.020**	-0.096***	-0.034***	-0.164***	-0.006	0.112***
	(0.008)	(0.010) ^{rp}	(0.009)	(0.012) ^{rp}	(0.009)	(0.016) ^{rp}
Country of origin (Portugal)						
France	-0.016	-0.178**	0.782***	0.864***	-0.549***	-0.956***
	(0.068)	(0.076) ^{rp}	(0.092)	(0.125) ^{rp}	(0.094)	(0.116) ^{rp}
Italy	0.052	-0.142	-0.364***	-0.547***	-0.713***	-1.105***
	(0.081)	(0.095) ^{rp}	(0.118)	(0.156) ^{rp}	(0.109)	(0.133) ^{rp}
USA	0.077	-0.163**	-0.417***	-0.447***	-1.747***	-2.493***
	(0.073)	(0.083) ^{rp}	(0.108)	(0.133) ^{rp}	(0.124)	(0.210) ^{rp}
Australia	-0.203**	-0.307***	-0.851***	-1.037***	-1.988***	-2.290***
	(0.096)	(0.110) ^{rp}	(0.155)	(0.191) ^{rp}	(0.202)	(0.288) ^{rp}
Chile	-0.302***	-0.455***	-0.069	0.069	-1.402***	-2.107***
	(0.089)	(0.098) ^{rp}	(0.142)	(0.157) ^{rp}	(0.148)	(0.233) ^{rp}
Grape variety (None)						
Cab. Sauv.	0.341***	0.492***	0.112	0.356***	0.037	-0.078
	(0.057)	(0.068) ^{rp}	(0.089)	(0.106) ^{rp}	(0.086)	(0.094) ^{rp}
Syrah	-0.059	-0.019	0.160	0.379***	-0.253***	-0.407***
	(0.070)	(0.083) ^{rp}	(0.098)	(0.119) ^{rp}	(0.097)	(0.112) ^{rp}
Red blend	0.136	0.147	0.314**	0.552***	0.077	-0.229
	(0.091)	(0.112) ^{rp}	(0.131)	(0.160) ^{rp}	(0.128)	(0.149) ^{rp}
Price	-0.048***	-0.132***	-0.043***	-0.138***	-0.067***	-0.148***
	(0.004)	(0.009) ^{rp}	(0.006)	(0.014) ^{rp}	(0.006)	(0.012) ^{rp}
ASC	-1.528***	-3.133***	-0.756***	-2.239***	-1.580***	-1.499***
	(0.114)	(0.156)	(0.142)	(0.178)	(0.137)	(0.145)
Log Likelihood	-5076.8	-4855.5	-2531.9	-2361.6	-2427.0	-2325.5
AIC	10179.7	9736.9	5089.8	4749.3	4880.0	4677.0
BIC	10261.5	9818.8	5162.6	4822.1	4952.8.4	4749.8

Notes: Elasticities are uninformative given that this experiment has unlabeled alternatives. BIC – Bayesian Information Criterion. AIC – Akaike Information Criterion. For the MIXL we used constrained triangular distributions and 500 Halton draws (e.g. Greene and Hensher, 2013). ^{rp} denotes random parameter. Standard errors in parenthesis; ***, **, * significance at 1%, 5%, 10% level respectively.

4.2. Inferred ANA and concordance with the stated approach

Existing literature (e.g. Hensher, Rose, & Greene, 2012) suggests that respondents' self-reported attribute attendance may not reveal their truthful behavior, thus we explore this inconsistency by also specifying latent class structures to endogenously estimate class membership (Table 5). ECLC significantly increases model fit, in terms of Log Likelihood and information criteria, compared to MNL or MIXL. This result is consistent across samples which suggests that our latent structure captures important observed choice variation even when compared to the MIXL.

Table 5 - ECLC model results with attribute non-attendance

Variables	North America	France	Portugal
Medals	0.616***	0.866***	0.671***
	(0.037)	(0.088)	(0.096)
Landscape	0.170*	0.479***	0.427***
	(0.091)	(0.138)	(0.100)
Alcohol	0.187***	0.281***	0.283***
	(0.028)	(0.021)	(0.021)
Country of origin (Portugal)			
France	0.337**	6.467***	-2.284***
	(0.157)	(0.919)	(0.193)
Italy	0.812***	-1.820***	-2.191***
	(0.234)	(0.191)	(0.167)
USA	0.112	-1.389***	-3.162***
	(0.243)	(0.180)	(0.248)
Australia	-0.362	-2.000***	-3.114***
	(0.354)	(0.307)	(0.313)
Chile	-1.477***	-1.554***	-2.461***
	(0.335)	(0.287)	(0.265)
Grape variety (None)			
Cab. Sauv.	1.974***	-0.397***	0.005
	(0.222)	(0.140)	(0.107)
Syrah	-0.129	-0.459**	-0.439**
	(0.154)	(0.197)	(0.197)
Red blend	0.046	0.314	0.764**
	(0.202)	(0.290)	(0.321)
Price	-0.281***	-0.417***	-0.355***
	(0.012)	(0.032)	(0.029)
ASC	-3.111***	-1.517***	-2.480***
	(0.235)	(0.212)	(0.249)
Log Likelihood	-4447.2	-2070.8	-1884.4
AIC	8934.4	4179.5	-3806.8
BIC	9060.3	4285.9	3913.2

Standard errors in parenthesis; ***, **, * significance at 1%, 5%, 10% level respectively.

We then analyze the concordance between both methods, comparing the stated and inferred ANA frequencies for each attribute and the different attribute combinations. For the inferred approach[6], several specifications were tested for each market, and we therefore chose to focus on the specification that gave better results in terms of AIC and BIC criteria to represent consumers behavior. Table 6 shows the non-attendance patterns as inferred by the ECLC models for the North American and European data. The pattern of results is different between the datasets. For the French and North American data, we find that the class involving a lexicographic behavior on price has the highest probability of membership, with 32.5% and 48.9% respectively. For the Portuguese, the class "ignore landscape and alcohol" reached the highest probability (22.7%), followed by the full attendance class, with 18.6%, which contrasts with the other datasets.

The model predicts that 6.9% of respondents in the French market ignore only the landscape, which matches well with what they stated (6%). The other reasonable match is the pattern of attendance to all attributes, in Portugal, with 18.6% making trade-offs between all attributes. In the other markets, the predicted proportion of full attenders is much lower than that revealed by the stated approach. Class eight presents a particular concern, as 8% of respondents in the North American market appear to have made their choices randomly (this behavior was also reported in other studies, such as Kragt, 2013). However, more than 85% of respondents attended to price in all samples (similar rates were found by Rose et al. 2012) which closely aligns with the stated rates. The same occurs for medals in the North American sample, being stated ignored by 29% of respondents and predicted by 33%. In the other cases, we found little concordance between the inferred and stated non-attendance summed probabilities, with the inferred non-attendance being more pronounced in the North American sample than in the European samples. This inconsistency between the two approaches has also been reported by other studies, such as Caputo et al. (2018), Kragt (2013), and Ortega and Ward (2016).

[6] We tested different combinations of non-attendance being equal in each market, such as complete attendance, random choice, ANA for single attribute and lexicographic preferences (only one of k attributes matters, and all k−1 coefficients are set to zero), but balancing improvements in model fit, we found the presented patterns best for each dataset. When dealing with ANA for COO and grape variety, we constrained all attribute levels to zero at the same time.

Table 6 - Attribute non-attendance patterns based on ECLC model results and contrasts with stated ANA

	Class no.			Inferred Probabilities from ECLC models			Stated ANA		
Class description	North America	France	Portugal	North America	France	Portugal	North America	France	Portugal
-Full attendance	1	1	1	0.028*	0.038***	0.186***	0.255	0.245	0.155
-Ignore only landscape	-	2	2	-	0.069**	0.152***	-	0.060	0.075
-Ignore only alcohol	-	3	3	-	0.260***	0.125***	-	0.120	0.065
-Ignore only COO	2	-	-	0.164***	-	-	0.040	-	-
-Ignore only grape variety	3	-	4	0.060***	0.159***	0.227***	0.023	0.035	0.060
-Ignore landscape and alcohol	-	-	-	-	0.043**	-	-	-	-
-Ignore landscape, alcohol, and grape variety	4	6	7	0.097***	0.107***	0.089***	0.013	0.015	0.030
-Ignore landscape, alcohol, and COO	5	-	-	0.043**	-	-	0.003	-	-
-Ignore awards, landscape, and alcohol		7			0.125***			0.015	
-Ignore landscape, alcohol, grape variety, and price	6	-	-	0.039**	-	-	0.003	-	-
-Only consider COO	-	4	5	-	0.076***	0.123***	-	0.005	0.005
-Only consider price	7	5	6	0.489***	0.325***	0.098***	0.023	0.010	0.010
-Ignore all (random choice)	8	-	-	0.080***	-	-	0.005	-	-

***, **, * significance at 1%, 5%, 10% level respectively

4.3. Impact on welfare estimates

Table 7 presents the WTP estimates for the MNL model assuming full attendance, and both the MIXL and the ECLC models accounting for ANA. Given that the overall fit is better for the latent class specifications, the WTP estimates from MNL and MIXL models might be queried. Overall, results suggest an upward bias for these welfare measures, in all samples. Besides the sign and significance differences between the models, different ranking orders of the WTP were also found. While the MNL and MIXL models show that consumers are willing to pay more for an awarded wine in both North American and Portuguese samples, the ECLC model reveal different preferences. In fact, North American consumers prefer Cabernet Sauvignon wines, followed by Italian wines and those with a gold medal. On the other hand, in general, the European consumers are more willing to pay for an awarded wine and produced in a certified landscape. Additionally, French consumers are willing to pay a high price premium of 15.5$ for French wines, and the Portuguese 2.15$ for Red blend wines.

Table 7 - WTP estimates, in US$7

	North America			France			Portugal		
	MNL	**MIXL**	**ECLC**	**MNL**	**MIXL**	**ECLC**	**MNL**	**MIXL**	**ECLC**
Medals	11.47***	9.91 (20.56)	2.19***	9.99***	8.24 (15.07)	2.07***	5.99***	7.21 (11.72)	1.89***
Landscape	1.13	1.23 (2.03)	0.60*	0.94	-0.08 (0.19)	1.15***	3.69***	1.96 (2.50)	1.20***
Alcohol	-0.62	-0.97 (1.53)	0.67***	0.33	-1.59 (1.92)	0.67***	0.61**	1.08 (1.54)	0.80***
COO									
France	3.90*	-1.80 (2.05)	1.20**	16.11***	-8.74 (12.42)	15.50***	-11.98***	-8.59 (12.27)	-6.44***
Italy	1.62	-1.53 (2.61)	2.88***	-8.71***	-5.79 (12.94)	-4.36***	-16.44***	-10.11 (11.11)	-6.18***
USA	3.01	-1.69 (2.25)	0.40	-8.63***	-4.35 (4.71)	-3.33***	-23.77***	-23.30 (27.52)	-8.92***
Australia	-7.18***	-3.33 (7.44)	-1.29	-17.30***	-9.80 (11.07)	-4.80***	-26.17***	-20.20 (19.03)	-8.78***
Chile	-1.70	-4.47 (8.01)	-5.25***	-1.96	0.67 (0.74)	-3.73***	-20.52***	-20.06 (21.52)	-6.94***
Grape variety									
Cab. Sauv.	9.65***	5.09 (6.05)	7.02***	6.29***	3.68 (6.18)	-0.95***	1.15	-0.79 (2.73)	0.01
Syrah	2.05	-0.20 (0.26)	-0.46	5.14**	3.77 (5.02)	-1.10**	-3.03**	-3.95 (16.78)	-1.24**
Red blend	9.17***	1.52 (1.93)	0.16	13.56***	5.26 (7.12)	0.75	4.41***	-2.13 (2.71)	2.15**

Notes: SD for random parameter specifications within parenthesis; ***, **, * significance at 1%, 5%, 10% level respectively

[7] WTP values for MNL and ECLC models were estimated as $WTP = -\frac{\beta_k}{\beta_{price}}$, while for the MIXL we calculated WTP based on unconditional estimates.

5. Conclusion

This paper analysis consumer behavior in light of the violation of the continuity axiom of consumer theory, and assuming bounded rationality via ANA. We explore the differences between what respondents' state to do and what we statistically infer they do regarding the profile of available information when making a choice. Our particular interest to capture ANA derives from the fact that failure to account for non-attendance may result in biased parameter estimates and poor model performance, and consequently incorrect conclusions and/or policy recommendations.

Previous studies show that a substantial proportion of respondents may ignore some attributes, our study provides evidence of this fact. We compared self-reported ANA at serial level to inferred ANA by estimating mixed logit models and equality constrained latent class models, using data collected from three different markets. In terms of model fit, the latent class specification outperforms the competing model based on continuous mixing distribution, suggesting that it may be a better model to accommodate ANA. The latent class model shows that only a relatively small proportion of respondents considered all attributes. On the other hand, in both stated and inferred approaches, and also in all datasets, we found that a large share of respondents attended to price, and they included this attribute in their choice decisions, which is crucial from an economic point of view for the computation of WTP. This result highlights the relevance of the price attribute in the wine consumer decision process, that might be a result of the complexity of choosing an experience and heterogeneous good as wine. It contrasts with high non-attendance levels at the price attribute observed in valuation of cultural or environmental amenities, for which the individuals reveal difficulty to attach a monetary value. Regarding all non-monetary attributes, we find discrepancies in the level of non-attendance provided by both approaches, with the results revealing that respondents think they have a higher attendance to some attributes than the predicted through model inference.

Concerning the implications for WTP measures, and despite the literature being inconclusive about the direction of bias that ANA may cause, our results show that accounting for non-attendance through inferred models produce much lower WTP values for wine attributes, as also different preference rankings. This finding reiterates previous results in the literature on agricultural and/or food choice (e.g. Kehlbacher, Balcombe, & Bennett, 2013; Scarpa et al., 2013). Additionally, the choice experiment helped to identify the most valued attributes by consumers in the three markets, taking into account the information/attribute processing heterogeneity which has implications for the decision making. Hence, our results contribute to develop more adequate strategies by policy makers and wineries. Based on WTP

results, we show that North American consumers have higher preferences for Cabernet Sauvignon wines, Italian and French wines, and awarded with a gold medal. The French and Portuguese consumers place more importance in awarded wines and are willing to pay for a wine produced in a certified landscape, which positions them in a consumer segment that appears to be more traditional, with potential to value the wine producing place when it is classified, for example, as world heritage. Our results also show the higher valuation that North American consumers place in grape variety, which is much higher when compared to the valuation by European consumers. These results highlight the importance of the information to include on the label and how, according to the target market.

Nevertheless, we are aware of the sensitivity of ECLC models to the assumptions about respondents who did not trade-off price, and to the fact that ECLC may overestimate the degree of ANA since it ignores taste heterogeneity (Hess et al. 2013). This is an interesting issue to further investigate non-attendance, adding the elicitation of the reasons why respondents apply such heuristic. This additional understanding of the reasons underlying ANA may contribute to improve the choice design and to define better specifications of choice models.

References

Alemu, M., Mørkbak, M., Olsen, S., & Jensen, C. (2013). Attending to the Reasons for Attribute Non-attendance in Choice Experiments. *Environmental Resource Economics*, *54*, 333–359. https://doi.org/10.1007/s10640-012-9597-8

Campbell, D., Hensher, D. A., & Scarpa, R. (2011). Non-attendance to attributes in environmental choice analysis: a latent class specification. *Journal of Environmental Planning and Management*, *54*(8), 1061–1076. https://doi.org/10.1080/09640568.2010.549367

Caputo, V., Van Loo, E. J., Scarpa, R., Nayga, R. M., & Verbeke, W. (2018). Comparing Serial, and Choice Task Stated and Inferred Attribute Non-Attendance Methods in Food Choice Experiments. *Journal of Agricultural Economics*, *69*(1), 35–57. https://doi.org/10.1111/1477-9552.12246

Carlsson, F., Kataria, M., & Lampi, E. (2010). Dealing with ignored attributes in choice experiments on valuation of Sweden's environmental quality objectives. *Environmental and Resource Economics*, *47*(1), 65–89. https://doi.org/10.1007/s10640-010-9365-6

Chalak, A., Abiad, M., & Balcombe, K. (2016). Joint use of attribute importance rankings and non-attendance data in choice experiments. *European Review of Agricultural Economics*, *43*(5), 1–24. https://doi.org/10.1093/erae/jbw004

Combris, P., Bazoche, P., Giraud-héraud, E., & Issanchou, S. (2009). Food choices: What do we learn from combining sensory and economic experiments? *Food Quality and Preference*, *20*(8), 550–557. https://doi.org/10.1016/j.foodqual.2009.05.003

Corsi, A. M., Mueller, S., & Lockshin, L. (2012). Let's See What They Have . . .: What Consumers Look For in a Restaurant Wine List Armando. *Cornell Hospitality Quarterly*, *XX*(X), 1–12. https://doi.org/10.1177/1938965511428448

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transportation Research Part B: Methodological*, *37*(8), 681–698. https://doi.org/10.1016/S0191-2615(02)00046-2

Greene, W. H., & Hensher, D. A. (2013). Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. *Applied Economics*, *45*, 1897–1902. https://doi.org/10.1080/00036846.2011.650325

Hensher, D. A., Collins, A. T., & Greene, W. H. (2013). Accounting for attribute non-attendance and common-metric aggregation in a probabilistic decision process mixed multinomial logit model: A warning on potential confounding. *Transportation*, *40*(5), 1003–1020. https://doi.org/10.1007/s11116-012-9447-0

Hensher, D. A., Rose, J. M., & Greene, W. H. (2005). The implications on willingness to pay of respondents ignoring specific attributes. *Transportation*, *32*(3), 203–222. https://doi.org/10.1007/s11116-004-7613-8

Hensher, D. A., Rose, J. M., & Greene, W. H. (2012). Inferring attribute non-attendance from stated choice data: implications for willingness to pay estimates and a warning for stated choice experiment design. *Transportation*, *39*, 235–245. https://doi.org/10.1007/s11116-011-9347-8

Hess, S., Stathopoulos, A., Campbell, D., O'Neill, V., & Caussade, S. (2013). It's not that I don't care, I just don't care very much: Confounding between attribute non-attendance and taste heterogeneity. *Transportation*, *40*(3), 583–607. https://doi.org/10.1007/s11116-012-9438-1

Kallas, Z., Escobar, C., & Gil, J. M. (2013). Analysis of consumers' preferences for a special-occasion red

wine: A dual response choice experiment approach. *Food Quality and Preference*, *30*(2), 156–168. https://doi.org/10.1016/j.foodqual.2013.05.008

Kehlbacher, A., Balcombe, K., & Bennett, R. (2013). Stated Attribute Non-attendance in Successive Choice Experiments. *Journal of Agricultural Economics*, *64*(3), 693–706. https://doi.org/10.1111/1477-9552.12021

Kragt, M. E. (2013). Stated and inferred attribute attendance models: A comparison with environmental choice experiments. *Journal of Agricultural Economics*, *64*(3), 719–736. https://doi.org/10.1111/1477-9552.12032

Lockshin, L., & Corsi, A. M. (2012). Consumer behaviour for wine 2.0: A review since 2003 and future directions. *Wine Economics and Policy*, *1*(1), 2–23. https://doi.org/10.1016/j.wep.2012.11.003

Mueller, S., Osidacz, P., Francis, I. L., & Lockshin, L. (2010). Combining discrete choice and informed sensory testing in a two-stage process: Can it predict wine market share? *Food Quality and Preference*, *21*(7), 741–754. https://doi.org/10.1016/j.foodqual.2010.06.008

Mueller, S., & Szolnoki, G. (2010). The relative influence of packaging, labelling, branding and sensory attributes on liking and purchase intent: Consumers differ in their responsiveness. *Food Quality and Preference*, *21*(7), 774–780. https://doi.org/10.1016/j.foodqual.2010.07.011

Ortega, D. L., & Ward, P. S. (2016). Information processing strategies and framing effects in developing country choice experiments: results from rice farmers in India. *Agricultural Economics (United Kingdom)*, *47*(5), 493–504. https://doi.org/10.1111/agec.12249

Palma, D., Ortúzar, J. D. D., Rizzi, L. I., Guevara, C. A., Casaubon, G., & Ma, H. (2016). Modelling choice when price is a cue for quality: a case study with Chinese consumers. *Journal of Choice Modelling*, *19*, 24–39. https://doi.org/10.1016/j.jocm.2016.06.002

Platania, M., Platania, S., & Santisi, G. (2016). Entertainment marketing, experiential consumption and consumer behavior: the determinant of choice of wine in the store. *Wine Economics and Policy*, *5*(2), 87–95. https://doi.org/10.1016/j.wep.2016.10.001

Revelt, D., & Train, K. (1998). Mixed Logit with repeated choices: Households' choices of appliance efficiency level. *The Review of Economics and Statistics*, *80*(4), 647–657. https://doi.org/10.1162/003465398557735

Rose, J. M., Hensher, D. A., Greene, W. H., & Washington, S. P. (2012). Attribute exclusion strategies in

airline choice: accounting for exogenous information on decision maker processing strategies in models of discrete choice. *Transportmetrica*, *8*(5), 344–360. https://doi.org/10.1080/18128602.2010.506897

Scarpa, R., Gilbride, T. J., Campbell, D., & Hensher, D. A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics*, *36*(2), 151–174. https://doi.org/10.1093/erae/jbp012

Scarpa, R., Zanoli, R., Bruschi, V., & Naspetti, S. (2013). Inferred and stated attribute non-attendance in food choice experiments. *American Journal of Agricultural Economics*, *95*(1), 165–180. https://doi.org/10.1093/ajae/aas073

Thiene, M., Scarpa, R., Galletto, L., & Boatto, V. (2013). Sparkling wine choice from supermarket shelves: the impact of certification of origin and production practices. *Agricultural Economics*, *44*, 523–536. https://doi.org/10.1111/agec.12036

Weller, P., Oehlmann, M., Mariel, P., & Meyerhoff, J. (2014). Stated and inferred attribute non-attendance in a design of designs approach. *Journal of Choice Modelling*, *11*, 43–56. https://doi.org/10.1016/j.jocm.2014.04.002.

Chapter 5

Dealing with ignored attributes through an inferred approach in wine choice experiments

Submitted for publication

Abstract

A growing number of studies have sought to understand the effects that rules and heuristics might have on discrete choice models' outputs, especially the impacts on willingness to pay measures. This paper presents a methodological approach to explore all possible strategies that respondents might use when making a choice of wine. Our results show that accounting for attribute non-attendance improves the goodness-of-fit of models and their predictive capability. The inferred attendance approach developed in this article discloses that a minority of respondents considers all attributes and there is a multitude of attribute attendance possibilities that the researcher should search for. The use of heuristics by wine consumers highlights one implication of this research. Wineries should develop more appropriate communication strategies for consumers, especially with regard to the information provided on the labels. Additionally, it provides evidence that ignoring attribute non-attendance leads to overestimates of welfare measures.

Keywords: attribute non-attendance; latent class models; discrete choice experiments; inferred attribute non-attendance; willingness to pay; wine choice

JEL classification: C25, C52, D11, Q18

1. Introduction

Discrete Choice experiments (DCE) are widely used to elicit individuals' valuation in different fields. Under this technique, respondents make repeated choices that are presented in sets formed by distinct alternatives corresponding to a product or service. Rooted in the continuity axiom, it assumes that respondents make trade-offs among all product's attributes, and considers that individuals are able to process information in a fully rational manner. However, a growing body of literature has demonstrated that individuals use simplifying strategies derived from heuristics when responding (Hensher, Rose and Greene, 2005, 2012; Hensher and Rose, 2009; Scarpa *et al.*, 2009; Campbell, Hensher and Scarpa, 2011). One of the most common simplifying strategies used is known as Attribute Non-attendance (ANA), where individuals ignore one or more attributes while making a choice. This violates the continuity axiom, and may arise for many reasons, such as cognitive burden, unfamiliarity with the good, or simply because an attribute or its levels are not relevant in the choice process (Alemu *et al.*, 2013).

The ability to identify ANA in DCE is important since when one attribute is ignored, any loss in this attribute is not compensated by an improvement in another. Consequently, not accounting for this heuristic could potentially lead to biased coefficients and policy orientations (Scarpa *et al.*, 2009; Campbell, Hensher and Scarpa, 2011). Thus, ANA is critical in understanding agents´ behavior and getting accurate estimates of utility parameters.

The study of ANA has been particularly prominent, and many studies have found evidence of its presence in several fields. While its roots can be found in transport economics (e.g. Hensher, Rose, and Greene 2005; Hensher 2006; Hess et al. 2013), and substantial literature exists in the fields of environment (e.g. Alemu et al., 2013; Carlsson, Kataria, & Lampi, 2010; Scarpa et al., 2009) and health economics (e.g. Hole, Kolstad, and Gyrd-Hansen 2013; Hole, Norman, and Viney 2016; Lagarde 2013), a growing literature, but still insufficient, appears in food or agricultural economics (e.g. Caputo et al. 2018; Thiene et al. 2013). In particular, literature has started to explore the role that ANA might play in improving the predictive capability of choice models, and its effects on willingness to pay (WTP) measures, compared to the specifications where all attributes are assumed to be relevant.

Two approaches have been used to identify ANA. Stated ANA, relies on self-assessment where respondents are directly asked whether or not they ignored one or more attributes (e.g. Carlsson et al., 2010; Hensher et al., 2005; Rose, Hensher, Greene, & Washington, 2012). Inferred ANA, in turn, makes use of analytical modelling techniques to infer non-attendance, without asking respondents their decision process, generally through a latent class approach (e.g. Campbell et al., 2011; Caputo et al., 2018; Hess

et al., 2013; Lagarde, 2013; Scarpa et al., 2009). Specifically, the most popular is the equality constrained latent class (ECLC) model, which assumes all classes having the same preference intensity, and only differ on the basis of a particular pattern of ignored attributes with no impact on utility.

In the specific case of wine, when making a purchase decision, the consumer is presented with a large variety of wines, with varying taste and attributes (Lockshin *et al.*, 2006), such as quality. Wine quality concept is widely considered of great complexity and it can be decomposed in numerous sub-components (Charters and Pettigrew, 2007), known as intrinsic and extrinsic properties (Jover, Montes and Fuentes, 2004). Extrinsic properties consist in information generally provided on the label of a bottle. However, consumers may have difficulty in processing and using the information provided to form an idea of the quality of the wine and consequently make a choice. Hence, choosing a wine bottle is a complex and multidimensional task, configuring a big challenge.

Understanding the drivers of wine consumers' purchasing decisions has been the object of a lively debate. Wine complexity leads researchers to try different alternatives to understand how consumers choose wines. DCE have been used to understand these choices (e.g. Lockshin et al. 2006; Mueller, Lockshin, et al. 2010; Perrouty, D´Hauteville, and Lockshin 2006; Sáenz-Navajas et al. 2013; Williamson et al. 2016). However, the number of studies assuming that some of these attributes are possibly ignored by respondents is very limited.

The aim of this paper is to explore consumer preferences for wine through the investigation of different ANA patterns and its effect on welfare measures and consequently on derived policy implications. Given that respondents may simultaneously ignore two or more attributes at the same time, it may be insufficient to consider the patterns where only a single attribute is ignored to fully address ANA. So, we investigate the effect of all ANA possible strategies used by wine consumers on (i) the detection of relevant ANA strategies, (ii) the welfare values found for wine, and (iii) the possible biases of ignoring ANA on the outcome of the different models. Particularly, we employ the 2^K approach where all possible combinations of ANA are tested, and we also explore simpler strategies such as lexicographic and single non-attendance. The final model results from the combination of the relevant patterns of these different models. Finally, we translate the evidence into WTP for wine and contrast it with the results from the multinomial logit (MNL) model in which full attendance is assumed.

This study contributes to the growing literature accounting for the use of heuristics, by investigating whether individuals have ignored the attributes in the choice sets. With the estimation of a 2^K model following a stepwise approach, we test all possible ANA response strategies in the case of a larger DCE,

which is often impossible to estimate. It also contributes to the study of heuristics in DCE in food/agricultural field, whose literature is limited and where ANA may be relevant but has been mostly ignored.

The remainder of this paper is structured as follows. Section 2 describes the survey design and introduces the data, followed by section 3 which presents the methodological procedure. Section 4 reports the results including WTP measures. Section 5 offers concluding remarks.

2. Survey, sample, and data collection

We constructed a choice experiment study on wine choice, using a European web panel. Six wine attributes describing each bottle of wine were used: medals, landscape, alcohol level, country of origin, grape variety and price. These attributes, as well as their respective levels (see Table 1), were determined through a literature review, advice from experts, focus groups, and through a pilot study. The stated choice exercise, generated by a D-efficient design using the Ngene software (ChoiceMetrics, 2014), is composed by ten choice sets each one with three alternatives, plus a non-purchase alternative (see Figure 1 for an example of a choice set). An online survey was employed by a market research company in August 2018. A total of 600 respondents completed the questionnaire.

Table 1 - Attributes and levels used in the choice sets

Attribute	Levels
Country of origin	Italy, Portugal, France, USA, Australia, Chile
Landscape	If it was produced in a certified landscape or not
Medals	If it has or not a gold medal
Alcohol level	12vol., 13.5vol., 15vol.
Grape variety	Cabernet Sauvignon, Syrah, Red blend, no information
Price	3.99€, 7.50€, 12.50€, 18€

To make the wines more realistic, five constant attributes always appear in all wines: a fictional brand, an arbitrary vintage (2016), volume, bottle style, and appellation of origin. We used a Bordeaux bottle style with the standard size (750ml). All wines have an appellation classification according to the country of origin system and equivalent to the "Protected Designation of Origin". We controlled for situation by focusing on a wine purchased for a formal dinner with guests at home.



Figure 1 - Choice set example

3. Methods

3.1. Latent Class Model to account for attribute non-attendance

The LCM identifies heterogeneous groups of respondents where individuals in the same group share the same preferences. The utility parameters, β_n , take a finite number C of values β_{n1}, β_{n2},..., β_{nC} with corresponding probabilities H_{n1}, H_{n2},..., H_{nC} of belonging to a particular latent class, c. Under this framework, two probabilities are estimated: (i) the class probabilities of choosing the alternative j in T choice situations conditional on the preferences within each class:

$$P_{njt|c} = \frac{\exp(\beta_c x'_{njt})}{\sum_{j=1}^{Jn} \exp(\beta_c x'_{njt})} \quad (1)$$

and (ii) the probability that respondent n belongs to a particular class, given by equation (2) (Greene & Hensher, 2003):

$$H_{nc} = \frac{\exp(\theta_c z_n)}{\sum_{c \in C} \exp(\theta_c z_n)}, \qquad c = 1, \ldots, C, \qquad \theta_C = 0 \quad (2)$$

where θ_c are class-specific parameters, with C^{th} vector normalized to ensure model identification, and z_n is a vector of alternative attributes and individual´s characteristics. This implies the maximization of the following log-likelihood function for the sample, equation (3):

$$\ln L = \sum ln\left(\sum_{c=1}^{C} H_{nc} \prod_{t=1}^{T} P_{njt|c}\right) \quad (3)$$

After obtaining the parameter estimates of θ_c, we can calculate the posterior respondent-specific estimate of the class probabilities ($\hat{P}_{c|n}$) using Bayes´ theorem, equation (4) (Greene & Hensher, 2003):

$$\hat{H}_{c|n} = \frac{\hat{P}_{nj|c}\, \hat{H}_{nc}}{\sum_{c=1}^{C} \hat{P}_{nj|c} \hat{H}_{nc}} \quad (4)$$

In our study, we employ ECLC models, where each class represents different ANA strategies used by respondents, where a non-attended attribute has its coefficient set to zero, while an attended attribute is assumed to take the same nonzero coefficient across classes (Hensher & Greene, 2010). This constrains the utility weights to be equal in all classes, in a similar way to multinomial logit models. To identify preference heterogeneity across classes, this assumption can be relaxed. However, this is not the aim in this study.

3.2. Exploring ANA patterns through different class models

The 2^K approach is a useful mechanism to give insights of the behavioral specification of the LCM to investigate the extent of attribute non-attendance profiles. Under this specification, for the K attributes there are 2^K pre-defined strategies employed by respondents. Following a stepwise approach (Lagarde, 2013) to run the 2^K LCM with six attributes we investigate 64 possible ANA strategies. To focus only on

ANA patterns, we constrained the estimated parameters to be equal across classes, and estimate different equality constrained latent class models (e.g. Scarpa et al. 2009; Hensher and Greene 2010; Lagarde 2013; Jourdain and Vivithkeyoonvong 2017). We start by estimating 7 classes, one for full attendance and six others (classes 2 to 7) for ANA where only one attribute can be ignored. We assume that if a simple ANA pattern got zero probability at one level it should be excluded at a different level, when combined within more complex strategies (e.g. if a consumer ignores the price attribute, then the combination between price and other attribute should also be irrelevant). Then, we run a new LCM with the old relevant rules plus the new ANA patterns derived from the combination between two attributes. In the third step we repeat this strategy, including ANA when three attributes at a time can be ignored. This procedure continues until we include all 64 patterns. The final ANA model from this procedure (which we denote as ECLC-2^K) only encompasses the relevant strategies, whose class-membership probabilities are different from zero. Despite this result, intuitively, we may think of other patterns. To investigate this hypothesis, we estimate a LCM with ANA for a single attribute (ECLC-ANA1) and a LCM with lexicographic preferences, or attribute attendance (AA) for a single attribute (ECLC-AA1) following the same strategy.

Finally, we logically combine the patterns from these 3 models mentioned above, and achieve another model (ECLC-final). The irrelevant patterns were always excluded until the final class model is reached. Model selection was based on the lowest value of the Akaike information criteria (AIC), and highest value of McFadden's Pseudo R^2 and predictive capability of correct choices. All models were estimated with Nlogit 6.0.

4. Results

4.1. ANA patterns

Table 2 reports the class membership probabilities of six LCMs accounting for full attendance (class 1) and non-attendance for one attribute (class 2), to briefly explore the extent to which respondents may have ignored some of the attributes (model 1 to model 6). Results show that 25% of the respondents might have ignored the alcohol attribute (model 3). However, for the other five attributes, the majority of respondents seem to have been prone to ignore them, suggesting a high incidence of non-attendance within our sample.

Table 2 - Preliminary results of LCM

Model	Attribute ignored in the second class	Average class membership (%)
Model 1	Medals	58.4
Model 2	Landscape	68.7
Model 3	Alcohol level	25.1
Model 4	Country of origin	73.0
Model 5	Grape	88.0
Model 6	Price	60.2

After these preliminary results, we proceed with the 2^K approach, and 6 successive ECLC models were estimated to test the simultaneous use of different non-attendance strategies by respondents, until all 64 patterns were tested (Appendix E). For simplification reasons, we assume that when an attribute was ignored, then its levels were also, otherwise the number of parameters would quickly proliferate under this method. Based on parsimony and practicality, as well on the significance and magnitude of utility coefficients, we excluded the non-significant class probabilities, and our 2^K model (i.e. ECLC-2^K) shows that respondents only used 6 of the 64 possible patterns (Table 3). This stepwise strategy suggests that our sample seem to rely on lexicographic behavior (only medals or COO attended), ignoring three of six attributes, and ignoring the combination of two attributes.

Then, the model accounting for ANA for a single attribute (ECLC-ANA1 in the Appendix E) and the model for lexicographic preferences (ECLC-AA1) were estimated to explore the employment of extreme ANA strategies (Table 3).

Herewith, and combining the relevant strategies with higher probability from these three models, we obtain the final model. Results of ECLC-final model indicate that 28% of the respondents only attend to medals, 26.6% only consider country of origin and only 14.1% attend to price when choosing wine in the specific occasion considered. However, it is striking to notice that only 10% make trade-offs between all attributes. Similar rates, 8.7%, were also found in the field of environmental economics (Campbell, Hensher and Scarpa, 2011). This contrasts with the results from the ECLC-AA1 and ECLC-2^K model in which 29% and 23.3%, respectively, of respondents consider all attributes, and 16.6% and 8.3% attend to medals, respectively. On the opposite, 8.8% of them looked at all attributes but the grape variety and 12.5% do not take into account grape and country of origin in their choices.

For the ECLC-final model, the aggregate membership probability, from summing up the probabilities of LC where each attribute was ignored, indicate that grape varietals were ignored by 90% of respondents, price by 54.6%, medals by 40.7%, COO by 54.6%, and alcohol and landscape by 68.7%. When compared to the descriptive analysis in Table 2, these results are quite similar, except the divergence in terms of attribute attendance specially for alcohol level (ANA rate increases from 25.1% to 68.7%) and medals (ANA decreases from 58.4% to 40.7%).

Overall, our findings suggest that limiting ANA to a single attribute or to the extreme lexicographic behavior may affect the conclusions about the incidence of non-attendance. This highlights the importance of using a procedure to account for mixing strategies where many attributes are not attended at same time.

Table 3 - Class memberships from the different ECLC models

Model	Description	Latent class probability (%)
ECLC-2^K	Full Attendance	23.3***
	ANA–GP (grape variety and price ignored)	40.8***
	ANA–MEDCOO (medals and country of origin ignored)	10.2***
	ANA–CS (country specific attributes ignored - country of origin, landscape and grape variety)	3.3**
	AA-MED (only medals attended)	8.3***
	AA-COO (only country of origin attended)	14.0***
ECLC-AA1	Full Attendance	29.0***
	AA-MED (only medals attended)	16.6***
	AA-ALC (only alcohol level attended)	13.3***
	AA-COO (only country of origin attended)	29.0***
	AA-P (only price attended)	12.1***
ECLC-final	Full attendance	10.0***
	AA-MED (only consider medals)	28.0***
	AA-P (only price attended)	14.1***
	AA-COO (only coo attended)	26.6***
	ANA-G (only grape ignored)	8.8***
	ANA-GCOO (grape and coo ignored)	12.5***

Note: ***, ** and * denote significance at 1%, 5%, 10% level respectively

4.2. Impact of ANA on model estimates and WTP

Table 4 reports the fit and the estimates of our MNL model which assumes full attribute attendance (i.e., passive bounded rationality) and ECLC models assuming the use of heuristics. Accounting for ANA results in a significant improvement in model fit and predictive capability over the MNL, suggesting that respondents made their choices based on very little information (as in Jourdain and Vivithkeyoonvong 2017; Lagarde 2013). The AIC and Log Likelihood decrease substantially and the McFadden´s Pseudo R^2 becomes higher in all ECLC models. These indicators show that the ECLC-2^K model slightly fits the data better than ECLC-final model. However, looking at the correctly predicted indicator, our proposed model (ECLC-final) performs best with the highest predictive capability, as it correctly predicted the actual choice outcome for more than 86 percent of the total number of observations, compared to 74.6 percent from ECLC-2^K. In terms of the sign and significance of coefficients, the models gave similar results except for grape variety (Cabernet Sauvignon and Syrah) and COO (specially from the new producers). As expected, the coefficient for price is negative and statistically significant indicating that consumer's utility decreases with increasing price.

Table 5 reports the WTP estimates for the MNL and ECLC models with best fit and better predictions, i.e. ECLC-2^K and ECLC-final. Almost all WTP estimates are statistically significant. Of greater interest are the differences in WTP magnitudes between models. With the clear exception of the "alcohol level" and "red blend" attributes the welfare measures are substantially lower when ANA is accounted for, meaning that WTPs from the MNL appear to be overestimated. Under this standard model, the highest WTPs are for Portuguese and French origins, while the ECLC put red blend wine on the top of preferences. Consumers seem prepared to pay a large premium for this type of wines (more than 10€), being the WTPs insignificant for both Cabernet Sauvignon and Syrah varieties under ECLC-final. These results contrast to the results from the ECLC-2^K, which suggests Chilean wines and Syrah varieties as important to wine consumers.

Additionally, when accounting for ANA, results show that people seem to put a relatively low weight on landscape, which decreases from 2.40€ under MNL to less than 0.90€ in both ECLC. In line with findings reported in other studies (e.g. Hole 2011; Scarpa et al. 2009; Sandorf, Campbell, and Hanley 2017; Heidenreich 2017; Erdem, Campbell, and Hole 2014; Scarpa et al. 2013; Hensher, Rose, and Greene 2005), these results show that ignoring ANA may lead to different (generally lower) values of welfare measures, and aside from the magnitudes and significance, even different ranking orders of the WTP estimates.

Table 4 - Estimates for the MNL and ECLC models

	MNL	ECLC–2^K	ECLC-final
Attribute	**β**	**β**	**β**
Medals	0.38***	0.21***	0.95***
	(0.05)	(0.04)	(0.04)
Landscape	0.16***	0.20***	0.30***
	(0.04)	(0.04)	(0.07)
Alcohol	0.04***	0.28***	0.24***
	(0.01)	(0.01)	(0.01)
COO (Italy)			
France	0.55***	1.09***	0.84***
	(0.05)	(0.06)	(0.07)
Portugal	0.749***	1.77***	1.50***
	(0.07)	(0.07)	(0.07)
USA	-0.37***	0.08	-1.01***
	(0.06)	(0.06)	(0.11)
Australia	-0.39***	-0.06	-2.10***
	(0.09)	(0.11)	(0.40)
Chile	0.05	1.05***	-0.69***
	(0.08)	(0.10)	(0.18)
Grape (none)			
Cabernet Sauvignon	0.23***	0.22**	-0.09
	(0.05)	(0.09)	(0.12)
Syrah	0.07	1.13***	-0.23
	(0.06)	(0.13)	(0.28)
Red Blend	0.42***	2.60***	4.00***
	(0.09)	(0.22)	(0.59)
Price	-0.07***	-0.26***	-0.36***
	(0.01)	(0.01)	(0.01)
Constant	-0.05	0.73***	0.60***
	(0.19)	(0.08)	(0.04)
LL	-7557.6	-6556.2	-6578.4
AIC	15141.2	13148.4	13192.8
Pseudo R^2	0.076	0.212	0.209
Correctly predicted (%)[8]	42.4	74.6	86.2

Note: ***, ** and * denote significance at 1%, 5%, 10% level respectively

Attribute levels given in parentheses are the base level for dummy coding

Standard errors in parentheses

[8] See Appendix E

Table 5 - Willingness to pay estimates for wine, for each model, in euros

Attribute	MNL WTP	ECLC-2^K WTP	ECLC-final WTP
Medals	5.72***	0.83***	2.68***
Landscape	2.40***	0.76***	0.84***
Alcohol	0.56***	1.07***	0.67***
COO (Italy)			
France	8.24***	4.23***	2.36***
Portugal	11.19***	6.88***	4.25***
USA	-5.48***	0.30	-2.86***
Australia	-5.90***	-0.23	-5.94***
Chile	0.80	4.08***	-1.93***
Grape variety (none)			
Cabernet Sauvignon	3.50***	0.85**	0.26
Syrah	1.07	4.41***	-0.66
Red blend	6.32***	10.11***	11.29***

Note: ***, ** and * denote significance at 1%, 5%, 10% level respectively.

Attribute levels given in parentheses are the base level for dummy coding.

5. Conclusion

The research in this paper has built upon the literature on inferred ANA. DCE conventionally assume behavior respecting the continuity axiom where respondents make trade-offs between all attributes. However, as shown in this study as well as in many other studies, this assumption may not always be suitable. The procedure we follow is based in the 2^K approach, combined with the most relevant lexicographic and single ignored attribute strategies. Thus, we estimate ECLC models, being aware of the limitations/implications, such as overestimation of the presence of ANA, on constraining all non-zero parameters to be equal across classes. However, the modelling framework we followed will imply a huge number of classes tested simultaneously if we assume different preference classes. The 2^K approach seems to be a very useful mechanism to explore ANA patterns, clearly for a small number of possible strategies. However, testing separately other patterns of ANA may reveal additional important strategies, which may contribute to increase the predictive power of model.

Accounting for ANA resulted in a considerable improvement in model fit and some differences in the significance and strength of the estimated coefficients. The different models suggest that wine consumers seem to behave in a lexicographic manner and made their choices on a basis of very little information, which demands appropriated communication strategies by wineries, for instance with respect to the information provided on the labels.

An issue of concern in DCE is the share of people ignoring the price or the monetary attribute. In this study, almost 55% of respondents ignore price, resulting in a large proportion of respondents whose wine value estimates cannot be accessed. With particular importance for the calculation of WTP values and compared to those obtained from the MNL specification, estimates from ECLC provide generally lower WTP values, and show that wine WTP are sensitive to attribute processing strategies. This empirical evidence illustrates the importance of accounting for the potential bias introduced when ANA behavior is ignored and reinforces past evidence which suggests that failure to identify and account for attribute processing rules tend to result in over-estimates of WTP.

Given the evidence in different fields and contexts, we believe that some concerns should be taken, at least, to be aware of respondents' propensity to ignore some attributes in the cases where it is neglected. This is particularly relevant in an economic appraisal context and for policy recommendations, and may be appropriate in the study of preferences for wine.

References

Alemu, M., Mørkbak, M., Olsen, S., & Jensen, C. (2013). Attending to the Reasons for Attribute Non-attendance in Choice Experiments. *Environmental Resource Economics*, *54*, 333–359. https://doi.org/10.1007/s10640-012-9597-8

Campbell, D., Hensher, D. A., & Scarpa, R. (2011). Non-attendance to attributes in environmental choice analysis: a latent class specification. *Journal of Environmental Planning and Management*, *54*(8), 1061–1076. https://doi.org/10.1080/09640568.2010.549367

Caputo, V., Van Loo, E. J., Scarpa, R., Nayga, R. M., & Verbeke, W. (2018). Comparing Serial, and Choice Task Stated and Inferred Attribute Non-Attendance Methods in Food Choice Experiments. *Journal of Agricultural Economics*, *69*(1), 35–57. https://doi.org/10.1111/1477-9552.12246

Carlsson, F., Kataria, M., & Lampi, E. (2010). Dealing with ignored attributes in choice experiments on

valuation of Sweden's environmental quality objectives. *Environmental and Resource Economics*, *47*(1), 65–89. https://doi.org/10.1007/s10640-010-9365-6

Charters, S., & Pettigrew, S. (2007). The dimensions of wine quality. *Food Quality and Preference*, *18*(7), 997–1007. https://doi.org/10.1016/j.foodqual.2007.04.003

Erdem, S., Campbell, D., & Hole, A. R. (2014). Accounting for Attribute-Level Non-Attendance in a Health Choice Experiment: Does it Matter? *Health Economics*, *24*, 773–789. https://doi.org/10.1002/hec

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transportation Research Part B: Methodological*, *37*(8), 681–698. https://doi.org/10.1016/S0191-2615(02)00046-2

Heidenreich, S. (2017). Decision heuristic or preference ? Attribute non - attendance in discrete choice problems. *Health Economics*, *27,* 157-171. https://doi.org/10.1002/hec.3524

Hensher, D. A. (2006). How Do Respondents Process Stated Choice Experiments? Attribute Consideration under Varying Information Load. *Journal of Applied Econometrics*, *21*, 861–878. https://doi.org/10.1002/jae.877

Hensher, D. A., & Greene, W. H. (2010). Non-attendance and dual processing of common-metric attributes in choice analysis : a latent class specification. *Empirical Economics*, *39*, 413–426. https://doi.org/10.1007/s00181-009-0310-x

Hensher, D. A., & Rose, J. M. (2009). Simplifying choice through attribute preservation or non-attendance: Implications for willingness to pay. *Transportation Research Part E*, *45*(4), 583–590. https://doi.org/10.1016/j.tre.2008.12.001

Hensher, D. A., Rose, J. M., & Greene, W. H. (2005). The implications on willingness to pay of respondents ignoring specific attributes. *Transportation*, *32*(3), 203–222. https://doi.org/10.1007/s11116-004-7613-8

Hensher, D. A., Rose, J. M., & Greene, W. H. (2012). Inferring attribute non-attendance from stated choice data: implications for willingness to pay estimates and a warning for stated choice experiment design. *Transportation*, *39*, 235–245. https://doi.org/10.1007/s11116-011-9347-8

Hess, S., Stathopoulos, A., Campbell, D., O'Neill, V., & Caussade, S. (2013). It's not that I don't care, I just don't care very much: Confounding between attribute non-attendance and taste heterogeneity.

Transportation, *40*(3), 583–607. https://doi.org/10.1007/s11116-012-9438-1

Hole, A. R. (2011). A discrete choice model with endogenous attribute attendance. *Economics Letters*, *110*(3), 203–205. https://doi.org/10.1016/j.econlet.2010.11.033

Hole, A. R., Kolstad, J. R., & Gyrd-Hansen, D. (2013). Inferred vs. stated attribute non-attendance in choice experiments: A study of doctors' prescription behaviour. *Journal of Economic Behavior and Organization*, *96*, 21–31. https://doi.org/10.1016/j.jebo.2013.09.009

Hole, A. R., Norman, R., & Viney, R. (2016). Response patterns in health state valuation using endogenous attribute attendance and latent class analysis. *Health Economics*, *25*, 212–224. https://doi.org/10.1002/hec

Jourdain, D., & Vivithkeyoonvong, S. (2017). Valuation of ecosystem services provided by irrigated rice agriculture in Thailand: a choice experiment considering attribute nonattendance. *Agricultural Economics*, *48*(5), 655–667. https://doi.org/10.1111/agec.12364

Jover, A. J. V., Montes, F. J. L., & Fuentes, M. del M. F. (2004). Measuring perceptions of quality in food products: the case of red wine. *Food Quality and Oreference*, *15*(5), 453–469. https://doi.org/10.1016/j.foodqual.2003.08.002

Lagarde, M. (2013). Investigating attribute non-attendance and its consequences in choice experiments with latent class models. *Health Economics*, *22*, 554–567. https://doi.org/10.1002/hec

Lockshin, L., Jarvis, W., D´Hauteville, F., & Perrouty, J.-P. (2006). Using simulations from discrete choice experiments to measure consumer sensitivity to brand, region, price, and awards in wine choice. *Food Quality and Preference*, *17*, 166–178. https://doi.org/10.1016/j.foodqual.2005.03.009

Mueller, S., Lockshin, L., Saltman, Y., & Blanford, J. (2010). Message on a bottle: The relative influence of wine back label information on wine choice. *Food Quality and Preference*, *21*(1), 22–32. https://doi.org/10.1016/j.foodqual.2009.07.004

Perrouty, J. P., D´Hauteville, F., & Lockshin, L. (2006). The Influence of Wine Attributes on Region of Origin Equity: An Analysis of the Moderating Effect of Consumer's Perceived Expertise. *Agribusiness*, *22*(3), 323–341. https://doi.org/10.1002/agr

Rose, J. M., Hensher, D. A., Greene, W. H., & Washington, S. P. (2012). Attribute exclusion strategies in airline choice: accounting for exogenous information on decision maker processing strategies in models of discrete choice. *Transportmetrica*, *8*(5), 344–360.

https://doi.org/10.1080/18128602.2010.506897

Sáenz-Navajas, M., Campo, E., Sutan, A., Ballester, J., & Valentin, D. (2013). Perception of wine quality according to extrinsic cues: The case of Burgundy wine consumers. *Food Quality and Preference*, *27*, 44–53. https://doi.org/10.1016/j.foodqual.2012.06.006

Sandorf, E., Campbell, D., & Hanley, N. (2017). Disentangling the influence of knowledge on attribute non-attendance. *Journal of Choice Modelling*, *24*, 36–50. https://doi.org/10.1016/j.jocm.2016.09.003

Scarpa, R., Gilbride, T. J., Campbell, D., & Hensher, D. A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics*, *36*(2), 151–174. https://doi.org/10.1093/erae/jbp012

Scarpa, R., Zanoli, R., Bruschi, V., & Naspetti, S. (2013). Inferred and stated attribute non-attendance in food choice experiments. *American Journal of Agricultural Economics*, *95*(1), 165–180. https://doi.org/10.1093/ajae/aas073

Thiene, M., Scarpa, R., Galletto, L., & Boatto, V. (2013). Sparkling wine choice from supermarket shelves: the impact of certification of origin and production practices. *Agricultural Economics*, *44*, 523–536. https://doi.org/10.1111/agec.12036

Williamson, P. O., Lockshin, L., Francis, I. L., & Loose, S. (2016). Influencing consumer choice: Short and medium term effect of country of origin information on wine choice. *Food Quality and Preference*, *51*, 89–99. https://doi.org/10.1016/j.foodqual.2016.02.018

Chapter 6

Discussion and conclusion

Given the structure by articles followed in this thesis, and to avoid repetition, this chapter provides the general main conclusions of each topic covered in the empirical work. This thesis sheds light on the possible challenges posed by preference elicitation for a complex good using discrete choice experiments. Modelling behavior is clearly a complex task, and the use of basic discrete choice models' specifications result in simplistic representations. The research here developed provides three broad contributions, that can be translated in methodological and practical implications.

Regarding the methodological issues, the thesis concludes that modelling consumer choice towards a highly differentiated product, requires accounting for distinct behavioral patterns and/or sources of heterogeneity. First, Chapter 2 focuses on the importance of unobserved heterogeneity, allowing for preference heterogeneity, scale heterogeneity, and both. In particular, several authors suggest that scale heterogeneity is being frequently neglected. By comparing different discrete choice models the understanding of consumers' unobserved heterogeneity is deepened, and evidence on the ability of the Scaled Multinomial Logit, the Mixed Logit, and the Generalized Mixed Logit models to fit the data is provided. The findings suggest that capturing scale heterogeneity but failing in capturing taste variation may be insufficient for empirical interest, with inferior model performance. Combining both sources of heterogeneity in the Generalized Mixed Logit model resulted in an improvement over the models accounting for the different sources of heterogeneity in a separated way. This shows the potential of this approach to better explain the choices of consumers who behave in a more extreme way, such as lexicographic preferences, where the option with a specific attribute is always chosen. This empirical assessment shows that the accommodation of both sources of unobserved heterogeneity plays an important role, and supports the evidence on the literature that the choice of wine is a complex task. It further highlights the fact that consumers may choose wines ignoring some information, using heuristics, or making random choices.

Given this evidence, and the increasing attention that the use of heuristics has received from choice researchers, the abandonment of the traditional full compensatory utility maximization framework is proposed, allowing for distinct behavioral patterns of choice. Then, in the second part of this empirical work, attribute non-attendance is considered and its impacts on DCE outputs, especially willingness-to-pay are explored. In this sense, this thesis contributes to deepen the understanding of the effects of

attribute processing heterogeneity on current choice modelling methodologies, both theoretically and empirically.

In Chapter 3, the literature on ANA modelling in DCE is reviewed in four areas of research: transportation, health economics, environmental economics and food choice. Existing methodologies for handling ANA are reviewed, and a common framework is proposed in order to facilitate the comparison of assumptions across models, in addition to proving a unique conceptual language. This review shows that non-attendance definitely occurs in the majority of the reviewed articles and that it matters in terms of the reliability of the discrete choice technique. On the one hand, because from a behavioral and theoretical perspective, individuals are assumed to behave in a compensatory manner, while ANA challenges the continuity axiom. On the other hand, because it may lead to biased DCE outputs. The vast majority of papers show biased willingness to pay estimates. The more problematic rates of non-attendance for monetary attributes appear in health and environmental fields, which we relate to involving a higher degree of consumer sensitivity when deciding over health or over environmental issues. This work shows the importance of non-attendance behavior in different contexts for the definition of more adequate policy orientations, as ignoring it leads to biased welfare estimates.

In terms of ANA identification, two approaches have been used to deal with attribute ignorance, one relying on self-reported non-attendance data and the other one based on model inference. In Chapter 4 these modelling approaches are compared, based on stated and inferred attribute non-attendance, in three different markets to ensure more accurate and robust results. Inferred ANA was derived from an equality constrained latent class model, where the class membership probabilities were estimated for pre-defined classes for each country ("Full attendance", "Ignore only landscape", "Ignore only alcohol", "Ignore only COO", "Ignore only grape variety", "Ignore landscape and alcohol", "Ignore landscape, alcohol, and grape variety", "Ignore landscape, alcohol, and COO", "Ignore awards, landscape, and alcohol", "Ignore landscape, alcohol, grape variety, and price", "Only consider COO", "Only consider price", "Ignore all/random choice"). These show that attendance differ quite substantially between classes and countries. Particularly, in the American and French markets only a relatively small proportion of respondents considered all attributes (less than 4%), contrasting with the Portuguese market where this class comprises almost 19% of respondents. In terms of attendance to price, we found a large share of respondents considering the price attribute in their choices, in all datasets and both at the inferred and stated approaches, which is important for the calculation of willingness to pay measures.

Regarding stated ANA, results show that respondents statements of non-attendance differ from inferred results, with little concordance between both approaches, which suggests that researchers should not truthfully believe and use respondents self-reported measures. The findings also suggest that respondents state that they have attended certain attributes, when they have not actually taken them into account according to model inference, which result in a higher rate of stated attribute attendance. Consequently, some attributes that appear to be important in the choice process are actually irrelevant and may be the target of incorrect policy definitions or marketing strategies.

These results immediately prompt researchers to misjudge respondents and think that they gave an incorrect or untruthful answer to the non-attendance question. However, this may lead to the development of another issue, not covered in this thesis, related to the reasons for ignoring attributes. In fact, ANA may be associated with behavioral reasons, but also because the questions might be incorrectly understood. On the other hand, the model selected to infer non-attendance may not have the most suitable specification to capture ANA (see Alemu et al. (2013) for more details). Thus, this discussion concerning the most appropriate way to deal with ANA may remain an open window for further developments, especially to investigate how to deal with non-attendance from respondents' statements that are considered as biased.

Given the possibility of biased results, in Chapter 5 a latent class model is estimated, assuming a stepwise approach to capture all possible combinations of non-attendance patterns employed by respondents, in a full study of ANA based on the 2^K Latent class model. Such model comprises a total of 64 classes, although many classes would have only a few members. The resulting model identified six relevant patterns of non-attendance. Only a few studies in the literature employ this approach, and searching for extreme behavior for ANA, such as ANA for single attributes and lexicographic behavior, has been the most frequent option in the literature. Furthermore, we have extended the stepwise approach by adding and testing other patterns considered as potentially relevant from the models capturing extreme behavior. This refinement, generated another model, which we refer to as the final model, with a much higher percentage of correctly predicted choices. Therefore, the process we followed in this article shows that there is a range of significant attribute attendance combinations that the researcher should test.

The evidence provided by this paper shows that it is important to consider different patterns of non-attendance, much further than just consider ANA for single attributes or ANA for lexicographic choices. . The modelling procedure followed in the paper allows the exploration of all ANA strategies, and it is shown that there is a gain in combining the relevant strategies from different models, as better predictive

capability is obtained. The results reinforce that wine consumers made their choices ignoring some attributes, and reveal that, in fact, some consumers use very little information and behave very close to a lexicographic manner when choosing wine.

In terms of welfare measures, the work here presented offers new evidence supporting the view that wrong assumptions about attribute processing heterogeneity, and specifically failure to account for it, would result in inadequate welfare estimates, which appear to be overestimated in our study. The most notable example is the change in willingness to pay estimates both in magnitude and sign in the valuation that American consumers attach to Red blend varieties under the MNL model (9.17$), when in the ECLC model this result falls to 0.16$. In France, the willingness to pay for Cabernet Sauvignon and Syrah wines has even changed the sign, while the significance remains. In all datasets, results show a significant drop in the valuation for medals when moving from the MNL and MIXL models to the ECLC model (from more than 5.99$ to less than 2.20$). Adding it to the existing literature in different contexts suggests that researchers should be concerned about the validity of standard specifications and results, as they significantly affect results and concomitantly the recommendations inferred.

Third, and regarding practical implications, the thesis is focused on the wine industry. Over the last decades it has been witnessed an intensive globalization and competition in this industry, which brings both challenges and opportunities to wine regions. Wine is a complex product with millions of consumers around the world. However, the study of consumers preferences for wine is relatively recent. The approach used in this thesis, with the employment of advanced techniques, such as DCE accounting for heuristics' use is *per se* an advance in the field. This thesis has shown that, in the context of wine behavior research, advanced model structures accounting for attribute processing heterogeneity have the potential to reach better performance than simpler approaches, providing further insights and different results.

From an empirical point of view, and reminding us of the context of wine choice in this work, the evidence of accounting for ANA suggests that price may act as an important and primary cue for quality, being considered by a prominent group of consumers. The results presented show that the country of origin, the gold medals, and the grape variety are the attributes which consumers are most willing to pay for when choosing wines for a special occasion. In the case of specific markets, North American consumers prefer the Cabernet Sauvignon variety, while Europeans prefer a Red blend. The origin is an important choice driver in the wine market, with consumers showing a dominant preference for wines from Old World producing countries in the markets under analysis. With the exception of the North American, Portuguese and French consumers reveal a clear preference for national wines.

Given the valuation attached to the landscape attribute in this study, there seems to exist some space for landscape certification to constitute a differentiating factor for wine for European consumers, at least for French and Portuguese. From the consumer's point of view, there is a need to link wine with something that guarantees its quality (Gergaud and Livat, 2007). Some authors (Angulo, Gil, Gracia, & Sánchez, 2000; Barreiro-Hurlé, Colombo, & Cantos-Villar, 2008) argue that the territory of origin is essential in choosing a product, being the landscape considered an important component of the origin of wine (Tempesta et al., 2010), and a relevant quality indicator for food-type products such as wine (Nelson, 1970). Thus, it appears to be a clear advantage to world heritage wine regions, such as the well-known Bordeaux and St-Emilion in France, Douro Valley in Portugal, or Prosecco in Italy. However, we do not exclude other types of landscape certification, being an example, the Globally Important Agricultural Heritage System (GIAHS) certification provided by the Food and Agricultural Organization (FAO). This finding suggests a joint effort of policy makers and wineries in an attempt to promote their wines in association with their landscapes as a differentiating attribute in the international market. It is also important to promote the certification of landscape labeling according to legal restrictions imposed by each type of certification where the vineyard is included.

Finally, important avenues for future research can be identified. As it has been already mentioned, while there is clear evidence that some respondents ignore some attributes, the knowledge about their reasons for non-attendance remain limited. The study of why attributes are ignored could help to better understand how respondents process information and consequently it will help to improve the experimental design, and to adopt more adequate choice models specifications. The investigation of the reasons for non-attendance requires the use of supplementary questions. Alternatively, the impact of respondents' characteristics on the probability of non-attendance, and how the survey characteristics can affect ANA can be analyzed. There is no doubt that ANA occurs. The results here obtained indicate that a combination of non-attendance with other heuristics, within a research framework accounting for multiple heuristics, despite being a challenge, may significantly contribute to a richer understanding of choice behavior. Given the live debate of which method to undertake to identify ANA, there is a very recent literature using eye-tracking data (visual inspection) to identify attribute attention. Including this method as a third way to capture ANA may be promising and may have potential to better understand the cognitive process in a choice experiment, and to improve model results.

On the other hand, it could be interesting to simultaneously capture the different behavioral issues covered in this thesis in a unified framework accounting for (a) differences in choice consistency (scale

heterogeneity), (b) preference heterogeneity, and (c) attribute non-attendance. In fact, the assumption that preferences across consumers vary only based on ANA behavior may be restrictive, and consequently, in the inferred approach, it may be important not only to account for ANA, but also to allow for consumer preference taste variation. This Latent class approach with random parameters will allow to allocate respondents to different preference groups combined with different attribute processing strategies. The preference heterogeneity may also be captured with behavioral factors, purchasing habits, lifestyle, and attitudinal questions, in a hybrid framework combining discrete choice models with latent variables. This approach will help to contribute to a better understanding of consumer's heterogeneity, which is interesting for policy makers and companies for the definition of their policies or marketing strategies.

Besides, extending this study to other countries could help to understand country specific wine valuation and to analyze whether the wine is a uniformly perceived product, whether its place of origin has any effect depending on the country of the respondent. It also contributes to the definition of specific strategies for wine producers, given the knowledge on the preferences by country and market, promoting a better alignment between wine consumers' preferences and the market supply. Additionally, a cross-national research may play an important role as wine is a globalized product, with firms interested in understanding the similarities and differences between market segments to develop better selling strategies.

Table 1 - Overview of the main findings of the thesis

Studying consumers heterogeneity		
Ch2	Aim	To model unobserved heterogeneity of wine consumers, by accommodating mixtures of preference heterogeneity and scale heterogeneity.
	Findings	Accounting for scale and taste heterogeneity is an important issue. Generalized mixed logit model performs better to explain consumers heterogeneity.
Studying heuristics		
Ch3	Aim	Review of ANA effects on DCE outputs, reflecting on the theoretical insights, econometric modelling, and developing a typology or organizing framework to classify choice models incorporating attribute non-attendance.
	Findings	Results show biased welfare measures in the majority of the reviewed articles, with more problematic non-attendance rates in the environmental and health economics fields. The organizing framework identifies two approaches to deal with ANA (stated and inferred) combined with modelling treatments (homogeneity or heterogeneity) for preference and attribute processing.
Ch4	Aim	To contrast the two methodologies to identify and deal with ANA; to explore the possible biases of ignoring ANA on WTP measures; and to identify the most important attributes for wine consumers.
	Findings	Evidence of ANA, only a small proportion of respondents considered all attributes. Price was one of the most attended attributes. Little concordance between stated and inferred ANA rates. Higher WTP in the stated approach, which suggests an overestimation of WTP. Old World wines are preferred, gold medals and grape variety. The landscape certification has a slight relevance, but it reveals potential to be a wine attribute.
Ch5	Aim	To identify all ANA possible strategies to detect the relevant ones used by wine consumers; to find the welfare values of wine attributes; and to find the ANA effects on the outcome of the different models.
	Findings	Results show potential benefits from the combination of relevant patterns from different models, which resulted in an increase of the predictive power of the final model. Only a small share of respondents attended to all attributes. Wine consumers made their choices based on very little information. WTP values are overestimated when ANA is neglected.

References

Alemu, M., Mørkbak, M., Olsen, S., & Jensen, C. (2013). Attending to the Reasons for Attribute Non-attendance in Choice Experiments. *Environmental Resource Economics*, *54*, 333–359. https://doi.org/10.1007/s10640-012-9597-8

Angulo, A. M., Gil, J. M., Gracia, A., & Sánchez, M. (2000). Hedonic prices for Spanish red quality wine. *British Food Journal*, *102*(7), 481–493. https://doi.org/10.1108/00070700010336445

Barreiro-Hurlé, J., Colombo, S., & Cantos-Villar, E. (2008). Is there a market for functional wines? Consumer preferences and willingness to pay for resveratrol-enriched red wine. *Food Quality and Preference*, *19*(4), 360–371. https://doi.org/10.1016/j.foodqual.2007.11.004

Gergaud, O., & Livat, F. (2007). How do consumers use signals to assess quality? In *American Association of Wine Economists* (Vol. 3).

Nelson, P. (1970). Information and Consumer Behavior. *Journal of Political Economy*, *78*(2), 311–329. https://doi.org/10.1086/259630

Tempesta, T., Arboretti, R., Corain, L., Salmaso, L., Tomasi, D., & Boatto, V. (2010). The importance of landscape in wine quality perception: An integrated approach using choice-based conjoint analysis and combination-based permutation tests. *Food Quality and Preference*, *21*(7), 827–836. https://doi.org/10.1016/j.foodqual.2010.04.007

Appendices

Appendix A

Questionnaire (e.g. for the North American Market)

Have you thought about how to choose a wine?

Section 1

Dear Sir/Madam,

We need your collaboration to fill in a survey for a research project being conducted by researchers from the University of Minho and the University of Trás-os-Montes and Alto Douro, in Portugal. The main goal is to analyze how consumers choose still wine and how they value its consumption.

Your answers are anonymous and confidential. Please respond as sincerely as possible. There are no right or wrong answers. It is very important that you complete the survey to the end.

Thank you for your collaboration!

Section 2

1. How old are you?
- □ Under 21 years
- □ Other, please indicate your age:_____

2. How often do you drink wine, on average?
- □ Daily
- □ A few times per week
- □ A few times per month
- □ Rarely and/or on special occasions
- □ Never/I am not a wine consumer

3. On average how many glasses of wine do you consume per week?
- □ 1 – 4
- □ 5 - 9
- □ 10 - 14
- □ 15 - 20
- □ More than 20

4. Are you the one who usually buys wine in your household?

□ Yes

□ No

5. How many bottles of wine do you usually purchase per month?

□ 1 – 3

□ 4 – 6

□ 7 – 9

□ 10 or more

6. Which range best represents the typical price you pay for a 750ml bottle of wine, for a formal consumption occasion (e.g. dinner with guests)?

□ Under 6,99$

□ 6$ - 9,99$

□ 10$ - 15,99$

□ 16$ - 19,99$

□ 20$ - 24,99$

□ Over 25 $

7. Which type of wine do you drink most often?

□ Red

□ White

□ Other

8. Where do you usually buy wine?

□ Local wine shop

□ Big wine chain store

□ Grocery store/supermarket

□ Online wine seller

□ Direct from the producer

□ Other:_____

9. How do you rate your wine knowledge profile?

□	□	□	□	□
New to wine	Know a little about wine	Somewhat knowledgeable about wine	Very knowledgeable about wine	Expert

10. Please, rate the following statements, from 1 = Completely disagree, 2 = Disagree, 3 = Partially disagree, 4 = Neither agree or disagree, 5 = Partially agree, 6 = Agree, 7 = Totally agree:

	1	2	3	4	5	6	7
I feel quite knowledgeable about wine.	○	○	○	○	○	○	○
Among my friends, I am one of the 'experts' on wine.	○	○	○	○	○	○	○
I rarely come across a wine that I have not heard of.	○	○	○	○	○	○	○
I know pretty much about wine.	○	○	○	○	○	○	○
I do not feel very knowledgeable about wine.	○	○	○	○	○	○	○
Compared to most other people, I know less about wine.	○	○	○	○	○	○	○
When it comes to wine, I really do not know a lot.	○	○	○	○	○	○	○
I have heard about most of the new wines that are around.	○	○	○	○	○	○	○

11. Please, rate the following statements, from 1 = Completely disagree, 2 = Disagree, 3 = Partially disagree, 4 = Neither agree or disagree, 5 = Partially agree, 6 = Agree, 7 = Totally agree:

	1	2	3	4	5	6	7
The wine I buy is important to me.	○	○	○	○	○	○	○
I am willing to spend over $25 on a bottle of wine.	○	○	○	○	○	○	○
I drink wine because I like the taste.	○	○	○	○	○	○	○
I am confident in my ability to choose wine in a liquor store.	○	○	○	○	○	○	○
I will still drink wine when price goes up.	○	○	○	○	○	○	○
The vintage of a wine is important to me.	○	○	○	○	○	○	○
I always match my food and wine.	○	○	○	○	○	○	○
I spend about one hour shopping for wine.	○	○	○	○	○	○	○
I often bring my own wine to restaurants, cafes and/or pubs.	○	○	○	○	○	○	○
I seek to further my wine knowledge.	○	○	○	○	○	○	○
I buy wine from wine specialty stores.	○	○	○	○	○	○	○
I enjoy aged wines and often cellar red wines for future consumption.	○	○	○	○	○	○	○
I have a collection (24 bottles or more) of wines I keep for future consumption.	○	○	○	○	○	○	○

12. How much money are you willing to pay for a bottle of red wine to consume during a formal consumption occasion (e.g. dinner with guests)?

_____

13. In general, for you and compared to other wines, the Portuguese wine is:

□ Poorer quality
□ Better quality
□ Similar
□ I do not know

Section 3

Imagine you are going to buy a 750ml bottle of a Controlled Designation of Origin wine, for a formal consumption occasion. Below we present you with 10 situations in which we ask you to choose between 3 alternative bottles of wine with different characteristics and prices. The characteristics are:

- Alcohol content: between 12%vol to 15%vol.

- Certified landscape (by UNESCO or Food and Agriculture Organization – FAO, etc.): reflects the environment in which the vines are cultivated. The alternatives are: a) wines from a non-certified landscape; b) wines produced in a place where the vineyard is located inside a certified landscape. This place may still be historically related to the traditional agricultural systems or old wine traditions, as a remarkable testimony of its culture and history.

- Awards/Medals: if the wine has a gold medal, or none.

- Grape variety: the wines may be Cabernet Sauvignon, Red Blend or Syrah, or the bottle of wine does not present that information.

- Country of origin.

- Gold medal or none.

- Price.

Imagine you are standing in front of a shelf of wine. You have 3 different bottles available. Which of the following wines would you buy to enjoy over a dinner with guests at home?

If none of these wines suit, please choose the option "I would purchase none of these".

Scenario 1

Wine 1



8,99$

Wine 2



12,99$

Wine 3



24,99$

Which wine do you prefer?

- ○ **Wine 1**
- ○ **Wine 2**
- ○ **Wine 3**
- ○ **I would purchase none of these**

Scenario 2

Wine 1



24,99$

Wine 2



8,99$

Wine 3



12,99$

Which wine do you prefer?

- ○ **Wine 1**
- ○ **Wine 2**
- ○ **Wine 3**
- ○ **I would purchase none of these**

Scenario 3

Wine 1



12,99$

Wine 2



17,99$

Wine 3



12,99$

Which wine do you prefer?

- ○ **Wine 1**
- ○ **Wine 2**
- ○ **Wine 3**
- ○ **I would purchase none of these**

Scenario 4

Wine 1



8,99$

Wine 2



24,99$

Wine 3



12,99$

Which wine do you prefer?

- ○ **Wine 1**
- ○ **Wine 2**
- ○ **Wine 3**
- ○ **I would purchase none of these**

Scenario 5

Wine 1



24,99$

Wine 2



12,99$

Wine 3



8,99$

Which wine do you prefer?

- ○ **Wine 1**
- ○ **Wine 2**
- ○ **Wine 3**
- ○ **I would purchase none of these**

Scenario 6

Wine 1



12,99$

Wine 2



8,99$

Wine 3



24,99$

Which wine do you prefer?

- ○ **Wine 1**
- ○ **Wine 2**
- ○ **Wine 3**
- ○ **I would purchase none of these**

Scenario 7

Wine 1	Wine 2	Wine 3
		
17,99$	8,99$	17,99$

Which wine do you prefer?

- ○ **Wine 1**
- ○ **Wine 2**
- ○ **Wine 3**
- ○ **I would purchase none of these**

Scenario 8

Wine 1



12,99$

Wine 2



12,99$

Wine 3



17,99$

Which wine do you prefer?

- ○ **Wine 1**
- ○ **Wine 2**
- ○ **Wine 3**
- ○ **I would purchase none of these**

Scenario 9

Wine 1	Wine 2	Wine 3
		
17,99$	17,99$	8,99$

Which wine do you prefer?

- ○ **Wine 1**
- ○ **Wine 2**
- ○ **Wine 3**
- ○ **I would purchase none of these**

Scenario 10

Wine 1



8,99$

Wine 2



24,99$

Wine 3



8,99$

Which wine do you prefer?

- ○ **Wine 1**
- ○ **Wine 2**
- ○ **Wine 3**
- ○ **I would purchase none of these**

Section 4

1. For you, the choices between the wines were:

□ Easy

□ Neither easy nor difficult

□ Difficult

2. Which of the following attributes did you ignore in your choices?

□ Country of origin

□ Awards/Medals

□ Alcohol content

□ Landscape

□ Grape variety

□ Price

□ I considered all attributes

3. Please, indicate which of the attributes did you consider 1=Always, 2= In most choices, 3=Seldom, 4=Never.

	1	2	3	4
Country of origin	○	○	○	○
Awards	○	○	○	○
Alcohol content	○	○	○	○
Landscape	○	○	○	○
Grape variety	○	○	○	○
Price	○	○	○	○

4. Please indicate the importance of other characteristics not included in the choice sets (from 1=Not all important, 2= Slightly important, 3= Moderately important, 4= Very important, 5=Extremely important)

	1	2	3	4	5
Region of origin	○	○	○	○	○
Brand	○	○	○	○	○
Vintage	○	○	○	○	○
Vines age	○	○	○	○	○
Producer size	○	○	○	○	○
Expert opinion/rating	○	○	○	○	○
Bottle shape	○	○	○	○	○

Section 5

1. In this section we ask you to respond to the following statements using a scale from 1 = Completely disagree, 2 = Disagree, 3 = Partially disagree, 4 = Neither agree or disagree, 5 = Partially agree, 6 = Agree, 7 = Totally agree:

	1	2	3	4	5	6	7
In general, a bottle of wine costs too much for me.	○	○	○	○	○	○	○
In general, I do not drink the wine after purchasing it.	○	○	○	○	○	○	○
In general, I feel that I've wasted money when I purchase wine.	○	○	○	○	○	○	○
The wine will not be very good for my health.	○	○	○	○	○	○	○
In general, I think the quality of wine will be poor.	○	○	○	○	○	○	○
The wine will be easily spoil.	○	○	○	○	○	○	○
I will not feel comfortable when drinking the wine.	○	○	○	○	○	○	○
I am worried about what others will think of me when I purchase wine.	○	○	○	○	○	○	○
I am worried that my friends might think if I bought the wrong wine.	○	○	○	○	○	○	○
I will not feel comfortable drinking the wine in public.	○	○	○	○	○	○	○

2. In this section we ask you to respond to the following statements using a scale from 1 = Completely disagree, 2 = Disagree, 3 = Partially disagree, 4 = Neither agree or disagree, 5 = Partially agree, 6 = Agree, 7 = Totally agree:

	1	2	3	4	5	6	7
Wine has more health enhancing properties than other alcohol.	○	○	○	○	○	○	○
Wine has better health properties than other alcoholic beverages.	○	○	○	○	○	○	○
Wine can reduce the risk of certain diseases.	○	○	○	○	○	○	○
I know how much alcohol is healthy to consume.	○	○	○	○	○	○	○
I think wine is a healthy alcoholic beverage.	○	○	○	○	○	○	○
I would drink more wine if it was healthy for me.	○	○	○	○	○	○	○
I read wine labels when I buy wine.	○	○	○	○	○	○	○
I would pay more for wine if it were health-enhanced.	○	○	○	○	○	○	○

Section 6

1. What is your gender?

Male ☐ Female ☐

2. What is your marital status?

Married ☐ Divorced ☐

Single ☐ Widowed ☐

3. What is the highest level of education you have completed?

Did not complete high school ☐ Bachelor´s degree ☐

High school ☐ Master´s degree ☐

College ☐ Doctoral degree ☐

4. Your personal behavior and decisions towards risk are (from 1 = nothing risky, 2 = Something risky, 3 = neutral, 4 = risky, 5 = completely risky):

	1	2	3	4	5
In job	○	○	○	○	○
In health	○	○	○	○	○
In finances	○	○	○	○	○
In life in general	○	○	○	○	○

5. What is your approximate household income per **month** (in $)?

_____

6. How many people are part of your household?

_____

Appendix B

Table 1 - Summary statistics of European sample descriptive variables

Variable	Sample	Variable	Sample
Socio-economic characteristics		**Consumption habits and buying behavior**	
Age groups (%)		Usual price paid (formal occasion) %	
21 – 30	15	Less than 1,99€	3
31 – 44	20	2€ - 4,99€	35
45 – 64	41	5€ - 9,99€	41
+ 65	24	10€ - 14,99€	14
Age (mean)	48.6	15€ - 19,99€	6
Gender (male=1; female=0)	0.53	More than 20 €	1
Education (%)		Wine consumption (%)	
High school	15	Daily	20
College	37	A few times per week	29
Bachelor's degree	33	A few times per month	29
Higher Education (Master, Doctoral degrees)	16	Rarely or on special occasions	23
Household size (mean)	2.3	Glasses of wine (per week) %	
Monthly income (thousand €)	3.55	Less than 3	48
		3 - 5	27
Wine knowledge (%)		+ 5	26
New to wine	19	Buy wine (1=yes; 0=no)	0.84
Know a little about wine	41	Number of bottles purchased (per month) %	
Somewhat knowledgeable about wine	32	Less than 3	48
Very knowledgeable about wine	7	3 - 5	29
Expert	1	6 - 10	15
		+ 10	8
		Maximum willingness to pay (formal occasion, average, €/bottle)	27.3

Appendix C

Annex to Chapter 3

Table 1 - DCE studies on ANA

Study	Main field	Number of Attributes	Alternatives	Choice sets/ (sample size)	Stated ANA		Inferred ANA		Implications of accommodating ANA
					Model	Treatment	Model	Treatment	
Hensher et al. (2005)	T	6	3 + SQ	16/(502)	MIXL	0-utility weights			Lower WTP, small model performance increase
Hensher (2006)	T	6	3 + SQ	16/(514)	MIXL	0-utility weights			VTTS decrease
Hensher, Rose,& Bertoia (2007)	T	5	2 + SQ	16/(514)	MIXL	0-utility weights			Higher VTTS
Puckett & Hensher (2008)	T	7	2 + SQ	4, 8 / (108 transporters, 102 shippers)	MIXL	0-utility weights			Lower VTTS
Puckett & Hensher (2009)	T	7	2 + SQ	4/(108)	ECLM	0-utility weights			Lower VTTS when accounting for ANA heterogeneity
Hensher & Rose (2009)	T	5	2 + SQ	16/(243)	MIXL	0-utility weights			Higher VTTS and model fit
Hensher & Greene (2010)	T	5	2 + SQ	16/(243)			LCM	0-utility weights	Higher VTTS
Hess & Hensher (2010)	T	5	2 + SQ	16/(205)	MNL, MIXL	Separate parameters by groups of attendance and non-attendance		Coefficient of variation for the conditional distribution	Higher model fit in the inferred approach
Rose et al. (2012)	T	4	3 + SQ	40, 20, 10 / (232)	MIXL	Excluding ignored attributes			Impacts on WTP
Hensher, Rose, & Greene (2012)	T	5	2 + SQ	16/(223)			ECLC	0-utility weights	Better model fit. Impacts on ANA depend on the sample
Hensher et al. (2013)	T	8	2 + SQ	16/(588)			RPLCM	0-utility weights and class heterogeneity	No VTTS differences

Study	Main field	Number of Attributes	Alternatives	Choice sets/ (sample size)	Stated ANA		Inferred ANA		Implications of accommodating ANA
					Model	Treatment	Model	Treatment	
Hess et al. (2013)	T	6, 5, 6	2 + SQ, 2 + SQ, 3	10, 8, 10 / (368, 996, 915)			ECLC, LCM, RPLCM	0-utility weights, Continuous distributed coefficients, Point value	Higher model fit for LC-MIXL
Greene & Hensher (2013)	T	7	2 + SQ	4/(108)			RPLCM		Improve model fit, Lower WTP
Collins et al. (2013)	T	4	3 + SQ + forced choice	40, 20, 10 / (213)			RPANA	Stated ANA as a covariate	Better model fit
Campbell et al. (2008)	E	5	2 + SQ	6/(434)	ECLM	0-utility weights, parametrization of scale parameter based on processing strategies			Better model fit and lower WTP
Scarpa et al. (2009)	E	5	2 + SQ	5, 6 /(n.a.)			ECLC, STAS	0-utility weights	Better model fit, lower WTP
Meyerhoff & Liebe (2009)	E	5	2 + SQ	36 = 6 blocks/ (621)	ECLM	0-utility weights			Small impact on model performance, similar WTP
Scarpa et al., (2010)	E	10	2 + SQ	12/(480)	MNL	Individual level and choice set ANA, 0-utility			Better model fit, higher WTP for serial ANA, lower WTP for choice task
Carlsson et al. (2010)	E	5	2 + SQ	6/(955)	MIXL	0-utility weights			No significant differences in WTP
Campbell et al. (2011)	E	5	2 + SQ	5, 6/(n.a.)			Scale-adjusted ECLC	0-utility weights	Lower WTP
Colombo et al. (2013)	E	8	2 + SQ	5/(441)	MIXL	Separate parameters by groups of attendance and non-attendance			WTP effect depends on the assumptions about ANA
Kosenius (2013)	E	5	3	36 = 6 blocks/ (307)	ECLM	0-utility weights			More efficient WTP
Kragt (2013)	E	4	2	4/(712)	MIXL	0-utility weights	ECLC	0-utility weights	Inferred approach presents a better fit, and lower WTP. No

Study	Main field	Number of Attributes	Alternatives	Choice sets/ (sample size)	Stated ANA		Inferred ANA		Implications of accommodating ANA
					Model	Treatment	Model	Treatment	
									differences in the WTP for the MIXL
Alemu et al. (2013)	E	7	2 + SQ	18 = 6 blocks/(968)	ECLM	0-utility weights	HH	Coefficient of variation for the conditional distribution	WTP decrease
Colombo & Glenk (2014)	E	5	2 + SQ	32 = 8 blocks (370)			LCM	0-utility weights	Improves models' performance, and affects WTP
Glenk et al. (2015)	E	3	2 + SQ	24/(748)			ECLC	0-utility weights	Inferred approach improves model fit; Impacts on WTP depending on the sample
Thiene et al. (2015)	E	5	3 + SQ	36 = 4 blocks/ (208)			LCM	0-utility weights	WTP decrease
Sandorf et al. (2017)	E	4	2 + SQ	Not reported /(397)			RPLCM	0-utility weights	WTP decrease
Jourdain & Vivithkeyoonvong (2017)	E	5	2 + SQ	12 = 2 blocks/(305)			ECLC	0-utility weights	Lower WTP; lower probabilities
Hole (2011)	H	6	2	8/(425)			EAA		Model fit improved and WTP decrease
Hole et al. (2013)	H	5	2	16/(571)	MIXL	0-utility weights	EAA/MEAA	0-utility weights	Better model fit, but similar WTP
Lagarde (2013)	H	6	2	16/(134)			ECLC	0-utility weights	Better model fit and higher WTP
Erdem et al. (2014)	H	6	2 + SQ	12/(594)			LCM	Including a vector of dummy variables	Gains in model fit with ALNA; WTP decrease
Hole et al. (2016)	H	5	3	15/(1027)			EAA, ECLC	0-utility weights	Improves model fit

Study	Main field	Number of Attributes	Alternatives	Choice sets/ (sample size)	Stated ANA		Inferred ANA		Implications of accommodating ANA
					Model	Treatment	Model	Treatment	
Heidenreich (2017)	H	5	2 + opt-out	32/(451)			LCM	0-utility weights	Overestimated welfare effects when ANA is a preference
Balcombe et al. (2011)	F	4	3	8/(1106)	MIXL	Unrestricted and restricted interactions. 0-utility weights			Better model fit
Scarpa et al. (2013)	F	7	2 + No buy	12/(295)	MIXL, MNL	Separate parameters by groups of attendance and non-attendance	ECLC	0-utility weights	Better fit in the ECLCM and lower WTP
Kehlbacher et al. (2013)	F	4	3 + don't know	6/(301)	MIXL	0-utility weights; Shrinking the means and variances of marginal utilities = 0			Improve model fit, and provides lower WTP
Thiene et al. (2013)	F	4	revealed and stated data	5 / (487 revealed + 1948 stated)			LCM	0-utility weights	Reduces differences between revealed and stated preferences
Balcombe et al. (2014)	F	7	4 + SQ	6/(341, 318)	MIXL, CA	Ranking as covariates, Rankings scaling parameters			Different WTP signs depend on the approach, better model performance
Chalak et al. (2016)	F	5	4 + opt-out	12/(597)	MIXL	Combine non-attendance with ranking data			
Ortega & Ward (2016)	F	5	3 + SQ	36 = 4 blocks (563 men + 560 women)	MIXL	0-utility weights	HH	Coefficient of variation for the conditional distribution	Mixed model fit results for both approaches; Inferred ANA show higher effect on men's WTP Lower general WTP
Caputo et al. (2018)	F	5	2 + SQ	16 = 2 blocks/(344, 257)	RP and EC	0-utility weights	ECLC	0-utility weights	WTP differ from serial-level and choice task-level ANA

Note: MIXL=Mixed logit model; LCM=Latent class model; RPLCM=Random parameter latent class model; EAA=Endogenous attribute attendance; MEAA=Mixed EAA; ECLC=Equality constrained latent class; ECLM=Error component logit model, CA=Contraction approach

Appendix D

Annex to Chapter 4

Table 1 - Summary statistics of descriptive variables

Variable	North America	France	Portugal
Socio-economic characteristics			
Age groups (%)			
21 - 30	26	16	13
31 - 44	70	12	27
45 - 64	73	53	38
+ 65	31	21	23
Age (mean)	42.7	45.6	43.3
Gender (male=1; female=0)	0.49	0.56	0.48
Education (%)			
High school	26	12	14
College	34	42	35
Bachelor's degree	11	30	37
Higher Education (Master, Doctoral degrees)	29	17	15
Household size (mean)	2.5	2.3	2.5
Consumption habits and buying behavior			
Usual price paid (formal occasion) %			
Under 6,99$	5	5	3
6$ - 9,99$	13	28	50
10$ - 15,99$	46	45	36
16$ - 19,99$	25	16	8
20$ - 24,99$	8	5	4
Over 25 $	5	2	1
Wine consumption (%)			
Daily	18	27.5	25
A few times per week	36	28.5	33
A few times per month	24	25.5	25
Rarely or on special occasions	23	18.5	19
Glasses of wine (per week) %			
Less than 3	54	46	37
3 - 5	23	24	31
6 - 10	16	16	18
+ 10	7	13	15
Buy wine (1=yes; 0=no)	0.89	0.84	0.82
Number of bottles purchased (per month) %			
Less than 3	58	50	40

Variable	North America	France	Portugal
3 - 5	24	32	30
6 - 10	13	12	20
+ 10	5	8	11
Market channel used to buy wine %			
Wine shop	62	27	20
Supermarket	53	83	95
Producer	5	21	12
On-line	9	7	2
Perceptual variables			
Wine knowledge (%)			
New to wine	17	18.5	23
Know a little about wine	43	46.5	51
Somewhat knowledgeable about wine	31	26	25
Very knowledgeable about wine	9	8	1
Expert	1	1	1
Easiness of choice (%)			
Difficult	10	13.5	16
Easy	44.8	36.5	32
Neither easy nor difficult	45.2	50	53

Table 2 – Attitudinal and perceptual variables towards wine

Variable (1=Completely disagree; 7=Totally agree)	**North America**	**France**	**Portugal**
Knowledge perception			
I feel quite knowledgeable about wine.	3.8	3.6	3.1
Among my friends, I am one of the 'experts' on wine.	3.4	3.2	2.9
I rarely come across a wine that I have not heard of.	3.1	3.5	3.1
I know pretty much about wine.	3.8	3.4	2.7
I do not feel very knowledgeable about wine.	4.2	4	4.3
Compared to most other people, I know less about wine.	4.3	3.8	3.7
When it comes to wine, I really do not know a lot.	4.4	3.8	4.2
I have heard about most of the new wines that are around.	3.4	3.4	3.2
Wine involvement			
The wine I buy is important to me.	4.6	4.3	4.3
I am willing to spend over $25 on a bottle of wine.	4	3.5	2.6
I drink wine because I like the taste.	5.3	4.9	5.3
I am confident in my ability to choose wine in a liquor store.	4.5	3.8	4
I will still drink wine when price goes up.	4.6	4.3	4.3
The vintage of a wine is important to me.	3.8	4	3.5
I always match my food and wine.	3.9	3.1	2.6
I spend about one hour shopping for wine.	3.1	4	3.7
I often bring my own wine to restaurants, cafes and/or pubs.	2.6	4	4.5
I seek to further my wine knowledge.	4.4	1.9	1.3
I buy wine from wine specialty stores.	3.8	2.5	2.3
I enjoy aged wines and often cellar red wines for future consumption.	3.7	4.7	4.4
I have a collection (24 bottles or more) of wines I keep for future consumption.	2.8	3.4	2.6
Perceived risk			
In general, a bottle of wine costs too much for me.	3	3.2	2.3
In general, I do not drink the wine after purchasing it.	2.7	2.8	1.9
In general, I feel that I've wasted money when I purchase wine.	2.6	2.3	1.9
The wine will not be very good for my health.	2.7	3.2	2.3
In general, I think the quality of wine will be poor.	2.8	2.7	2.1
The wine will be easily spoil.	2.8	3.5	2.3
I will not feel comfortable when drinking the wine.	2.4	2.7	1.9
I am worried about what others will think of me when I purchase wine.	2.5	3.2	1.8
I am worried that my friends might think if I bought the wrong wine.	2.7	3.7	2.4
I will not feel comfortable drinking the wine in public.	2.8	2.7	1.7

Variable (1=Completely disagree; 7=Totally agree)	**North America**	**France**	**Portugal**
Health perception			
Wine has more health enhancing properties than other alcohol.	5	4.7	5.2
Wine has better health properties than other alcoholic beverages.	5.1	4.7	5.3
Wine can reduce the risk of certain diseases.	4.8	4.4	4.9
I know how much alcohol is healthy to consume.	4.7	4.1	5.1
I think wine is a healthy alcoholic beverage.	4.3	4.3	4.7
I would drink more wine if it was healthy for me.	4.8	4	4
I read wine labels when I buy wine.	4.9	5	5
I would pay more for wine if it were health-enhanced.	4	4.3	4.4

Table 3 - Self-reported attribute non-attendance for all ANA combinations

	North America		France		Portugal	
ANA pattern	**Number of respondents**	**Proportion of total**	**Number of respondents**	**Proportion of total**	**Number of respondents**	**Proportion of total**
Full attendance	**102**	**0.255**	**49**	**0.245**	**31**	**0.155**
Ignoring 1 attribute	**121**	**0.303**	**74**	**0.37**	**71**	**0.355**
Country of origin	16	0.040	14	0.07	15	0.075
Medals	21	0.053	4	0.02	8	0.04
Alcohol level	23	0.058	24	0.12	13	0.065
Landscape	27	0.068	12	0.06	15	0.075
Grape variety	9	0.023	4	0.02	12	0.06
Price	25	0.063	16	0.08	8	0.04
Ignoring 2 attributes	**81**	**0.203**	**44**	**0.22**	**54**	**0.27**
COO+medals	2	0.005	3	0.015	6	0.03
COO+alcohol	7	0.018	0	0.0	0	0.0
COO+landscape	4	0.010	2	0.01	3	0.015
COO+grape	3	0.008	3	0.015	0	0.0
COO+price	7	0.018	3	0.015	7	0.035
Medals+alcohol	7	0.018	2	0.01	3	0.015
Medals+landscape	8	0.020	3	0.015	4	0.02
Medals+grape	3	0.008	4	0.02	2	0.01
Medals+price	4	0.010	5	0.025	3	0.015
Alcohol+landscape	13	0.033	7	0.035	9	0.045
Alcohol+grape	3	0.008	4	0.02	6	0.03
Alcohol+price	6	0.015	3	0.015	3	0.015
Landscape+grape	8	0.020	2	0.01	4	0.02
Landscape+price	2	0.005	2	0.01	2	0.01
Grape+price	4	0.010	1	0.005	2	0.01
Ignoring 3	**56**	**0.140**	**22**	**0.11**	**31**	**0.155**
COO+medals+alcohol	2	0.005	1	0.005	0	0.0
COO+medals+landscape	3	0.008	1	0.005	2	0.01
COO+medals+grape	1	0.003	0	0.1	0	0.0
COO+medals+price	5	0.013	3	0.015	1	0.005
Medals+alcohol+landscape	10	0.025	3	0.015	3	0.015
Medals+alcohol+grape	4	0.010	0	0.0	2	0.01

ANA pattern	North America		France		Portugal	
	Number of respondents	Proportion of total	Number of respondents	Proportion of total	Number of respondents	Proportion of total
Medals+alcohol+price	3	0.008	1	0.005	2	0.01
Alcohol+landscape+grape	5	0.013	3	0.015	6	0.03
Alcohol+landscape+price	1	0.003	0	0.0	3	0.015
Landscape+grape+price	0	0.000	0	0.0	0	0.0
COO+alcohol+landscape	1	0.003	0	0.0	0	0.0
COO+alcohol+grape	2	0.005	0	0.0	1	0.005
COO+alcohol+price	4	0.010	1	0.005	3	0.015
COO+landscape+grape	1	0.003	2	0.01	4	0.02
COO+landscape+price	1	0.003	1	0.005	1	0.005
COO+grape+price	6	0.015	4	0.02	1	0.005
Medals+landscape+grape	2	0.005	0	0.0	1	0.005
Medals+landscape+price	1	0.003	0	0.0	1	0.005
Alcohol+grape+price	0	0.000	1	0.005	0	0.0
Medals+grape+price	4	0.010	1	0.005	0	0.0
Ignoring 4	**25**	**0.063**	**7**	**0.035**	**10**	**0.05**
COO+medals+alcohol+landscape	3	0.008	1	0.005	0	0.0
COO+medals+alcohol+grape	2	0.005	0	0.0	0	0.0
COO+medals+alcoholprice	4	0.010	0	0.0	1	0.005
Medals+alcohol+landscape+grape	4	0.010	4	0.02	4	0.02
Medals+alcohol+landscape+price	1	0.003	0	0.0	0	0.0
Alcohol+landscape+grape+price	1	0.003	0	0.0	0	0.0
Alcohol+COO+grape+price	4	0.010	0	0.0	1	0.005
Landscape+COO+grape+price	0	0.000	1	0.005	0	0.0
Medals+landscape+COO+grape	3	0.008	0	0.0	1	0.005
Landscape+alcohol+grape+COO	1	0.003	1	0.005	2	0.01
Medals+alcohol+grape+price	0	0.000	0	0.0	0	0.0
Medals+landscape+grape+price	0	0.000	0	0.0	1	0.005
Landscape+alcohol+COO+price	0	0.000	0	0.0	0	0.0
Medals+COO+grape+price	2	0.005	0	0.0	0	0.0
Medals+landscape+COO+price	0	0.000	0	0.0	0	0.0
Lexicographic choices	**13**	**0.033**	**4**	**0.02**	**3**	**0.015**
COO	1	0.003	1	0.005	1	0.005

	North America		France		Portugal	
ANA pattern	**Number of respondents**	**Proportion of total**	**Number of respondents**	**Proportion of total**	**Number of respondents**	**Proportion of total**
Medals	0	0.0	0	0.0	0	0.0
Alcohol	1	0.003	0	0.0	0	0.0
Landscape	2	0.005	1	0.005	0	0.0
Grape	0	0.0	0	0.0	0	0.0
Price	9	0.023	2	0.01	2	0.01
Ignoring all attributes	**2**	**0.005**	**0**	**0.0**	**0**	**0.0**
TOTAL	**400**	**1**	**200**	**1**	**200**	**1**

Appendix E

Annex to Chapter 5

Table 1 - ANA patterns tested in each model and average class membership probabilities

Class	Coefficient estimated in the class[b]						ECLC-ANA1	ECLC-ANA2	ECLC-ANA3	ECLC-ANA4	ECLC-ANA5	ECLC-ANA6
1	β_{aw}	β_{land}	β_{alc}	β_{coo}	β_{grape}	β_{price}	0	-	-	-	-	-
2	β_{aw}	β_{land}	β_{alc}	β_{coo}	β_{grape}	0	36.4	0	-	-	-	-
3	β_{aw}	β_{land}	β_{alc}	β_{coo}	0	β_{price}	3.3	8.6	9.5	9.5	10.5	10.5
4	β_{aw}	β_{land}	β_{alc}	0	β_{grape}	β_{price}	30.9	0	-	-	-	-
5	β_{aw}	β_{land}	0	β_{coo}	β_{grape}	β_{price}	20.15	0.3	0	-	-	-
6	β_{aw}	0	β_{alc}	β_{coo}	β_{grape}	β_{price}	0	-	-	-	-	-
7	0	β_{land}	β_{alc}	β_{coo}	β_{grape}	β_{price}	9.2	0	-	-	-	-
					2 attributes							
8	β_{aw}	β_{land}	β_{alc}	β_{coo}	0	0		16.88	12.6	12.9	11.7	11.6
9	β_{aw}	β_{land}	β_{alc}	0	β_{grape}	0		25.9	0	-	-	-
10	β_{aw}	β_{land}	0	β_{coo}	β_{grape}	0		7.8	0	-	-	-
11	β_{aw}	0	β_{alc}	β_{coo}	β_{grape}	0		-	-	-	-	-
12	0	β_{land}	β_{alc}	β_{coo}	β_{grape}	0		0	-	-	-	-
13	β_{aw}	β_{land}	β_{alc}	0	0	β_{price}		15.4	0	-	-	-
14	β_{aw}	β_{land}	0	β_{coo}	0	β_{price}		2.6	0.1	0	-	-
15	β_{aw}	0	β_{alc}	β_{coo}	0	β_{price}		-	-	-	-	-
16	0	β_{land}	β_{alc}	β_{coo}	0	β_{price}		0	-	-	-	-
17	β_{aw}	β_{land}	0	0	β_{grape}	β_{price}		8.5	0	-	-	-
18	β_{aw}	0	β_{alc}	0	β_{grape}	β_{price}		-	-	-	-	-
19	0	β_{land}	β_{alc}	0	β_{grape}	β_{price}		8.1	6.8	7	5.6	5.5
20	β_{aw}	0	0	β_{coo}	β_{grape}	β_{price}		-	-	-	-	-
21	0	β_{land}	0	β_{coo}	β_{grape}	β_{price}		5.8	3.9	4.7	4.9	4.9
22	0	0	β_{alc}	β_{coo}	β_{grape}	β_{price}		-	-	-	-	-
					3 attributes							
23	β_{aw}	β_{land}	β_{alc}	0	0	0			0	-	-	-
24	β_{aw}	β_{land}	0	β_{coo}	0	0			0	-	-	-
25	β_{aw}	0	β_{alc}	β_{coo}	0	0			0	-	-	-
26	0	β_{land}	β_{alc}	β_{coo}	0	0			-	-	-	-

Class	Coefficient estimated in the class[b]						ECLC-ANA1	ECLC-ANA2	ECLC-ANA3	ECLC-ANA4	ECLC-ANA5	ECLC-ANA6
27	β_{aw}	β_{land}	0	0	β_{grape}	0			26.7	28.6	2.1	-
28	β_{aw}	0	β_{alc}	0	β_{grape}	0			5.2	3.4	1.6	-
29	0	β_{land}	β_{alc}	0	β_{grape}	0			-	-	-	-
30	β_{aw}	0	0	β_{coo}	β_{grape}	0			9.2	2.6	0.6	-
31	0	β_{land}	0	β_{coo}	β_{grape}	0			-	-	-	-
32	0	0	β_{alc}	β_{coo}	β_{grape}	0			-	-	-	-
33	β_{aw}	β_{land}	0	0	0	β_{price}			0	-	-	-
34	β_{aw}	0	β_{alc}	0	0	β_{price}			10.5	9.7	11.6	11.7
35	0	β_{land}	β_{alc}	0	0	β_{price}			-	-	-	-
36	β_{aw}	0	0	β_{coo}	0	β_{price}			4.3	3.6	3.2	-
37	0	β_{land}	0	β_{coo}	0	β_{price}			-	-	-	-
38	0	0	β_{alc}	β_{coo}	0	β_{price}			-	-	-	-
39	β_{aw}	0	0	0	β_{grape}	β_{price}			11.1	10.7	5.3	5.3
40	0	β_{land}	0	0	β_{grape}	β_{price}			-	-	-	-
41	0	0	β_{alc}	0	β_{grape}	β_{price}			-	-	-	-
42	0	0	0	β_{coo}	β_{grape}	β_{price}			-	-	-	-
	4 attributes											
43	β_{aw}	β_{land}	0	0	0	0				-	-	-
44	β_{aw}	0	β_{alc}	0	0	0				-	-	-
45	0	β_{land}	β_{alc}	0	0	0				-	-	-
46	β_{aw}	0	0	β_{coo}	0	0				-	-	-
47	0	β_{land}	0	β_{coo}	0	0				-	-	-
48	0	0	β_{alc}	β_{coo}	0	0				-	-	-
49	β_{aw}	0	0	0	β_{grape}	0				-	-	-
50	0	β_{land}	0	0	β_{grape}	0				-	-	-
51	0	0	β_{alc}	0	β_{grape}	0				-	-	-
52	0	0	0	β_{coo}	β_{grape}	0				7.6	3.9	3.8
53	β_{aw}	0	0	0	0	β_{price}				-	-	-
54	0	β_{land}	0	0	0	β_{price}				-	-	-
55	0	0	β_{alc}	0	0	β_{price}				-	-	-
56	0	0	0	β_{coo}	0	β_{price}				0	-	-

Class	Coefficient estimated in the class[b]						ECLC-ANA1	ECLC-ANA2	ECLC-ANA3	ECLC-ANA4	ECLC-ANA5	ECLC-ANA6
57	0	0	0	0	β_{grape}	β_{price}				-	-	-
5 attributes												
58	β_{aw}	0	0	0	0	0					15.9	15.5
59	0	β_{land}	0	0	0	0					9.2	-
60	0	0	β_{alc}	0	0	0					0	-
61	0	0	0	β_{coo}	0	0					5.8	5.9
62	0	0	0	0	β_{grape}	0					3	-
63	0	0	0	0	0	β_{price}					5.1	5.2
Full ANA (random choices)												
64	0	0	0	0	0	0						-

[b] For simplification and to reduce computational burden, we explore all ANA patterns with a general coefficient for COO and grape variety attributes. After finding the relevant patterns, we estimate all models including the dummies for these attributes' levels. As usual, we always include the Full Attendance pattern, even if it appears to be as irrelevant.

Note: 0=ignored attribute

Table 2 - Nlogit output for predicted choices

ECLC-2^K

```
+---------------------------------------------------------+
| Cross tabulation of actual y(ij) vs. predicted y(ij)  |
| Row indicator is actual, column is predicted.         |
| Predicted total is N(k,j,i)=Sum(i=1,...,N) Y(k,j,i).  |
| Predicted y(ij)=1 is the j with largest probability.  |
+---------------------------------------------------------+

--------+-------------------------------------------------------------------
NLOGIT Cross Tabulation for 4 outcome Multinomial Choice Model
XTab_Frq|          ALT1          ALT2          ALT3          NONE         Total
--------+-------------------------------------------------------------------
    ALT1|       1110.00       392.000       89.0000       .000000       1591.00
    ALT2|       250.000       1305.00       129.000       .000000       1684.00
    ALT3|       532.000       133.000       1089.00       .000000       1754.00
    NONE|       .000000       .000000       .000000       971.000       971.000
   Total|       1892.00       1830.00       1307.00       971.000       6000.00
```

$$\text{Correctly predicted} = \sum \frac{diag}{total} = \frac{1110 + 1305 + 1089 + 971}{6000} = 0.7458$$

ECLC-FINAL

```
+---------------------------------------------------------+
| Cross tabulation of actual y(ij) vs. predicted y(ij)  |
| Row indicator is actual, column is predicted.         |
| Predicted total is N(k,j,i)=Sum(i=1,...,N) Y(k,j,i).  |
| Predicted y(ij)=1 is the j with largest probability.  |
+---------------------------------------------------------+

--------+-------------------------------------------------------------------
NLOGIT Cross Tabulation for 4 outcome Multinomial Choice Model
XTab_Frq|          ALT1          ALT2          ALT3          NONE         Total
--------+-------------------------------------------------------------------
    ALT1|       1257.00       245.000       89.0000       .000000       1591.00
    ALT2|       61.0000       1494.00       129.000       .000000       1684.00
    ALT3|       306.000       .000000       1448.00       .000000       1754.00
    NONE|       .000000       .000000       .000000       971.000       971.000
   Total|       1624.00       1739.00       1666.00       971.000       6000.00
```

$$\text{Correctly predicted} = \sum \frac{diag}{total} = \frac{1257 + 1494 + 1448 + 971}{6000} = 0.8617$$