Universitat Politècnica de Catalunya
Institut de Sostenibilitat

Doctoral Thesis:

"Consumer Preference Heterogeneity Towards Olive Oil Virgin Extra: Hypothetical and Non-hypothetical Choice Experiments"

By

AHMED YANGUI

Under the supervision of

Thesis director: Professor José María Gil
Co-director: Dr. Montserrat Costa-Font

PhD Program: Sustainability
Main Subject: Agricultural Economics

Barcelona, 2014

﴿ بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ ﴾

﴿ وَمَا تَشَاءُونَ إِلَّا أَنْ يَشَاءَ اللَّهُ إِنَّ اللَّهَ كَانَ عَلِيمًا حَكِيمًا ﴾

Table of contents

List of Tables

List of Figures

List of Appendix

ABSTRACT

The present dissertation aim a contributing to both agricultural economics and marketing literature by addressing specific issues related to discrete choice models and choice experiments. More precisely, this thesis focuses on two main issues: 1) new tools to tackle with preference heterogeneity; and 2) new response formats to allows researchers to take into account the information provided by no chosen profiles. These two issues have generated three studies, which form the main core of this thesis: two are related to issue 1) (Chapters 2 and 3), while the third one is related to the issue 2) (Chapter 4).

In the first one, we evaluate consumers' preferences heterogeneity using a methodological framework with two novelties over past studies: 1) it accounts for both preference heterogeneity around the mean and the variance of random parameters; and 2) it considers both socio-demographic characteristics of consumers as well as their attitudinal factors. Estimated coefficients and moments of Willingness-to-Pay (WTP) distributions are compared with those obtained from alternative Random Parameter Logit (RPL) models. Results suggest that the proposed framework significantly increase the good-of-Fit and provides more useful insights for policy analysis. The most important attribute affecting consumers' preferences towards extra virgin olive oil are the price and the product's origin. The consumers perceive the organic olive oil attribute negatively, as they think that it is not worth paying a premium for a product that is healthy in nature.

The second paper aims at investigating the role of psychological factors in consumer's purchase decision process. The paper hypothesizes that differences in consumers' personality traits, such as food-related personality traits, purchasing habits and lifestyles, affect consumers' preferences for extra virgin olive oil. The methodological framework is based on the specification of an extended hybrid choice model (HCM), which was estimated following a two-step procedure. In the first step, a structural equation model was estimated to test hierarchical relationships between latent variables to explain purchasing intentions towards an organic olive oil. In the second step, the resulting latent variables were introduced in a random parameter logit (RPL) model to investigate the main determinants of consumers' choices related to extra virgin olive oil. The results from this study reinforce the need to include the psychological characteristics of consumers to better explain how individuals make food choices and to better understand the decision maker's process. Interestingly, Catalan consumers perceive a disutility from the organic attribute compared to other production system alternatives (conventional and PDO), while subjective norms and a higher perception

of behavioural control only partially mitigate this effect. Environmental or health concerns seem to not be relevant to consumers' choices related to organic olive oil as the conventional olive oil is already perceived as a healthy product per se.

In the third paper, we compares the ability of hypothetical and non-hypothetical choice experiment respect to incentive compatible ranking conjoint analysis and incentive compatible sequential best worst scaling. The comparison done in terms of estimated partworths, internal and external predictive power, estimated WTP, and participants' response consistency. In general, the results reveal higher preferences regularity between the respondents across the different treatments implying not statistically difference in the marginal participants' WTP. Additionally, the participants behave similarly whether there are asked to choose or to state their most preferred through the two ranking elicitation mechanism. The best worst scaling (BWS) format has been revealed to outperform the other formats in terms of predictive power as its cognitive process seems to better fits the natural tendency of humans at identifying the extreme values.

RESUMEN

Esta tesis doctoral trata de realizar algunas contribuciones relevantes para el campo de la Economía y el Marketing Agroalimentario en el diseño de los experimentos de elección y en la estimación de los modelos de elección discreta. Concretamente, la tesis aborda dos cuestiones principales: 1) la utilización de nuevas herramientas de estimación que tengan en cuenta la heterogeneidad de las preferencias de los consumidores; 2) el diseño de nuevos formatos de respuesta en los que se tenga en cuenta la información proporcionada por las opciones de productos no elegidas en los conjuntos de elección. Estas dos cuestiones se han abordado en los tres capítulos centrales en los que se estructura la tesis. Los dos primeros capítulos están relacionados con la heterogeneidad de las preferencias (capítulo 2 y 3), mientras que el tercero está relacionado con la segunda cuestión (capítulo 4).

En el Capítulo 2, el análisis de la heterogeneidad de las preferencias se ha basado en la utilización de dos herramientas novedosas en relación a la literatura existente: 1) la consideración explícita de la heterogeneidad de la preferencia en torno tanto a la media como a la varianza de los parámetros aleatorios; y 2) la inclusión no sólo de las características socio-demográficas de los consumidores sino también de las lexicográficas. Los coeficientes estimados y los momentos de la distribución de la Disposición a Pagar (DAP) han sido comparadas con las obtenidas en el modelo más utilizado en la literatura, el modelo Logit de parámetros aleatorios (RPL). Los resultados indican que el marco metodológico propuesto permite incrementar significativamente la bondad del ajuste, por lo que se obtienen estimaciones más precisas sobre las que orientar las posibles estrategias de marketing. Desde el punto de vista aplicado, el precio y el origen del producto son los atributos que más influyen en las preferencias de los consumidores hacia el aceite de oliva virgen extra. Los consumidores perciben negativamente el atributo oliva ecológico ya que piensa que no merece la pena pagar un sobreprecio para un producto que ya es considerado como sano en sí mismo.

El Capítulo 3 se dedica especialmente a una cuestión que en el capítulo anterior sólo se consideró parcialmente, y es el efecto de las variables psicográficas y las actitudes en el proceso de toma de decisiones de compra por parte de los consumidores. La hipótesis de partida es que los rasgos de la personalidad de los consumidores así como sus hábitos de compra y estilos de vida, determinan sus preferencias; en este caso, hacia el aceite de oliva virgen extra. El marco metodológico adoptado se basa en la especificación de un modelo de elección híbrida extendido (HCM), que se estimó siguiendo un procedimiento a dos etapas.

En la primera etapa, se estimó un modelo de ecuaciones estructurales para determinar las relaciones jerárquicas entre las variables latentes con el objetivo de explicar las intenciones de compra de los consumidores hacia el aceite de oliva ecológico. En la segunda etapa, las variables latentes resultantes fueron introducidas en un modelo Logit de parámetros aleatorios (RPL) con el fin de estudiar los factores determinantes de la elección del aceite de oliva virgen extra por parte de los consumidores. Los resultados de este Capítulo refuerzan la idea de la necesidad de tener en cuentas factores de personalidad y actitudes a la hora de analizar las preferencias de los consumidores y sus procesos de toma de decisiones. Consistente con el capítulo anterior, el atributo ecológico no es relevante a la hora de elegir un aceite de oliva en relación a otros sistemas de producción (convencional y Denominación de Origen Protegida), mientras que las normas subjetivas y una mayor percepción de control del comportamiento sólo atenúan parcialmente este efecto. Las preocupaciones ambientales o de salud parecen no tener ningún impacto relevante en las elecciones de los consumidores del aceite de oliva ecológico.

En el Capítulo 4 se hace una revisión de los diferentes formatos de respuesta que se han venido utilizando en la literatura en el análisis conjunto, prestando atención especial a los formatos de ranking (RCA) y de escalas mejor/peor (BWS). En un contexto no hipotético se comparan estos dos formatos con el experimento de elección tradicional, tomando como referencia el experimento tradicional hipotético. La comparación se hace teniendo en cuenta los parámetros estimados, la capacidad predictiva tanto interna como externa, la consistencia de las respuestas de los participantes y los valores calculados de la Disposición a Pagar. En general, los resultados revelan una cierta consistencia de los resultados, al menos en términos de la Disposición a Pagar. Asimismo, los dos métodos basados en un ranking generan resultados similares, tanto si se considera únicamente la opción preferida como la totalidad del ranking. Sin embargo, globalmente, el método de escalas mejor/peor (BWS) parece comportarse mejor en relación al resto de formatos en términos de capacidad predictiva, lo que parece confirmar que este tipo de formato se adapta mucho mejor a la tendencia natural de los seres humanos a la identificación de los valores extremos.

*I would like to extend my gratitude to my parents for providing continued support and encouragement for each new adventure I decide to undertake. I attribute much of my success to my father **Abdel Kader** and my mother **Mounira**, as they instilled a strong work ethic. Thank you for telling me what I'm capable of. For giving me the support that I needed to build a dream to chase after. And for believing that I have the talent to reach my goals. Through your guidance, I learned that there are no limits to what I can accomplish.*

My dear wife, you don't have an idea how thankful I am of having you in my life. Choosing you as my wife is the best decision that I've ever made. You are such an amazing wife. I simply thank ALLAH for giving me such a good and perfect partner in my life. It is your love that made me changed. Thank you for all you've done for me and thank you for loving me. I hope that in this dissertation you can find a minimum sentiment of thankfulness. I love you.

Acknowledgment

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First of all, I humbly dedicate this piece of work to ALLAH. Because of you, this academic journey has not only been made possible but was also fruitful. Thank you for gives me all valuable things that I have and guides my way, not only during the realization of this thesis but during my whole life.

I would like to sincerely express my great gratitude and special appreciation to my supervisors José M Gil and Montserrat Cost-Font for their valuable advice, guidance, wonderful encouragement and patience throughout the research and thesis preparation without forgetting their good human relationships.

Thanks also to the Center for Agro-food Economy and Development (CREDA-UPC-IRTA) for support financially and personally.

I would like also to thank La Agencia Española de Cooperación Internacional para el Desarrollo (AECID) and the Mediterranean Agronomic Institute of Zaragoza (IAMZ) for the pre-doctoral scholarship.

Special thanks to my friend Dr. Faiçal Akaichi from Scotland's Rural College, who greatly enriched my knowledge and who help me a lot in the fulfilling of this thesis. From deep inside thank you for your support and help.

I would like to thank all the persons in the CREDA group for their nice treatment and friendship during my stay, many thanks to all of you with a special thanks to BouAli, Amr, Islam, Ossama, Teresa, Lluc, Cristina, Zein, Feliu, Martin and Anna.

It is always impossible to personally thank everyone who has helped successful completion of this study. To those of you without specifically name, my sincere thanks go to all of you, my friends and colleagues in Spain and Tunisia. Really, this is a great opportunity to express my acknowledgement, respect and gratefulness to you for your friendship, encouragement and assistance.

I specially wish to express my sincere thankfulness to my beloved parents for your tremendous love and many prayers that have been always very valuable and for teaching me how to be a strong person unconsciously by making me gets back up whenever I stumble

I am the most grateful to my wife Rim who overflows me with love and inspiration day by day; I share this accomplishment with you. Thank you for your encouragement, for believing in my capabilities, moral supports in all the good and hard times and patience during these years.

Special thanks to families BACCAR and YANGUI members especially for their continuous support and love.

Chapter 1

Introduction

Increasing competition in the agro-food sector, accentuated by economic globalization, has led companies to promoting innovative strategies to enhance the agricultural products, differentiating them through perceived signs of quality, sustainability or territoriality, etc. Moreover, the Directorate General for Health and Consumer Protection has explicitly stated its desire to adapt the Common Agricultural Policy (CAP) to be able to integrate the interests and preferences heterogeneity of the society in the planning, design, and implementation of agro-food policies (European commission, 2003). Therefore, an extensive effort has been and should continue to be generated mainly addressed to: 1) develop adequate and experimental tools to provide more realistic consumer's values for new agro-food attributes; and 2) provide increasing evidence of consumers' preferences heterogeneity as the basis for market segmentation, targeting and positioning.

One of the most common method used to estimate the economic value of differentiated food products or innovative attributes associated to them has been the estimation of consumer' Willingness to Pay (WTP). In fact, estimates of consumer' WTP for new value added traits have become important determinants of new product adoption (Lusk, 2003). WTP is the maximum amount a person is willing to pay to get a product or a service. Moreover, from this measure, market potential can be assessed by comparing WTP with market prices. If the consumer WTP is higher than the market value, the consumer will buy the product.

A number of methods have been used to elicit consumer's preferences and WTP for food products. Among them, the most commonly used have been contingent valuation (CV), experimental auctions (EA), and conjoint analysis (CA)/choice experiment (CE). The three methods are based on consumers' stated preferences.

Despite its popularity on valuing food quality attributes as well as environmental amenities and its flexibility, as it is not too costly and it is easier to implement than other experimental methods, CV has been criticized for a variety of reasons. It is characterized as a deeply flawed methodology (Diamond and Hausman, 1994) mainly for two reasons. First, CV tends to overstate the amount consumers are willing to pay for specific food attributes as no serious budget constraint is considered, generating a significant divergence from the actual behavior of consumers (Lusk, 2003; Carson and Hanemann, 2005). The second reason is related to its lower internal consistency of results provided due to the vulnerable reliability between the survey response and the economic theory implications which can lead to generate violations of economic theory (Diamond and Hausman, 1994; Carson and Hanemann, 2005).

Under the maintained hypothesis of truthful responses, people have little incentive to expand cognitive effort on decisions involving hypothetical stated preferences, making elicited values noisier and systematically biased (Lusk and Shogren, 2007). As a consequence, the use of experimental auctions (EA) has gained recognition among applied economists to elicit consumer valuations for both new public and private goods. Lusk and Shogren (2007) outlined that the main advantages of EA over the other value elicitation methods is that they put the participants in an active market environment where they can incorporate market feedback, and where there are real economic consequences to stating preferences that differ from what they actually want. In addition, the advantage of EA is the exchange mechanism. It use real money and real goods to create a market where the participant's attention is focused on the valuation task, which creates incentives for people to think carefully about what they will actually pay for the good. Hence, the WTP values obtained from EA can be more precise than other hypothetical elicitation methods.

However, EAs have showed some potential limitations: 1) it is more costly than other elicitation methods, generating geographical and regional restrictions in the sampling; 2) it has been shown that the amount of money participants receive as compensation for participation in the auction may generate biases in their bids; 3) it is quite common to observe null bids, generally due to the lack of interest by the bidders for the auctioned good (Lusk and

Hudson, 2004); and 4) (and the most important for the purpose of this study) EAs only allow for the valuations of changes in just one attribute of the good of interest (e.g. conventional vs organic production) when in a real shopping setting consumers have to take decisions among a complex bundle of food attributes (and attribute levels).

This is why in recent years multiple Conjoint Analysis, and specially the Choice Experiment response format, has attracted increasing attention by food economists. CEs overcome the binary discrete choice of CV and the one attribute change valuation in the EA format by allowing for the valuations of changes in one or more of the attributes of the good of interest. In general, respondents in CEs are asked to select their preferred alternative from a given set of alternatives (the choice set), and are typically asked to perform a sequence of such choices giving rise to a panel of discrete choices. Experimental design theory is used to construct the alternatives, which are defined in terms of both attributes and attributes' levels (Louviere and Street, 2000 and Street and Burgess, 2007). Contrary to the EA, which has been considered an unfamiliar market mechanism for most consumers (Alfnes et al., 2006), the CE task closely mimics consumers' typical shopping experience as they have to choose one product from several competing options (Louviere and Street, 2000; Lusk and Hudson, 2004; Alfnes et al., 2006). The CE not only allows researchers to investigate trade-offs between competing product attributes such as PDO, organic, price, package size, etc., which cannot be easily done with the CV method (Lusk and Hudson, 2004, Carson and Hanemann, 2005), but also to estimate the cross-price elasticities between novel and existing products (Lusk and Hudson, 2004; Ding et al., 2005). Finally, the CE has strongest foundations with economic theory as it is consistent with both the Lancaster's microeconomic approach (Lancaster, 1966) and the Random Utility Theory (RUT) (McFadden, 1974). CE assumes that individuals are rational and make choices to maximize their utility (derived from the characteristics or attributes that a good possesses, rather than directly from the good per se) taking into account their budget constraint.

Despite these interesting advantages, the CE also presents certain limitations. First, several studies have found that subjects' responses may be inconsistent across choice questions or are influenced by the complexity of the choice task due to the number of hypothetical products that the individual must evaluate during the experiment (Lusk and Hudson, 2004). Second, in many applications, the CE is hypothetical, that is, it does not consider person's budget constraint, which can overestimate, as in the CV method, participants' WTP. Third, also in many applications, it is informationally inefficient as it does not provide information about relative preferences among the no chosen profiles. Finally, it is difficult to incorporate observed and unobserved consumer characteristics derived from, for example, socio-demographic characteristics, personality traits, perceptions, etc., as other explanatory variables into discrete choice models. In any case, theoretical and empirical research during the last decade has tried to overcome most of the above mentioned limitations (Ding et al. 2005; Ding, 2007; Lusk et al. 2008; Chang et al., 2009; Dong, 2010; Lancsar et al. 2013, Akaichi et al., 2013). This is also the main objective of this thesis. More precisely, this thesis focuses on two main issues: 1) new tools to tackle with preference heterogeneity; and 2) new response formats to allows researchers to take into account the information provided by no chosen profiles. These two issues have generated three studies, which form the main core of this thesis: two are related to issue 1) (Chapters 2 and 3), while the third one is related to the issue 2) (Chapter 4).

In relation to the first issue, initially, discrete choice models were estimated through fix parameter multinomial models in which the deterministic portion of utility was assumed not to vary across individuals (preference homogeneity). Furthermore, it was assumed that choices were independent from each other, which was clearly a heroic assumption in the presence of repeated sequential choices such as in choice experiments. Failure to account for preference heterogeneity may not only result in poor model performance (i.e., generating an incorrect standard error and biased parameter estimates) but also affect elasticities, willingness-to-pay measures and substitution patterns, all of which could lead to problems in the reliability of model results (Hynes et al., 2008; Hess et al., 2010). Although some form of conditional taste heterogeneity can be accommodated in fixed parameter multinomial logit (Scarpa et al., 2005), unconditional heterogeneity is both more appealing and informative in consumer analysis.

Therefore, methods that account for preference heterogeneity have received a significant amount of attention in recent literature (Campbell et al 2010; Greene and Hensher 2013). Among the most relevant we can cite: 1) the use of segmentation strategies (Shen 2010); 2) the inclusion of interaction effects to explain sources of heterogeneity (Mtimet and Albisu 2006); 3) the use of random parameter estimates, assuming preference coefficients to be randomly distributed across individuals (Revelt and Train 1998); and 4) the combination of interaction effects and random parameters (Hensher and Greene 2003), or segmentation strategies and random parameters (Greene and Hensher 2013).

Among these approaches, the estimation of a Random Parameter Logit Model (RPL) (Revelt and Train 1998) to obtain willingness-to-pay (WTP) estimates for food attributes has become increasingly popular. The RPL model relies on the relaxation of the three main limitations of conventional logit models: 1) it allows for random preference variation across individuals through the distribution of random parameters; 2) it relaxes the assumption of independence from irrelevant alternatives (IIA), and 3) it allows for correlation among unobserved factors over time (Train 2003). However, the RPL also has some limitations. Lenk and DeSarbo (2000) and Scarpa and Thiene (2005) showed that although the RPL model provides an interesting way to account for preference heterogeneity, it might be inadequate if different groups of individuals with different group-specific preferences exist as in this case individual partworths across the participants could be very similar. To overcome such limitation, Greene et al. (2006) extended the Random Parameter Logit model (GHR–RPL) to account for heterogeneity around both the mean and the variance of the parameter distributions.

Chapter 2 provides the first empirical application in the agro-food sector of such methodology. More precisely, we intend to account for preference heterogeneity in two ways: (i) by identifying further behavioral information associated with the mean of the random parameter distribution by the parameterization of its heterogeneity through attitudinal factors, and (ii) by providing more information about the variance, allowing it to expressed as a function of individual specific characteristics. The performance of the extended RPL is evaluated against the traditional RPL model taking into account just the heterogeneity around the mean. We also illustrate the implications of each model on the moments of the WTP distribution.

Discrete choice modelling based on the Random Utility Theory (RUT) defines individuals' utility as a function of product attribute levels and their socio-demographic characteristics covariates. However, in the last decade, the literature has highlighted the relevance of individuals' psychological factors, personality traits and attitudes in individuals' decision making process (Lusk, 2010; Chen, 2007; Yáñez et al., 2010). However, these variables cannot be directly measured but inferred from observed variables. Therefore, our second study, Chapter 3, focuses on the incorporation of latent variables such as consumer's personality traits, lifestyles, and purchase habits as explanatory variables in discrete choice models. This Chapter tries to cover a gap in the existing literature as only very few studies have investigated the potential effect of purchase habits, food-related personality traits and lifestyle orientation on consumer's behavior (Chen, 2007; Eertmans et al., 2005).

To achieve this objective, the Hybrid choice model (HCM) represents a promising new class of models which merge classic choice models with structural equations models (SEM) for latent variables (LV) (Ben-Akiva et al., 2002). Regardless of their conceptual appeal, to the best of our knowledge, up to date there is not studies that applied HCM in agro food marketing. This Chapter extends previous HCM applications by, first, estimating a random parameter logit model (RPL) into a panel data context (taking into account the heterogeneity around the mean) and, second, estimating the relationships between latent variables based on the Theory Planned Behavior (TPB).

In relation to the second issue, a key objective of discrete choice experiments is to obtain sufficient quantity of high quality choice data to estimate the appropriate choice models to be used to explore various relevant issues related to consumer decision making. However, the basic choice experiment consist in asking respondents to choose the most preferred of various alternatives offered in a choice set, thereby obtaining one choice observation per set. Generally, there are two ways to increase the number of observations: 1) increasing the sample size; or 2) increasing the number of choice sets evaluated by each respondent. The former clearly has cost implications, especially when real incentives are involved, while the latter can increase the complexity of the choice task to respondents (Lancsar et al., 2013).

A third way to increase the number of observations has been to develop new CA question formats. In this context, the Ranking Conjoint Analysis (RCA) and Best Worst Scaling (BWS) have been proposed in the literature. In the first one, respondents are asked to express preferences for several alternatives existing in the choice set by ranking them from the most to the least preferred, while in the second respondents are asked to choose not only the best option in each choice set, but also the worst option in relation to the remaining options (Louviere, 2008). As can be observed, both approaches provide information about the relative preference of respondents in relation to the non-chosen profiles of the basic choice experiment. However, according to Marley and Louviere (2005), best worst tasks seem to be easier for people than to complete a traditional ranking task. It takes advantage of a person's natural tendency to identify and respond more consistently to extreme options.

In this context, Chapter 4 in this thesis is addressed to compare the ability of three CA response formats (CE, RCA, and BWS) in terms of estimated partworths, predictive power, estimated WTP, and participants' response consistency in two contexts: 1) when only the most preferred option is considered (RCA and BWS are coded as traditional CE); and 2) when the full ranking informations from RCA and BWS are considered. The three CA response formats are implemented in a non hypothetical setting. For comparison purposes, results from a hypothetical CE will be considered as a benchmark. None of the published studies on CA have compared, at the same time, the performance of the three CA formats (CE, RCA and BWS) which is one of the main novelties of this thesis.

From an empirical point of view, this thesis is focused in assessing consumers' preferences towards the consumption of the olive oil. The olive oil is an important element of the Mediterranean diet and a valuable crop for Southern European countries in terms of both income and cultivated area. It has become a highly differentiated product. First, the International Olive Oil Council (2013) has classified olive oils in three main categories: extra virgin, virgin, and (refined) olive oil. Second, the olive oil has been one of the most relevant products in terms of geographical differentiation (Protected Denomination of Origin - PDO). Finally, the organic attribute has gained recognition among public authorities and farmers.

Spain is the first producer and exporter country of extra-virgin olive worldwide. Additionally, olive oil constitutes a fundamental component of the Spanish diet. As a consequence, the vast majority of Spanish consumers are knowledgeable about this product, and all of them are aware about market prices and product characteristics. Catalonia is the second region within Spain in terms of total olive oil consumption (with a per capita consumption of 9.93 litres in 2011).

The market share of certified olive oil is still in its earlier stages. The market value for organic olive oil was 11 million Euros in 2012 (MAGRAMA, 2013) but it only represented around 0.4% of the total olive oil market. Catalonia, again, occupies the second position in relation to the consumption of organic olive oil (13% of the Spanish total consumption in value) after Madrid. In relation to the PDO olive oil, from 2000 to 2012, the number of extra virgin olive oils registered as PDO increased from 7 to 28 DOP, most of them located in Andalusia and Catalonia. However, PDO olive oil only represents 1.5% of the total olive oil market (MEC, 2012).

Data for this thesis have been gathered from two experiments. The first one is related to the first issue discussed above. Information from this data set has been used in Chapters 2 and 3. The second experiment was designed to collect information for Chapter 4. Both experiments were conducted in Catalonia. As mentioned above, Catalonia is the second region in terms of olive oil consumption. Moreover, the population of Catalonia is quite heterogeneous, with an adequate combination of urban (Barcelona is the second largest town in Spain) and rural environments, which seems to be adequate for the purpose of this thesis.

Our first experiment consists of the implementation of a hypothetical CE to analyze preference heterogeneity. A sample of 401 consumers of extra virgin olive oil was recruited and surveyed in September 2009. The survey was divided in four blocks. The first block was designed to elicit information on respondents' buying and consumption habits concerning different types of olive oil. The second block was designed to obtain information about different attributes considered by respondents when buying extra-virgin olive oil, with special attention paid to attitudes towards the organic attribute. The third block addressed the choice experiment, where four main attributes were identified price, production system, origin of the product, and the origin of the brand, each defined with three levels. The last block was designed to obtain information about the socio-demographic characteristics and respondents' lifestyles.

The second experiment was carried out in May-June 2013. The experiment involved 220 subjects randomly distributed over four treatments: hypothetical choice experiment (HHCE), non-hypothetical choice experiment (NHCE), non-hypothetical ranking conjoint analysis (NHRCA), and non-hypothetical best worst scaling (NHBWS). During each treatment, the participants did to two main tasks. The first task consists of the main CE (HHCE, NHCE, NHRCA or NHBWS, depending on the treatment). Each treatment of the experiment was conducted over 5 sessions throughout both different days of the week and different hours of day. Each session includes a maximum of 10-15 persons. After the two tasks, the participants fulfilled a short questionnaire aimed at collecting socio-demographic and lexicographic characteristics of respondents as well as on attitudes and olive purchasing and consumption habits.

References

Akaichi, F., R. M. Nayga, Jr. and J. Gil. (2013) “Are Results from Non-Hypothetical Choice-Based Conjoint Analysis and Non-Hypothetical Recoded-Ranking Conjoint Analysis Similar?” Forthcoming in American Journal of Agricultural Economics 95: 949-963.

Alfnes, F., Guttormsen, A.G., Steine, G., and Kolstad. K. (2006) “Consumers’ willingness to Pay for the Color of Salmon: A Choice Experiment with Real Economic Incentives.” American Journal of Agricultural Economics 88, 1050-1061.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J. Bhat, C.a, Bierlaire, M., Bolduc, D., Börsch-Supan, A., Brownstone, D., Bunch, D.S., Daly A., De Palma, A., Gopinath, D., Karlstrom, A., and Munizaga, M.A., (2002) “Hybrid choice models: Progress and challenges”, Marketing Letters 13, 163-189.

Campbell D, Doherty E, Hynes S, Rensburg TV, 2010. Combining discrete and continuous mixing approaches to accommodate heterogeneity in price sensitivities in environmental choice analysis. Agricultural Economics Society Annual Conf, Edinburg (UK), Mar 29-31.

Carson, R., and Hanemann W.M. (2005) “Contingent Valuation” Chapter 17 in Handbook of Environmental Economics, eds. Mäler K.G. and Vinvent J.R, Elsevier B.V.

Chang, J.B., Lusk, J.L., and Norwood, B. (2009) “How closely do hypothetical surveys and laboratory experiments predict field behavior?” American Journal of Agriculture Economics 96, 518-534.

Chen, M.F. (2007) “Consumer attitudes and purchase intentions in relation to organic foods in Taiwan: Moderating effects of food-related personality traits”, Food quality and preference 18, 1008-1021.

Diamond, A.P., and Hausman, J.A. (1994) “Contingent Valuation: Is some number better than no number?” The Journal of Economic Perspectives 8, 45-64.

Ding, M. (2007) "An incentive-aligned mechanism for conjoint analysis" journal of Marketing Research 44, 214-223.

Ding, M., Grewal, R., and Liechty, J. (2005) "Incentive-Aligned Conjoint Analysis." Journal of Marketing Research 42, 67–83.

Dong, S., Ding, M., and Huber, J. (2010) "A simple mechanism to incentive-align conjoint experiment" International Journal of Research in Marketing 27, 25-32.

Eertmans, A., Victoir, A., Vansant, G., & Van den Bergh, O. (2005). Food-related personality traits, food choice motives and food intake: Mediator and moderator relationships. Food Quality and Preference 16, 714-726.

European Commision (2013) "consumer demand" http://ec.europa.eu/agriculture/organic/consumer-confidence/consumer-demand_en.

European Commission, (2003) "Consumer Interests in the Common Agricultural Policy efficiency and equity" Directorate General Health and Consumer Protection. Bruselas.

Greene WH, Hensher DA, 2013. Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. Applied Economics 45, 1897-1902.

Greene, W.H., Hensher, D.A., and Rose, J. (2006) "Accounting for heterogeneity in the variance of unobserved effects in mixed logit models", Transportation Research B, 40, 75–92.

Hensher DA., Greene WH, 2003. Mixed logit models: state of practice. Transportation 30, 33-176.

Hess, S., and Rose. J.M. (2009) "Allowing for intra-respondent variations in coefficients estimated on repeated choice data", Transportation Research Part B 43, 708-719.

Hynes S, Hanley N, Scarpa R, 2008. Effects on welfare measures of alternative means of accounting for preference heterogeneity in recreational demand models. American Journal of Agricultural Economics 90, 1011–1027.

International Olive Oil Council (2013) “Designation and definitions of olive oils” http://www.internationaloliveoil.org/estaticos/view/83-designations-and-definitions-of-olive-oils?lang=en_US

Lancaster, K. J. (1966), “A new approach to consumer theory”, The journal of political economy 74, 132-157.

Lancsar, E., Louviere, J., Donaldson, C., Currie, G., and Burgess, L. (2013) “Best worst discrete choice experiment in health: Methods and an application” Social Science and Medicine 76, 74-82.

Lenk P, DeSarbo W, 2000. Bayesian inference for finite mixtures of generalized linear models with random effects. Psychometrika 65, 93-119.

Louviere, J.J. and Street, D. (2000). “Stated-preference Methods” in David A Hensher & Kenneth J Button (eds), Handbook of Transport Modelling, Pergamon Press, Amsterdam, Netherlands, 131-143.

Louviere, J.J., Street, D., Burgess, L., Wasi, N., Islam, T., and Marley, A.A.J. (2008) “Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information” Journal of Choice Modelling1, 128-163.

Lusk, J. (2003) “ Effect of cheap talk on consumer willingness to pay for golden rice” American Journal of Agricultural Economics 85, 840-856.

Lusk, J.L, and Hudson, D. (2004) “Willingness to pay estimates and their relevance to agribusiness decision making” Review of Agricultural Economics 26, 152-169.

Lusk, J.L. (2010) “Experimental auction markets for studying consumer preferences” chapter 12 in Consumer-driven innovation in food and personal care of products, Woodhead publishing limited.

Lusk, J.L., and Shogren, J.F. (2007) “Experimental Auctions: Methods and applications in economic and marketing research” Cambridge University Press, Cambridge, UK.

Lusk, J.L., Fields, D., and Prevatt. W. (2008) “An Incentive Compatible Conjoint Ranking Mechanism.” American Journal of Agricultural Economics 90, 487–98.

Marley, A.A.J., and Louviere, J.J. (2005) “Some probabilistic models of best, and best-worst choices” Journal of Mathematical Psycology 49, 464-480.

McFadden, D. (1974), “Conditional logit analysis of qualitative choice behavior”, En Zarembka, P. (ed.) Frontiers in econometrics. Academic press, Nueva York.

Ministerio de Agricultura, Alimentación y Medio ambiente (MAGRAMA) (2013) “Agricultura Ecológica Estadísticas 2012” Madrid.

Ministerio de Economía y Competitividad (MEC) (2012) “Información sectorial de alimentos: Aceite de oliva”

Mtimet N, Albisu LM, 2006. Spanish wine consumer behavior: a choice experiment approach. Agribusiness 22, 343-362.

Revelt D, Train K, 1998. Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. Rev Econo Stat 80, 647-657.

Scarpa, R., Philippidis, G., and Spalatro, F. (2005) “Product-country images and preference heterogeneity for mediterranean food products: a discrete choice framework”, Agribusiness 21, 329-349.

Shen J, 2010. Latent class model or mixed logit model? A comparison by transport mode choice data. Applied Economics 42, 2915-2924.

Street, D., and Burgess, L.B. (2007) “The Construction of Optimal Stated Choice Experiments: Theory and Methods”, 1st ed, John Wiley and Sons, Hoboken, New Jersey.

Train, K. (2003) "Discrete Choice Methods with Simulation" Cambridge: Cambridge University Press.

Yañez, M.F., Raveau, J. and Ortúzar, J.D. (2010) "Inclusion of latent variables in Mixed Logit Models: Modelling" Transport research part A 44, 744-755.

Chapter 2

Revealing additional preference heterogeneity with an extended random parameter logit model: The case of extra virgin olive oil

2.1. Introduction

Preference elicitation methods have been extensively used by economists and market researchers to determine consumers' willingness to pay (WTP) for specific product attributes. Discrete choice modeling is a preference elicitation method that has been widely used in previous research (Lusk and Shroeder 2004; Ding et al. 2005; among many others). Initially almost all researchers applying discrete choice models assume that error variances are homogeneous over individuals by the application of Multinomial Logit (MNL), Condidtional Logit (CL), and Nested Logit (NL). However, evidence of preference heterogeneity in both revealed preference data (Hensher 2008) and stated preference data (Hess and Rose 2009) is increasing. Failure to account for preference heterogeneity may not only result in poor model performance (i.e., generating an incorrect standard error and biased parameter estimates) but also affect elasticities, willingness-to-pay measures and substitution patterns, all of which could lead to problems in the reliability of model results (Hynes et al 2008; Hess et al 2010).

Therefore, methods that account for preference heterogeneity have received a significant amount of attention in recent literature (Campbell et al 2010; Greene and Hensher 2013). Among the most relevant we can cite: 1) the use of segmentation strategies (Shen 2010); 2) the inclusion of interaction effects to explain sources of heterogeneity (Mtimet and Albisu 2006); 3) the use of random parameter estimates, assuming preference coefficients to be randomly distributed across individuals (Revelt and Train 1998); and 4) the combination of interaction effects and random parameters (Hensher and Greene 2003), or segmentation strategies and random parameters (Greene and Hensher 2013).

Among these approaches, the estimation of a Random Parameter Logit Model (RPL) (Revelt and Train 1998) to obtain willingness-to-pay (WTP) estimates for food attributes has become increasingly popular. The RPL model relies on the relaxation of the three main limitations of conventional logit models: 1) it allows for random preference variation across individuals through the distribution of random parameters; 2) it relaxes the assumption of independence from irrelevant alternatives (IIA), and 3) it allows for correlation among unobserved factors over time (Train 2003).

However, the RPL also has some limitations. Indeed, although the RPL model uses continuous (i.e., normal, log-normal or triangular) distributions for individual tastes to account for preference heterogeneity, it does not identify the heterogeneity source. Additionally, the RPL estimated parameters ($\beta_{ki} = \beta_k + \sigma_k \vartheta_{ki}$), depend on the so-called individual specific heterogeneities (ϑ_{ki}), which follows a normal distribution with zero mean and a standard deviation of one (Hensher and Greene 2003), and σ_k, the standard deviation of the distribution of the estimated parameter around its mean. Therefore, if the product $\sigma_k \vartheta_{ki}$ is small, the estimated parameters across individuals will be very similar. Furthermore, Lenk and DeSarbo (2000) and Scarpa and Thiene (2005) have shown that although the RPL model provides an interesting way to account for preference heterogeneity, it might be inadequate if different groups of individuals with different group-specific preferences exist.[1] To overcome such limitations, Greene et al. (2006) extended the Random Parameter Logit model (GHR–RPL) to account for heterogeneity around both the mean and the variance of the parameter distributions and illustrated the implications on the moments of the WTP.

The traditional approach to identify the preference heterogeneity is to directly include the individual observable socio-demographic characteristics in the utility function. However, Morey and Russmon (2003) showed that this procedure could be very restrictive as it assumes that some segments, which seem to have the same characteristics, have the same preferences. Although improving the model goodness-of-fit, Scarpa and Thiene (2011) also argued that

[1] The use of the LCM could also lead to inefficient estimates because it might oversimplify the population's preferences, especially when a small number of classes are defined and the distribution of preferences is continuous within classes (Allenby and Rossi 1998).

this approach is relatively poor in giving some insight about the source of heterogeneity. McFadden (1986) and Ben Akiva et al. (2002) discussed the important role of attitudes and beliefs to understand and estimate individuals' preferences, and to what extent they were conceptually important in choice decision protocols. However, in spite of the potential benefit of introducing attitudinal factors to explain individuals' preferences heterogeneity, up to our knowledge, only a few papers have addressed this issue (Moore, 2008; Stolz el al., 2011a; Stolz el al., 2011b).

The aim of this paper is to assess consumer's preferences heterogeneity for extra-virgin olive oil[2] in Catalonia (North-East Spain). Spain is the first producer and exporter country of extra-virgin olive world-wide. Additionally, olive oil constitutes a fundamental component of the Spanish diet. As a consequence, the vast majority of Spanish consumers are knowledgeable about this product, and all of them are aware of market prices and product characteristics. Catalonia is the second region within Spain in terms of total olive oil consumption (with a per capita consumption of 9.93 liter in 2011). Moreover, Catalonia has a quite heterogeneous population with an adequate combination of urban (Barcelona is the second largest town in Spain) and rural environments which seems to be adequate for the purpose of this study.

To tackle with this objective, the methodological framework adopted is based on the estimation of the GHR-RPL model. More precisely, we intend to account for preference heterogeneity in two ways: (i) by identifying further behavioral information associated with the mean of the random parameter distributions by the parameterization of its heterogeneity through attitudinal factors such as health awareness, environment awareness, organic olive oil trust, subjective norms, organic olive oil purchasing intention and knowledge, and (ii) by providing more information about the variance, allowing it to be expressed as a function of

[2] Extra-virgin indicates that the olive oil has been produced by using mechanical means only, without any chemical treatment and contains no more than 0.8% free acidity. It is considered as the highest quality olive oil.

individual specific observed characteristics. The performance of the GHR–RPL with attitudinal factors is evaluated against three alternative RPL models: 1) the conventional RPL model without accounting for heterogeneity around the mean (RPL1); 2) the RPL model taking into account the heterogeneity around the mean of the random parameters as a function of socio-demographic characteristics (RPL2); 3) same than RPL2 but in this case heterogeneity is a function of attitudinal factors (RPL3). The paper also illustrates the implications of each model on the moments of the WTP distribution. Finally, from an empirical point of view, specific attention is placed on the organic attribute of the extra-virgin olive oil.

This paper is structured as follows. In the next section, we outline the methodological framework used in this paper. Section 3 describes the experimental design and the empirical model. The main results and discussion are presented in section 4. The paper ends with some concluding remarks.

2.2. The extended random parameter logit model (GHR–RPL)

The choice information used in Random Utility Modeling (RUM) can come from the observations of actual choices in a real setting (revealed preferences) or from choices made in hypothetical settings (stated preferences) (Louviere and Hensher 1982; Louviere 2001). From the latter type of choice information, choice experiments (CE) are derived. The CE is in accordance with both the Random Utility Theory (RUT) (McFadden 1974) and the Lancaster's consumer theory (Lancaster, 1966). The RUT assumes that decision makers are rational and that individuals make choices to maximize their utility, taking into account budget constraints. In parallel, Lancaster's consumer theory presumes that the utility of a defined good can be segregated into product attribute utilities and proposes that consumers make choices based on attribute preferences. Therefore, the utility is derived from the attributes and attribute levels. The respondents are asked to make repeated choices between hypothetical alternatives described by combinations of attributes and their levels and are asked to choose their preferred alternative.

McFadden (1974) proposed an econometric framework to estimate discrete choice models based on random utility models. The individual utility of a particular option can be expressed as follows:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \tag{1}$$

where V_{ij} is a deterministic or observed component that is a function of alternative product characteristics (X_{ij}) and ε_{ij} is the stochastic or non-observed component. Individual "i" will choose alternative "j" if it provides him a higher utility than any k^{th} available alternative. The probability of consumer "i" choosing alternative j out of the total set of options is expressed as follows:

$$P_{ij} = Prob\left[U_{ij} > U_{ik}\right] = Prob\left[V_{ij} + \varepsilon_{ik} > V_{ik} + \varepsilon_{ik}\right] \ \forall\, j \neq k \ \in C_n \tag{2}$$

where C_n is the choice set and the observed component V_{ij} is expressed as follows:

$$V_{ij} = \beta_0 + \sum_k \beta_k * X_{kj} + \beta_{price} * P_j \tag{3}$$

where β_0 is an alternative-specific constant for alternative j, β_k the marginal utility of attribute X_{kj} and β_{price} is the marginal utility of the price P_j of alternative j for consumer i.

Different assumptions about the stochastic component generate different models. If the stochastic component ε_{ij} has a type I extreme value distribution, we obtain the familiar Multinomial Logit model as a conditional logit model, where the probability of consumer i choosing option j from a specific choice set (C_n) is expressed as follows:

$$P_{ij} = \frac{e^{\mu V_{ij}}}{\sum_{k=1}^{J} e^{\mu V_{ik}}} \qquad \forall\, j \in C_n \tag{4}$$

This model is based on the following quite restrictive assumptions: 1) the consumers are assumed to be homogeneous, which implies that all coefficients for all attributes considered in the utility function are assumed to be the same across the sample; 2) the property of independent of irrelevant alternatives (IIA) holds; and 3) errors are independent over time (Hensher et al 2005 and Van Loo et al 2011).

To overcome some of these restrictive assumptions, several alternatives have been proposed in the literature. One of the most used in the literature is the Random Parameter Logit (RPL) model. The RPL model is based on the assumption that β_k parameters in (2) are distributed across individuals according to a statistical distribution. This model accounts for preference heterogeneity among individuals and it is flexible enough to accommodate alternative specifications, although it does not explain the source of this heterogeneity (Train 2003)

Under the RPL model, the probability that an individual “i” chooses alternative “j” in a particular choice set C_n is expressed as follows:

$$Prob_i\{j\ is\ chosen\} = \int L_{ij}(\beta_{ij}) f(\beta_i/\theta)\, d\beta_i, with\ k\ \in\ C_n \tag{5}$$

where $f(\beta_i/\theta)$ is the density function of the β_i coefficients; θ refers to the moments of the parameter distributions (the mean and the standard deviation of β_i) and

$$L_{ij}(\beta_{ij}) = \frac{e^{V_{ij}}}{\sum_{k=1}^{J} e^{V_{ik}}} = \frac{e^{\beta_{ij}X_{ij}}}{\sum_{k=1}^{J} e^{\beta_{ik}X_{ik}}} \tag{6}$$

As mentioned above, the model is flexible enough to specify any distribution for the estimated parameters) (such as normal, log-normal, triangular, uniform, etc.).

When the ith individual (i=1….N) faces a choice among J alternatives (j=1….J) in each of the T choice sets (t= 1….T), the utility of individual “i” associated with each alternative in each choice situation can be expressed as follows (Train 2003):

$$U_{ijt} = \beta_{ij}X_{ijt} + \varepsilon_{ijt} \tag{7}$$

The simplest specification treats the coefficients as varying among the individuals but being constant for the choice situation of each person. Under the RPL model, the individual parameter estimates β_{ij} are expressed as follows:

$$\beta_{ij} = \beta_j + \sigma_j\vartheta_{ij} \tag{8}$$

where β_j is the sample mean for alternative j, σ_j is the standard deviation of the distribution of the partworth around the mean and ϑ_{ij} are individual specific heterogeneities with a mean of zero and a standard deviation of one (Hensher and Greene 2003). The product characteristics (X) are observable, and we can estimate β_j and σ_j and test (Hensher et al 2005) which alternative parametric distribution for β_j and σ_j (e.g., normal, log-normal, uniform or triangular) provides the best approximation of sample preferences.

Greene et al (2006) suggested extending the RPL model (hereinafter GHR–RPL) to capture additional alternative unobserved variation, by first estimating deep parameters to account for heterogeneity around the mean of the distribution and, second, adding further behavioral information associated with the variance of the random parameter distribution, through the parameterization of its heteroscedasticity. Hence, equation (8) can be re-written as follows:

$$\beta_{ij} = \beta_j + \delta_j Z_i + \sigma_{ij}\vartheta_{ij} \tag{9}$$

where the vector Z_i is a set of choice-invariant characteristics that produce individual heterogeneity in the mean of the randomly distributed coefficients; δ_j are parameters that capture the mean shift; σ_{ij} is specified as $\sigma_{ij} = \sigma_j exp[\omega_j hr_i]$, where ω_j are parameters that capture the variance heterogeneity of the random parameters in the systematic utility; and hr_i are individual specific characteristics.

The individual choice probabilities can be approximated using the three-step procedure suggested by Train (2003): 1) for any given value of θ, draw a value of β_i from $f(\beta_i/\theta)$ and label it β_i^r with r = 1….R[3]; 2) calculate the logit formula $L_{ij}(\beta_i^r)$ with this draw; and 3) repeat steps 1 and 2 many times and average the results. This average is the simulated probability:

$$\widehat{P_{ij}} = \frac{1}{R}\sum_{r=1}^{R} L_{ij}(\beta_i^r) \tag{10}$$

where R is the number of draws. The simulated probabilities are inserted into the log-likelihood function to obtain a simulated log-likelihood (SLL):

$$SLL = \sum_{i=1}^{I}\sum_{j=1}^{J} d_{ij}\, Ln\widehat{P_{ij}} \tag{11}$$

where d_{ij}=1 if the individual “i” chooses “j” and zero otherwise. The simulated maximum likelihood estimator (SML) is the value of θ that maximizes SLL.

The willingness to pay (WTP) for product attributes is the price change associated with a unit increase in a given attribute and can be calculated as the negative ratio of the partial derivative of the utility function with respect to the attribute of interest, divided by the derivative of the utility function with respect to the variable “Price” (Van Loo et al 2011):

$$WTP_{Attribute} = -\frac{\frac{\partial U_{ijt}}{\partial Attribute}}{\frac{\partial U_{ijt}}{\partial Price}} = -\frac{\beta_{Attribute}}{\beta_{Price}} \tag{12}$$

The mean and standard deviations of the WTP are derived by generating a distribution of 1000 WTP estimates using the parametric bootstrapping method proposed by Krinsky and Robb (1986). This approach significantly reduces the problem related to potential changes of sign caused by the extreme values of the behavioral WTPs distributions. In other words, we are increasing the probability of prices to be randomly distributed across the individuals following an unconstrained distribution.

[3] Halton draws are used because they have been shown to provide a more efficient distribution of draws for numerical integration, in comparison to random draws (Bhat 2003; Train 2003).

To test for differences in WTP distributions derived from the different estimated RPL models, the nonparametric combinatorial test mentioned in Poe et al., (2005) will be used. This test consists of calculating all differences between the WTPs estimated from two RPL models for all possible combinations of the bootstrapped values. The proportion differences that are negative are considered as the *p*-value associated with the one side test that the WTP estimated from the first model overestimate the WTP estimated from the second model.

2.3. Empirical application

2.3.1. Survey and sample characteristics

Data for this study were collected from a survey of a representative sample of the Catalonian population with quotas by postal code. A total of 425 persons participated in face-to-face interviews, 401 of which participated in the choice experiment. The data collection was conducted in September 2009 during different shopping hours and at different types of food retail stores. The questionnaire used was divided into four sections. The first section was designed to elicit information on respondents' buying and consumption habits concerning different types of olive oil. The second section was designed to obtain information about different attributes considered by respondents when buying extra-virgin olive oil, with special attention paid to attitudes towards the organic attribute. The third section addressed the choice experiment. The last section was designed to obtain information about the socio-demographic characteristics and lifestyles of the respondents. Attitudes were measured using eleven-point Likert scales (from 0 to 10, where 0 indicates total disagreement and 10 indicates total agreement)[4].

From the 401 respondents who completed the survey, 40% came from Barcelona (the main town) and 60% came from elsewhere in the Catalan region. Approximately 80% of respondents were women, consistent with Gil et al (2002), as the objective population was

[4] This scale is very comprehensive for respondents in Spain as it coincides with the traditional grading system at schools.

made up of those responsible for shopping within households. The average age of the respondents was 49 years old (with a standard deviation (SD) of 15.39). With respect to the education level, 27.3% of the respondents had completed only primary studies, 46.8% had completed secondary studies or professional education, and nearly 25.6% had obtained a university degree. Finally, 70% of the respondents were married, and the average household size was approximately 3 members. All respondents bought olive oil regularly. In fact, most of the respondents used to purchase olive oil weekly or every two weeks and nearly 30% purchased it monthly or quarterly. (Refined) Olive oil and conventional extra-virgin olive oil are oil types most commonly bought; only 9.25% of respondents buy extra-virgin olive oil with a protected denomination of origin designation and less than 1% buys organic olive oil. Finally, the mean price paid for one liter of conventional extra-virgin olive oil was 3.42 euro (SD=0.80).

2.3.2. The choice experiment design

To implement the choice experiment, attributes and attribute levels were first selected on the basis of a three-step process: 1) a literature review of consumers' extra-virgin olive oil purchase and consumption habits; 2) two focus groups (of 8 people each) to identify main consumption patterns and attitudes toward extra-virgin olive oil, with special focus on the organic attribute; and 3) observation in retail outlets of real prices and informal interviews with consumers about their reasons for choosing a specific olive oil. As a result, four main attributes were identified: price, production system, origin of the product, and origin of the brand (see Table 2.1). To avoid the level effect between attributes (De Wilde et al., 2009), each attribute was defined as having three levels.

Taking into account the number of attribute levels, a total 81 (3^4) hypothetical bottles of extra-virgin olive oil were obtained. This led to a large number of choice sets affecting respondents' decisions and a consequent decrease in response reliability (Chung et al 2010).

To reduce the number of combinations that participants had to evaluate, an orthogonal factorial design was generated, resulting in 9 product profiles and 9 choice sets. Each choice set consisted of three alternatives plus the “none of them” option. We employed the strategy proposed by Street and Burgess (2007) to obtain a 100% efficient main effects design. Figure 2.1 shows one of the choice sets offered to respondents.

Table 2.1 Attributes and attributes levels in the CE for extra-virgin olive oil

Attributes	Levels
Production system	Conventional (CONV)
	Protected Denomination of Origin (PDO)
	Organic (ORG)
Origin	Spain (ESP)
	Catalonia (CAT)
	Imported (IMP)
Brand	Spanish manufacturer (BESP)
	Catalonia manufacturer (BCAT)
	Private label (PRIV)
Price	3.70 €/l
	6 €/l
	7.5 €/l

2.3.3. The empirical models

As mentioned in the introduction, the methodological approach followed in this paper has been the estimation of the Greene et al (2006) GHR–RPL model that accounts for heterogeneity around both the mean and the variance of the distributions of the estimated parameters. Results from this model are going to be compared with three alternative RPL models: RPL1, which does not account for heterogeneity; 2) RPL2, which account for heterogeneity around the mean as a function of respondents’s socio-demographic characteristics; and 3) RPL3, similar than RPL2 but in this case heterogeneity around the mean is specified as a function of respondents’ attitudinal factors.

Appendix 2.1 shows the description of the deterministic components of respondents’ utility that are common for the four models. Except for the price, which is assumed to be continuous, the rest of the variables are considered categorical and coded as either dichotomous of effect-coded dummy variables.

RPL models allow a higher level of flexibility in specifying some coefficients to be fixed or randomly distributed across respondents. In this study, and based on the Wald test statistic, four parameters estimates associated with PDO, ORG, CAT, and PRICE are defined to be random and following unconstrained[5] normal distributions. In contrast to the approach taken by Revelt and Train (1998), the price coefficient is not assumed to be invariant across individuals. As noted by Train and Weeks (2005), assuming a fixed price coefficient implies that the standard deviations of unobserved utility are the same for all observations. Therefore, estimation practices that ignore this source of variation may lead to erroneous interpretation and policy conclusions (Scarpa et al 2008).

Appendix 2.1 also includes the main respondents' socio-demographic characteristics that are considered to generate potential sources of heterogeneity (gender, age, education level, town size, and olive oil purchasing frequency). Finally, six attitudinal factors related to consumers' perceptions about organic olive oil (health awareness, environment awareness, trust, subjective norms, organic olive oil purchasing intention and knowledge) were also included in the utility functions (Table 2.2). These six attitudinal factors were defined by 18 items using a set of scales defined in the literature and measured through the application of a Confirmatory Factor Analysis (CFA). The internal consistency reliability, measured by Cronbach's α (Chen 2007), was greater than 0.7 in all cases. The variance extracted was greater than 50 percent in all cases indicating that latent variables were adequately represented by the defined items.

[5] Greene el al (2006) commented that the impact of accommodating heterogeneity around the mean and variance of random parameter distributions does not guarantee an advantage to using any constrained distributions.

Table 2.2 Attitudinal factors results from the confirmatory factor analysis (CFA)

Índ	Factors	Means	Standard Deviation	Variance	Cronbach 's Alpha	References
	Health Awareness (HL)			81.96	0.898	Adapted from Alemán et al. (2006), and Roitner-Schobesberger et al. (2007)
HL_1	The consumption of organic olive oil reduces human exposure to chemical residues.	6.867	1.764			
HL_2	Organic olive oil is healthy for children.	6.862	1.660			
HL_3	The product is suitable for a healthy diet.	7.088	1.636			
	Environment Awareness (ENV)			91.27	0.957	
EV_1	The production of organic olive oil helps indirectly to reduce water pollution by waste chemicals and pesticides.	6.923	1.680			
EV_2	The production of organic olive oil helps indirectly to conserve agricultural soil.	6.933	1.716			
EV_3	The production of organic olive oil improves environmental sustainability	6.893	1.809			
	Trust (TRT)			69.79	0.860	Adapted from Krystallis and Chryssohoidis (2005), and Roitner-Schobesberger et al. (2007);
TR_1	I trust the product because of its certification by an organization or regulatory board of organic farming.	6.447	1.601			
TR_2	I trust the product because it is sold exclusively in specialty stores.	6.668	1.646			
TR_3	I have confidence in the information provided on the product label.	6.202	1.710			
TR_4	I have confidence that a product certified as organic really is organic.	6.103	1.866			
	Purchase intention (PINT)			76.91	0.858	Adapted from Lea and Worsley (2005)
PI_1	If I have more information and confidence, I buy organic olive oil.	5.923	2.179			
PI_2	I buy more if the product is cheaper.	5.770	2.219			
PI_3	If organic olive oil is more readily available, I most often buy it.	5.655	2.246			
	Knowledge (KNW)			87.63	0.861	
KN_1	Lack information about the benefits of organic products.	6.905	1.834			
KN_2	Lack of information about the label that identifies products as organic.	6.872	1.889			
	Subjective norms (SBN)			86.61	0.926	Chen (2007)
SN_1	My kids prefer organic olive oil.	2.342	2.475			
SN_2	My family prefers organic olive oil.	2.465	2.422			
SN_3	Persons who are important to me prefer organic olive oil.	2.578	2.436			

2.4. Results

The four models mentioned above were estimated by the Simulated Maximum Likelihood (SML) method. Table 2.3, which compares the goodness-of-fit measures of the four models, shows that the GHR-RPL model outperforms the other three RPL models and provides the best fit to the data (the McFadden R-square and the Akaike Information Criterion are optimized in the GHR-RPL model). Furthermore, the values of the likelihood ratio statistic also provide evidence of the superior goodness of fit of the GHR–RPL model.

Results also indicate that attitudinal factors are more relevant than socio-demographic characteristics to explain the consumers' heterogeneity, as discussed by Scarpa and Thiene (2011) (RPL3 outperforms RPL2). Finally, Figures 2.2, 2.3, 2.4, and 2.5 show the estimated marginal utilities from the 4 RPL models[6]. The four utility distributions have very similar shapes. That is, the introduction of additional sources of heterogeneity even in the GHR–RPL model makes both the mean and the standard deviation to change but not the global shape of the distribution.

Table 2.3 Goodness of fit of alternative estimated models

Measures	RPL 1	RPL2	RPL3	GHR-RPL
Log likelihood (LL)	-3075.640	-3018.199	-2982.072	-2928.414
χ^2 (df)	---	$114.882^{***}(28)$	$72.254^{***}(4)$	$107.316^{***}(24)$
McFadden R^2	0.383	0.395	0.402	0.413
CAIC	1.71536	1.699	1.67671	1.66023
AIC	1.71538	1.69926	1.67691	1.66080
N. parameters	12	40	36	60

[6] A kernel density function has been used to graph the non-parametrically distribution of the marginal utility of the respondents in both models.

Table 2.4 shows the estimated parameters for the four RPL models[7]. In all cases, the no-option coefficient is negative and significant, indicating that most of the respondents tried to participate in the choice experiment by choosing one of the proposed olive oil alternatives instead of the no-option alternative. Table 2.4 also shows that all parameter estimates associated with the attribute levels considered in the utility function are statistically significant and with the expected sign, with the only exception of the estimated parameters associated to the levels of the attribute Origin of Brand: "BCAT" (Catalonian Manufacturer Label), which is not significant in any of the four models, and "PRIV" (Private Label), which is only significant in RPL3 and GHR-RPL, where heterogeneity around the mean is defined as a function of attitudinal factors.

As mentioned above, accounting for mean and variance heterogeneity of random parameters estimates has been proved to be relevant as the GHR–RPL model clearly outperforms the other models being significant a large number of specific parameters associated with mean and variance heterogeneity. Therefore, in the next paragraphs we concentrate in explaining results obtained from such model.

Table 2.4 Estimated coefficients

Parameters	RPL1		RPL2		RPL3		GHR-RPL	
CONV ^{a}	0.425	(---)	0.273	(---)	0.537	(---)	0.495	(---)
PDO^b	0.254***	(0.041)	0.237**	(0.105)	0.283***	(0.044)	0.303***	(0.039)
ORG^b	-0.679***	(0.054)	-0.510***	(0.133)	-0.820***	(0.058)	-0.798***	(0.053)
ESP^a	0.107	(---)	0.133	(---)	0.072	(---)	0.04	(---)
CAT^b	0.503***	(0.046)	0.493***	(0.106)	0.556***	(0.047)	0.581***	(0.040)
IMP	-0.610***	(0.042)	-0.626***	(0.042)	-0.628***	(0.043)	-0.621***	(0.045)
BESP^a	-0.047	(---)	-0.018	(---)	0.074	(---)	0.077	(---)
BCAT	-0.009	(0.039)	-0.024	(0.039)	0.005	(0.039)	0.006	(0.052)
PRIV	-0.056	(0.038)	-0.042	(0.038)	-0.079**	(0.039)	-0.083*	(0.059)
PRICE^b	-0.907***	(0.043)	-1.006***	(0.067)	-0.923***	(0.038)	-0.987***	(0.029)
No-option	-6.528***	(0.178)	-6.831***	(0.191)	-6.888***	(0.198)	-6.933***	(0.112)

[7] In relation to interaction terms, for space limitation purposes, Table 4 just shows the statistically significant interaction parameters.

Table 2.4 Estimated coefficients (Continued)

Standard deviations of parameter distributions								
PDO	0.326***	(0.055)	0.339***	(0.052)	0.436***	(0.047)	0.265***	(0.053)
ORG	0.732***	(0.057)	0.800***	(0.057)	0.803***	(0.064)	0.683***	(0.172)
CAT	0.681***	(0.050)	0.682***	(0.051)	0.710***	(0.052)	0.542***	(0.074)
PRICE	0.803***	(0.041)	0.795***	(0.033)	0.769***	(0.039)	0.700***	(0.046)
Heterogeneity in mean (Attitudinal factors)								
PDO-ENV	---		---		---		-0.153*	(0.079)
PDO-KNW	---		---		-0.131***	(0.044)	-0.136***	(0.050)
PDO-SBN	---		---		-0.207***	(0.045)	-0.230***	(0.042)
ORG-ENV	---		---		---		0.177*	(0.096)
ORG-KNW	---		---		0.128***	(0.057)	0.164***	(0.051)
ORG-SBN	---		---		0.306***	(0.055)	0.301***	(0.053)
CAT-PINT	---		---		0.138**	(0.056)	0.141**	(0.062)
CAT-SBN	---		---		-0.096**	(0.046)	-0.124***	(0.039)
PRICE-TRT	---		---		---		-0.163***	(0.045)
PRICE-PINT	---		---		0.206***	(0.032)	0.204***	(0.030)
PRICE-KNW	---		---		---		-0.063**	(0.030)
PRICE-SBN	---		---		0.097***	(0.036)	0.139***	(0.030)
Heterogeneity in mean (socio-demographic factors)							Heterogeneity in variance	
PDO-UNIV	---		---		---		0.636***	(0.109)
PDO-TS	---		-0.154**	(0.071)	---		---	
PDO-GEN	---		---		---		-0.538***	(0.167)
PDO-MONTH	---		---		---		0.275***	(0.153)
PDO-QUART	---		---		---		-0.454***	(0.213)
PDO-AGE	---		0.183**	(0.086)	---		1.213***	(0.200)
ORG-MONTH	---		0.218***	(0.089)	---		---	
ORG-QUART	---		0.160***	(0.087)	---		---	
ORG-AGE	---		-0.326***	(0.109)	---		-0.377**	(0.181)
CAT-UNIV	---		---		---		0.635***	(0.084)
CAT-MONTH	---		---		---		-0.489***	(0.134)
CAT-QUART	---		0.136***	(0.076)	---		0.230**	(0.103)
CAT-AGE	---		---		---		0.321***	(0.099)
PRICE-UNIV	---		0.120***	(0.044)	---		0.128***	(0.045)
PRICE-TS	---		0.186***	(0.071)	---		-0.248***	(0.091)
PRICE-AGE	---		---		---		0.205***	(0.068)

[a] This represents the base level.
[b] Random parameters following normal distributions
Notes: 1) See Tables 1 and 2, and Appendix 1 for variable definitions. 2) Values in parentheses represent the parameters' standard errors. 3) (***), (**) and (*) indicate that the corresponding parameter is statistically significant at the 1%, 5% or 10% level, respectively.

In line with Gracia and Magistris (2007), consumer's preferences in Catalonia towards the organic olive oil are positively affected by their more positive attitude towards environmental benefits provided by the organic production system. Equally important is the effect of subjective norms associated with the consumption of organic olive oil in mitigating the disutility related to its consumption (Chen, 2007). An interesting result that arises from this study is that consumer's attitudes towards health benefits provided by organic olive oil (HL) and trust (TRT) do not seem to have a significant effect on consumers' marginal utilities towards the organic olive oil. This may be can be related to the consumer's positive perception about the healthiness of the extra virgin olive oil regardless the type of production system (PDO, organic, or conventional) (Calatrava, 2002 and Vega-Zamora et al., 2011). In any case, results also suggest that more information on the properties of the organic olive oil could be relevant to increase the consumers knowledge about this product and then to increase the probability to buy it (the coefficient of the interaction ORG-KNW is positive and significant), in line with the results found in of Gracia and Magistris (2008). Moreover, the negative and significant coefficient of the interaction ORG-AGE to explain variance heterogeneity suggests this information should be mainly address to younger consumers.

Contrary to the organic attribute, Catalonian consumers show a strong preference for PDO extra virgin olive oil. Scarpa and Del Guidice (2004) arrived to the same conclusion in Italy. This results is consistent with prior expectations as results from the survey indicated that 9.26% of Catalonian consumers use to buy PDO extra virgin olive oil while less than 1% buy, occasionally organic olive oil. PDO extra virgin olive oil is very knowledgeable among Catalonian and Spanish consumers. There exist 28 PDO brands in Spain, five of them are located in Catalonia. Additionally, the production of this type of olive oil continues to grow being the domestic market its main destination and, to a lesser extent, the EU (Ruiz-Castillo, 2008). In any case, such positive preference is not homogeneous among Catalonian consumers. PDO olive oil is highly preferred by the older population with higher education levels and showing a higher purchasing frequency of olive oil. Results from Table 2.4 also suggest that consumer's preferences in Catalonia towards PDO olive oil are negatively correlated with factors affecting attitudes towards organic olive oil such us subjective norms, consumers' concerns related about the environmental benefits of organic olive oil and knowledge.

The price coefficient is significant and has a negative sign, becoming the most restrictive factor for purchasing extra virgin olive oil (Menapace et al., 2011; Parras-Rosa et al., 2008; Scarpa and Del Giudice, 2004). Moreover, the corresponding standard deviation is significant indicating relevant Catalonian consumers' preferences heterogeneity. The estimated parameters show that this negative utility is mitigated, to a certain extent, in consumers who are more likely to purchase organic olive oil. Furthermore, the statistically significance of the variance heterogeneity coefficients show that price heterogeneity within the model varies taking into account the effect of some socio-demographic factors such as the university education level (UNIV), age (AGE) and town size (TS). However, while the effect of two first variables is positive that of the last one is negative, indicating that older, more educated consumers and those living in rural areas are less sensitive with respect to price.

Apart from price, the origin is the most important attribute affecting consumers' preferences toward extra-virgin olive oil. This finding is consistent with previous studies on the importance of geographical origin in consumer decision making (Menapace el al., 2011; Schnettler et al., 2008; Scarpa et al., 2004; Scarpa et al., 2005). Catalan olive oils are preferred over other Spanish or imported oils, while olive oil produced in other Spanish regions is preferred over imported olive oil. The positive preference associated with the Catalan olive oil increases for consumers who are more likely to buy organic olive oil. This result was also found by Cicia et al. (2002) who evaluated the preferences of regular consumers of organic food towards the purchasing intention of extra virgin olive oil. They concluded that regular organic food consumers pay more attention to the origin of the product which is taken as a proxy of organic olive oil quality. Furthermore, and consistent with the results we discussed above a about the marginal utilities associated to PDO olive oil, older consumers with higher education level have a stronger preference for the local origin as well as for respondents who buy extra-virgin olive oil less frequently. In fact, results from the survey indicated that this consumer segment used to buy olive oil in large quantities from local cooperatives.

Finally, the lack of significance associated with the local brand attribute level (BCAT) indicates that in the case of extra-virgin olive oil, respondents are more interested in the origin of the product than in the origin of the brand, although this result could be related to the fact that many consumers do not acknowledge the origin of the brand (that is, whether the manufacturer is located or not in Catalonia). Results also show that, on average, consumers do not value private labels for this specific product.

Table 2.5 presents the moments of the WTP distributions derived from the four estimated models, as well as their confidence intervals. Results for the GHR–RPL model indicate that Catalonian consumers are willing to pay a 60% premium for a Catalan olive oil over an olive oil from another Spanish region and a 30% for a PDO extra virgin olive oil over the conventional counterpart. In contrast, the mean WTP for the organic attribute and imported olive oil are both negative. That is, consumers reveal that they have to be rewarded to shift from the conventional to organic olive oil, as well as, from purchasing olive oil of national origin to imported olive oil.

Table 2.5 Willingness to pay for the attribute levels

	RPL1		**RPL2**		**RPL3**		**GHR-RPL**	
	WTP (SD)[a]	**CI**[b]	**WTP**	**CI**	**WTP**	**CI**	**WTP**	**CI**
PDO	0.281 (0.048)	[0.191, 0.375]	0.238 (0.104)	[0.038, 0.438]	0.308 (0.048)	[0.216, 0.403]	0.306 (0.041)	[0.227, 0.388]
ORG	-0.748 (0.075)	[-0.910, -0.610]	-0.510 (0.133)	[-0.785, -0.264]	-0.886 (0.066)	[-1.018, -0.755]	-0.808 (0.056)	[-0.919, -0.70]
CAT	0.557 (0.051)	[0.458, 0.656]	0.488 (0.109)	[0.272, 0.712]	0.603 (0.051)	[0.499, 0.702]	0.590 (0.041)	[0.507, 0.675]
IMP	-0.671 (0.051)	[-0.778, -0.575]	-0.625 (0.058)	[-0.744, -0.519]	-0.678 (0.050)	[-0.775, -0.580]	-0.631 (0.047)	[-0.728, -0.547]

[a] Standard deviation
[b] Confidence interval at 5% significance level.

Additionally, results displayed in Table 2.6 from the non-parametric combinatorial test reveal that it is not possible to reject at the 5% level of significance that the estimated WTPs obtained from the four models are statistically similar for all attributes' levels except for the organic attribute level. With respect to the latter main differences exist when we compare the RPL2 model with the rest. This result indicates that when the heterogeneity around the mean is specified as a function of socio-demographic factors, the model tends to underestimate WTP values although in our study only significant differences have been found in relation to the organic attribute.

Table 2.6 Hypothesis test of equality WTPs across the treatments

Hypothesis	WTP_{PDO}	WTP_{ORG}	WTP_{OCAT}	WTP_{OIMP}
		p-values		
RPL1 vs RPL2	**0.357**	**0.064**	**0.279**	**0.280**
RPL1vs RPL3	**0.347**	**0.081**	**0.263**	**0.426**
RPL1 vs GHR-RPL	**0.341**	**0.255**	**0.311**	**0.288**
RPL2 vs RPL3	**0.256**	**0.005**	**0.172**	**0.203**
RPL2 vs GHR-RPL	**0.253**	**0.022**	**0.194**	**0.433**
RPL3 vs GHR-RPL	**0.494**	**0.184**	**0.417**	**0.247**

2.5. Concluding remarks

The main aim of this paper has been to assess the consumers' preferences towards extra-virgin olive oil in Catalonia. The methodological approach to tackle with this issue has been the extended Random Parameter Logit model (GHR-RPL) proposed by Greene et al. (2006) which overcomes some of the assumptions inherent in conventional discrete choice models by taking into consideration potential sources of heterogeneity in both the mean and the variance of the random parameter distributions. Both attitudinal factors and individuals' socio demographic characteristics have been considered as potential sources of preference heterogeneity around the means and the variances of random parameters estimates, respectively. Implications on the goodness-of-fit measures as well as on the moments of the willingness to pay haves been investigated by comparing the estimates from the GHR-RPL model against alternative specification of RPL models.

Results presented in this paper suggest a number of points. First, accounting for mean and variance heterogeneity in the random parameters generates better goodness-of-fit than other specification alternatives. The generated distribution from the GHR-RPL model is similar to other models but the estimated moments are different. Moreover, the specification of the heterogeneity around the mean and the variance of the estimated parameters as a function of both respondents' socio-demographic characteristics and their attitudinal factors, respectively, significantly increases the model statistical adequacy, and provides useful insights for policy analysis.

In terms of the empirical results, our study suggests that the most important attribute that affects consumers' preferences toward extra-virgin olive oil is the price, followed by the origin of the product, especially the local origin. The brand (label) has not been proved to be significant in this specific case. Noticeably, this study suggests that Catalan consumers perceive the organic olive oil attribute negatively, as they think that it is not worth paying a premium for a product that is healthy in nature (i.e. perceived as one of the main representative products of the so-called Mediterranean diet). On the other hand, the PDO certification is highly appreciated by the Catalonian.

When heterogeneity is taking into account, these results can be shaded. In fact, the negative utility associated to organic olive oil is moderated by individuals' environmental awareness or by the influence of subjective norms. On the contrary, attitudes towards the health benefits of organic olive oil do not seem to have a significant effect. The Catalonian origin of extra virgin olive oil is much more appreciated by older consumers with higher education level and being low-frequency buyers. Finally, the negative perception of price is less significant for higher educated and aged respondents.

Moments of WTP distributions have been calculated using a parametric bootstrapping approach to reduce the potential changes of sign that can take place in extreme values of such distributions. Results suggest that consumers are willing to pay a higher premium for a local olive oil followed by a PDO certified olive oil. The organic and the imported attributes are negatively evaluated, indicating that consumers would accept compensation or a price reduction for an olive oil with such characteristics.

In any case, this study has been based on the use of generic alternatives and hypothetical responses to choice experiment questions. Lusk and Schroeder (2004) showed that hypothetical choices could overestimate the marginal willingness to pay for extra-virgin olive oil. Therefore, results from this study could be extended to non-hypothetical environments in which consumers face choices involving real products and real money in a series of choice scenarios.

References

Allenby GM, Rossi, PE, 1998. Marketing models of consumer heterogeneity. J Econ 89: 57–78.

Ben-Akiva M, McFadden D, Train K, Walker, J, Bhat CA, Bierlaire M, Bolduc D, Börsch-Supan A, Brownstone D, Bunch DS, Daly A, De Palma A, Gopinath D, Karlstrom A, Munizaga MA, 2002. Hybrid choice models: Progress and challenges. Marketing Letters 13: 163-174.

Bhat CR, 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Trans Res Part B-Metho 37: 837-855.

Calatrava J, 2002. Actitudes del consumidor español respecto a los productos ecológicos. Consejería de Agricultura y Pesca de la junta de Andalucía.

Campbell D, Doherty E, Hynes S, Rensburg TV, 2010. Combining discrete and continuous mixing approaches to accommodate heterogeneity in price sensitivities in environmental choice analysis. Agricultural Economics Society Annual Conf, Edinburg (UK), Mar 29-31.

Chen MF, 2007. Consumer attitudes and purchase intentions in relation to organic foods in Taiwan: Moderating effects of food-related personality traits. Food qua and pref 18: 1008-1021.

Chung C, Boyer T, Han S, 2010. How many choice sets and alternatives are optimal? Consistency in choice experiments. Agribusiness 27: 114-125.

Cicia G, Del Guidice T, Scarpa R, 2002. Consumers' perception of quality in organic food: A random utility model under preference heterogeneity and choice correlation from rank-orderings. Brit Food J 104: 200-213.

De Wilde E, Cooke ADJ, Janiszewski C, 2009. Attentional contrast during sequential judgements: A source of the number-of-levels effect. J Mark Res 45: 437-449.

Ding M, Rajdeep G, John L, 2005. Incentive-Aligned Conjoint Analysis. J Mark Res 42: 67-82.

Gil JM, Tamburo LG, Sánchez M, 2002. Seguridad alimentaria y comportamiento del consumidor en España. Gobierno de Aragón departamento de Agricultura, Zaragoza, Spain. 45 pp.

Gracia A, Magistris T, 2007. Organic food product purchase behaviour: a pilot study for urban consumers in the south of Italy. Span J Agr Res 5: 439-451.

Gracia A, Magistris T, 2008. The demand for organic foods in the south of Italy: A discrete choice model. Food Policy 33: 386-396.

Greene WH, Hensher DA, 2013. Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. Applied Economics 45: 1897-1902.

Greene WH, Hensher DA, Rose J, 2006. Accounting for heterogeneity in the variance of unobserved effects in mixed logit models. Transportation Research B 40: 75–92.

Hensher DA, 2008. Empirical approaches to combining revealed and stated preference data: Some recent developments with reference to urban mode choice. Res in transp econ 23: 23-29.

Hensher DA., Greene WH, 2003. Mixed logit models: state of practice. Transportation 30: 33-176.

Hensher DA, Rose JM, Greene WH, 2005. Applied Choice Analysis: A Primer. Cambridge Univ Press, Cambridge. UK.

Hess S, Rose JM, 2009. Allowing for intra-respondent variations in coefficients estimated on repeated choice data. Transp Res Part B 43: 708-719.

Hess S, Rose JM, Bain S, 2010. Random scale heterogeneity in discrete choice models. The transportation research board 89th annual meeting, Washington DC (USA), Jan 10-14. pp. 25.

Hynes S, Hanley N, Scarpa R, 2008. Effects on welfare measures of alternative means of accounting for preference heterogeneity in recreational demand models. Amer J Agr Econ 90: 1011–1027.

Krinsky I, Robb AL, 1986. On Approximating the Statistical Properties of elasticities. Rev Econ and Stat 64: 715–19.

Lancaster KJ, 1966. A new approach to consumer theory. J Poli Econ 74: 132-157.

Lenk P, DeSarbo W, 2000. Bayesian inference for finite mixtures of generalized linear models with random effects. Psychometrika 65: 93–119.

Louviere JJ, 2001. Choice experiments: An overview of concepts and issues. Edward Elgar publishing press, Massachusetts.

Louviere JJ, Hensher DA. 1982. The design and analysis of simulated choice or allocation experiments in travel choice modeling. Transp Res Record 890: 11-17.

Lusk JL, Shroeder TC, 2004. Are choice experiment incentive compatible? A test with quality differentiated beef steaks. Ame J Agr Econ 86: 467-482.

McFadden D, 1986. The Choice Theory Approach to Market Research. Marketing Science 5: 275-297.

McFadden D, 1974. Conditional logit analysis of qualitative choice behavior. (New York Academic press, New York, USA.

Menapace L, Colson G, Grebitus C, Facendola M, 2011. Consumers' preferences for geographical origin labels: evidence from the Canadian olive oil market. Eur Rev Agr Econ 38: 193–212.

Moore R, 2008. Using attitudes to characterize heterogeneous preferences. American Agriculture Economics Association Annual Meeting, Orland, July 27-29.

Morey E, Rossmann KG, 2003. Using Stated-Preference Questions to investigate Variations in Willingness to Pay for Preserving Marble Monuments: Classic Heterogeneity, Random Parameters, and Mixture Models. J Cult Econo 27: 215-229.

Mtimet N, Albisu LM, 2006. Spanish wine consumer behavior: a choice experiment approach. Agribusiness 22: 343-362.

Parras-Rosa M, Vega-Zamora M, Gutiérrez-Salcedo M, 2008. La demanda de aceites de oliva ecológico. Una aproximación al diferencial de precios. IV Cong int de la RED SIAL, Argentine, october 27-31.

Poe GL, Giraud KL, Loomis JB, 2005. Computational Methods for Measuring the Difference of Empirical Distributions. Ame J Agr Econ 87: 353-365.

Revelt D, Train K, 1998. Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. Rev Econo Stat 80: 647-657.

Ruiz-Castillo B, 2008. Las denominaciones de orígenes protegidos y el aceite de oliva en España. Distribución y Consumo 102: 57-68.

Scarpa R, Del Guidice T, 2004. Market segmentation via mixed Logit: extra-virgin olive oil in urban Italy. J Agr Food Ind Org 2: 1-18.

Scarpa R, Thiene M, 2005. Destination choice models for rock climbing in the Northeastern Alps: a latent-class approach based on intensity of preferences. Land Economics 81: 426–444.

Scarpa R, Thiene M, 2011. Organic Food Choices and Protection Motivation Theory: addressing the psychological sources of heterogeneity. Food Qua Pref 22: 532-541.

Scarpa R, Thiene M, Train K, 2008. Utility in willingness to pay space: A tool to address confounding random scale effects in destination choice to the Alps. Ame J Agr Econ 90: 994-1010.

Schnettler B, Ruiz D, Sepulveda O, Sepulveda N, 2008. Importance of the country of origin in food consumption in a developing country. Food Qua Pref 19: 372–382.

Shen J, 2010. Latent class model or mixed logit model? A comparison by transport mode choice data. Applied Economics 42: 2915-2924.

Soler F, Gil JM, Sánchez M, 2002. Consumer's acceptability of organic food in Spain. Results from an experimental auction market. Brit Food J 104: 670-687.

Stolz H, Stolz M, Hamm U, Janssen M, Ruto E, 2011a. Consumer attitudes towards organic versus conventional food with specific quality attributes. NJAS-Wageningen J Life Sc 58: 67-72.

Stolz H, Stolz M, Janssen M, Hamm U, 2011b. Preferences and determinants for organic, conventional and conventional-plus products, the case of occasional organic consumers. Food Qua Pref 22: 772-779.

Street D, Burgess LB, 2007. The Construction of Optimal Stated Choice Experiments: Theory and Methods. Wiley-interscience press, New Jersey, USA.

Train K, 2003. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, UK.

Train K, Weeks M, 2005. Discrete Choice Models in Preference Space and Willing-to Pay space. In: Applications of Simulation Methods in Environmental and Resource Economics (Alberini A, Scarpa R, eds.). Dordrecht (The Netherlands), pp. 1–16.

Tsakiridou E, Mattas K, Tzimitra-Kalogianni I, 2006. The influence of consumer characteristics and attitudes on the demand for organic olive oil. J Int Food & Agri Mark 1: 23-31.

Van Loo EJ, Caputo V, Nayga JrRM, Meullenet JF, Ricke SC, 2011. Consumers' willingness to pay for organic chicken breast: Evidence from choice experiment. Food Qua Pref 22: 603-613.

Vega-Zamora M, Parras-Rosa M, Murgado-Armenteros EM[a], Torres-Ruiz FJ, 2013. The influence of the term organic of organic food purchasing behavior. Procedia-Social and Beh Sc 8: 660-671.

	Alternative "A"	*Alternative "B"*	*Alternative "C"*	*Alternative "D"*
System of production	***Extra-virgin olive oil with PDO***	***Conventional extra-virgin olive oil***	***Organic extra-virgin olive oil***	***None of them***
Origin of olive oil	***Spain***	***Catalonia***	***Imported***	
Brand	***Spanish Manufacturer***	***private label***	***Catalonia Manufacturer***	
Price	***3.70 €/liter***	***7.50 €/liter***	***6 €/liter***	

Figure 2.1 Example of choice sets

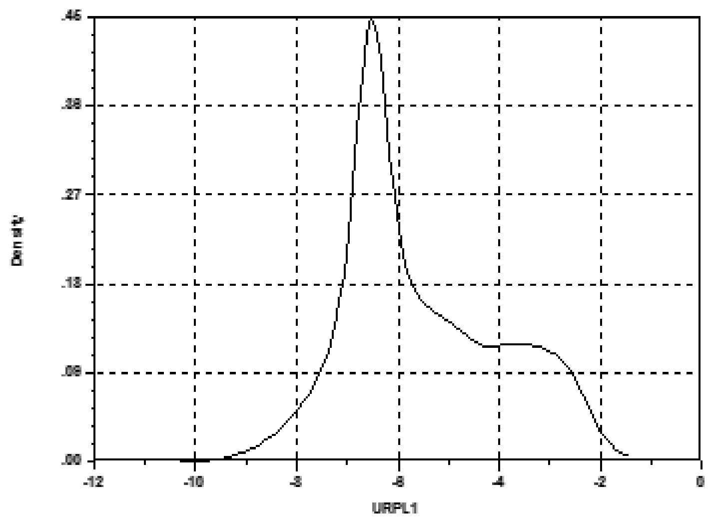

Figure 2.2 Kernel density estimates for marginal utility distribution RPL1
Mean = -5.535 STD = 1.543

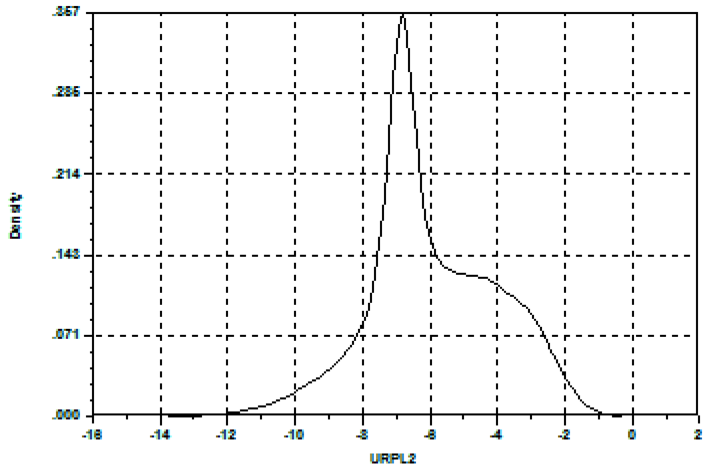

Figure 2.3 Kernel density estimates for marginal utility distribution RPL2
Mean = -5.947 STD = 1.906

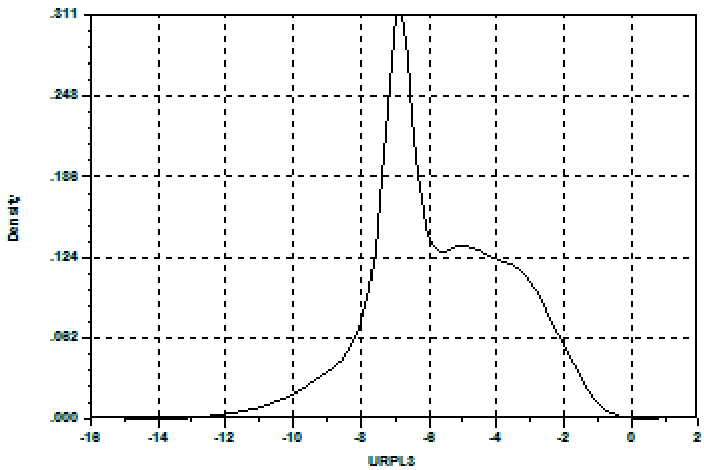

Figure 2.4 Kernel density estimates for marginal utility distribution RPL3
Mean = -5.693 STD = 2.045

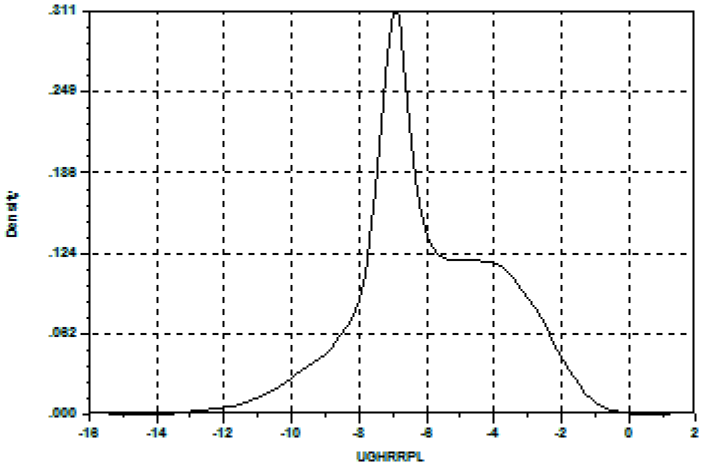

Figure 2.5 Kernel density estimates for marginal utility distribution GHR-RPL
Mean = -5.979 STD = 2.110

Appendix 2.1 Explanatory variables included in the estimated models

Empirical model factors		
Production System	PDO	Defined as an effect-coded dummy variable, it takes the value 1 if the olive oil is produced under a Protected Denomination of Origin (PDO), 0 if it is organic and -1 if it is conventional.
	ORG	Defined as an effect-coded dummy variable, it takes the value 1 if the olive oil is organic, 0 if it is PDO and -1 if it is conventional.
Origin of olive oil	CAT	Defined as an effect-coded dummy variable, it takes the value 1 if it is produced in Catalonia, 0 if it is imported and -1 if it produced in other Spanish region.
	IMP	Defined as an effect-coded dummy variable, it takes the value 1 if the olive oil is imported, 0 if is produced in Catalonia and -1 if it is produced in another Spanish region.
Brand	BCAT	Defined as an effect-coded dummy variable, it takes the value 1 if the olive oil is sold under a Catalonian manufacturer's brand, 0 if it has a private label and -1 if it is sold under another Spanish manufacturer's brand.
	PRIV	Defined as an effect-coded dummy variable, it takes the value 1 if the olive oil is sold under a private label, 0 if it is sold under a Catalonian manufacturer's brand and -1 if the manufacturer's brand is from another Spanish region.
Price	PRICE	A continuous variable
No option	NOP	A dummy variable that takes the value 1 if the respondent has chosen the alternative "none of them" and 0 otherwise.
Individual specific characteristics		
Gender	GEN	A dummy variable that takes the value 1 if the respondent is a woman and 0 otherwise.
Age	AGE	A dummy variable that takes the value 1 if the respondent have more o equal 50 years old and 0 otherwise.
Town size	TS	A dummy variable that takes the value 1 if the respondent lives in a town with over 10,000 inhabitants and 0 otherwise.
Education level	SEC	Defined as an effect-coded dummy variable, it takes the value 1 if the respondents have completed secondary school, 0 if the respondent has a university degree, and -1 otherwise.
	UNIV	Defined as an effect-coded dummy variable, it takes the value 1 if the respondent has a university degree, 0 if the respondent has completed secondary school, and -1 otherwise.
Purchase frequency	MONTH	Defined as an effect-coded dummy variable, it takes the value 1 if the respondent purchases olive oil monthly, 0 if is the respondent purchases olive oil every three months or more, and -1 if is the respondent purchases olive oil weekly.
	QUART	Defined as an effect-coded dummy variable, it takes the value 1 if the respondent purchases olive oil every three months or more, 0 if the respondent purchases olive oil monthly and -1 if is the respondent purchases olive oil weekly.

Chapter 3

The effect of personality traits on consumers' preferences for extra virgin olive oil

3.1. Introduction

As an essential element of the Mediterranean diet, olive oil is a chief food product for most Mediterranean countries. Its relevance corresponds to its ancient tradition as well as its social and agro-environment dimension. The quality and sensory characteristics of extra virgin olive oil are the result of different factors, such as the environment and cultural practices. As a result, a large variety of olive oil types and specifications exist currently in the international market, which spurs strong competition among market agents to seek new horizons and marketing strategies. Olive oil quality specifications (origin,…) provide confidence to consumers regarding its attributes, which makes them more dependent on labelling information (Scarpa and Del Giudice, 2004). Taking into account the increasing amount of information available on olive oil labels, consumers can focus on different pieces of information and develop heterogeneous perceptions that lead them to different purchase behaviours (Menapace et al., 2011; Philippidis el al. 2002). Thus, a better understanding of the consumers' selection process that underlies their motivations to buy olive oil has become a key factor for strategic success.

However, gaining this understanding is not an easy task, as an individual's final choices depend not only on the extrinsic and intrinsic attributes of olive oils but also on non-food effects (Chen, 2007). In this context, some studies have highlighted the importance of human psychological factors to enhance the behavioural representation of the selection process (Ben Akiva et al., 2002; Johansson et al., 2006; Scarpa and Thiene, 2011). However, only a few studies have investigated the potential effect of personality traits on shaping consumer behaviour (Ajzen, 2005; Chen, 2007; Eertmans et al., 2005).

This paper aims to investigate the effects of food-related personality traits, lifestyle orientations and purchase habits on shaping an individual's purchase intentions towards olive oil in Catalonia (North-East Spain). Special attention is paid to the organic attribute. Spain is the first producer and exporter country of extra-virgin olive worldwide. Additionally, olive oil constitutes a fundamental component of the Spanish diet. As a consequence, the vast majority of Spanish consumers are knowledgeable about this product, and all of them are aware of market prices and product characteristics. The market value for organic olive oil was 11 million Euros in 2012 (MAGRAMA, 2013). Catalonia is the second region within Spain in terms of total olive oil consumption (with a per capita consumption of 9.93 litres in 2011). It also occupies the second position in relation to the consumption of organic olive oil (13% of the Spanish total consumption in value) after Madrid. Moreover, the population of Catalonia is quite heterogeneous, with an adequate combination of urban (Barcelona is the second largest town in Spain) and rural environments, which seems to be adequate for the purpose of this study.

To achieve the mentioned objective, a discrete choice modelling approach that accounts for preference heterogeneity has been adopted. This type of model has received a significant amount of attention in the recent literature (Campbell et al., 2010; Greene and Hensher, 2013). Different methodological approaches have been adopted to account for this heterogeneity preference: 1) the use of segmentation strategies (Shen 2010); 2) the inclusion of interaction effects to explain sources of heterogeneity (Mtimet and Albisu, 2006); 3) the use of random parameter estimates, assuming preference coefficients to be randomly distributed across individuals (Revelt and Train, 1998); and 4) the combination of interaction effects and random parameters (Hensher and Greene, 2003) or segmentation strategies and random parameters (Greene and Hensher, 2013). In all cases, the heterogeneity of preferences is assumed to be a function of the observed variables.

Recently, Ben Akiva et al. (2002) introduced the hybrid choice model (HCM). The HCM model extends the normal discrete choice modelling by defining an individual's utility function as a function of observed explanatory variables, such as product attributes and respondents' socio-economic characteristics, and including latent variables that can reflect consumers' psychological factors, personality traits or attitudes. Previous empirical applications of the HCM, mainly in the field of transport economics (Bolduc et al., 2008; Yáñez et al., 2010), have shown that 1) the inclusion of latent variables significantly improved the goodness-of-fit of the model and 2) psychological factors better contributed to capture a consumer's preference heterogeneity. One of the main contributions of this study is that it constitutes one of the first attempts to apply the HCM approach to food marketing.

Traditionally, the HCM model has involved two steps. In the first step, latent variables (i.e., food-related personality traits, lifestyles or purchase habits, among others) are derived from observed indicators via a "multiple-indicator, multiple cause" model (MIMIC) used to relate latent individual traits to observable determinants. In the second step, the latent variables are incorporated into the discrete choice model as explanatory variables to estimate a multinomial logit model, which ignores the effect of consumers' heterogeneity. In this context, this paper extends the existing literature in two ways. First, this paper does not merely obtain latent variables from observed indicators; it estimates hierarchical relationships between latent variables using a structural equation model (SEM), which provides a better insight into the consumers' decision-making process. Second, the estimation of the HCM requires integrating the variation of the latent variables within the basic framework of multinomial choice models (Ashok et al., 2002). Estimating a random parameter logit (RPL) model considering the latent variables as random parameters solves this problem (Yáñez et al., 2010). Thus, this study will apply the HCM in a panel data context constructed from the repeated choice data set while considering sample heterogeneity.

The paper is structured as follows. The second section includes the conceptual framework and hypothesis definition of the SEM model. In the third section, we illustrate the theoretical specification of the HCM. In the fourth and fifth sections, we explain the empirical setting and the results. We conclude by addressing some limitations and providing recommendations for further research.

3.2. A conceptual model for organic olive oil purchasing intention

One of the main advantages of the HCM is that it allows users to better characterise the structure of the selection process (Rungie et al., 2011). However, instead of merely defining some latent variables (that represent personal traits or individuals' lifestyles) from the observed indicators of individuals, this study also considers the potential structural relationships that can exist among latent variables. To tackle with this issue, a conceptual model has been specified based on the theory of planned behaviour (TPB) (Azjen, 2005). The TPB considers that the intention to perform behaviours can be predicted with high accuracy from attitudes towards the behaviour, subjective norms and perceived behavioural control. Figure 3.1 shows the conceptual model used in this study from which the following hypothesis will be tested.

Attitudes are formed based on an individual's beliefs about the specific food product considered, which reflects the extent to which a person positively or negatively evaluates the behaviour in question. For example, organic food is perceived as more healthy, natural, nutritious and sustainable than its conventional counterpart (Stolz et al., 2011). This perception results in a more positive attitude towards organic food, which is believed to be positively related to the intention of purchasing organic food (Chen, 2007)

Hypothesis 1. The consumer's positive attitude towards organic olive oil positively correlates with his (her) intention to purchase organic olive oil.

A large number of variables can influence people's beliefs and attitudes towards organic olive oil, such as personality traits, ethnicity, emotion, mood, etc. (Ajzen, 2005). Personality traits play an important role in predicting and explaining human behaviour. Chen (2007) showed that food-related personality traits, defined as food involvement (the level of importance of food in a person's life), exert a positive effect on a consumer's attitude towards organic foods. Bell and Marshall (2003) argued that the level of food involvement was a significant discriminating factor between food items in sensory evaluations. Therefore, the following hypothesis can be proposed:

Hypothesis 1a. Consumers that show a higher level of food involvement are expected to have a more positive attitude towards organic olive oil.

Food-related personality traits compromise people in food-related activities, such as their procurement, preparation, cooking, etc. (Goody, 1982). The recent literature shows that cooking skills play a significant role on dietary changes to promote healthy eating (Van den Horsk et al., 2010). Hence, due to the importance of olive oil in the Spanish diet, cooking skills are hypothesised to positively affect the attitude towards organic olive oil. Thus, the following hypothesis can be proposed:

Hypothesis 1b. Consumers with better cooking skills are expected to have a more positive attitude towards organic olive oil.

An individual's lifestyle is reflected in his/her personality and self-concepts, which are determined by his/her interests, opinions, activities, etc. Moreover, attitudes, behavioural tendencies and habits are derived from changes in lifestyles (Chen, 2009). Shaharudin et al. (2010) showed that consumers' lifestyles were related to their attitude towards the purchasing of organic food. Moreover, Krishnan (2011) confirmed that consumers' lifestyles strongly related to their purchasing brands.

Hypothesis 1c. Consumers with ordered lifestyles have a more positive attitude towards organic food.

Over the last decade, food scares (BSE, dioxins, foot and mouth disease, etc.) have reshaped food consumer behaviour to a certain extent. Consumers are now more concerned about food safety issues (Chen, 2007). Moreover, according to Chen (2009), a healthy consumption lifestyle, attitudes towards organic food and the intention to purchase organic food appear to significantly correlate. Therefore, the following hypothesis can be proposed:

Hypothesis 1d. The healthiness of a consumer's lifestyle positively correlates with his (her) positive attitude towards organic foods.

According to the TPB, perceived behaviour control represents an individual's perceived ease or difficulty in performing a particular behaviour. It is assumed to be determined by the total set of accessible control belief (Ajzen, 2005). In the framework of organic food, perceived control would include the effects of both external (such as time spent, availability, recognition (labelling), confidence, etc.) and internal variables (such as skills, knowledge, abilities, habits, etc.), which consumers believe can influence their judgment of risks and benefits associated with such products (Ajzen, 2005; Chen, 2007).

Hypothesis 2. When consumers perceive more behavioural control over purchasing organic food, the intention to purchase it will increase.

Hypothesis 2a. A better knowledge about certification (labels) and the benefits of organic olive oil increases consumers' behavioural control.

Consumer's habits (persistence in doing what somebody is accustomed to do) can simplify behaviour, as many decisions become routine and can be adopted with minimal conscious control. However, this factor is not easily measured (Ajzen, 2005). In this study, we have assessed the effect of "purchasing habits" by considering two latent variables, "price involvement" and "quality involvement". The first variable involves the impact of price and price promotions, and the second involves the impact of food quality on consumers' purchasing habits because both have been shown to be important in consumer buying behaviour (Mann et al., 2012; Menapace et al., 2011).

Hypothesis 2b. As consumers are less sensitive to price and price promotions, their perceived behavioural control increases.

Hypothesis 2c. As food quality plays a more important role in consumers' food habits, their perceived behavioural control increases.

Finally, the third component of consumer intention is subjective norms. It reflects the degree of social pressure (surrounding the consumer: family, friends, etc.) felt by individuals with regard to their behaviour (Chen, 2007).

Hypothesis 3. Social pressure positively affects a consumer's purchasing intention related to organic olive oil.

3.3. Theoretical background of hybrid choice model

As mentioned above, the sequential estimation method of the HCM involves two steps: the definition of latent variables and their introduction to the discrete choice model as further explanatory variables (Ashok et al., 2002). In this paper, we have also explicitly considered the potential structural relationships among latent variables following the conceptual model specified above, which has been estimated as a structural equation model (SEM) (Jöreskov and Sörbomm, 1996). This section is structured following the two steps of the modelling process.

3.3.1. Structural equation model specification

The SEM consists of three main types of relationships (Jöreskov and Sörbomm, 1996). First, the identification of latent variables requires the definition of several observed indicators introduced as questions within a stated preference survey. Therefore, a measurement model is identified after performing a confirmatory factor analysis. The outcome of the measurement model relates observed indicators to the exogenous latent variables:

$$x = \boldsymbol{\Lambda_x}\ \boldsymbol{\xi} + \boldsymbol{\delta} \tag{1}$$

where x is a q×1 vector of observed exogenous variables; $\boldsymbol{\Lambda_x}$ is a q×n matrix of coefficients of the regressions of x on $\boldsymbol{\xi}$, which is an n×1 random vector of latent independent variables; and $\boldsymbol{\delta}$ is a q×1 vector of error terms in x. Furthermore, $\boldsymbol{\delta}$ is assumed to not correlate with $\boldsymbol{\xi}$.

Conversely, observed indicators are related to endogenous constructs:

$$y = \boldsymbol{\Lambda_y}\ \boldsymbol{\eta} + \boldsymbol{\varepsilon} \tag{2}$$

where y is a p×1 vector of observed indicators; $\boldsymbol{\Lambda_y}$ is a p×m matrix of coefficients of the regressions of y on $\boldsymbol{\eta}$, which is an m×1 random vector of latent dependent variables; and $\boldsymbol{\varepsilon}$ is a p×1 vector of error terms in y. Furthermore, ε is assumed to not correlate with η.

The third equation specifies the causal relationships that exist among both exogenous and endogenous latent constructs.

$$\eta = \beta\ \boldsymbol{\eta} + \boldsymbol{\Gamma}\ \boldsymbol{\xi} + \boldsymbol{\zeta} \tag{3}$$

where β is an m×m matrix of coefficients of the $\boldsymbol{\eta}$ vector of dependent variables in the structural relationships; $\boldsymbol{\Gamma}$ is an m×n matrix of coefficients of the $\boldsymbol{\xi}$ vector of independent variables in the structural relationship; and $\boldsymbol{\zeta}$ is a m×1 vector of errors.

The Full SEM model is estimated with robust maximum likelihood (RML), due to a potential lack of normality.

3.3.2. Integrating latent variables into discrete choice model specification

The application of the HCM implies the design of a choice experiment, which is based on both the random utility theory (RUT) (McFadden, 1974) and the Lancaster consumer theory (Lancaster, 1966). The RUT assumes that the utility provided by alternative j to individual i (i=1,…N) is given by the following:

$$U_{ij} = V_{ij} + \varepsilon_{ij}, \tag{4}$$

where V_{ij} is a deterministic component, which is a function of alternative product characteristics (Lancaster, 1966) and individuals' latent characteristics; and ε_{ij} is the stochastic or non-observed component. The deterministic component, V_{ij}, can be represented as follows:

$$V_{ij} = \beta_{ik} * X_{kj} + \beta_{il} * \eta_{il} \tag{5}$$

where X_{kj} is the vector of attributes related to alternative j; β_{ik} is the vector of marginal utilities of the individual i related to the k attributes in alternative j; η_{il} is the vector of latent characteristics corresponding to the i-th individual; and β_{il} is the vector of marginal effects of η_{il} on the utility function of the i-th individual.

The probability of consumer i choosing the alternative j out of the total set of options is defined as follows:

$$P_{ij} = Prob\left[U_{ij} > U_{ik}\right] = Prob\left[V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik}\right] \ \forall j \neq k \ \in C_n \tag{6}$$

where C_n is the choice set. Assuming that the stochastic component, ε_{ij}, follows the type I Extreme Value distribution, the probability of consumer i choosing option j from a specific choice set (C_n) is defined as follows:

$$P_{ij} = \frac{\exp(\mu V_{ij})}{\sum_{k=1}^{J} \exp(\mu V_{ik})} \qquad \forall j \in C_n \tag{7}$$

The sequential estimation method of the HCM requires integrating over the variation of latent variables within the basic framework of multinomial choice models (Ashok et al., 2002). Yañez et al. (2010) showed that this integration could be attained by estimating an RPL model that considers the latent variables as random parameters.

Under the RPL model, the probability that individual "i" chooses alternative "j" in a particular choice set C_n is given by the following:

$$Prob_i\{j \ is \ chosen\} = \int L_{ij}(\beta_{ij}) f(\beta_i/\theta)\, d\beta_i, with \ j \in C_n \qquad (8)$$

where $f(\beta_i/\theta)$ is the density function of the β_i coefficients and θ refers to the moments of the parameter distributions, which can take any specified form, such as normal, lognormal, triangular, uniform, etc. Moreover,

$$L_{ij}(\beta_{ij}) = \frac{\exp(V_{ij}(X_{ij}, \eta_{il}, \beta_i))}{\sum_{k=1}^{J} \exp(V_{ij}(X_{ij}, \eta_{il}, \beta_i))} \qquad (9)$$

Furthermore, the parameter estimates β_{ij} are defined as follows to capture additional non-observed variations and to better explain preference heterogeneity among individuals (Hensher, 2005):

$$\beta_{ij} = \beta_j + \delta_j Z_i + \sigma_j \vartheta_{ij} \qquad (10)$$

where β_j is the sample-mean for the alternative j; ϑ_{ij} is the individual specific heterogeneity, with mean zero and standard deviation equal to one (Hensher and Greene, 2003); and Z_i is a set of choice invariant characteristics that produce individual heterogeneity in the means of the randomly distributed coefficients, such as individual-specific characteristics.

Because the resulting model is specified to include both fixed and random coefficients, the simulated maximum likelihood (SML) technique provides a faster and easier way to estimate the individual choice probabilities (Ben Akiva et al., 2002). According to Train (2003), the simulation proceeds in three steps for any given value of θ. First, a value of β_i is drawn from $f(\beta_i/\theta)$ (β_i^r with r = 1….R[8]). Second, the logit, $L_{ij}(\beta_i^r)$, is calculated from this draw. Finally, steps 1 and 2 are repeated, and the obtained results are averaged. This average is the simulated probability:

$$\widehat{P_{ij}} = \frac{1}{R}\sum_{r=1}^{R} L_{ij}(\beta_i^r) \tag{11}$$

where R is the number of draws. The simulated probabilities are inserted into the log-likelihood function to give a simulated log-likelihood (SLL):

$$SLL = \sum_{i=1}^{I}\sum_{j=1}^{J} d_{ij}\, Ln\widehat{P_{ij}} \tag{12}$$

where d_{ij}=1 if i chooses j and d_{ij}=0 otherwise. The maximum simulated likelihood estimator, (MSLE), is the value of θ that maximises SLL.

3.4. Methods and empirical setting

3.4.1. The survey

The data used in this study were obtained from a survey carried out on a representative sample of the Catalonian (North-East Spain) population with quotas by postal code. Information was gathered from 401 persons. Participants were recruited using two filters: 1) they have bought olive oil virgin extra in the last three months; and 2) they are responsible for shopping within the household. Face-to-face interviews were conducted in September 2009 at different shopping hours and different types of food retail stores. The questionnaire consisted of four major blocks. The first block was designed to elicit information on respondents' purchasing and consumption habits about different types of olive oil. The second and third blocks were reserved for the measurement scales and the indicators related to the three main determinants of the TPB and to obtain information about socio-demographic characteristics and consumers' personality traits and lifestyles. All indicators were measured using eleven-

[8] Halton draws were used because they have been shown to provide more efficient distributions for numerical integration compared to random draws (Bhat, 2003).

point Likert scales (from zero to 10, where zero indicates total disagreement and 10 total agreement)[9]. The last block included the choice experiment task.

3.4.2. The choice experiment design

Four attributes (price, production system, the origin of the product and the origin of the brand) with three levels each were used in the experiment design (Table 3.1). The attribute and attribute levels were determined based on a three-step qualitative study: 1) a literature review regarding the consumer behaviour of organic and/or extra virgin olive oil; 2) four focus groups of eight people each to identify main consumption patterns and attitudes towards extra virgin olive oil, with special attention to the organic attribute; and 3) observation in retail outlets to identify real prices and informal interviews about reasons of choosing a specific product.

Table 3.1 Attributes and attributes levels in the Choice Experiment

Attributes	***Levels***
Production system	Conventional Protected Denomination of Origin (PDO) Organic
Origin	Spain Catalonia Imported
Brand	Spanish manufacturer Catalonia manufacturer Private label
Price	**3.70 €/l** **6 €/l** **7.5 €/l**

Considering the number of attribute levels, a total of 81 (3^4) hypothetical bottles of extra virgin olive oil were obtained. Therefore, an orthogonal factorial design was generated that resulted in nine product profiles and nine choice sets. We followed Street and Burgess (2007) to obtain a 100% efficient main effects design. Each respondent was forced to choose among

[9] Respondents can easily understand this scale, as the grading system at Spanish schools is based on a similar system.

three different types of extra virgin olive oil (alternatives A, B and C) in each choice set. These types were generated from different attribute level combinations, plus a fourth alternative (alternative D) that reflected the “no buying” scenario. Figure 3.2 provides an example of a choice set.

3.5. Results and discussions

3.5.1. Sample characteristics

As mentioned above, a total of 401 respondents completed the survey. Consistent with MAGRAMA (2008), approximately 80% of respondents were women. Approximately 70% of the respondents were married, and the average age of the sample was 49 years old (with a standard deviation of 15.39). The average household size was three members. Furthermore, 35% were households with one or more members younger than 18 years old, and only 14% contained members with children under six years old. Regarding the education level, 27.3% of respondents only completed primary school, while 46.8% completed secondary studies or professional education. Finally, regarding the geographic distribution of the sample, 40% came from Barcelona (the Catalonian capital), while 60% came from the rest of the Catalonian region.

Consistent with Jiménez-Guerrero et al. (2012), some descriptive results of the survey suggest that most respondents usually purchase olive oil virgin extra , but only 9.25% of the respondents search for protected denomination of origin (PDO) extra virgin olive oil. Olive oil is normally purchased weekly or every two weeks, although a significant percentage of respondents (nearly 30%) purchase it monthly or quarterly (in many cases directly from the farmer producer or the cooperative). The consumption of organic olive oil is marginal (less than 0.6% buy it regularly). Among the reasons for not buying organic olive oil, respondents highlighted the high price, the lack of availability in the supermarket where they are accustomed to shopping or the lack of information about organic food.

3.5.2. The structural equation model (SEM): Consumer's purchasing intentions

Following the traditional procedure to estimate an SEM, a confirmatory factor analysis (CFA) was first carried out for the entire set of constructs. Six "personality latent variables" (ordered life style, healthy life style, price involvement, food quality involvement, food involvement and cooking skills) and five "behavioural latent factors" (attitude, behavioural control perception, purchase intention, knowledge and subjective norms) were obtained (Appendix 3.1 and Appendix 3.2). Standardised factor loading estimates were all significant and above the recommended value of 0.7 (Hair et al., 1999). The main parameters to test the robustness of the construct, following Kline (2005), appear to show good results for almost all constructs. The internal consistency of reliability of each construct reached an acceptable Cronbach alpha of over 0.7, and the composite reliabilities were greater than 0.7, except for the factor "Healthy life style", which was 0.6. Nevertheless, we chose to retain this factor in our model.

The SEM was estimated in the second step. Table 3.2 summarises the estimation results and the main goodness-of-fit measures. The model meets the accepted goodness-of-fit criteria according to Hair et al. (1999), and Kline (2005): 1) the normed Chi-square (NC) is smaller than 3; 2) the value for the root mean square error of approximation (RMSEA) is 0.065 (lower than 0.8); 3) regarding the incremental fit index, the comparative-fit-index (CFI) is 0.952, which exceeds the value guidelines in the literature (0.90); 4) the normed-fit-index (NFI), non-normed-fit-index (NNFI) and relative fit-index (RFI) are all above 0.9, indicating that the conceptual model adequately fits the data; and 5) the adjusted R^2 values are reasonably high for this type of model.

The results from Table 3.2 indicate that both consumers' social pressure (subjective norms) and their perceived behaviour control positively affect consumers' intentions to purchase organic olive oil. Therefore, the second and third hypotheses are supported by our data, which is consistent with Chen (2007). However, the first hypothesis is not supported. Attitudes are negatively related to organic olive oil purchasing intention. This result is not surprising and is related to the consumer's positive perception in Spain (and we also assume in Catalonia) about the healthiness of the extra virgin olive oil, irrespective of the type of production system (organic or conventional) (Calatrava, 2002 and Vega-Zamora et al., 2011).

Table 3.2 Results from the Structural Equation Model (SEM) to explain consumer's purchasing intentions towards organic olive oil

Structural relationships	*Parameter Estimate*	*Std error*	R^2	*Goodness of fit statistics*
Attitude → Food Involvement	0.299***	0.0653	0.329	
Attitude → Healthy Life Style	-0.0784	0.0701		
Attitude → Ordered Life Style	0.384***	0.0825		$\chi 2$ = 2021.270
Attitude → Cooking Skills	0.033	0.0575		df = 741
Perceived Behavioural Control → Knowledge	0.248***	0.0655	0.318	NC = 2.727 < 3 RMSEA = 0.0658 < 0.08
Perceived Behavioural Control → Price Involvement	0.234***	0.0549		CFI = 0.952 > 0.90 NFI = 0.926 > 0.90
Perceived Behavioural Control → Quality Involvement	0.491***	0.0532		NNFI = 0.946> 0.90 IFI = 0.952 > 0.90
Purchase intention → Subjective Norm	0.167***	0.0351	0.623	RFI = 0.918 > 0.90
Purchase intention → Attitude	-0.127***	0.0388		
Purchase intention → Perceived Behavioural Control	0.772***	0.0559		

Notes : ***p<0.01;**p<0.05; *p<0.1

Furthermore, the results show that only the variables "food involvement" and "ordered lifestyle" do positively affect attitudes, which supports hypotheses 1.a and 1.c. According to Chen (2007) and Bell and Marshall (2003), consumers with higher food involvement personality traits hold a more positive attitude towards organic food and have capabilities to better discriminate between healthier foods. Additionally, the results reveal that an ordered lifestyle seems to enhance an individual's attitude towards organic olive oil. Gracia and Magistris (2008) arrived at the same result, which suggested that consumers attempting to follow an ordered life are more likely to develop environmentally friendly attitudes and follow a healthier diet in which olive oil plays an important role.

On the contrary, the relationships among attitudes, cooking skills and healthy lifestyles are not significant (hypotheses 1.b and 1.d are not supported). In both cases, this result is related to the perception of conventional olive oil as a healthy product, which plays an important role in the Mediterranean diet. Organic olive oil is not perceived as healthier than its conventional counterpart.

The results indicate that "knowledge", "food quality involvement" and "price involvement" significantly and positively affect a consumer's perceived behavioural control, which supports hypotheses 2.a and 2.c but rejects hypothesis 2.b. Although the standardised corresponding factor loading of "price involvement" was significantly different from zero, its positive coefficient led us to reject its associated hypothesis (2.b). This finding is consistent with Eertmans el al. (2005), who stated that the price negatively related to a healthy diet. Moreover, respondents consider price to be the main barrier to purchasing organic food, as mentioned above.

3.5.3. The choice model: consumer's preferences for olive oil attributes

The second step in the HCM consists of estimating an RPL model that incorporates latent variables (LV) obtained from the SEM. The estimated utility function includes all attribute levels defined as coded effects, except the price attribute, which is introduced as a continuous variable as well as latent variables. Socio-demographic variables, such as gender (GEND), age (AGE) and town size (TS), are defined as dummy variables (1 representing women, age less than 50 years and town size over 10000 inhabitants, respectively). The Education Level contains three categories; thus, two dummy variables were defined: university degree (UNIV) and completed secondary school (SECOND), which were defined as coded effects for which the primary school comprised the base level. Finally, all random parameters were assumed to be normally distributed according to the Wald test (Hensher el al., 2005).

Table 3.3 shows the estimated parameters from the RPL model. The no-option coefficient is negative and significant, which indicates that most of the respondents attempted to participate in the choice experiment by choosing one of the proposed olive oil alternatives instead of the no-option. The results also reveal that the organic attribute generates a disutility to consumers, while the most preferred olive oil is the one produced under a Protected Designation of Origin (PDO). This result, as mentioned above, can be explained by the fact that olive oil is already perceived as a healthy product in Spain, as it occupies a prominent position in the Mediterranean diet. In line with Calatrava (2002), the organic attribute does not add any additional value to Spanish consumers. This finding contradicts the results reported in other studies, such as Gracia and Magistris (2008) for Italy, Soler et al. (2002) and Vega-Zamora et al. (2013) for Spain or Tsakiridou el al. (2006) for Greece. However, consumers were only forced choose between organic olive oil and its conventional counterpart in all mentioned studies, whereas we have considered the trade-offs not only with other olive oil attributes but also with other attribute levels within the production system (i.e., PDO) in our study. Moreover, Catalan and Spanish consumers are not sufficiently concerned about environmental issues. Therefore, environmental concerns are not a key factor in consumer’s food choices, especially in the case of olive oil (Vega-Zamora et al., 2011).

Contrary to the organic attribute, Catalonian consumers show a strong preference for PDO extra virgin olive oil. PDO extra virgin olive oil is well known among Catalonian and Spanish consumers. Twenty-eight PDO brands exist in Spain, and five of them are located in Catalonia. Additionally, the production of this type of olive oil continues to grow; the domestic market and, to a lesser extent, the EU are its main destination (Ruiz-Castillo, 2008).

As expected and consistent with the previous literature, the price parameter is negative and significant (Menapace et al., 2011; Vega-Zamora et al., 2011). The local origin of olive oil plays an important role in shaping consumer’s preferences in Catalonia. Catalan olive oils are preferred over other Spanish or imported oils, while olive oil produced in other Spanish regions is preferred over imported olive oil, as in Jiménez-Guerrero et al. (2012). In contrast, the brand did not significantly impact consumers' utilities, which indicates that respondents are more interested in the origin of olive oil virgin extra than in the origin of the brand, although this result could be related to the fact that many consumers do not acknowledge the origin of the brand (that is, whether the manufacturer is located or not in Catalonia). The results also show that consumers do not value private labels for this specific product on average.

Interestingly, almost all personal trait latent variables (except ordered lifestyle) significantly affected the respondents' utilities towards extra virgin olive oil (Table 3.3). In line with previous results, we note that the sign of the variable "healthy lifestyle" is negative and significant. Consistent with previous results about the organic attribute, a healthy lifestyle is not related to the selection of olive oil, although healthy lifestyles may be conducive of healthier food choices (Losasso et al., 2012). In Catalonia, olive oil is perceived as a key determinant of the traditional Mediterranean diet, which is independent of the healthy or unhealthy typology of consumer's diets. A similar conclusion can be reached when we observe the negative sign of the variable "cooking skills", which reflects the general acceptance of this product when preparing food. Indeed, people with higher cooking skills seem to look for alternatives to olive oil virgin extra.

The other three variables, Food Involvement", "Price Involvement" and "Quality Involvement", positively affect the consumer's utility associated with extra virgin olive oil. A large number of extra virgin olive oil references are available in Catalonian markets, which can accommodate a broad range of food consumers' strategic behaviours. People looking for good prices can easily accommodate their preferences either by buying directly from the producer or cooperative (30% of our sample) or by choosing a promoted product at the retail outlet. Nevertheless, those looking for quality can also easily accommodate their expectations.

Table 3.3 Estimated parameters from the Random parameter Logil (RPL)

Parameters	***RPL***	***Standard error***
Conventional (CONV)[1]	1.280	----
Denominated Origin Protected (DOP)	0.251***	0.039
Organic (ORG)	-1.531***	0.253
Spanish origin (OSP)[1]	0.178	----
Catalan origin (OCAT)	0.490***	0.036
Imported origin (OIMP)	-0.668***	0.045
Spanish manufacturer (MSP)[1]	0.074	----
Catalan manufacturer (MCAT)	-0.005	0.050
Private brand (PRB)	-0.069	0.055
Price	-0.868***	0.027
No option (NOP)	-3.265***	0.818
Orderly lifestyle (OLS)	-0.240	0.515
Healthy lifestyle (HLS)	-0.820**	0.282
Price Involvement (PIN)	1.587***	0.430
Quality involvement (QIN)	1.505**	0.537
Food involvement (FIN)	1.022**	0.463
Cooking-Skills (COS)	-2.408***	0.435
	standard deviations	Standard error
DOP	0.410***	0.032
ORG	0.733***	0.049
OCAT	0.765***	0.034
Price	0.794***	0.030
OLS	0.261***	0.024
HLS	0.549***	0.035
PIN	0.012	0.012
QIN	0.504***	0.041
FIN	Fixed Parameter	----
COS	0.149**	0.049
Parameter-Variable	*Heterogeneity in mean*	Standard error
ORG-ATT	0.276***	0.039
ORG-BCP	-0.093**	0.041
ORG-SBN	0.190***	0.033
OLS-SECOND	-0.511**	0.239
OLS-UNIV	-0.353	0.323
OLS-GEND	-0.854*	0.469
OLS-TS	1.804***	0.449
HLS-SECOND	0.661***	0.155
HLS-GEND	1.002***	0.243
HLS-TS	-2.070***	0.284
HLS-AGE	1.198***	0.230
PIN-UNIV	-0.881**	0.290
PIN-GEND	-1.198**	0.375
PIN-TS	0.779**	0.347
PIN-AGE	-0.491*	0.278
QIN-SECOND	1.820***	0.287
QIN-UNIV	-0.761**	0.382
QIN-GEND	-1.646**	0.501
QIN-TS	-1.072*	0.583
FIN-SECOND	-1.635***	0.247
FIN-UNIV	0.730**	0.352
FIN-GEND	0.964**	0.384
FIN-AGE	-2.241***	0.384
COS-SECOND	0.405*	0.233
COS-UNIV	1.425***	0.298
COS-GEND	1.009**	0.427
COS-AGE	2.689***	0.413
Goodness-of-fit		
L-likelihood	-2903.046	
R2 adjs	0.41527	

Notes : ***p<0.01;**p<0.05; *p<0.1; [1] Base level; (SE): Standard Error; Gender (GEND), age (AGE) and town size (TS), are defined as dummy variables (1, representing women, age lower than 50 years, and town size over 10000 inhabitants, respectively). Education is defined by two dummy variables: university degree (UNIV) and completed secondary school (SECOND)

Table 3.3 (middle part) shows that the standard deviation of all relevant attributes and personal traits are significant, which indicates the heterogeneity in the preferences of relevant Catalonian consumers. The negative effect of healthy lifestyles on consumers' preferences is not homogeneous across the sample. In fact, the negative coefficient becomes positive if we consider the interaction effect for women and younger people. The negative effect is mitigated for respondents that have completed secondary school but increases for people living in larger towns. Its negative effect on cooking skills is mitigated for women and well-educated people.

The positive effect of food involvement on consumer's utility increases for women and the highest educated population, but it becomes negative for younger respondents. The positive effect of "Price Involvement" is mitigated for women and the better-educated population but significantly increases for people living in larger towns. The positive effect on the consumer's quality involvement when shopping is mitigated in larger towns and, practically, disappears in the case of women.

Finally, behavioural latent variables affect the utility assigned to the organic attribute. However, this attribute negatively affects the utility of consumers, as mentioned above. The interaction parameters found at the lower part of Table 3.3 indicates that this negative effect is partially mitigated in consumers affected by subjective norms or with a positive attitude towards organic food. Nevertheless, the organic attribute does not seem to play a significant role in the extra virgin olive oil market.

3.6. Conclusions

The use of limited information models, such as conventional choice models, could be problematic if the decision making process is strongly conditioned by a consumer's personality traits and lifestyles. In this paper, an HCM was applied to better understand the consumer's behavioural process related to the purchase of olive oil virgin extra in Catalonia. Special attention was paid to the organic attribute of the oil. This approach has been proven to be flexible enough to investigate the effect of consumers' food-related personality traits, lifestyles and purchasing habits on consumer's purchase intention towards organic olive oil as well as the main determinants of consumer choices when buying olive oil virgin extra .

The results from this study suggest that almost all personal trait latent variables significantly affect the respondents' utilities towards extra virgin olive oil. A "healthy lifestyle" is significantly but negatively associated with extra virgin olive oil utility, which shows that olive oil preferences in Catalonia respond more to dietary traditions than to healthy food choices. Nevertheless, this result was not homogeneous across the sample. In fact, the negative effect of "healthy lifestyle" was mitigated in women. This result shows that this population segment cares more about diet and the impact of food on health and thus bases its food choices on healthy reasons.

Food-related activities (cooking skills) are more related to social and self-activities than to healthy food measures. Extra virgin olive oil is normally used in Catalonia for salads, boiled vegetables or grilled food. In this context, people with higher cooking skills attempt to use alternative products to traditional olive oil. Additionally, the variables "price involvement" and "quality involvement" also significantly and positively affect the respondents' utilities towards extra virgin olive oil. However, the impact of these two variables is not homogeneous. Significant differences were found in people living in larger towns. While the overall positive effect of "Price Involvement" increases in larger towns, this initial positive effect of "Quality Involvement" is significantly mitigated.

The results also suggest that Catalan consumers perceive a disutility from the organic attribute compared to other production system alternatives (conventional and PDO). The price is not a relevant factor to explain this result, as organic olive oils are cheaper than PDO olive oils on average. Environmental or health concerns seem to not be relevant to consumers' choices related to olive oil. The organic attribute is not perceived as a significant quality cue, whereas people looking for quality select PDO extra virgin olive oil. This result suggests that the traditional marketing strategies that have been used in Catalonia to promote the consumption of olive oils based on environmental or health issues should be changed. Strategies that attempt to reinforce the "local" attribute should be encouraged.

Nevertheless, the results from this study reinforce the need to include the psychological characteristics of consumers, such as attitudes, food-related personality traits, purchase habits and lifestyle orientation, to better explain how individuals make food (olive oil, in this case) choices and to better understand the decision maker's process. These findings are likely to encourage a more widespread application of the HCM in the agro-food marketing field. From a methodological point of view, more research should be addressed to provide new tools to simultaneously estimate the HCM while considering heterogeneity across individuals.

References

Ajzen, I. (2005). *Attitude, personality and behavior*. (2nd ed.). England: Berkshire.

Ashok, K., Dillon, W., & Yuan, S. (2002). Extending discrete choice models to incorporate attitudinal and other latent variables. *Journal of Marketing Research*, 39, 31-46.

Bell, R., & Marshall, D. W. (2003). The construct of food involvement in behavioral research: Scale development and validation. *Appetite*, 40, 235-244.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J. Bhat, C.a, Bierlaire, M., Bolduc, D., Börsch-Supan, A., Brownstone, D., Bunch, D.S., Daly A., De Palma, A., Gopinath, D., Karlstrom, A., & Munizaga, M.A. (2002). Hybrid choice models: Progress and challenges. *Marketing Letters*, 13, 163-175.

Bhat, C. R. (2003). Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B-Methodological*, 37, 837-855.

Bolduc, D., Boucher, N., & Alvarez-Daziano, R. (2008). Hybrid choice modelling of new technologies for car choice in Canada. *Transportation Research Record*, 2082, 63-71.

Calatrava, J. (2002). Actitudes del consumidor español respecto a los productos ecológicos. Consejería de Agricultura y Pesca de la junta de Andalucía Tomado de Junta de Andalucía.

Campbell, D., Doherty, E., Hynes, S., & Rensburg, T.V. (2010). Combining discrete and continuous mixing approaches to accommodate heterogeneity in price sensitivities in environmental choice analysis. Paper presented at the Agricultural Economics Society Annual Conference, Edinburg, 29 March.

Chen, M. F. (2007). Consumer attitudes and purchase intentions in relation to organic foods in Taiwan: Moderating effects of food-related personality traits. *Food quality and preference*, 18, 1008-1021.

Chen, M.F. (2009). Attitude toward organic foods among Taiwanese as related to health consciousness, environmental attitudes, and the mediating effects of a healthy lifestyle. *British Food Journal*, 111, 165-178.

Eertmans, A., Victoir, A., Vansant, G., & Van den Bergh, O. (2005). Food-related personality traits, food choice motives and food intake: Mediator and moderator relationships. *Food Quality and Preference*, 16, 714-726.

Goody, J. (1982). *Cooking, cuisine and class: A study in comparative sociology*, New York: Cambridge University Press.

Gracia, A., & Magistris, T. (2008). The demand for organic foods in the south of Italy: A discrete choice model. *Food Policy,* 33, 386-396.

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.

Hair, J., Anderson, E., Tathan, R., & Black, W. (1999). *Multivariate Data Analysis with Readings*. (5th ed.). New Jersy: Englewood Cliffs.

Hensher, D. A., Rose, J.M., & Greene, W.H. (2005). *Applied Choice Analysis: A Primer*. New York: Cambridge University Press.

Hensher, D.A., & Greene, W.H, (2003). Mixed logit models: state of practice. *Transportation*, 30, 133-176.

Jiménez-Guerrero, J.F., Gásquez-Abad, J.C., Mondéjar-jiménez, J.A., & Huertas-García, R. (2012). Consumer preferences for olive oil attributes: a review of the empirical literature using a conjoint approach. *Olive oil-constituents, quality, health properties and bioconversions.* Publisher: In Tech.

Johansson, M.V., Heldt, T., & Johansson, P. (2006). The effect of attitudes and personality traits on mode choice. *Transport research part A*, 40, 507-525.

Jöreskov, K. & Sörbom, D. (1996). *Lisrel 8: User's Reference Guide. Scientific Software*. (2nd ed.). USA.

Kline, R. (2005). *Principles and practices of structural equation modeling*. New York: the Guildford press.

Krishnan, J. (2011). Lifestyle – A tool for understanding buyer behavior. *International Journal of Economics and Management*, 5, 283-298.

Lancaster, K. J. (1966). A new approach to consumer theory. *The journal of political economy,* 74, 132-157.

Losasso, C., Cibin, V., Cappa, V., Roccato, A., Vanzo, A., Andrighetto, I., & Ricci, A. (2012). Food safety and nutrition: Improving consumer behavior. *Food Control,* 26, 252-258.

MAGRAMA, (2008). *Guía de buenas prácticas para la producción y comercialización de alimentos ecológicos*. Ministerio de Medio Ambiente y Rural y Marino.

Mann, S., Ferjani, A. and Reissig, L. (2012). What matters to consumers of organic wine? *British Food Journal*, 114, 272-184.

McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior*. New York: Academic press.

Menapace, L., Colson, G., Grebitus, C., & Facendola, M. (2011). Consumers' preferences for geographical origin labels: evidence from the Canadian olive oil market. *European Review of Agricultural Economics,* 38, 1-20.

Mtimet, N., & Albisu, L.M. (2006). Spanish wine comsumer behavior: a choice experiment approach. *Agribusiness*, 22, 343-362.

Philippidis, G., Kakaroglou, I., & Sanjuan A. (2002). Territorial product association in Greece: The case of olive oil. Paper presented at the Xth EAAE Congress, Zaragoza, 28-31 August.

Revelt, D., & Train, K. (1998). Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. *Review of Economics and Statistics,* 80, 647-657.

Ruiz-Castillo, B. (2008). Las denominaciones de origen protegidos y el aceite de oliva en España. *Distribución y Consumo,* 57-68.

Rungie, C.M., Coote, L.V., & Louviere, J.J. (2011). Structural Choice Modelling: Theory and Applications to Combining Choice Experiments. *Journal of Choice Modelling*, 4, 1-29.

Scarpa, R. and Del Guidice, T. (2004). Market segmentation via mixed Logit: extra-virgin olive oil in urban Italy. *Journal of Agriculture and Food Industrial Organization*, 2, 1-20.

Scarpa, R., and Thiene, M. (2011). Organic Food Choices and Protection Motivation Theory: addressing the psychological sources of heterogeneity. Food Quality and Preference, 22, 532-541.

Shaharudin, M.R., Pani, J.J., Mansor, S.W. and Elias, S.J. (2010). Factors affecting purchase intention of organic food in Malaysia's Kedah state. *Cross-Cultural Communication*, 6, 105-121.

Soler, F., Gil., J.M., & Sánchez, M. (2002). Consumer's acceptability of organic food in Spain. Results from an experimental auction market. *British Food Journal*, 104, 670-687.

Stolz, H., Stolze, M., Janssen, M., & Hamm, U. (2011). Preferences and determinants for organic, conventional and conventional-plus products: the case of occasional organic consumers. *Food Quality and Preference*, 22, 772-779.

Street, D. & Burgess, L.B. (2007). *The Construction of Optimal Stated Choice Experiments: Theory and Methods*. New Jersey: Hoboken.

Train, K., (2003). *Discrete Choice Methods with Simulation*. New York: Cambridge University Press.

Tsakiridou, E., Mattas, K., and Tzimitra-Kalogianni, I. (2006). The influence of consumer characteristics and attitudes on the demand for organic olive oil. *Journal of International Food & Agribusiness Marketing*, 1, 23-31.

Van den Horst, K. et al. (2010). Ready-meal consumption: associations with weight status and cooking skills. *Public Health Nutrition*, 14, 239-245.

Vega-Zamora, M., Parras-Rosa, M., Torres-Ruiz, F.J., & Murgado-armenteros, E.M. (2011). Los factores impulsores e inhibidores del consumo de alimentos ecológicos en España. El caso del aceite de oliva. *ITERCIENCIA*, 36, 178-184.

Yañez, M.F., Raveau, J. and Ortúzar, J.D. (2010). Inclusion of latent variables in Mixed Logit Models: Modelling and forecasting. *Transport research part A*, 44, 774-753.

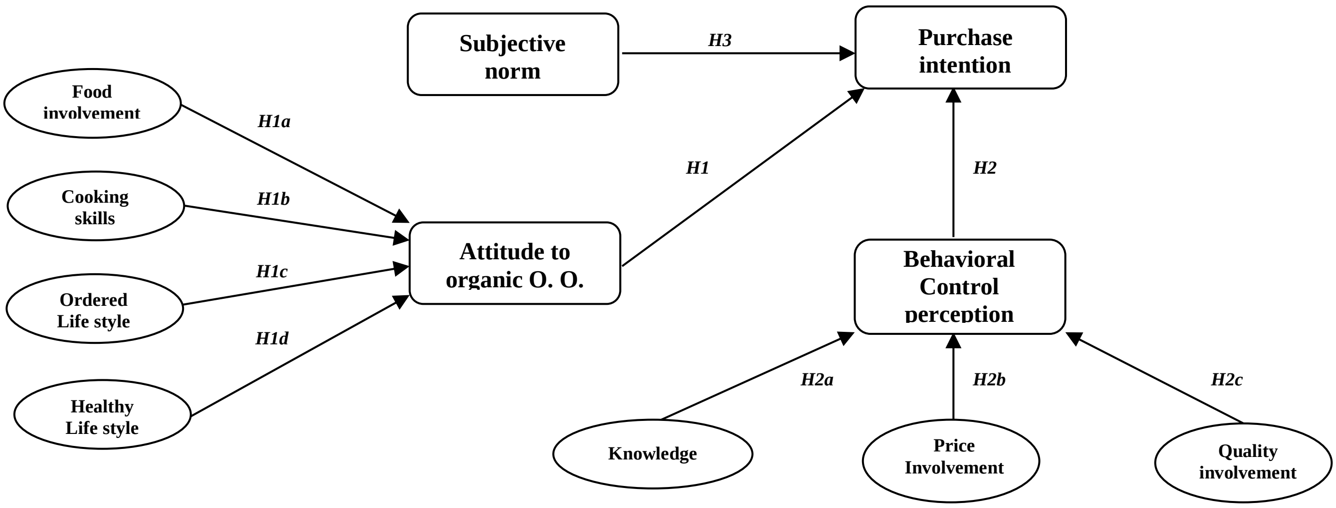

Figure 3.1 A conceptual model to understand organic olive oil purchase intention.

	Alternative "A"	***Alternative "B"***	***Alternative "C"***	***Alternative "D"***
System of production	*Extra-virgin olive oil with PDO*	*Conventional extra-virgin olive oil*	*Organic extra-virgin olive oil*	*None of them*
Origin of olive oil	*Spain*	*Catalonia*	*Imported*	
Brand	*Spanish Manufacturer*	*private label*	*Catalonia Manufacturer*	
Price	*3.70 €/liter*	*7.50 €/liter*	*6 €/liter*	

Figure 3.2 Example of a choice set

Appendix 3.1 Confirmatory factor Analysis results of personality traits factors

Índ	*Factores and ítems*	*Mean (SD)*	*Standarized Factor loadings (SE)*	*Varianze*	*Cronbach 's Alpha*	*Composite Reliability (variance extracted)*	*Referencies*
	Ordered Lifestyle			74.40%	0.82	0.819 (0.602)	Gil et al. (2000)
OLS_1	I try to reduce stress.	6.888 (1.892)	1.372*** (0.108)				
OLS_2	I try to lead an ordered life and methodical.	7.308 (1.571)	1.325*** (0.0674)				
OLS_3	I try to equilibrate between my work and my personal life.	7.317 (1.679)	1.304*** (0.104)				
	Healthy lifestyle			56.75%	0.57	0.559 (0.302)	Gil et al. (2000)
HLS_1	I try to control salt intake.	6.720 (2.74)	1.097*** (0.157)				
HLS_2	I eat frequently fruits and vegetables.	7.312 (2.180)	1.062*** (0.117)				
HLS_3	I try to not eat precooked foods.	8.180 (1.621)	1.489*** (0.121)				
	Price involvement			75.68%	0.88	0.885 (0.663)	Soler and Gil (2002)
PIN_1	I usually buy more the product in promotions	7.040 (2.159)	1.995*** (0.0906)				
PIN_2	I usually pay attention in the promotions.	7.135 (2.177)	2.072*** (0.0929)				
PIN_3	I remember the price paid in the last time.	6.343 (2.397)	1.415*** (0.126)				
PIN_4	I compare the prices of different bands available.	6.723 (2.160)	1.696*** (0.104)				
	Quality involvement			77.64%	0.83	0.840 (0.636)	Soler and Gil (2002)
QIN_1	I buy the product independently to their price.	5.535 (2.433)	1.656*** (0.117)				
QIN_2	It is relevant for me paying more if the product has more quality.	6.553 (1.813)	1.635*** (0.0851)				
QIN_3	Pay more if the product has a guaranteed quality.	6.683 (1.793)	1.578*** (0.0927)				
	Food involvement			68.08%	0.83	0.846 (0.584)	Adapted from Chen (2007) and Candel (2001)
FIN_1	Mainly, I eat to have good health.	7.947 (1.599)	0.942*** (0.0804)				
FIN_2	Eating is a pleasure.	8.248 (1.404)	1.065*** (0.0754)				
FIN_3	The food accounts a significant part of the family's traditions.	8.190 (1.486)	1.334*** (0.0664)				
FIN_4	The food is a link to provide information about other cultures.	8.015 (1.651)	1.314*** (0.0981)				
	Cooking skills			58.87%	0.76	0.767 (0.456)	Candel (2001)
COS_1	I like cooking.	6.697 (2.430)	1.522*** (0.120)				
COS_2	I like to watch food programs on TV.	6.082 (2.797)	1.895*** (0.126)				
COS_3	I like to subscribe to cooking magazines.	3.750 (3.091)	2.191*** (0.125)				
COS_4	I like to offer food as gifts.	5.650 (2.531)	1.69*** (0.128)				

Notes : ***p<0.01;**p<0.05; *p<0.1; SD: Standard Deviation; SE: Standard Error.

Appendix 3.2 Confirmatory factor Analysis results of Behavioral factors

Índ	*Factor*	*Means (SD)*	*Standarized Factor loadings (SE)*	*Varianze*	*Cronbach 's Alpha*	*Composite Reliability (variance extracted)*	*Referencies*
	Attitude			81,96	0.97	0.948 (0.755)	Adapted from Alemán et al. (2006), and Roitner-Schobesberger et al. (2007)
ATT_1	The consumption of organic olive oil reduces human exposure to chemical residues.	6.867 (1.764)	1.502*** (0.110)				
ATT_2	Organic olive oil is healthy for children.	6.862 (1.660)	1.178*** (0.0678)				
ATT_3	The product is suitable for a healthy diet.	7.088 (1.636)	1.324*** (0.0666)				
ATT_4	The production of organic olive oil helps indirectly to reduce water pollution by waste chemicals and pesticides.	6.923 (1.680)	1.553*** (0.0579)				
ATT_5	The production of organic olive oil helps indirectly to conserve agricultural soil.	6.933 (1.716)	1.648*** (0.0563)				
ATT_6	The production of organic olive oil improves environmental sustainability	6.893 (1.809)	1.662*** (0.0626)				
	Behavioral Control Perception			69,79	0.87	0.816 (0.443)	Adapted from Krystallis and Chryssohoidis (2005), and Roitner-Schobesberger et al. (2007);
BCP_1	I trust the product because of its certification by an organization or regulatory board of organic farming.	6.447 (1.601)	1.306*** (0.108)				
BCP_2	I trust the product because it is sold exclusively in specialty stores.	6.668 (1.646)	1.293*** (0.0840)				
BCP_3	I have confidence in the information provided on the product label.	6.202 (1.710)	1.35*** (0.0930)				
BCP_4	I have confidence that a product certified as organic really is organic.	6.103 (1.866)	1.441*** (0.109)				
BCP_5	The product is not available in the usual supermarkets where I normally do my shopping.	7.270 (1.843)	0.758*** (0.124)				
BCP_6	Seek the product, me generates high cost in terms of time and money.	6.728 (1.862)	0.622*** (0.114)				
	Purchase intention			76,91	0.858	0.875 (0.701)	Adapted from Lea and Worsley (2005)
PI_1	If I have more information and confidence, I buy organic olive oil.	5.923 (2.179)	1.938*** (0.221)				
PI_2	I buy more if the product is cheaper.	5.770 (2.219)	1.856*** (0.100)				
PI_3	If organic olive oil is more readily available, I most often buy it.	5.655 (2.246)	1.912*** (0.116)				
	Knowledge			87,63	0.861	0.876 (0.780)	
KN_1	Lack information about the benefits of organic products.	6.905 (1.834)	1.586*** (0.118)				
KN_2	Lack of information about the label that identifies products as organic.	6.872 (1.889)	1.705*** (0.116)				
	Subjective norms			86,61	0.926	0.934 (0.825)	Chen (2007)
SBN_1	My kids prefer organic olive oil.	2.342 (2.475)	2.059*** (0.104)				
SBN_2	My family prefers organic olive oil.	2.465 (2.422)	2.382*** (0.0710)				
SBN_3	Persons who are important to me prefer organic olive oil.	2.578 (2.436)	2.215*** (0.0885)				

Notes : ***p<0.01;**p<0.05; *p<0.1; SD: Standard Deviation; SE: Standard Error.

<u>Chapter 4</u>

Are ranking preferences information methods comparable with the choice experiment information in predicting actual behavior?

4.1. Introduction

Since its introduction conjoint analysis (CA) has become one of the most popular marketing research tools (Lusk et al., 2008; Campbell and Lorimer, 2009). In CA's tasks, participants are provided with at least two product profiles and are asked to rate/rank them or select the profile they prefer most. The most widely used CA format to elicit consumers' preferences for market and non-market goods is choice experiment (CE). In CE, respondents are shown a set of combinations of attributes (i.e., profiles) and are asked to indicate which of the combinations or profiles they would purchase. CE gained popularity thanks to its ability to mimic the real market setting where consumers who are faced with competing products purchase the product that fits most their preferences. However, it is informationally inefficient, since it only allows the observation of the most preferred option (Lusk et al., 2008; Louviere et al., 2008; Lanscar et al., 2013). According to Lancsar et al. (2013), there are three ways to gain more insights about individual preferences in CE: 1) increasing the sample size; and/or 2) asking the respondents to evaluate more choice sets; or 3) increasing the number of options per choice set.

In contrast with CE, participants in a ranking conjoint analysis (RCA) are provided with a set of product concepts and they are asked to rank them from the most to the least preferred. The use of RCA as an alternative to CE is becoming popular since it provides information not only about the most preferred product concept but also about consumers' preferences for all the product concepts included in a choice set, which could lead to a more efficient preference estimates (Chang et al., 2009; Louviere et al., 2008; Lusk et al., 2008).

In the same line, Louviere et al. (2004) introduced another CA format named best worst scaling (BWS). The BWS approach consists in asking respondents to firstly choose the best and the worst option, then the second best and the second worst options from the remaining options and so on until a complete preference ordering of all the options is obtained. BWS tasks seem to be easier to handle by respondents due to human skills at identifying extremes (Helson, 1964; Flynn and Marley, 2012). As in the RCA, the additional choice information obtained from BWS has been showed to improve the statistical efficiency of choice models, especially, when it is combined with an appropriate experimental design (Lancsar et al., 2013).

Despite the wide application of the aforementioned CA formats (i.e. CE, RCA and BWS) over the last two decades, few researchers, however, have compared their performance in terms of the estimated marginal partworths, the predictive power of the derived models, and the reliability of the Willingness to Pay (WTP) values deduced from the estimated partworths. Most of the past literature focused on assessing the incentive compatibility of CA and proposing modified CA formats to incentivize subjects to truthfully reveal their preferences (Lusk and Schroeder, 2004; Ding et al., 2005; Ding et al., 2007; Lusk et al., 2008). Others papers compared the validity of the estimates obtained from different CA formats (Boyle et al., 2001; Caparros et al., 2008; Chang et al., 2009; Pignone et al., 2011; Akaichi et al., 2013). For instance, Caparros et al. (2008) pointed out that CE and RCA provide similar results, when a similar experimental design is used for both CA formats. Akaichi et al. (2013) confirmed this result for small choice sets; however, they found discrepancies between respondents' preferences in CE and RCA when large choice sets are used.

It is noteworthy that the aforementioned studies compared CE and RCA considering information only on the most preferred option. In other words, the authors did not consider respondents' preferences for all the options included in each choice set. Furthermore, the previous studies, surprisingly, did not assess the comparability of BWS to CE and RCA, although BWS's superiority in terms of realism and ease of its implementation. To fill this gap, this paper stands out by: 1) comparing the performance of CE, RCA and BWS in terms of estimated partworths, predictive power and estimated WTP; and 2) considering, in the estimation of partworths in RCA and BWS, not only the information on the most preferred option but also the information on the other options.

It is worth noting that the majority of the studies that have assessed the comparability of CA formats reported results obtained from economic experiments conducted in hypothetical settings (with the exception of Chang et al., 2009; and Akaichi et al., 2013). Additionally, the experiments involving BWS as the preference elicitation method have been always implemented in hypothetical settings (Louviere et al. 2008; Scarpa et al. 2011; Lancsar et al. 2013). However, due to the skepticism surrounding the validity of values obtained from hypothetical CA experiments (Lusk and Schroeder 2004; Ding, et al. 2005; Alfnes et al., 2006; Lusk and Shogren, 2007; Dong et al., 2010), we conducted the CE, the RCA and the BWS in a non-hypothetical setting (further explanation is given in the section dedicated to the experimental design). In this study we also conducted a hypothetical CE to be used as the benchmark.

To compare the external validity of the three CA formats, we have included a non-hypothetical holdout choice task in the experimental design in the CE, RCA and BWS. Finally, one of the main assumption underlying stated preference methods is that respondents know their preferences and these preferences are stable and coherent (Brown et al., 2008). According to Hoeffler and Ariely (1999), preferences' consistency or stability is positively correlated with choice experience and cognitive choice effort. For instance, in repeated choices, respondents are expected to be more precise and consistent in their decisions due to the learning effect (Brouwer et al., 2010). On the contrary, when they face hard choice tasks (e.g. too many choice sets or too many options per choice set), respondents are less precise and consistent in their choices compared with those facing an easy choice scenario (Brouwer et al., 2010). In this study, we compared the consistency of respondents' answers in CE, RCA and BWS to find out which of the three CA formats provide more consistent subjects' responses. To tackle this issue, one of the choice sets faced by respondents was repeated at the end of the choice experiment.

To sum up, our study stands out by assessing the comparability of non-hypothetical CE (NHCE), non-hypothetical RCA (NHRCA) and non-hypothetical BWS (NHBWS) in terms of estimated partworths, internal and external predictive power, estimated WTP, and participants' responses consistency, in two contexts: 1) the additional information (i.e. respondents' preference for all the options) obtained in RCA and BWS is not taken into account and, hence, only the most preferred option is considered (RRCA and RBWS)[10]; and 2) the additional information is included and the econometric models for NHRCA and NHBWS are estimated. This will allow us to assess the comparability of the three CA formats before and after considering the additional information.

This study is structured into five sections. In the next section, the experiment design and the experimental procedures are described. The econometric model used to estimate the partworths is outlined in section 3. The results are discussed in the fourth section and we finish by drawing some concluding remarks.

4.2. Experiment design

In this study, four treatments were carried out, hypothetical CE (HCE), NHCE, NHRCA and NHBWS. To assess the comparability of these CA formats a representative sample of 220 real consumers was recruited. Participants were randomly and equally assigned to the four treatments. Olive oil is the food product used in our experiment. The main attributes and attribute levels were first identified based on the literature review and the information collected from two focus groups of high and low experienced consumers of olive oil. Four attributes of olive oil were considered. Three of them have three levels: type of olive oil

[10] RRCA and RBWS stand for Recoded Ranking Conjoint Analysis and Recoded Best Worst Scaling, relatively. The data obtained in RCA and BWS were recoded as choice data (i.e only considering the option ranked first or the option chosen as the best option and recode it as the most preferred option).

(virgin extra, virgin, and olive oil)[11], origin (Andalucía, Catalonia, and rest of Spain) and price (2.20 €/liter, 3.50 €/liter, and 4.80 €/liter, which account for 85% of the price distribution in retail outlets). Brand is the fourth attribute and has two levels (Manufacturer label and Private label).

Given these attribute levels, a full factorial design of 54 (3^3*2) combinations (i.e. one-liter bottles of olive oil) was generated. Presenting respondents with 54 combinations, however, could place a high level of cognitive burden on respondents. To reduce the number of combinations that participants have to evaluate, we followed Street and Burgess (2007) and we generated an orthogonal fractional factorial design of 9 combinations. These 9 combinations were considered as the first option in each choice set. Since participants were provided with choice sets of 5 options each (plus a no-choice option), the other four options were obtained using the following generators (1000), (1111), (2121), and (2122) (Street and Burgess, 2007). This resulted in a 100% efficient main-effects design.

Participants, in each treatment, performed two choice tasks (i.e. main task and holdout task). In the main task and depending on the treatment (HCE, NHCE, NHRCA or NHBWS) were successively offered a total of 10 choice sets (i.e. first they received the 9 choice sets obtained in the efficient design. Then the fifth choice set was again given to participants to assess the consistency of their decisions). In each choice set, participants were asked to mark their most preferred option (or to rank all the options in RCA and BWS). The holdout task consists in a single-choice card of 10 options (including a no-choice option) which are different from the options provided to participants in the main task. In the holdout task, participants were required to choose the most preferred option of the 10 options included in

[11] The three types of olive oil were defined according to the International Olive Council (IOC). **Extra virgin olive oil:** virgin olive oil which has a free acidity, expressed as oleic acid, of not more than 0.8 grams per 100 grams; **Virgin olive oil:** virgin olive oil which has a free acidity, expressed as oleic acid, of not more than 2 grams per 100 grams; **Olive oil** is the oil consisting of a blend of refined olive oil and virgin olive oils fit for consumption as they are. It has a free acidity, expressed as oleic acid, of not more than 1 gram per 100 grams. The three olive oils defined have other characteristics of which correspond to those fixed for this category in the IOC standard.

the choice set. Each treatment of the experiment was conducted in 5 sessions throughout different days of the week and different hours of the day. A number of 10-15 persons participated in each session. After finishing the two tasks, participants were asked to complete a short questionnaire about their socio-demographic and lexicographic characteristics as well as their attitudes toward olive oil.

At the beginning of the experiment, participants were informed that they would receive 15 Euros in cash at the end of the experiment. Additionally, participants in NHCE, NHRCA and NHBWS were informed that they would be participating in non-hypothetical tasks and, hence, it is in their best interest to reveal their actual preferences. We then explained them who the CA mechanism works. In the next section details about the experimental procedure of each treatment are presented.

Hypothetical (HCE) and non-hypothetical choice experiment (NHCE)

In HCE, we asked participants to assume that each choice set is a real shopping situation. Participants were informed, however, that they are not required to actually buy the chosen products and pay the corresponding price. In each choice set, participants were asked to indicate the option they prefer most bearing in mind their real purchase habits (Appendix 4.1). They were also informed that if they did not like any one of the provided olive oil combinations, they simply can choose the no-choice option. The NHCE experiment was similar (Appendix 4.1), but participants were informed that each choice set is a real shopping scenario. Therefore, participants could receive the option they had selected and pay its posted price. After finishing the main task, participants in both treatments were given a choice set of 10 options (i.e. holdout task) and were then asked to choose the option they prefer most (Appendix 4.4).

After completing the two tasks and the survey, we asked for a volunteer among the participants to randomly draw a number between 1 and 2 to determine the binding task. If the binding task is the main task, participants in the HCE, receive the 15 Euros and the experiment finishes. In the NHCE, another volunteer is selected to randomly draw one of the 9 choice sets[12] to determine which of the choice set will be the binding one. Then, each participant obtains the option she/he has chosen in the binding choice set and receives 15 Euros minus the price indicated in that option. In case, the participant chose the no-choice option, she (he) receives the 15 Euros and do not buy any product. If the binding task is the holdout task, regardless the type of treatments (HCE or NHCE), each participant has to buy the chosen option and pays the corresponding price. If the chosen option is the no-choice option, the participant receives the 15 Euros and did not buy any product.

Non-hypothetical rank conjoint analysis (NHRCA)

The same 10 choice sets were presented to each participant, who was asked to rank the options in each choice set from the most to the least preferred option (Appendix 4.2). In case participants do not like any one of the presented alternatives, she (he) could choose the no-choice option. The non-hypothetical nature of the experiment was also revealed to participants since the beginning. After completing all the choice sets in the main task, participants were given a choice set of 10 options (i.e. holdout task) and were then asked to choose the option they prefer most. After completing the main and the holdout task, a volunteer among participants was asked to randomly draw the binding task. If the main task was chosen as the binding task another volunteer was approached to draw the binding choice set. Following Lusk et al. (2008), to ensure that the ranking treatment is incentive compatible, participant had to purchase the binding product with a probability proportional to the rank she (he) assigned to each one of the options. Particularly, each participant who did not choose the no-choice option draws a number from 1 to 50 to select the biding product. If the number drawn is between 1 and 17, participant should purchase the most preferred option and pay its

[12] The last choice set (the number 10) was the same as the fifth choice set. Therefore, to allow that all the choice sets have the same probability to be drawn we removed the tenth choice set.

price. If the drawn number is between 18 and 30 the second most preferred option will be the biding product. If the drawn number is between 31 and 40, participant should purchase the third option in her (his) preference ranking and pay its corresponding price. If the number drawn is between 41 and 47, participant buys her (his) fourth most preferred option. Finally if the drawn number is between 48 and 50, participant has to buy the least preferred option. If the binding task was the holdout task, the procedure was similar to the one implemented in HCE and NHCE treatments.

Non-hypothetical best worst scaling (NHBWS)

Similar to the previous treatments, the same 10 choice sets were presented to each participant in the main task. In NHBWS, each participant was asked to choose the most preferred or the best option, followed by the worst option of the four remaining options, followed by the second best option of the three remaining options, followed by the second worst option of the two remaining options (Appendix 4.3). Hence, a complete ranking of the five options can be deduced (i.e. the option ranked first, second, third, fourth and fifth are the first best option, the second best option, the remaining option, the second worst option and the first worst option, respectively). Similar to the other three treatments and after finishing the main task, participants in NHBWS were given a choice set of 10 options (i.e. holdout task) and were then asked to choose the option they prefer most. Once participants finish the main and the holdout task, similar procedure to the one applied in the NHRCA was used to determine the binding task and the binding product.

4.3. Methodological approach

Based on the random utility theory (McFadden, 1973) and the Lancaster's (1966) theory (1966), the i^{th} individual's utility function U_{ijs} towards an option j from a choice set s can be decomposed into a deterministic component V_{ijs} and a stochastic component ε_{ijs}.

$$U_{ijs} = V_{ijs} + \varepsilon_{ijs} \tag{1}$$

The deterministic component is commonly specified as linear in parameters and includes variables that represent the attributes of the alternative and the characteristics of the respondents. In our empirical specification, and taking into account the experiment design, the deterministic component is given by (socio-demographic variables have been eliminated for simplicity):

$$V_{ijs} = \beta_0 NoBuy + \beta_1 EVOO_{ijs} + \beta_2 OO_{ijs} + \beta_3 BrManf_{ijs} + \beta_4 CAT_{ijs} + \beta_5 RSp_{ijs} + \beta_6 price_{ijs} \quad (2)$$

In Eq. (2) the attributes levels (the extra virgin olive oil (EVOO), the olive oil (OO), the Manufacturer Brand (BrManf), the Catalonian origin (CAT) and the "Rest of Spain" origin (RSp)) were effect coded (-1, 0, 1)[13], except for the price that was coded as a linear variable. The parameter "NoBuy" represents the no-choice option and has been coded as a dummy variable that takes the value 1 when the option was chosen by the participant; and 0, otherwise. To estimate the utility function (3), in NHCE and HHCE, the random parameter logit model (RPL) was used to account for respondents' preferences heterogeneity. In the RPL the marginal utility of the attribute levels is not assumed to be constant, but varies across the sample according to a continuous probability distribution function. As shown by Train (2003), the probability of consumers *i* to choose the option *j* in the choice set *s* is as follows:

$$Prob_i\{j \text{ is chosen}\} = \int L_{ij}\left(\beta_{ij}\right) f(\beta_i/\theta) d\beta_i \quad (3)$$

where $f(\beta_i/\theta)$ is the density function of the coefficients β_i. θ refers to the moments (the mean and standard deviation) of the parameter distributions and

$$L_{ij}(\beta_{ij}) = \frac{e^{V_{ij}}}{\sum e^{V_{ik}}} = \frac{e^{\beta_{ij} X_{ij}}}{\sum e^{\beta_{ik} X_{ik}}}, with\ k \in C_s \quad (4)$$

However, according to Lusk et al. (2008), the estimation of the partworths in NHRCA and NHBWS was carried out using the rank-order random parameter logit (RO-RPL) model.

[13] The attribute levels virgin olive oil (VOO), private brand (BrPRV), and Andalucía (AND) were considered as the baseline for the attributes: type of olive oil, brand, and origin, respectively.

This model assumes that the probability of a particular ranking of the options presented in a choice is the product of the multinomial choice probability for always choosing the best of the remaining options. That is, the probability (L_{ij}) that an individual *i* rank the five options as A> B> C> D> E from a choice set of five options (A, B, C, D, E) will be modeled as the product of the probability of choosing A as the best option from the choice set (A, B, C, D, E), the probability of choosing B as the best option among the remaining options (B, C, D, E), the probability of choosing C as the best option among the remaining options (C, D, E), and the probability of choosing D as the best option among the remaining options (D, E). therefore, L_{ij} is given by:

$$L_{ij}(ranking\ A, B, C, D, E) = \frac{e^{V_{iA}}}{\sum_{j=A,B,C,D,E} e^{V_j}} * \frac{e^{V_{iB}}}{\sum_{j=B,C,D,E} e^{V_j}} * \frac{e^{V_{iC}}}{\sum_{j=C,D,E} e^{V_j}} * \frac{e^{V_{iD}}}{\sum_{j=D,E} e^{V_j}} \qquad (5)$$

It is worth noting that the estimation of RPL and RO-RPL takes into account the unobserved effect of possible correlations between the attributes (Hensher et al., 2005). In fact, we assume that all the partworths β_{ij} of our empirical model are random and follow a normal distribution with mean β and variance-covariance matrix Ω, as they are not independently distributed.

Taking into account the between-subjects nature of our experimental design, it was necessary to test for the regularity of preferences across treatments. As in Lusk and Schroeder (2004), Caparros et al. (2008) and Lanscar et al. (2013), it is important to investigate whether differences in parameter estimates across samples are indeed due to the underlying preferences or to difference in variance. The null hypothesis of the test is the equality of preferences across treatments (i.e.$\mu_{HCE}\beta_{HCE} = \mu_{NHCE}\beta_{NHCE}$, with μ is scale parameter). The test statistic is a likelihood-ratio type ($-2(LL_j - \sum LL_i)$), and it is distributed as a χ^2 with K(M-1) degrees of freedom. LL_j is the log likelihood values of the pooled data (e.g. HCE plus NHCE data), LL_i is the log likelihood values of the estimated model for each treatment, K is the number of restrictions, and M is the number of treatments (Louviere et al., 2000). If the hypothesis is rejected, comparing the estimated WTP for each treatment would be appropriate because the error variance is constant within each sample, and it will be simplified when calculating the marginal WTPs.

Willingness to pay

WTP estimates were calculated by dividing the estimated partworth associated with the attribute's level by the estimated partworth of the price attribute with a negative sign. To test the statistical of possible differences of the estimated WTP for each attribute across treatments, the non-parametric complete combinatorial test proposed by Poe et al. (2005) was used. This test first requires the generation of a distribution of 1000 WTP estimates using, for example, the parametric bootstrapping method proposed by Krinsky and Robb (1986). The complete combinatorial test is then applied to compare the 1000 bootstrapped WTP values in one treatment (HCE) with the 1000 bootstrapped WTP values in the other treatment (e.g. NHCE).

Consistency, internal and external validity

As aforementioned, to assess participants' responses consistency and their stability across treatments, the fifth choice set was repeated at the end of the main task[14]. To measure the consistency of participants' responses in each treatment, we calculated the proportion of participants who gave the same response in the fifth and the tenth choice set. The response is counted as a hit it is found to be the same in the fifth and the tenth choice sets. Then, the hit rate is calculated by dividing the total number of hits by the total number of participants in each treatment. To compare the hit rates across treatments, the Z-test was used. It is expected that: 1) in HCE, participants are less consistent in their choices than in NHCE due to the lack of economic incentive to reveal the real preferences; 2) in NHBWS, participants are more

[14] In line with Brower et al. (2010), respondents felt significantly more confident and certain about their choice at the end of the choice experiment than they were at the beginning. Therefore, repeating the fifth choice set at the end of the experiment instead the first one could increase the reliability of the test.

consistent than participants in NHRCA due to the fact that BWS task is relatively easier for participants than RCA task (Flynn and Marley, 2012); and 3) compared with HCE and NHCE, participants' responses are less consistent in NHRCA and NHBWS. The reason behind this hypothesis is that respondent's task in NHRCA and NHBWS is harder since she (he) has to state her (his) preferences for all the option in the choice sets (Ben-Akiva et al. 1992).

To assess the internal validity of the estimated parameters we have used the estimated partworths to predict participant's response in the main task. The hit rate is calculated by comparing the predicted participants' decisions, using the maximum utility approach, to their real response in the fifth choice set in each treatment. Finally, as regard to the external validity, the estimated partworths from the main task are used to predict participants' responses in the holdout task. The predicted and the actual decision in the holdout task are compared to determine the hit rate. A Z-test was used to assess the difference between hit rates across treatments for both internal and external validity.

4.4. Results

Table 4.1 (the first four columns) shows the number of times participants choose or ranked first each one of the options presented in the choice sets. As it can be observed, there exists a significant similarity between the choice experiment treatments (HCE vs. NHCE) as well as between the ranking treatments (RRCA vs. RBWS). In the HCE, respondents chose the first option as the best choice around 20% of the time, 14% of the time the second option, 17% of the time the third option, 17% of the time the fourth option, 16% of the time the fifth option and 16% of the time the no-choice option. These percentages are statistically similar to those obtained in NHCE (i.e. 16%, 15%, 17%, 18%, 16% and 18% for the first to the fifth option plus the no-choice option, respectively). A similar result was found when comparing the ranking treatments. In the RRCA, respondents chose the first option 23% of the time, the second option 15% of the time, the third option 21% of the time, the fourth option 17% of the time, the fifth option 17% of the time and the no-choice option the 6% of the time as the most preferred option; these percentages are not statistically different from those found in the RBWS. However, Table 4.1 also shows that there are significant differences in participant choices between CE treatments (HCE and NHCE) and ranking treatments recoded as traditional choice (RRCA and RBWS). Results in the right side of Table 4.1 (the last two columns) show the percentage of times participants rank, for example, the first option in the choice sets as the first, the second, the third, the fourth or fifth most preferred option. As it can be observed no significant differences were found between ranking treatments. In general, the results show that when a similar experimental design is used for all treatments, respondents seem to behave similarly in the rankings treatments whether taking into account all the ranking information or only the option ranked first.

The second way of comparing results from the four treatments is to analyze the means and the standard deviations of the estimated partworths (Table 4.2). In the first four columns we report the estimates corresponding to the four treatments where the dependent variable has been coded in a similar way. Particularly, in HCE and NHCE, the dependent variable takes the value of 1 when the option is the chosen one and 0 otherwise; while in RRCA and RBWS the dependent variable is coded as 1 when the option is ranked first and 0 otherwise. In the last two columns, we present the results of the estimation of the ranking models taking into account the full ranking information. Results reveal that the no-choice option parameter is negative and significant across treatments, indicating that the majority of the respondents opted for choosing or ranking the real options instead of selecting the no-choice option. Additionally, results in Table 4.2 shows that all random parameters have the expected sign and are statistically significant as well as their standards deviations. However, the lower magnitude of the standards deviations, in most of cases, suggests the non-existence of significant high preference heterogeneity between respondents.

Table 4.1 Comparison of choices across the treatments

Treatments							
Options	% of time participants choosing each option as the best choice				**Rank**	% of time participants ranking the first option as	
	HHCE	**NHCE**	**RRCA**	**RBWS**		**NHRCA**	**NHBWS**
1	19.79	16.16	23.23	23.63	**1**	23.23	23.63
2	13.73	14.54	15.35	13.93	**2**	17.17	14.94
3	16.76	17.37	21.21	17.17	**3**	15.95	19.79
4	17.17	17.97	16.96	17.97	**4**	16.56	13.93
5	16.16	16.36	17.17	20.20	**5**	21.01	20.60
No option	16.36	17.57	6.06	7.07	**No option**	6.06	7.07
Test for equality proportions	*Chi-square: 2.30 p-value: 0.806*		*Chi-square: 4.20 p-value: 0.520*			*Chi-square: 4.34 p-value: 0.501*	
	Chi-square: 62.02 p-value: 0.000						

Interestingly, the results reveal that participants' preferences for the attributes of olive oil in the four CA formats are similar in terms of sign and significance although their estimated values are different. In fact, our findings (Table 4.2) show that the extra virgin is most preferred olive oil for consumers, while the olive oil is the least preferred. We also found that the local olive oil (Catalonian olive oil) is preferred over the olive oil from the other locations, while olive oil produced in Andalusia (i.e. South of Spain) is preferred over olive oil produced in the rest of Spain. This result is consistent with the findings of Jiménez-Guerrero et al. (2012) who found that the origin of olive oil is key attribute for Spanish consumers. Furthermore, we found that consumers seem to prefer the manufacturer brand over the private label. Finally, our results show that price is the main obstacle for buying olive oil.

Table 4.2 RPL and RO-RPL models estimates for elicitation methods

	Only choice or first Rank data in ranking treatments				Full ranking information	
Traitment	**HHCE**	**NHCE**	**RRCA**	**RBWS**	**NHRCA**	**NHBWS**
Models	**RPL**	**RPL**	**RPL**	**RPL**	**RO-RPL**	**RO-RPL**
			Random parameter estimates			
NoBuy	-4.822***	-3.340***	-5.050***	-5.015***	-4.641***	-4.437***
(SE)	(0.472)	(0.353)	(0.465)	(0.464)	(0.321)	(0.288)
EVOO	1.555***	0.925***	0.794***	0.988***	0.665***	0.687***
	(0.347)	(0.211)	(0.172)	(0.200)	(0.072)	(0.068)
VOO[1]	-0.202	0.096	0.088	-0.112	0.063	-0.031
	(----)	(----)	(----)	(----)	(----)	(----)
OO	-1.352***	-1.021***	-0.883***	-0.876***	-0.728***	-0.656***
	(0.276)	(0.243)	(0.177)	(0.237)	(0.068)	(0.075)
BrManf	0.347***	0.213**	0.158**	0.187**	0.131**	0.145***
	(0.091)	(0.088)	(0.074)	(0.085)	(0.043)	(0.037)
BrPrv[1]	-0.347	-0.213	-0.158	-0.187	-0.131	-0.145
	(----)	(----)	(----)	(----)	(----)	(----)
AND[1]	-0.657	-0.334	0.113	-0.394	0.150	0.027
	(----)	(----)	(----)	(----)	(----)	(----)
CAT	1.308***	0.911***	0.553***	0.940***	0.274***	0.612***
	(0.192)	(0.193)	(0.135)	(0.164)	(0.065)	(0.073)
RSp	-0.650***	-0.577***	-0.666***	-0.545***	-0.424***	-0.639***
	(0.178)	(0.173)	(0.139)	(0.132)	(0.063)	(0.075)
Price	-1.602***	-1.213***	-0.697***	-1.089***	-0.447***	-0.834***
	(0.165)	(0.112)	(0.114)	(0.113)	(0.051)	(0.058)
			Standards deviations of random parameters			
EVOO	2.495***	1.772***	1.460***	2.195***	1.048***	1.129***
	(0.312)	(0.220)	(0.179)	(0.252)	(0.082)	(0.084)
OO	2.946***	1.715***	1.365***	2.638***	0.930***	1.324***
	(0.436)	(0.212)	(0.193)	(0.320)	(0.079)	(0.097)
BrManf	0.141	0.283**	0.224**	0.410**	0.350***	0.156***
	(0.131)	(0.114)	(0.086)	(0.138)	(0.050)	(0.048)
CAT	2.140***	1.329***	0.689***	1.143***	0.630***	0.597***
	(0.283)	(0.183)	(0.138)	(0.189)	(0.080)	(0.061)
RSp	0.779***	0.992***	0.502**	0.639***	0.596***	0.489***
	(0.192)	(0.181)	(0.168)	(0.165)	(0.097)	(0.067)
Price	1.620***	0.497***	1.153***	1.128***	0.945***	0.894***
	(0.201)	(0.077)	(0.115)	(0.133)	(0.052)	(0.055)
Number of observations	2970	2970	2970	2970	7155	7110
Log-likelihood	-523.2198	-591.4546	-619.1057	-575.7293	-1745.575	-1070.695

[1] base line; (***) (**) (*) Statistically significant at 1%, 5%, and 10% level.

A likelihood ratio test has been used to test for significant differences among the estimated partworths across the 4 treatments (again considering the first choice (in HCE and NHCE) /the first rank (in NHRCA and NHBWS) or the full ranking information (in NHRCA and NHBWS). The results of likelihood ratio test are displayed in Table 4.3. Results show that we could reject the null hypothesis of equality of the estimates obtained in HCE and NHCE (LR= 49.05; p < 0.005). Furthermore, we compare the results from the three non-hypothetical treatments taking into account only the most preferred option. In this case, the null hypothesis is clearly rejected (LR = 163.89; p < 0.005). The final two tests compare the estimates from the two ranking treatments either coded as a traditional choice experiment or considering the full rank information. In both cases, the null is also rejected (LR = 66.64; p < 0.005 and LR = 66.87; p < 0.005, respectively). Therefore, the results pointed out that the differences between the estimates across all treatments are statistically significant. Consequently, are these discrepancies in the partworths' values is going to affect the comparability of the four conjoint analysis formats in terms of internal and external validity, as well as participants' WTP.

Table 4.3 Results from preference regularity tests across the treatments

Test for preference regularity	Number of observations	Log Likelihood	Likelihod Ratio (LR)	Degrees of freedom	p-value
All treatments	5940	-1139.1138			
HHCE	2970	-523.2198			
NHCE	2970	-591.4546			
H_0:test of equality between hypothetical and non-hypothetical CE			49.05	12	P<0.005
All treatments	11880	-2413.8433			
HHCE	2970	-523.2198			
NHCE	2970	-591.4546			
RRCA	2970	- 619.1057			
RBWS	2970	- 575.7293			
H_0:test of equality between hypothetical and non-hypothetical first choice option			208.66	72	P<0.005
All treatments	8910	-1868.23			
NHCE	2970	-591.45			
RRCA	2970	- 619.10			
RBWS	2970	- 575.72			
H_0:test of equality between non-hypothetical first choice option			163.89	36	P<0.005

Table 4.3 Results from preference regularity tests across the treatments (continued)

All treatments	5940	-1228.15			
RRCA	2970	- 619.10			
RBWS	2970	- 575.72			
H_0:test of equality between non-hypothetical NHRCA and NHBWS			66.64	12	P<0.005
All treatments	14265	-3457.67			
NHRCA	7155	-1745.57			
NHBWS	7110	-1678.66			
H_0:test of equality between non-hypothetical NHRCA and NHBWSD			66.87	12	P<0.005

The results of responses' consistency, internal and external validity analysis are displayed in Table 4.4. The results reveal that, in general terms, the consistency of participants' responses is relatively high. The hit rate ranges from 76.36% to 87.27% when only the most preferred option is considered (i.e. HHCE, NHCE, RRCA, and RBWS). Results from the one tailed Z-test for consistency, using the hit rates show that there are no statistical differences (at the 5% level of significance) between treatments. Consistent with Brouwer et al. (2010), the high consistency or stability of participants' responses could be due to the learning effect related to repeated choice task. Additionally, we found that the consistency of participants' responses in HCE and NHCE is statistically similar. This similarity maybe the result effort made by the experimenter to incentivize participants to reveal their real preferences independently of whether they are taking part in a hypothetical or a non-hypothetical CE. Interestingly, when the full ranking information is taken into account, to estimate the partworths in the ranking treatments, the consistency' hit rate decreases to 49.09% and 45.45%, in the NHRCA and NHBWS, respectively. This result supports the hypothesis that the stability of ranking information decreases when the number of options to be ranked increases and this seems to be due to the higher cognitive effort spent in RCA and BWS compared with CE (Ben-Akiva et al. 1992). Finally, results from the Z-test highlights that there are no significant differences between NHRCA and NHBWS regarding participants' response consistency when only the first choice or the full ranking information is considered.

Table 4.4 Consistency, internal and external validity tests across treatments

	Consistency				Internal validity			External validity		
Treatments	**Nº of choices**	**Nº of correct Predictions**	**Hit rate (%)**	***p-value***	**Nº correct of predictions**	**Hit rate (%)**	***p-value***	**Nº correct of predictions**	**Hit rate (%)**	***p-value***
HHCE vs	55	48	87.27	0.103	35	63.63	0.153	13	23.63	0.013
NHCE	55	43	78.18		40	72.72		24	43.63	
HHCE vs	55	48	87.27	0.069	35	63.63	0.153	13	23.63	0.049
RRCA	55	42	76.36		40	72.72		21	38.18	
HHCE vs	55	48	87.27	0.069	35	63.63	0.421	13	23.63	0.000
RBWS	55	42	76.36		34	61.81		29	52.72	
HHCE vs	55	48	87.27	0.000	35	63.63	0.272	13	23.63	0.032
NHRCA	55	27	49.09		38	69.09		22	40	
HHCE vs	55	48	87.27	0.000	35	63.63	0.344	13	23.63	0.000
NHBWS	55	25	45.45		37	67.27		34	61.81	
NHCE vs	55	43	78.18	0.410	40	72.72	0.5	24	43.63	0.280
RRCA	55	42	76.36		40	72.72		21	38.18	
NHCE vs	55	43	78.18	0.410	40	72.72	0.111	24	43.63	0.170
RBWS	55	42	76.36		34	61.81		29	52.72	
NHCE vs	55	43	78.18	0.001	40	72.72	0.337	24	43.63	0.349
NHRCA	55	27	49.09		38	69.09		22	40	
NHCE vs	55	43	78.18	0.000	40	72.72	0.266	24	43.63	0.028
NHBWS	55	25	45.45		37	67.27		34	61.81	
RRCA vs	55	42	76.36	0.5	40	72.72	0.111	21	38.18	0.062
RBWS	55	42	76.36		34	61.81		29	52.72	
RRCA vs	55	42	76.36	0.001	40	72.72	0.337	21	38.18	0.422
NHRCA	55	27	49.09		38	69.09		22	40	
RRCA vs	55	42	76.36	0.000	40	72.72	0.266	21	38.18	0.006
NHBWS	55	25	45.45		37	67.27		34	61.81	
RBWS vs	55	42	76.36	0.001	34	61.81	0.211	29	52.72	0.090
NHRCA	55	27	49.09		38	69.09		22	40	
RBWS vs	55	42	76.36	0.000	34	61.81	0.274	29	52.72	0.167
NHBWS	55	25	45.45		37	67.27		34	61.81	
NHRCA vs	55	27	49.09	0.351	38	69.09	0.418	22	40	0.011
NHBWS	55	25	45.45		37	67.27		34	61.81	

As regard to the internal validity (central part of table 4.4), results are quite similar across treatments and when using part or the full ranking information. The hit rate ranges from 62% to 73% indicating that, in the main task, the estimated partworths correctly predicted 62% to 73% of participants' responses. Results from the one tailed Z-test show that no significant differences have been found between the four CA formats in terms of internal validity.

Regarding the external validity, the estimated parameters in the hypothetical CE correctly predicted only 23.63% of participants' responses in the holdout task, which is significantly lower than the predictive power of the estimates obtained in the non-hypothetical CA formats. Incentivizing participants to truthfully reveal their preferences seems to enhance the predictive power of the estimated models (Chang et al. 2009; Lusk et al. 2008). Consistent with the findings of Akaichi et al. (2013), we found that the external validity of the estimates in the non-hypothetical CA formats is similar when the responses in NHRCA and NHBWS are coded as choice data taking into account only the option ranked first. Although, the external validity in RBWS is higher than NHCE and RRCA (i.e. the external validity is 52.63%, 38.18%, and 43.63% in RBWS, RRCA, and NHCE, respectively), however this gain in external predictive power is not statistically significant at 5%.

Perhaps the most striking result in our paper is that when all the ranking information is considered in the estimation of the partworths, the external validity is significantly higher in NHBWS than in NHRCA and NHCE. However the external was found to be similar in NHCE and NHRCA. Therefore, our results suggest two important findings: 1) considering all the ranking information is important to appropriately compare choice and ranking CA formats; and 2) NHBWS seems to outperform the other CA formats. Leading us to suggest the use of the NHBWS instead of NHRCA which was the recommended CA formats to use by previous studies.

Since one of the main reasons of using choice experiments is to estimate consumers' willingness to pay for specific food attributes. The results of the comparability of the four CA formats in terms of WTP are displayed in Table 4.5. As it can be observed, consumers are willing to pay a price premium ranging from 0.76 Euro and 1.48 Euro for the extra virgin olive oil; however, they are willing to pay a lower price for the olive oil with respect to virgin olive oil (baseline level). In relation to the origin, respondents were willing to pay a price premium (varies between 0.61 and 0.86 euro across treatments) for the olive oil from Catalonia. Finally, on average, consumers are willing to pay an average premium of about 0.2 Euro for the manufacturer brand with respect to the private brand.

Results displayed in Table 4.6 show a strong similarity of the estimated WTP values across the CA formats. Therefore, the statistical differences we have found among partworths estimates are indeed underlying differences in the error variances due to difference cognitive effort spent by respondents in the different CA formats. Carlsson and Martinsson (2001) and Lusk and Schroeder (2004), also arrived to a similar conclusion when they compared hypothetical and non-hypothetical CE. Two potential explanations could be given: 1) the inclusion of the no-choice option freed participants from being forced to choose a real option for which they have to pay the corresponding price (Hensher 2010); and 2) the effect of a well-script presentation explaining the objectives and the characteristics of each treatment at the beginning of every session has contributed to reduce the hypothetical bias (Hensher, 2010; Murphy et al., 2005). In relation to the non-hypothetical treatments, Akaichi et al. (2013) arrived to a similar result when they compared the NHCE and the NHRCA (considering only information about the most preferred option).

Table 4.5 Estimated Willingness to Pay for each attribute level

	Only choice or first Rank data in ranking treatments				Full ranking information	
Traitment	**HHCE**	**NHCE**	**RRCA**	**RBWS**	**NHRCA**	**NHBWS**
Models	**RPL**	**RPL**	**RPL**	**RPL**	**RO-RPL**	**RO-RPL**
EVOO	0.970*	0.762*	1.139*	0.907*	1.488*	0.823*
	[0.55; 1.38]	[0.38; 1.14]	[0.51; 1.76]	[0.52; 1.28]	[1.01; 1.96]	[0.63; 1.01]
OO	-0.8439*	-0.842*	-1.266*	-0.804*	-1.630*	-0.786*
	[-1.15; -0.53]	[-1.24; -0.43]	[-1.86; -0.66]	[-1.23; -0.37]	[-2.10; -1.15]	[-0.98; -0.58]
BrManf	0.216*	0.176*	0.227	0.171*	0.294*	0.174*
	[0.10; 0.33]	[0.02; 0.32]	[-0.001; 0.45]	[0.01; 0.33]	[0.08; 0.50]	[0.08; 0.26]
CAT	0.816*	0.751*	0.793*	0.863*	0.613*	0.734*
	[0.54; 1.08]	[0.43; 1.07]	[0.36; 1.21]	[0.56; 1.16]	[0.3; 0.92]	[0.54; 0.92]
RSp	-0.406*	-0.475*	-0.956*	-0.501*	-0.950*	-0.766*
	[-0.63; -0.18]	[-0.75; -0.19]	[-1.41; -0.49]	[-0.74; -0.26]	[-1.28; -0.61]	[-0.95; -0.57]

Values in brackets correspond to Confidence Intervals;
An * indicates that the value is statistically different from zero at the 5% siginicance level.

To sum up, this is study is the first attempt to compare the four most used CA formats (i.e. HCE, NHCE, NHRCA and NHBWS). Our results suggest that, independently of whether the partworths are estimated considering only the option ranked first or the full ranking information, the four CA formats provide similar results in terms of the sign and the significance of estimated partworths as well as the estimated WTP values. Therefore, if the estimation of consumers' WTP is the main objective of using CA, then the use of any one of the four CA format is appropriate. Nonetheless, the use of HCE might be preferred due to the simplicity and lower cost of its implementation. However, if practitioners are interested in assessing consumers' preferences and the use of the estimated partworths for prediction sakes (e.g. to predict the change in consumers' choices when the price changes), then NHBWS should be used due to its higher predictive power especially when all the ranking information is included in the estimation.

Table 4.6 Hypothesis test of equality WTPs across the treatments

Hypothesis	EVOO	OO	BrManf	CAT	RSp
	p-value of Complete combinatorial test				
HHCE vs NHCE	0.471	0.439	0.488	0.486	0.458
HHCE vs RRCA	0.486	0.479	0.425	0.483	0.499
HHCE vs RBWS	0.439	0.438	0.398	0.483	0.467
HHCE vs NHRCA	0.478	0.499	0.349	0.387	0.394
HHCE vs NHBWS	0.462	0.471	0.413	0.466	0.436
NHCE vs RRCA	0.449	0.454	0.420	0.453	0.442
NHCE vs RBWS	0.404	0.372	0.382	0.457	0.419
NHCE vs NHRCA	0.442	0.435	0.326	0.357	0.363
NHCE vs NHBWS	0.425	0.404	0.405	0.437	0.497
RRCA vs RBWS	0.456	0.416	0.448	0.498	0.463
RRCA vs NHRCA	0.456	0.448	0.380	0.388	0.395
RRCA vs NHBWS	0.487	0.456	0.487	0.480	0.446
RBWS vs NHRCA	0.487	0.454	0.424	0.411	0.428
RBWS vs NHBWS	0.479	0.463	0.472	0.474	0.438
NHRCA vs NHBWS	0.482	0.495	0.385	0.414	0.368

4.5. Conclusions

Focusing on a market good, the olive oil, this study, aimed to compare HCE, NHCE, NHRCA and NHBWS in terms of consistency, estimated partworths, internal and the external validity and estimated WTP. This study expands the work of Akaichi et al (2013) in three ways: 1) it includes an additional ranking CA format (i.e. BWS) which has been proved to be easily understood by consumers when they are asked to rank several combinations of attributes; 2) the BWS is conducted in a non-hypothetical setting adopting the approach used by Lusk et al. (2008) for RCA; 3) in the ranking treatments (NHRCA and NHBWS), the full ranking information is used to estimate the partworths, while in Akaichi's et al. (2013), only the information provided by the first ranked option was considered.

Results from this study suggest a number of points. First, the estimated partworths from the four CA formats are similar in terms of sign and significance although they have different values. For instance, all the CA format showed that responded preferred the local extra virgin oil. Additionally, the results show that there are not significant differences across the treatments in terms of estimated WTP. This could be used as an argument in favor of the use of hypothetical choice experiment especially, when the main reason for using the CE is to assess consumers' WTP.

Second, using NHRCA or NHBWS could be advantageous if the practitioner is interested in determining not only the most preferred option but also consumers' preferences for all the combinations of attributes included in a choice set. In fact, carrying out CE, RCA and BWS in non-hypothetical context and considering only the most preferred option in the estimation of partworths, were found to give similar results in terms of partworths and WTP. However, the use of RCA and BWS might be preferred since they provide practitioner with information all the options included in a choice set.

Finally, our results showed that the NHBWS seems to outperform the other CA formats when the full ranking information is used. As aforementioned, this superiority might be a result of the fact that BWS better fits natural human skills at identifying extreme values that decreases the cognitive burden on participants which in turn could increase the predictive power of the estimates. Therefore, our findings affirm the superiority of BWS compared with the other formats of CA when it is conducted in non-hypothetical settings. It noteworthy that more research work is, however, needed to see to what extent our results can be generalized. For instance, we think it important to know how our findings will be affected if large choice sets are used. Furthermore, additional empirical applications with alternative food products should be carried out to find out whether our findings are sensitive to the type product (e.g. food vs. non-food product or durable vs. fresh food) used in the experiment.

References

Akaichi, F., R. M. Nayga, Jr. and J. Gil. "Are Results from Non-Hypothetical Choice-Based Conjoint Analysis and Non-Hypothetical Recoded-Ranking Conjoint Analysis Similar?" American Journal of Agricultural Economics 95, 949-963.

Alfnes, F., Guttormsen, A.G., Steine, G., and Kolstad. K. (2006) "Consumers' willingness to Pay for the Color of Salmon: A Choice Experiment with Real Economic Incentives." American Journal of Agricultural Economics 88, 1050-1061.

Ben-Akiva, M., Morikawa, T., and Shiroishi, F. (1992) "Analysis of the reliability of preferences ranking data" Journal of Business Research 24, 149-164.

Boyle, K.J., Holmes, T.P., Teisl, M.F., and Roe., B. (2001) "A Comparison of Conjoint Analysis Response Formats." American Journal of Agricultural Economics 83, 441–54.

Brouwer, R., Dekker, T., Rolfe, J., and Windle, J. (2010) "Choice certainty and consistency in repeated choice experiments" Environmental Resource Economics 46, 93-109.

Brown, T.C., Kingsley, D., Peterson, G.L., Flores, N.E., Clarke, A., and Birjulin, A. (2008) "Reliability of individual valuations of public and private goods: choice consistency, response time, and preference refinement" Journal of Public Economic 92, 1595-1606.

Campbell, D., and Lorimer, V.S. (2009) "Accommodating attribute processing strategies in stated choice analysis: do respondents do what they say they do?" presented at European Association of Environmental and Resources Economist Annual Conference, Amsterdam, June.

Caparrós A., Oviedo, J.L., and Campos., P. (2008) "Would you Choose your Preferred option? Comparing Choice and Recoded Ranking Experiments." American Journal of Agricultural Economics 90, 843-855.

Carlsson, F., and Martisson, P. (2001) "Do hypothetical and actual marginal willingness to pay differ in choice experiments?" Journal of Environment Economics and Management 41, 179-192.

Chang, J.B., Lusk, J.L., and Norwood, B. (2009) "How closely do hypothetical surveys and laboratory experiments predict field behavior?" American Journal of Agriculture Economics 96, 518-534.

Ding, M. (2007) "An incentive-aligned mechanism for conjoint analysis" journal of Marketing Research 44, 214-223.

Ding, M., Grewal, R., and Liechty, J. (2005) "Incentive-Aligned Conjoint Analysis." Journal of Marketing Research 42, 67–83.

Dong, S., Ding, M., and Huber, J. (2010) "A simple mechanism to incentive-align conjoint experiment" International Journal of Research in Marketing, Vol. 27, pp. 25-32.

Flynn, T., and Marley, A.J. (2012) "Best worst scaling: Theory and methods" Sydney, CenSoC Working Paper Nº 12-002.

Helson, H. (1964). "Adaptation-Level Theory" New York: Harper & Row.

Hensher, D.A. (2010) "Hypothetical bias, choice experiment and willingness to pay" Transport Reseach Part B 44, 735-752.

Hensher, D.A., Rose, J.F., Greene, W.H. (2005) "Applied Choice Analysis: A Primer." Cambridge University Press, Cambridge.

Hoeffler, S., and Ariely, D. (1999) "Constructing stable preferences: a look into dimensions of experience and their impact on preferences stability" Journal of Consumer Psycologic 8, 113-126.

Jiménez-Guerrero, J.F., Gásquez-Abad, J.C., Mondéjar-jiménez, J.A. and Huertas-García, R. (2012:233) "consumer preferences for olive oil attributes: a review of the empirical literature using a conjoint approach" chapter 12 in "olive oil constituents, quality, health properties and bioconversions", edited by Boskou Dimitrios, Publisher: In Tech.

Krinsky, I., and Robb, A.L. (1986) "On Approximating the Statistical Properties of elasticities." Review of Economics and Statistics 64,715–19.

Lancaster, K. J. (1966), "A new approach to consumer theory", The journal of political economy, 74, 2, 132-157.

Lancsar, E., Louviere, J., Donaldson, C., Currie, G., and Burgess, L. (2013) "Best worst discrete choice experiment in health: Methods and an application" Social Science and Medicine 76, 74-82.

Louviere, J., Street, A., Burgess, L., Wasi, N., Islam, T., and Marley, A. (2004) " Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information" Sydney, CenSoC Working Paper Nº 04-005.

Louviere, J.J. and Street, D. (2000). "Stated-preference Methods" in David A Hensher & Kenneth J Button (eds), Handbook of Transport Modelling, Pergamon Press, Amsterdam, Netherlands, 131-143.

Louviere, J.J., Street, D., Burgess, L., Wasi, N., Towhidul, I., Anthony, A.J.M. (2008) "Modelling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information" Journal of Choice Modelling, Vol. 1, pp. 128-163.

Lusk, J.L., and Schroeder., T.C. (2004) "Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks." American Journal of Agricultural Economics 86, 467–82.

Lusk, J.L., and Shogren, J.F. (2007) "Experimental auction: Methods and applications in economic and marketing research" Cambridge: Cambridge University Press.

Lusk, J.L., Fields, D., and Prevatt. W. (2008) "An Incentive Compatible Conjoint Ranking Mechanism." American Journal of Agricultural Economics 90:487–98.

McFadden, D. (1973), "Conditional logit analysis of qualitative choice behavior", En Zarembka, P. (ed.) Frontiers in econometrics. Academic press, Nueva York.

Murphy, J.J., Geoffrey, J.A., Stevens, T.H., and Weatherhead, D. (2005) "Meta-Analysis of Hypothetical Bias in Stated Preference Valuation." Journal of Environmental and Resource Economics 30, 313–325.

Pignone, M.P., Brenner, A.T., Hawley, S., Sheridan, S.L., Lewis, C.L., Jonas, D.E., and Howard., K. (2011) "Conjoint Analysis Versus Rating and Ranking for Values Elicitation and Clarification in Colorectal Cancer Screening." Journal of General Internal Medicine (published on line August 26th 2011).

Poe, G.L., Giraud, K.L., and Loomis, J.B. (2005) "Computational Methods for Measuring the Difference of Empirical Distributions." American Journal of Agricultural Economics 87, 353-365.

Scarpa, R., Nataro, S., Louviere., J., and Raffaelli, R. (2011) "Exploring scale effects of best/worst Rank ordered choice data to estimate benefits of tourism in alpine grazing commons" American Journal of Agricultural Economics, Vol. 93, pp. 813-828.

Street, D., and Burgess, L.B. (2007) "The Construction of Optimal Stated Choice Experiments: Theory and Methods", 1st ed, John Wiley and Sons, Hoboken, New Jersey.

Train, K., (2003). "Discrete Choice Methods with Simulation". Cambridge University Press, Cambridge.

Appendix 4.1 An example of a choice set presented in the HCE and NHCE

Choice set 1			**Identification number:** ………			
	OLIVE 1	**OLIVE 2**	**OLIVE 3**	**OLIVE 4**	**OLIVE 5**	**OPTION "NONE"**
T.O. Oil : Brand: Origin: Price:	Virgin Private label Andalucía 3.50	Olive oil Private label Andalucía 3.50	Aceite de oliva Manufacturer Rest of Spain 4.80	Virgin Extra Manufacturer Cataluña 4.80	Virgin Extra Manufacturer Cataluña 2.20	None of them
	☐	☐	☐	☐	☐	☐
Please indicate your most preferred olive oil (taking into account your actual purchase behavior) (Please tick in the appropriate box). In case of none of the five oils resembles for your usual purchase, please choose "None"						

Appendix 4.2 An example of a choice set presented in the NHRCA

Choice set 1			**Identification number: ………**			
	OLIVE 1	**OLIVE 2**	**OLIVE 3**	**OLIVE 4**	**OLIVE 5**	**OPTION "NONE"**
T.O. Oil : Brand: Origin: Price:	Virgin Private label Andalucía 3.50	Olive oil Private label Andalucía 3.50	Aceite de oliva Manufacturer Rest of Spain 4.80	Virgin Extra Manufacturer Cataluña 4.80	Virgin Extra Manufacturer Cataluña 2.20	None of them
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	
Please Rank the olives oil 1, 2, 3, 4, and 5 from the most preferred to the least preferred (taking into account your actual purchase behavior) from 1 to 5. (1= most preferred and 5 = least preferred). Or mark the option "None of them" in the case none of five oils resembles your usual behavior.						

Appendix 4.3 An example of a choice set presented in the NHBWS

Choice set 1			**Identification number:** ………			
	OLIVE 1	**OLIVE 2**	**OLIVE 3**	**OLIVE 4**	**OLIVE 5**	**OPTION "NONE"**
T.O. Oil :	Virgin	Olive oil	Aceite de oliva	Virgin Extra	Virgin Extra	None of them
Brand:	Private label	Private label	Manufacturer	Manufacturer	Manufacturer	
Origin:	Andalucía	Andalucía	Rest of Spain	Cataluña	Cataluña	
Price:	3.50	3.50	4.80	4.80	2.20	
	B 2B W 2W	B 2B W 2W	B 2B W 2W	B 2B W 2W	B 2B W 2W	☐

Please from the five olives oil (taking into account your actual purchase habits), first indicates your best option marking the box "B". Of remaining four olives indicates your worst option marking the box "W", of the remaining three olives please indicates your second best option marking the box "2B" and from the two remaining olives please indicates your second worst option marking the box "2W". In the case none of five oils resembles your usual behavior please mark the option "None of them"

Appendix 4.4 The choice set of Holdout task

Identification number: ………						
	OLIVE 1	**OLIVE 2**	**OLIVE 3**	**OLIVE 4**	**OLIVE 5**	**OPTION "NONE"**
T.O. Oil : Brand: Origin: Price:	Virgin Extra Private label Andalucía 3.50	Virgin Private label Andalucía 2.20	Olive oil Private label Cataluña 4.80	Virgin Extra Manufacturer Cataluña 3.50	Virgin Private label Rest of Spain 4.80	None of them
	☐	☐	☐	☐	☐	
	OLIVE 6	**OLIVE 7**	**OLIVE 8**	**OLIVE 9**	**OLIVE 10**	
T.O. Oil : Brand: Origin: Price:	Olive oil Private label Rest of Spain 3.50	Virgin Manufacturer Cataluña 2.20	Virgin Extra Manufacturer Rest of Spain 2.20	Virgin Manufacturer Andalucía 4.80	Olive oil Manufacturer Andalucía 3.50	
	☐	☐	☐	☐	☐	☐

Please indicate your most preferred olive oil. In case of none of the five oils resembles for your usual purchase, please choose "None"

Chapter 5

Conclusions

Preference elicitation methods have been extensively used by economists and market researchers to determine consumers' preferences willingness to pay (WTP) for specific product attributes. Choice experiment is a preference elicitation method that has been widely used in previous research. To get more reliable information, past literature has mainly focused in both the experiment design and the estimation procedures.

In relation with the latter, the earliest applications assumed that error variances were homogeneous over individuals, estimating a multinomial logit or a conditional logit (CL) model. However, consumer heterogeneity is one of the most fundamental concepts in marketing strategy and planning and it is the base for market segmentation. Moreover, failure to account for preference heterogeneity may not only result in poor model performance (i.e., generating an incorrect standard error and biased parameter estimates) but also affect elasticities and willingness-to-pay measures, all of which could lead to problems in the reliability of model results. Accordingly, substantial efforts have been devoted to incorporate heterogeneity in discrete choice models.

In this thesis, we have applied two new procedures to account for preference heterogeneity. The first procedure (Chapter 2) relies on extending the random parameter logit model framework (GHR-RPL) to account for heterogeneity around both the mean and the variance of the parameter distributions. This procedure contributes to deeply understand the sources of preference heterogeneity. We have assessed the effect of socio-demographic characteristics of consumers as well as the behavioral factors such as consumers attitudes, subjective norms, trust, purchase intention, etc. We have also illustrated the implications on the moments of the WTP.

The second procedure is based on the notion of including new variables to account for preference heterogeneity (Chapter 3). In fact, the literature has focused mainly in treating preference heterogeneity as derivatives of differences in observed consumers' socio-economic characteristics. However, less attention had been paid to the potential role of unobserved variables such as psychological or behavioral factors, which are difficult to measure, on shaping consumers' choices. We have estimated a hybrid choice model (HCM) which allows extending the normal discrete choice modelling by defining an individual's utility function as a function of observed explanatory variables, such as product attributes and respondents' socio-economic characteristics covariates, and including latent variables that can reflect consumers' psychological factors, personality traits or attitudes. To the best of our knowledge, this thesis constitutes the first attempt to introducing HCM in agro-food markets. Moreover, this thesis extends the existing literature in two ways. First, this paper does not merely obtain latent variables from observed indicators but it estimates hierarchical relationships between latent variables using a structural equation model (SEM), which provides a better insight into the consumers' decision-making process. Second, we have estimated a random parameter logit (RPL) model considering the latent variables as random parameters as the best way to integrate latent variables in a multinomial choice framework.

In relation to the experiment design, past literature has applied alternative incentive-compatible conjoint analysis (CA) formats to elicit consumers' preferences and WTP for market (and non-market) goods. The most relevant are: 1) choice experiments (CE), where consumers who are faced with competing products purchase the product that fit most their preferences; 2) ranking conjoint analysis (RCA) where consumers are also provided with a set of product concepts but they are asked to rank them from the most to the least preferred; and 3) best worst scaling (BWS) where respondents are asked to firstly choose the best option in each choice set, then the worst option, then the second best and the second worst options from the remaining options and so on until a complete preference ordering of all the options is obtained.

Few researchers, however, have compared the performance of alternative CA formats in terms of estimated marginal partworths, the predictive power of the derived models, and the reliability of the Willingness to Pay (WTP) values deduced from the estimated partworths. Chapter 4 compares the three mentioned above CA formats in a non-hypothetical setting. The hypothetical CE has been used as the benchmark. This thesis has expanded previous literature in three main ways: 1) it includes an additional ranking based CA, the best worst scaling (BWS), which has been proved to be easily understood by consumers when ranking preferences and which has been re-coded as a RCA; 2) it is the first attempt to design the BWS incentive compatible; 3) in the ranking treatments, the full ranking information is used to estimate the partworths, while in previous literature only the information provided by the first ranked option had been considered

Throughout the thesis, the olive oil market has been chosen as the case study. Spain is the first producer and exporter country of extra-virgin olive worldwide. Additionally, olive oil constitutes a fundamental component of the Spanish diet. As a consequence, the vast majority of Spanish consumers are knowledgeable about this product, and all of them are aware of market prices and product characteristics. Geographically, this thesis has focused on Catalonian consumers. Catalonia is the second more important region in terms of olive oil consumption. Two different experiments have been carried out to tackle with the three issues (objectives or chapters) mentioned above. In both cases, a representative sample of Catalonian consumers has been surveyed.

Although in each of the chapters we have introduced some conclusions, in the following lines we want to highlight a few general conclusions from the whole thesis. Comparing the results from Chapters 2 and 3, we can highlight that the two estimation procedures used in this thesis (GHR-RPL and HCM) coincide with the idea that socio-demographic factors only provide a partial explanation of preference heterogeneity. Behavioral factors, personality traits, lifestyle, and purchase habits, on the contrary, better contribute to capture consumer's preference heterogeneity, apart from significantly improving the goodness-of-fit of the models. This result has to be taken into account by policy makers and marketing managers to design their policies or marketing strategies. Moreover, future research should not neglect these variables when eliciting consumers' preferences for food attributes.

From an empirical point of view, it seems that the most important attribute that affects consumers' preferences toward extra-virgin olive oil are the price, followed by the origin of the product, especially the local origin. The organic attribute is not perceived as a significant quality cue. Environmental and health concerns, and a healthy lifestyle seem to not be relevant to consumers' choices. The most relevant quality cue is the PDO attribute. This result suggests that the traditional marketing strategies that have been used to promote the consumption of olive oils, especially organic olive oils, based on environmental or health issues have to be reoriented. May be strategies trying to reinforce the local attribute could be encouraged.

Furthermore, we have shadowed light about the significant effect of consumer's purchasing habits and attitudes on food choices. In this thesis, two purchasing attitudes have been defined: "price Involvement" and "Quality Involvement". Both have a positive effect on the consumers' utility. Given the large number of food choices that consumers have to make every day, people looking for good prices (price involvement) can easily accommodate their preferences either by buying directly from the producer or cooperative or by choosing a promoted product at the retail outlet. Something similar can be said for consumers looking for quality as the market offers a reasonable product mix. In any case, purchasing habits and attitudes strongly depend on the availability of the product in the market, the frequency and intensity of promotional activities and personal characteristics such as shopping frequency, impulsivity, etc.

In relation to the experiment design (and the election of the CA format), results from this thesis suggest comparisons across different CA formats only can be undertaken if the same design is used. We have also highlighted that estimated parthworths from hypothetical and non-hypothetical experiments are similar in terms of sign and significance. As a consequence if the interest of the researcher is only to assess consumers' preferences and consumer' WTP hypothetical setting could be adequate. However, from the external validity point of view, the non-hypothetical experiments clearly outperform the hypothetical one.

The use of ranking based conjoint analyses should be encouraged. Their predictive power is similar to the traditional choice based conjoint analysis but they provide additional information on consumers' preferences for all the combination of attributes included in a choice set. Finally, the best worst scaling (BWS) format has been revealed to outperform the other formats in terms of predictive power as its cognitive process seems to better fits the natural tendency of humans at identifying the extreme values.

In any case, this dissertation is just the starting point of a future research career. In this context, we would like finishing this dissertation by outlining some potential issues for future research that have arisen from this study. First, we would like insisting about the need to incorporate psychological factors in discrete choice models to allow researchers to reveal consumers' preference heterogeneity and to better understand the consumers' decision making process. The use of the HCM model should be encouraged as a useful tool to tackle with this issue. In this thesis we have used a two-step approach, while further research should be addressed to simultaneously identify latent variables and to include them into the discrete choice model, providing more efficient estimators of the parameters involved.

Secondly, the principal axiom of the random utility theory (RUT) is the rationality of the consumers (consumers maximize their utility function taking into account all factors or attributes available as well as their budget). However, this is not true. Several choice experiment studies have found that in a given choice set, respondents sometimes appear to ignore one or several attributes (Kehbacher et al., 2013; Hensher and Rose, 2009). Therefore, their behavior is inconsistent with the basic assumption of the RUT which generates biased willingness-to-pay (WTP) estimates. Further research should be carried out on this issue. While some research has been done on non-attendance attributes, the literature has merely accounted for attributes that consumers have or have not taken into account when making choices. However, no attempt has been done to take into account the ranking of non-ignored attributes in the choice experiment. Additionally, we could investigate to what extent the different CA formats are sensitive to non-attendance attributes and what are the effects in terms of estimated partworths, internal and external predictive power and WTP measures.

Finally, we would like mentioning that, as have been observed, each chapter has been written as it was a journal paper. The first one (chapter 2), titled “Revealing additional preference heterogeneity with extent Random parameter logit model: The case of extra virgin olive oil for Catalan consumers” is under review in the *Spanish Journal of Agricultural Research*. The second one (chapter 3) titled “The effect of personality traits on consumers' preferences for extra virgin olive oil” is under review in *Food Quality and Preference*. The third paper (chapter 4) titled “Are ranking preferences information methods comparable with the choice experiment information in predicting actual behavior?” is under review in the *European Review of Agricultural Economics*.

References

Hensher, D.A., and Rose, J.M. (2009) “Simplifying choice through attribute preservation or non-attendance: Implication for willingness to pay” Transportation Research Part E, 45, pp. 583-590.

Kelhbacher, A., Balcombe, K., and Bennett, R. (2013) “Stated attribute non-attendance in successive choice experiment” Journal of Agriculture Economics, 64, pp. 693-706.