PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE

SCHOOL OF ENGINEERING

MODELLING WINE CONSUMER PREFERENCES USING HYBRID CHOICE MODELS: INCLUSION OF INTRINSIC AND EXTRINSIC ATTRIBUTES

DAVID ESTEBAN PALMA ARANEDA

Thesis submitted to the Office of Graduate Studies in partial fulfilment of the requirements for the Degree of Doctor in Engineering Sciences

Advisors:

JUAN DE DIOS ORTUZAR SALAS

LUIS IGNACIO RIZZI CAMPANELLA

Santiago de Chile, June 2016

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE

ESCUELA DE INGENIERIA

MODELLING WINE CONSUMER PREFERENCES USING HYBRID CHOICE MODELS:

INCLUSION OF INTRINSIC AND EXTRINSIC ATTRIBUTES

DAVID E. PALMA ARANEDA

Members of the Committee:

JUAN DE DIOS ORTUZAR SALAS

LUIS IGNACIO RIZZI CAMPANELLA

CRISTIAN ANGELO GUEVARA CUE

EDUARDO AGOSIN TRUMPER

CARLOS CORNEJO ALARCON

STEPHANE HESS

CRISTIAN VIAL EDWARDS

Thesis submitted to the Office of Research and Graduate Studies in partial fulfilment of the requirements for the Degree of Doctor in Engineering Sciences.

Santiago de Chile, June, 2016

To Gilda, for her love, patience and care.

ACKNOWLEDGMENTS

I would like to begin by thanking my two advisors: Juan de Dios Ortúzar and Luis Ignacio Rizzi. They had the patience and good will to guide me through all the work put on this thesis, lending me their knowledge and experience, and always pushing me to go further. Even more, they trusted and supported my ideas, teaching me to face research as a creative and challenging work.

I would also like to thank Angelo Guevara and Stephane Hess, who were very kind on lending their time and effort on shaping raw ideas (and data) into actual models and interesting results; I have learnt much from them. I am also very thankful for the time spent at Leeds, where I got the chance to work and share with such a wonderful and talented group of people. I learnt much during my stay, both professionally and personally.

I would like to thank Gerard Casaubon for his help throughout this thesis, and even before that, during my research as a Master student. I am thankful for him introducing me to the world of wine and sensory science. I also want to thank Professors Eduardo Agosín and Carlos Cornejo, who have always been supportive and extremely helpful at every stage of my research work. Their help has been and invaluable addition during my time as a postgraduate student.

I would like to thank all the crew at Centro de Aromas y Sabores, who always made me feel welcome and helped me dealing with the subtleties of sensory sciences. Thanks are specially due to Cristian Castro, for all his help with this project. I am thankful to everyone at the Centre, and greatly appreciate all of your help.

I would also like to thank Juan Francisco Rodríguez and Felipe Bascuñán from WineClr. Their support for my research was key on stressful times. I am in debt with Benjamín Lizana too, with whom I had the pleasure of working together through an important part of the thesis. Without his help, support and hard work this project would not have been completed.

Thanks are due to all those who partially financed this research and my postgraduate studies: CONICYT through CONICYT-PCH / Doctorado Nacional / 2015 - 21151075; CONICYT + PAI / *Tesis de Doctorado en la Empresa* national tender, 2014 - 781413005; the Chilean Fund for the Development of Scientific and Technological Research (FONDECYT) through Project 1121058 and 1150590; the Millennium Institute in Complex Engineering Systems (ICM: P05-004F; FONDECYT: FB016); the vice presidency of research (VRI) of Universidad Católica de Chile; and Anakena and Concha y Toro S.A. wineries.

Throughout all the time preparing my thesis many people helped me in different ways. Apparently casual and inconsequential conversations are often the seed for new ideas, or the way to see an old problem from a fresh perspective. Many people have taught me that caring for others' work and sharing one owns is the fastest way to breakthrough. I would like to thank Jaime Allen, Felipe González, Carlos Bueno, Lidia Bonet, Alejandra Lizana, Sebastián Gebhardt, Cristóbal Sirhan, Manuel Ojeda, Romain Crastes, Jiwei Zheng, Chiara Calastri, Mauro Capurso, Thijs Dekker and many more.

I would also like to thank all who helped me during the last experiment of this thesis: Francisca León, Paulina Pérez and Jorge Navarro. Your help is much appreciated.

I would like to thank my family for their constant support and encouragement in following my interests. I will always owe much of what I am and do to them.

And finally, I would like to thank my beautiful and clever wife. This work would have been impossible without her love, care, patience and refreshing ideas.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

MODELACION DE PREFERENCIAS DE CONSUMIDORES DE VINO USANDO MODELOS HIBRIDOS DE ELECCION DISCRETA: INCLUSION DE ATRIBUTOS INTRINSECOS Y EXTRINSECOS

Tesis enviada a la Dirección de Investigación y Postgrado en cumplimiento parcial de los requisitos para el grado de Doctor en Ciencias de la Ingeniería

DAVID E. PALMA ARANEDA

RESUMEN

La industria alimentaria enfrenta un mercado crecientemente competitivo, forzando a los productores a buscar constantemente oportunidades para diferenciar sus productos de sus competidores. La diferenciación requiere conocer y ocuparse de las preferencias de los consumidores, pero medir preferencias de consumidores no es tarea fácil.

Esta tesis apunta a validar los modelos de elección discreta como una herramienta apropiada y útil para medir las preferencias de consumidores de alimentos y bebidas. Basándonos en el *Total Food Quality Model* (Grunert, 2005), identificamos tres capacidades críticas que cualquier modelo apropiado y útil debiera tener: (i) modelación de la heterogeneidad de preferencias, (ii) consideración del efecto positivo del precio como un indicador de calidad (asociación entre precio y calidad) y (iii) modelación conjunta de la influencia de atributos intrínsecos y extrínsecos en la decisión de compra. Los atributos extrínsecos son aquellos observables por los consumidores antes de comprar el producto por primera vez (por ejemplo el empaque, el precio y la publicidad), mientras que los

intrínsecos son aquellos que sólo pueden ser percibidos después de consumir el producto (por ejemplo, sabor y aroma). Nos ocupamos de cada punto a través de un análisis de tres experimentos distintos, usando el vino como caso de estudio. El vino es un producto ideal para estudiar dada su complejidad social, química y sensorial, destacando todas las particularidades de los alimentos y bebidas.

El primer experimento modela la heterogeneidad de preferencias en una forma útil para aplicaciones de *marketing*. Mediante un estudio de preferencias declaradas (PD) descubrimos actitudes de los consumidores que explican la heterogeneidad en sus preferencias, pero también encontramos que es difícil ligar estas actitudes a características observables de los consumidores. Esta última falta de correlación torna difícil extrapolar conclusiones a nivel de la población.

El segundo experimento busca medir separadamente el efecto positivo del precio (como indicador de calidad) de su efecto negativo como limitante dada la restricción presupuestaria de los consumidores. Hacemos esto mediante el uso de variables latentes. El enfoque usado funciona como es de esperar, permitiendo una estimación más precisa de los parámetros de precio y ayudando a evitar problemas de sesgo de especificación que tornan los parámetros estimados inconsistentes.

El tercer experimento mide conjuntamente el efecto de atributos intrínsecos simples y extrínsecos. Usamos un modelo en cuya estructura el efecto de los atributos intrínsecos en la compra está medida por el nivel de agrado del consumidor con el producto. El modelo también funciona adecuadamente, permitiendo calcular incluso disposiciones al pago por atributos intrínsecos lo que es bastante novedoso.

Descubrimos que armonizando la estructura de los modelos de elección con una versión simplificada del *Total Food Quality Model* podíamos alcanzar mejoras significativas. Más precisamente, logramos una mejor comprensión de las preferencias del consumidor y su proceso de compra en general. Estas mejoras hacen de los modelos de elección discreta una herramienta adecuada y útil para modelar preferencias de consumidores de alimentos y bebidas.

Miembros de la Comisión de Tesis Doctoral:

Juan de Dios Ortúzar Salas
Luis Ignacio Rizzi Campanella
Cristián Angelo Guevara Cué
Carlos Cornejo Alarcón
Stephane Hess
Cristián Vial Edwards

Santiago, Junio de 2016

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

MODELLING WINE CONSUMER PREFERENCES USING HYBRID CHOICE MODELS: INCLUSION OF INTRINSIC AND EXTRINSIC ATTRIBUTES

Thesis submitted to the Office of Research and Graduate Studies in partial fulfilment of the requirements for the Degree of Doctor in Engineering Sciences by

DAVID E. PALMA ARANEDA

ABSTRACT

The food and beverages industry faces an increasingly competitive market, forcing producers to constantly search for opportunities to differentiate their products from their competition. Differentiation requires producers to know and tend for consumer preferences, but measuring consumer preferences is a difficult task.

This thesis aims to validate discrete choice models as a proper and useful tool for measuring consumer preferences for food and beverages. Based on the *Total Food Quality Model* (Grunert, 2005), we identify three critical properties that any proper and useful modelling technique should have: (i) treat preference heterogeneity, (ii) consider the positive effect of price as a cue for quality (price – quality association), and (iii) jointly model the influence of intrinsic and extrinsic attributes on the purchase decision. Extrinsic attributes are the ones observable by consumers before the first purchase of the product (e.g. packaging, price, advertising), while intrinsic attributes are those that can only be perceived after consuming the product (e.g. taste and aroma). We deal with each point by means of three different experiments, using wine as a case study. Wine is an ideal product

to study given its social, chemical and sensory complexity, highlighting all particularities of foods and beverages.

The first experiment aims to model preference heterogeneity in a way that is useful for marketing applications. Using a Stated Choice (SC) experiment we find consumer attitudes to explain preference heterogeneity. But we also find that attitudes are hard to link with consumers' observable characteristics. This latter lack of correlation makes it difficult to extrapolate conclusions at the population level.

The second experiment seeks to separately measure the positive effect of price as a cue for quality, from its negative effect as a strain to the consumers' budget restrain by using a latent variable approach. The approach works as expected allowing for a more precise estimation of the price parameters and helping to avoid misspecification problems that may render the estimated coefficients inconsistent.

The third experiment jointly measures the effect of (simple) intrinsic and extrinsic attributes. We use a model structure where the effect of intrinsic attributes on the purchase decision is mediated by the consumers' liking of the product. The model also works adequately, even allowing us to calculate willingness to pay for intrinsic attributes, a novel result.

We find that harmonizing the structure of discrete choice models with a simplification of the *Total Food Quality Model* leads to significant improvements. More precisely, it leads to a better understanding of consumers' preferences and their choice process in general.

These improvements allow turning choice models into appropriate and useful tools for modelling preferences for food and beverages products.

Member of the Doctoral Thesis Committee:

Juan de Dios Ortúzar Salas

Luis Ignacio Rizzi Campanella

Cristián Angelo Guevara Cué

Carlos Cornejo Alarcón

Stephane Hess

Cristián Vial Edwards

Santiago, June 2016

1 INTRODUCTION

This chapter offers a general overview of the thesis' main subjects. It includes a brief introduction to the subject (section 1.1), an enumeration of the thesis's objectives and hypothesis (section 1.2), a summary of the methodology (1.3), and it closes with a brief literature review of choice models applied to food and beverages (section 1.4).

1.1 The choice of food and beverages

Product differentiation is a primary objective of almost any producer in any market. By differentiating their products, producers do not directly compete with others achieving the benefits of a monopoly (Chamberlin, 1933). But for differentiation to be effective, the product's differentiating characteristics must be perceived and valued by consumers (Kotler & Keller, 2012, p. 289). This poses a difficult question: what do consumers perceive and value in a given product category?

The question about perception and value is formulated from the perspective of consumers. Therefore, for a product to differentiate, it is not enough to improve its objective attributes, such as the chemical composition of a new food or beverage, or the technical components of a new gadget. Instead, differentiation lies in consumers' perceptions and preferences. If a product's attributes –no matter how revolutionary or technically superior they might be- are not perceived and valued by consumers, then they are as good as non-existent. Furthermore, it is possible to conceive a product that can be differentiated from its competitors without even changing its physical characteristics, but by changing its perception and valuation by consumers.

The question, then, becomes how to understand –and measure– consumers' perceptions and preferences. If a certain attribute is perceived by consumers and matches their preferences, then it will have value to them. But if attributes and preferences do not match, then there is no value in the product, no differentiation and no additional profit for the producer.

Measuring consumers' perceptions and preferences is never an easy task, and in the case of food and beverages it poses some particular complications. Food and beverages are experience goods (Nelson, 1970), meaning that consumers cannot perceive all their attributes before consumption. Unlike a car, the attributes of which can be read in a manual and the product experienced during a test drive, food and beverages must be consumed to achieve full perception about them. This makes the purchase decision of food and beverages a two-stage process, influencing the way perceptions and preferences should be measured.

Grunert (2005) proposes a conceptualization of food and beverages consumption process, effectively describing a two-stage behavioural model. The first stage is concern with a product's first purchase. As consumers lack full information about a product the first time they face its purchase decision, they must rely on an expectation of quality (or *expected quality*) to decide their purchase. This means that consumers must look for cues of quality in a product before purchasing it. These cues can be any attributes than can be perceived before consumption. Such attributes are called *extrinsic*, and typical examples of them are packaging, price and advertising of the product. This decision with incomplete information is the first stage of the decision process.

After buying the product consumers can eat or drink it, allowing them to perceive its *intrinsic* attributes such as taste and aroma. After tasting the product, consumers achieve full information about it, perceiving its *experienced quality*. But the *experienced quality* does not depend only on the product's intrinsic attributes, but also on the consumers' expectation of quality, that is, on *expected quality*. Furthermore, the consumption context can also influence the *experienced quality*. To see this, consider drinking a certain wine during a pleasant meal or during a tense situation; probably the *experienced quality* of the product will be quite different.

As part of any subsequent purchase occasions (i.e. re-purchase) consumers will have more information than during the first purchase. Those times, consumers will once again be influenced by extrinsic attributes, but also by their previous *experienced quality*. However, the influence of *experienced quality* in new purchases will be determined by how precisely consumers can recall it. In other words, if consumers cannot remember how much they liked a particular product, then the influence of *experienced quality* on the repurchase decision may be lower.

The influence of extrinsic and intrinsic attributes in the *expected* and *experienced quality* is not homogenous among consumers, as preferences for both types of attributes vary between consumers. For instance, while some consumers may prefer vivid colours in packages, others may favour more opaque colouring; and the same is valid for the intrinsic attributes: while some people may prefer sweet products, others may rather consume sour ones. Furthermore, in the case of intrinsic attributes, important variability exists at the perception level (Lawless 1980), making it hard to determine how consumers perceive products.

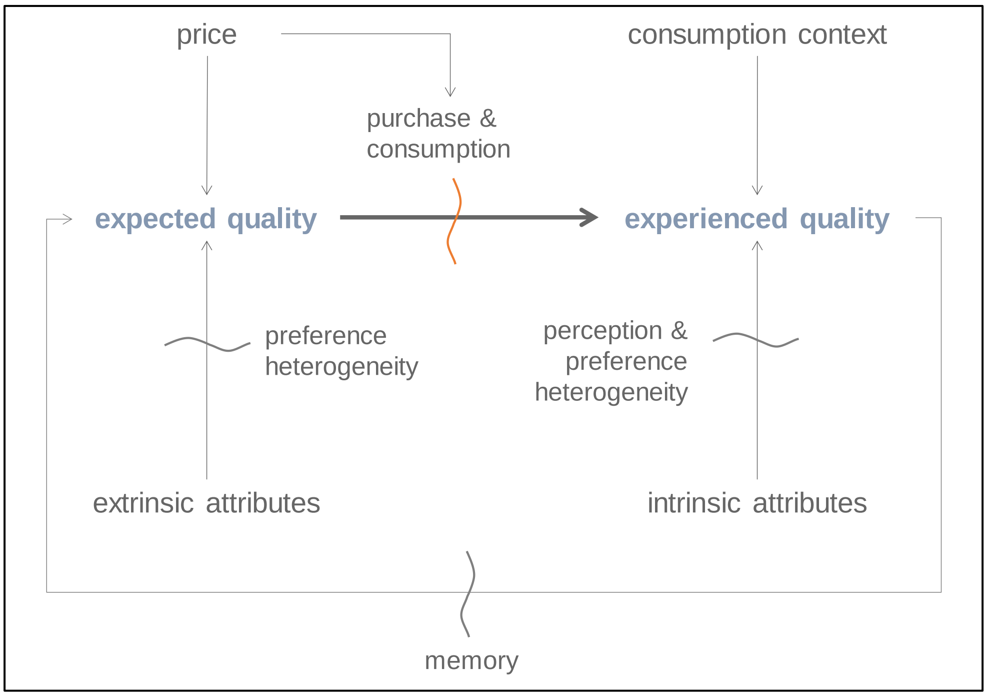


Figure 1-1 - Behavioural model of food and beverages purchase decisions (based on Grunert 2005)

Finally, the behavioural model allows for price influencing the purchase decision in two opposite ways. First of all, price can be used as a cue for quality, just like any other extrinsic attribute. So, consumers may assume that higher prices imply higher quality, a reasonable assumption in perfect markets according to Scitovsky (1945). But second, all consumers are subject to a budget restriction, pushing them to prefer lower prices. Therefore, price has two opposite effects: one positive as a cue for quality, and one negative as a strain to the budget constraint. Figure 1-1 presents a graphical representation of the behavioural model.

Based on the model proposed by Grunert (2005), there are three requirements that any mathematical implementation of the purchase process should allow for:

(i) Preference heterogeneity, that is, the model should not assume that all consumers have the same preferences;

(ii) Separate measurement of the effects of price as a cue for quality and as a strain to the consumers' budget restriction; and

(iii) Measurement of intrinsic and extrinsic attributes to compare their relative importance.

While these three requirements do not guarantee an ideal model, they capture the core of the behavioural model. Therefore, any method to study consumer preferences for food and beverages should fulfil the previous requirements to be considered an adequate and useful tool.

Among all food and beverages, wine is a particularly complex product. Not only it has deep cultural and religious symbolism (Stanislawski 1975), but wine also has an incredible rich and complex sensory variety (Ferreira et al. 2007), to the point of being overwhelming to new consumers (Charters & Pettigrew 2003). These characteristics make of wine a very interesting product to study, as it concentrates all the characteristics that define an experience product. Additionally, wine is an important part of the Chilean economy, as exports during 2015 were valued at USD 1,428 million (ODEPA, 2016).

1.2 Objectives and hypothesis

The main objective of this thesis is to show that discrete choice models are a useful and adequate tool to understand and measure consumers' perceptions and preferences for food and beverage products. To do this, we use wine as a case study.

We base our analysis in Grunert (2005)'s behavioural model, and therefore pursue the inclusion of preference heterogeneity, the double effect of price, and extrinsic and intrinsic attributes in our modelling. Table 1-1 summarises the thesis objectives and Table 1-2 its hypotheses. Each objective is associated with one experiment and each hypothesis is tested on at least one experiment.

Table 1-1 – Main and specific objectives

Main objective	Show that discrete choice models are a useful and adequate tool to understand and measure consumers' perception and preferences for food and beverage products.
Specific objectives	i. Consider preference heterogeneity in a way that eases extrapolation of results to the general population ii. Separately measure the positive effect of price as a cue for quality and its negative effect as a strain to the consumer's budget restriction iii. Measure the effect of both intrinsic and extrinsic attributes in a way that allows comparing their relative importance

Table 1-2 - Hypotheses

Hypotheses	i.	The level of involvement correlates with consumers' preferences and can be explained, to a reasonable extent, by consumers' socio-demographic characteristics and consuming habits.
	ii.	Consumers use price as a cue for quality, leading to price having a double effect in utility. Measuring only the net effect of price can lead to inconsistent estimates.
	iii.	The impact of intrinsic attributes on the consumer's choice is mediated by their liking of the product.

Throughout this document, we discuss how preference heterogeneity, the double effect of price, and the joint measurement of extrinsic and intrinsic attributes can be handled with discrete choice models. We report, discuss and analyse results from a small qualitative investigation and four different choice experiments.

The small qualitative research piece aimed to identify the motives for drinking among Chilean wine consumers. Drinking motives are relevant, as they may influence the way consumers choose. Through in-depth interviews and focus groups, we discovered four motives for wine drinking among Chilean wine consumers. Though a paper was written reporting these results, a quantitative validation of its results is still pending. These results are presented and discussed in APPENDIX I: Tell me why you like to drink wine: Drinking motives as a basis for market segmentation.

The first choice experiment provided data to study heterogeneity in consumers' preferences. We attempted to explain preference heterogeneity based on the consumers' level of involvement with wine. For this we modelled consumers' involvement as a latent variable. Results from this experiment are discussed in Chapter 2.

The second experiment was designed to differentiate the consumers' use of price as a cue for quality from its effect as a strain to the budget constraint. We did this by explicitly modelling consumers' perception of quality and price's positive effect on it. Then, we separately modelled the purchase decision as a trade-off between quality and price, therefore capturing the negative effect of price as a strain to the budget constraint. This modelling approach is based on the behavioural model's structure (Figure 1-1). This experiment is discussed in Chapter 3.

The third experiment was designed to jointly measure the impact of extrinsic and intrinsic attributes on the consumers' choice of wine. The effect of the intrinsic attributes on choice is mediated by how much consumers like the product, so their effect is not as direct as that of the extrinsic attributes. Besides this, we also measured the positive and negative effects of price separately, by modelling consumers' perception of quality and purchase decision as two different but related processes, in a similar fashion to the procedure used in the second experiment. This experiment is discussed on chapter 4.

An additional fourth experiment is reported in APPENDIX II: Modelling wine consumer choices in an incentive-compatible experiment. The results from this experiment are still preliminary, and a paper has not been written yet, but due to its relevance and scale we decided to include it in this document. This experiment, unlike the previous ones, is incentive-compatible, providing revealed preference data under controlled conditions. In this experiment we study the complete purchasing process, that is: selection of wine in a shelf based on extrinsic attributes, tasting, and re-purchase. We modelled all stages jointly, allowing us to compare the effect of intrinsic and extrinsic attributes, and allowing for purchasing more than just one bottle of wine by each participant.

1.3 Methodology

In this section we briefly discuss the foundations of discrete choice models and give a short review of their applications in the food and beverages' literature.

1.3.1 Choice models

Discrete choice models (DCM) were developed in the 1970's mainly in the area of transportation research (McFadden 1973) and are now heavily used in many different areas, such as urban planning, market research, telecommunications, housing, security, management and public policy. DCM require finding or defining a choice context where consumers must pick one alternative from a choice set of available alternatives.

One possible way to gather such data is through a stated preference (SP) experiment, consisting in a series of simulated choices (i.e. individuals are asked what they would choose in hypothetical, but realistic, situations). Alternatives (i.e. products) are defined by a set of attributes. Consumers, on the other hand, are assumed to have preferences for these attributes, which can vary across individuals. It is typically further assumed that consumers exhibit compensatory behaviour, that is, if a given alternative has one attribute with a low rating and another with a high one, it can still be chosen by compensating on these. The interaction of attributes and preferences determine the amount of utility that each consumer attains from each alternative. Consumers are assumed to choose the alternative that maximizes their utility.

However, as the modeller -unlike the consumer- does not have full information about the alternatives, individual utility is assumed to have a random error component. Depending on the assumed distribution of these error components, several discrete choice models can be generated (Ortúzar & Willumsen 2011, Chapter 7).

Using DCM to study consumer preferences has several advantages. First of all, it is an indirect method to assess the importance of a product's attribute. Direct and indirect methods are not equivalent (Louviere & Islam 2008), and asking consumers directly for their preferences may render biased results (Mueller et al. 2010a). Secondly, the methodology requires consumers to do nothing else than what they would normally do when buying (i.e. choosing among a set of alternatives).

There are several DCM models available to analyse discrete choice data. The most popular ones belong to the logit family (McFadden 1973; Ortúzar & Willumsen 2011; Train 2009); in particular, the Multinomial Logit (MNL) model assumes utility to be additive on the attributes, and the error component to also be additive and independent identically distributed (iid) Extreme Value type I. This last assumption allows deriving a closed form for the probability of choosing a given alternative

(McFadden 1973). Assuming that all respondents answer the same choice situations, the basic MNL formulation can be expressed as following in practice:

$$U_{int} = X_{it}\beta \quad + \varepsilon_{int} \tag{1.1}$$

$$P_{int} = \frac{e^{X_{it}\beta}}{\sum_j e^{X_{jt}\beta}} \tag{1.2}$$

where U_{int} is the utility of alternative i for choice situation t for individual n; X_{it} is a vector of attributes; ε_{int} is the iid Extreme Value type I error component; P_{int} is the probability that individual n chooses alternative i in choice scenario t; e is the exponential function; and β is a vector of parameters to be estimated. This formulation is lacking a scale parameter (λ) inversely related to the unknown standard deviation of the errors ε_{int} as it cannot be identified and is typically normalised to one.

The MNL model has too many restrictions (it does not allow for correlation among alternatives, heteroskedasticity and taste variations among individuals). For this reason, it has been almost completely replaced by the Mixed Logit (ML) model (McFadden & Train 2000; Train 2009, chapter 6), which has concentrated most attention and research in the last years, thanks to its flexibility.

One version of the ML (the random parameters model), assumes that the parameters can vary among individuals (i.e. β_n) and that come from a common distribution $\beta \sim f(\theta)$, with θ a vector of *population parameters* of the distribution (e.g. mean and standard deviation). This assumption, however, makes the calculation of probabilities somewhat more involved, now having the following form:

$$P_{int} = \int \frac{e^{X_{it}\beta}}{\sum_j e^{X_{jt}\beta}} f(\beta|\theta) d\theta \quad (1.3)$$

Probabilities of this form are usually estimated using Monte Carlo methods (Train 2009, chapters 9 and 10). This method basically consists in drawing a large number (*K*) of random points from $f(\theta)$, each called β_k, and then calculating $\sum_{\beta_k} \frac{e^{X_{it}\beta_k}}{\sum_j e^{X_{jt}\beta_k}} / K$, which is a consistent estimator of the integral.

Another common form of the ML model (the error components model) consists in including additional random error terms to allow passing from the simple covariance matrix associated with the MNL model to any one desired (Train & McFadden 2000). In particular, in the case of SP data it is good practice to allow for an error component that correlates all responses by the same individual. This error term (ν_{in}) is different between individuals and alternatives, but common across responses of the same individual.

The error component is many times assumed to follow a Normal distribution with mean zero and a standard deviation σ to be estimated. This extra error component leads to the following form of the utility and probabilities.

$$U_{int} = X_{it}\beta_n + \nu_{in} + \varepsilon_{int} \quad (1.4)$$

$$P_{int} = \iint \frac{e^{X_{it}\beta + \nu_{in}}}{\sum_j e^{X_{jt}\beta + \nu_{in}}} f(\beta|\theta) \phi(\nu_{in}|0, \sigma) d\theta \, d\nu \quad (1.5)$$

where $\phi()$ is the Normal probability density function. The integral is usually estimated using Monte Carlo methods (Ortúzar & Willumsen 2011, Chapter 8).

1.3.2 MIMIC models

Multiple indicators and multiple causes (MIMIC) models are a particular kind of structural equations models (SEM, Bollen 1989). They are characterized by the presence of one or more *latent variables*, their *indicators* and their *causes*. Latent variables are unobservable variables –usually psychological constructs- that are of interest to the researcher, but as they are unobservable, they cannot be measured directly. Indicators are observable variables that depend on the latent variable, i.e. the value of indicators is at least partially determined by the latent variable. Finally, the causes of the latent variable are (usually observable) variables that are able to explain the value of the latent variable.

For an easier understanding of the concepts behind a MIMIC model, let us consider an application in the field of education. The government wants to increase student's skills, but they do not know where to invest to maximize the impact. First of all, a student's skill is not directly observable, therefore they must use standardized test to measure it. Only after measuring the skill level of their students, they can attempt to explain these values based on different possible observable variables: their household income, their parent's level of education, the number of books at their homes, etc. In this example, the skill level is the latent variable (an unobservable psychological construct), the standardized tests and their individual items are the indicators (as more skilful students are expected to answer correctly more often), and socio-demographic characteristics of the students are assumed to be the cause of the latent variable (as the latent variable is explained by these characteristics). Such a model can be represented as in Figure 1-2.

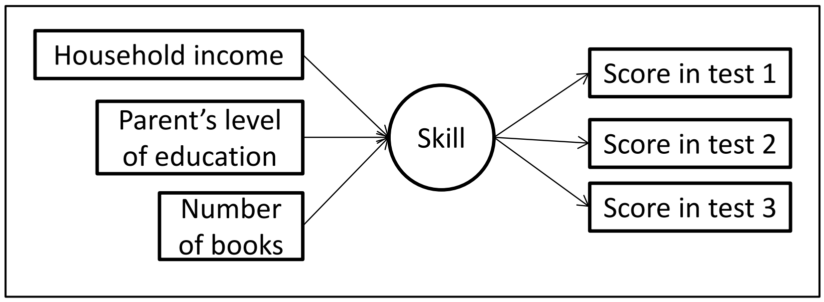


Figure 1-2 - Example of a graphical representation of a MIMIC model

The relationship between the cause and latent variable is called *structural equation*, and is usually assumed to be a linear model. In the example of Figure 1-2, the structural equation could be written as:

$$skill_n = \beta_{income} income_n + \beta_{parents} parentsLvlEduc_n + \beta_{books} nBooks_n + \varepsilon_n \qquad (1.6)$$

where $skill_n$ is the skill level of student *n*; $income_n$ is the income of his/her household; $parentsLvlEduc_n$ is his/her parents' level of education; $nBooks_n$ is the number of books in his/her household; ε_n is an independent identically distributed Normal error; and $\beta_{income}, \beta_{parents}$ and β_{books} are parameters to be estimated.

The relationship between the latent variable and its indicators is called *measurement equation*. These equations can have many different forms, but in this thesis they are usually assumed to be ordered logit models (Train 2009, chapter 7.3). In the case of

Figure 1-2, this would imply that the score in each test can have a finite number of values, e.g.: 1, 2, 3, 4, 5, 6 or 7. The probability of student n obtaining a score s on test 1 depends on his/her skill through the following equations:

$$l_n^{test1} = \lambda_{test1} skill_n + \epsilon_n^{test1} \tag{1.7}$$

$$\begin{aligned} P(score1_n = s) &= P(\tau_{s-1}^{test1} < l_n^{test1} < \tau_s^{test1}) \\ &= \frac{1}{1 + e^{\lambda_{test1} skill_n - \tau_s^{test1}}} - \frac{1}{1 + e^{\lambda_{test1} skill_n - \tau_{s-1}^{test1}}} \end{aligned} \tag{1.8}$$

where l_n^{test1} is an auxiliary latent variable, ϵ_n^{test1} is an Extreme Value type I error component; and $score1_n$ is the score obtained by student n in test 1. λ_{test1} is a parameter to be estimated measuring the correlation between the latent variable *skill* and its indicator, the test score. $\tau_1^{test1}, \ldots, \tau_7^{test1}$ are thresholds to be estimated, except for the extreme ones (i.e. 1 and 7) that have their values fixed to $-\infty$ and $+\infty$, respectively for identification purposes.

MIMIC models can be estimated simultaneously using maximum likelihood methods.

1.3.3 Hybrid choice models

Hybrid choice models (HC, Ortúzar and Willumsen, 2011, section 8.4.3; Bolduc & Alvarez-Daziano 2010) are a mixture of MIMIC and discrete choice models. Basically, they are choice models where at least one of the utility attributes is a latent variable, estimated using a MIMIC model. Figure 1-3 illustrates how the example of Figure 1-2 can be inserted in a HC model.

The MIMIC and choice model components of an HC model can be estimated using maximum likelihood techniques.

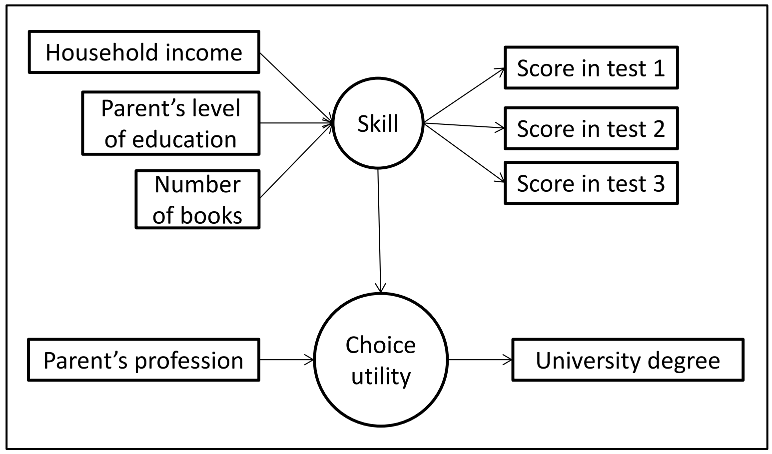


Figure 1-3 - Example of hybrid choice model

1.4 Choice models in the food and beverages literature

DCM started to become common on the food and beverages preferences literature during the last decade. Scarpa et al. (2005) measured consumers' willingness to pay (WTP) for place of origin on table grapes, oil and oranges; Angulo & Gil (2007) estimated WTP for certified beef; Gracia & de Magistris (2008) studied demand for organic food and how it is affected by knowledge about it; Jaeger & Rose (2008) measured the impact of consuming context on the choice of packed fruit; Barreiro-Hurlé et al. (2010) studied the impact of labelling foods with nutrition information on consumers' choice of healthy alternatives; Grisolía et al. (2012) studied how WTP for sea urchin in Spain (where the product was not usually consumed) changed after tasting a well-prepared dish of it; Adamowicz & Swait (2013) studied how habit influences grocery shopping; and O'Neill et al. (2014) explained preference heterogeneity for bought and house-cooked food based on consumers' attitudes.

Despite its relatively recent popularity, DCM had been applied long before that by Guadagni (1983), who studied the choice of coffee brands using scanner data. Furthermore, several authors (Moskowitz & Schiler 2006; Krystallis et al. 2010; Ortúzar 2010) support the use of these models in the food and beverage preference studies.

In the particular case of wine, there are already several applications. Angulo et al. (2000) estimated a hedonic price function using price ranges (and therefor a logit model) and data from the Spanish market, finding region of origin and year of harvest to be the main determinants of price. Lockshin et al. (2006) studied the Australian market, discovering that price influences choice probability differently depending on the wine's initial price: if the wine's initial price is low, an increase of price increases the probability of choosing it; but for more expensive wines an increase of price reduces the probability of choice (i.e. an inverse U-shape relation between price and choice utility). They also found the intensity of this behaviour to be influenced by consumers' characteristics. Mtimet & Albisu (2006) measured WTP for designation of origin in the Spanish market, finding the same inverse U-shape for utility as a function of price. Barreiro-Hurlé et al. (2008) estimated WTP for a hypothetical resveratrol-enriched (functional) wine, finding it to be positive and significant. Mueller et al. 2010b studied the impact of back-label information on wine consumers' choice, using latent classes to account for preference heterogeneity.

They did not find correlation between consumers' preferences and their socio-demographic characteristics. Jarvis et al. (2010) studied wine consumers' preferences for label information, comparing the performance of images, content description and metaphors, finding that images and content description performed better. They also used latent classes to capture preference heterogeneity, finding three classes, but could not correlate them with socio-demographic or attitudinal data. Palma *et al.* (2013) measured preferences for extrinsic attributes among Chilean wine consumers, finding a small effect of price, probably due to the double effect of price as a cue for quality and as a strain to the budget constraint. Costanigro et al. (2014) studied the WTP for low-sulphite wine using best-worst scaling and a logit model. They concluded that the sulphite's effect –as well as the effect of being an organic product– were marginal. Williamson et al. (2016) studied the influence on choice of advertising country of origin as a quality cue, both in the short and medium term (10 days after exposure). The advertising significantly increased the chance of selecting a wine from the advertised origin in the medium term.

2 UNDERSTANDING WINE CONSUMERS' PREFERENCES: A WEB-BASED STATED CHOICE SURVEY STUDY

2.1 Introduction

Wine purchase, as any other food and beverage consumption, can be conceptualized as a two-stage process (Grunert, 2005). First, consumers decide whether or not to buy a product, without actually tasting it. This decision is based on an expectation of the product's quality, which is constructed by consumers from available cues, as they cannot taste or smell the product at this point. Second, the consumer actually tastes the product and fully appreciates it. This process induces a dichotomous classification of a product's attributes: those that can be appreciated before buying are called extrinsic (e.g. price, packaging, advertising, etc.), and those that can only be appreciated after purchase are called intrinsic (e.g. colour, taste and aroma).

This research focuses on the first stage of the purchase process: the role of extrinsic attributes in the purchase of wine, incorporating preference heterogeneity. To make the study more manageable, we focus on a specific wine-drinking context, that of an informal dinner with friends. This allows us to apply a discrete choice modelling approach, simplifying both data requirement and modelling complexity as in Lockshin et al. (2006), Mtimet & Albisu (2006), Mueller et al. (2010) and Jarvis et al. (2010).

Discrete choice models (DCM) have become common in the food and beverages preferences literature lately (see Ortúzar 2010 for a review, and Adamowicz & Swait 2013 and O'Neill et al. 2014 for more recent examples). There are also several applications on wine (Lockshin et al. 2006, Mtimet & Albisu 2006, Barreiro-Hurlé et al. 2008, Mueller et al. 2010, Jarvis et al. 2010).

To account for heterogeneity in consumers' preferences, we use a hybrid modelling approach that mixes demographic and consuming-behaviour data with the respondents' stated choices. We estimate structural equation models to uncover latent variables that feed the DCM, giving rise to Hybrid Choice Models (HCM). These latent variables aim to measure consumers' involvement with wine -a construct often studied in wine literature (Brunner & Siegrist 2011)- using a novel short questionnaire. To understand how much considering heterogeneity can improve the understanding of wine purchasing, the HCM will be compared to a simpler DCM that only considers heterogeneity regarding the price of wine.

To accomplish our goals, we designed a stated choice survey where respondents faced a hypothetical choice between different wines for an informal dinner with friends. The survey was designed to provide also the information required for the estimation of the HCM. The survey was implemented through the web and was answered by members of a wine club.

To the authors' knowledge, this is the first experiment of this sort in Latin America and particularly in Chile, a relevant New World wine producing country. In the last few years, wine consumers in developing nations have become increasingly involved and more knowledgeable about wine; this has shifted industry focus on local markets from mass production to premium quality. Yet, quality is often understood only from the expert's perspective. In this research, we attempt to understand expected quality from the consumer’s standpoint, as derived from extrinsic attributes.

2.2 Materials and Methods

Here the sample, the models and the experiment are described.

2.2.1 Description of the sample used

The experiment was performed on two stages. On the first stage, 842 participants answered a web survey concerning their socio-demographics, wine consuming habits and attitudes. Afterwards, a subset of 274 individuals also completed a Stated Choice (SC) survey. All participants were clients of a Chilean wine specialty store. Table 2-1 summarizes the main characteristics of each sample. Despite the difference in size, both samples are very similar. The only significant differences are the purchase frequency, the likelihood of keeping a wine stock at home and the level of education. Both samples represent the richer end of the wine consumers’ spectrum, as 80% of respondents belong to the richest 20% of Chilean households (Ministerio de Desarrollo Social 2012).

Table 2-1 - Main characteristics of first and second stage samples

Category	Item	1st	2nd
Sample size	Number of individuals	842	274
Consumption habits	Weekly number of consuming occasions *	2.70	2.64
	Drink wine at lunch on weekdays (%)*	12.00	14.00
Purchasing habits	Number of purchases during a month	3.45	3.14
	Number of bottles per purchase*	7.78	7.95
	Buy bottles of more than 50 US$ (%)*	22.00	20.00
	Keeps a stock of wine at home (%)	88.00	92.00
Use of distribution channels (Likert scale from 0/never to 3/always)	Supermarket*	2.31	2.31
	Specialty store*	1.85	1.85
	Internet*	1.36	1.39
Attitudes Level of agreement with each phrase on a 1 to 7 Likert scale (this scale is used by the Chilean educational system, therefore respondents are totally familiar with it)	I know a lot about wine*	4.68	4.84
	I like trying new wines*	6.33	6.44
	There are expensive wines I don't like*	5.27	5.49
	Wine is a family tradition for me*	5.14	5.08
	Choosing wine at the supermarket can be difficult*	3.87	4.03
	Wine is for weekends*	3.10	2.92
	Wine is a social drink*	5.28	5.34
	When I want to make sure that I am buying a good wine, I choose an expensive one*	3.67	3.26
Demographics	Female (%)*	24.00	28.00
	Age*	41.80	42.94
	Number of people in household*	3.19	3.18
	Number of adults in household*	2.49	2.48
	Highest level of formal education (3=university)	2.90	2.99
	Monthly income (1000 US$)*	5.47	5.00
** Answers from both samples are statistically equivalent. The Kolgomorov-Smirnov two-sided test and the Chi-square test, both at 5% significance were used to compare answers.*			

2.2.2 Discrete choice models

In DCM, alternatives (i.e. bottles of wine) are defined as a set of attributes. Consumers, on the other hand, are assumed to have preferences for these attributes, and these preferences can vary across individuals. The most popular discrete choice model belongs to the logit family (Ortúzar & Willumsen 2011; Train 2009); in particular, the mixed logit (ML) model has concentrated most attention and research in the last years, thanks to its flexibility (McFadden & Train 2000). The ML model assumes that each alternative *i* provides a particular level of utility U_{int} for consumer *n* in choice situation *t*, when selected. This utility is assumed to depend linearly on the alternative's attributes x_{kit} (k enumerating attributes), and the consumer's preferences β_{kn}, through a linear function. As the modeller does not possess as much information as the consumer, an additional random error term ε_{itn}, representing all those attributes perceived only by the consumer, is added to the consumer's utility. Therefore, the utility of each alternative can be expressed as following.

$$U_{int} = \sum_k \beta_{kn} x_{kit} + \varepsilon_{itn} \tag{2.1}$$

In the ML model, ε_{itn} is assumed to distribute IID Extreme Value Type-I. As preferences are heterogeneous within a sample of consumers, the ML model allows β_{kn} to have a probability distribution $f(\beta_{kn}|\theta_k)$ on the population, that depends on a set of parameters θ_k. Then, the probability of choosing alternative *j*, given that the consumer chooses the alternative with higher utility, is given by (Train 2009):

$$P_{jnt} = \int \frac{e^{\sum_k \beta_{kn} x_{kjt}}}{\sum_i e^{\sum_k \beta_{kn} x_{kit}}} f(\beta_{kn}|\theta_k) d\beta_{kn} \tag{2.2}$$

Another way of considering heterogeneity in consumers' preferences is through systematic taste variations (Ortúzar and Willumsen 2011, page 279), by interacting product's attributes with consumer characteristics in the utility function, in the following way:

$$U_{int} = \sum_k \beta_{kn} x_{kit} + \sum_l \sum_k \gamma_{lk} x_{kit} z_{ln} + \varepsilon_{itn} \tag{2.3}$$

where z_{ln} represents characteristic *l* of consumer *n* and γ_{lk} represents the weight of the interaction in the utility. Even though the traditional way of including interactions is using socio-demographic characteristics such as age or sex in the *z* variables, it is also possible to use latent variables (i.e. non observable characteristics) such as consumers' attitudes. As attitudes are indirectly measured, they cannot be considered as deterministic exogenous variables. Therefore, if z_{ln} now represents the expected value of latent variable *l* for individual *n*, an additional random error term ω_l must be considered. In such a case, the utility function could be expressed as follows:

$$U_{int} = \sum_k \beta_{kn} x_{kit} + \sum_l \sum_k \gamma_{lk} x_{kit} (z_{ln} + \omega_l) + \varepsilon_{itn} \tag{2.4}$$

If all β_{kn} and γ_{lk} are considered common across the sample (i.e. not random parameters) and ω_l is considered to follow a random distribution $g(\omega_l|\vartheta_l)$ with parameters ϑ_l, the probability of choosing alternative *j* in the choice situation *t* by consumer *n* is

$$P_{jnt} = \int \frac{e^{\sum_k \beta_k x_{kjt} + \sum_l \sum_k \gamma_{lk} x_{kjt} (z_{ln} + \omega_l)}}{\sum_i e^{\sum_k \beta_k x_{kit} + \sum_l \sum_k \gamma_{lk} x_{kit} (z_{ln} + \omega_l)}} g(\omega_l|\vartheta_l) d\omega_l \tag{2.5}$$

When $(z_{ln} + \omega_l)$ is estimated through a Structural Equations Model or a MIMIC model (Bollen, 1989), and these latent variables are used on a DCM, the composition of both models is called a Hybrid Choice Model (HDC; Yañez et al. 2010). HDC models can be estimated sequentially or simultaneously (Raveau et al. 2010). Sequential estimation means that the MIMIC model is estimated first, and then its output is used on a second stage as input for the DCM estimation. Simultaneous estimation, instead, makes use of Full Information Maximum Likelihood to estimate both models in a single process. Both estimation methods assure parameters' consistency, even though the second one is more efficient.

2.2.3 Experimental design of the stated choice survey

We set up an SC survey on an on-line survey platform, and sent it to the 837 consumers who had previously answered the first (descriptive) survey. In the SC survey each person faced six hypothetical choice scenarios with four alternatives each. All scenarios included a non-purchase alternative (Figure 2-1).

There is ample literature identifying the most relevant attributes of wine from the consumers' standpoint (Lockshin & Corsi 2012 present a review and Schnettler & Rivera 2003; Jiménez et al. 2006; Mora et al. 2010; and Cerda et al. 2010 studied the subject in Chile). Nevertheless, we complemented our literature research with our own qualitative study of local consumers (not reported in this document).

Figure 2-1 - Example of a choice exercise

Six attributes were selected to be included in the experiment: Label design, Grape variety, Alcoholic content, Price, Discount and Advice. To improve realism, Type of wine (i.e. if Varietal, Reserve or Grand Reserve) was added as an additional seventh attribute, though it directly correlated with Price through a deterministic rule (all wines with prices lower than 7 US$, were labelled as Varietal; wines from 7 to 20 US$ were labelled Reserve, and wines above 20 US$ were labelled Grand Reserve). Although seven is not considered a large number of attributes in typical DCM studies, it is bordering the limit in the Chilean case (Caussade et al. 2005). Table 2-2 presents the levels for all attributes. A maximum of four levels was allowed to keep the number of choice situations for each person from growing excessively, while maintaining level balance.

Table 2-2 - Attributes' levels in the SC experiment

	Label Design	Grape Variety	Alcohol content	Advice	Price	Discount
1	Delicate	Cabernet Sauvignon	8.5° G.L.	None	100%	0%
2	Contrast	Merlot	11.0° G.L.	Salesman	120%	10%
3	Natural	Carménère	12.5° G.L.	Friend	130%	20%
4	-	Syrah	14.5° G.L.	Critic	160%	-

As measuring particular brands' effects was not one of our objectives, we used a fixed fictional brand for all alternatives. Consumers were made aware that the brand was fictional. Although Alcoholic content is not considered as one of the most relevant attributes of wine (Cohen et al. 2008), we included it as it was of much current interest to Chilean wine makers. Context (consuming occasion) is also an important attribute (Martínez-Carrasco et al. 2006), so it was fixed for all exercises as “an informal dinner with friends”.

We pivoted prices to avoid consumers discarding alternatives because they were too expensive or cheap, which would violate the compensatory behavior assumption. Consumers were asked to declare the maximum amount of money they were willing to spend on a bottle of wine for an informal dinner with friends, before being presented with the choice exercise. This value was scaled using the percentages on the Price column of Table 2, and discounts were later applied over the scaled price.

The four red grape varieties included are those most common in Chile (ODEPA 2012). The alcohol content was made to vary enough to consider 8.5° G.L, a level that was inexistent in the Chilean market at the time of the study. The levels used for *Label design* were taken from Orth & Malkewitz (2008), where five classes of wine label designs are identified, but only three of them were considered to describe the Chilean market well enough. These are delicate (muted, sleek and delicate), contrast (stark, not harmonic) and natural (representative, archetypical). To measure the effect of the design classes, and not of a particular label design, three different labels were constructed for each level of the Label design attribute, and they were assigned randomly as needed. All labels were designed following Orth & Malkewitz (2008) parameters, by a professional designer. In Figure 1, from left to right, the first and second are of class contrast, while the third and fourth are natural and delicate, respectively.

A D-efficient balanced design (Rose & Bliemer 2009; Ortúzar & Willumsen 2011, section 3.4) was built using N-gene (http://choice-metrics.com/). A simple logit model structure was assumed with priors from a pilot study with 19 participants (whose answers were kept in the final data set). The design was divided in two blocks of six scenarios each, with every respondent randomly assigned to one block. To avoid order bias, both the exercise and the order of presentation of the scenarios were randomized.

A ranking of grape varieties was asked for at the beginning of the survey while the ranking of recommendations was known from the previous (first stage) survey. These rankings were exploded generating additional fictional choices where the only difference among alternatives were Grape variety or Advice. To see this better, consider the following respondent’s Grape variety ranking: (i) Cabernet Sauvignon, (ii) Carménère, (iii) Merlot and (iv) Syrah. In this case three fictional choices would be created. In the first, four alternatives were available, all of which had the same base level attributes except for Grape variety, each with a different level. The first fictional choice would be choosing Cabernet Sauvignon, as it was the respondent's first preference. The second choice would have only three alternatives available (omitting the Cabernet Sauvignon wine) and the respondent would choose Carménère. In the third choice, only Merlot and Syrah wines would be available and the respondent would choose Merlot. An analogous procedure was used to model the Advice (recommendation) ranking.

2.3 Results

Two models were estimated: (i) a ML model, assuming preferences are homogenous but price sensitivity is not (i.e. with a random price parameter) (ii) a HDC model, capturing consumers' heterogeneity through the inclusion of two latent variables. Both models included a pseudo panel effect correlating all observations from the same respondent: an error component was added to the utilities of each alternative for each individual (i.e. error components were common across all alternatives j of each individual, but different between individuals and alternatives with different positions in the choice set). These errors are normally distributed with mean zero and standard deviation σpanel to be estimated.

Estimation was based on 3382 observations from 272 respondents. Socio-demographic and attitudinal data was available from the first stage web survey. 2076 observations came from choice exercises, due to 127 consumers answering 6 exercises and 146 answering 9. The remaining 1306 observations came from ranking explosion of grape variety and advice, due to 19 respondents who only provided their advice ranking (2 choices each), one who only provided his/her grape variety ranking (3 choices) and 253 who provided both rankings.

2.3.1 ML model

The ML model estimated coefficients, t-tests and other relevant goodness-of-fit statistics are shown in Table 2-3 (those coefficients equal to 0 are either the base or statistically equivalent to it). The Price coefficient was specified as normally distributed.

Even though Price had the expected negative sign, it presented a low level of significance ($p = 0.14$) and a large standard deviation. Even more, the estimated parameter implies that approximately 46% of the sample could have a positive coefficient for Price, i.e. these people would prefer a more expensive wine *ceteris paribus*. However, if individual parameters were estimated (a possibility of ML models) we may find that the actual proportion is much smaller (see the discussion about this issue in Sillano and Ortúzar 2005).

2.3.2 HDC model

The wine attributes interacted with two latent variables associated with the individuals: *Sophistication* and *Social cohesion*. The model was estimated sequentially using the Lavaan package (Rosseel 2012) in R (R Core Team 2014), and Python Biogeme (Bierlaire, 2003). The latent variables were estimated using a MIMIC model, based on data from all respondents in the first survey (i.e. the one measuring consumers' characteristics). This allowed estimating the MIMIC model with a larger dataset (i.e. 842 instead of 274 individuals) than would have been possible if we had been restricted to those who answered the SC experiment. This is why we estimated the HDC model sequentially and not simultaneously.

Table 2-3 - Estimation results for the ML model

Attribute	**Level**	**Coefficient**	**t-test**
Grape Variety	Cabernet Sauvignon	0	
	Merlot	-0.468	-6.04
	Carménère	0	
	Syrah	-0.259	-3.48
Label design	Delicate	0	
	Contrast	0	
	Natural	0.159	3.00
Advice (recommendation)	None	0	
	Salesman	0	
	Critic	0.665	10.71
	Friend	0.517	8.66
Alcohol content	° G.L.	0.064	4.25
Price	Mean effect	-0.016	-1.09
	Standard deviation	0.178	6.20
Type of wine	Varietal	0	
	Reserve or Grand Reserve	1.040	2.77
Discount	10%	0.444	5.27
	20%	0.628	6.64
Others	Constant	1.340	5.93
	Panel effect std. deviation	0.366	6.64
Number of observations (number of individuals)		3382 (274)	
Number of estimated parameters		13	
Log likelihood		-4406.45	
Akaike information criterion (AIC)		8838.90	
Corrected Rho squared		0.138	

Figure 2-2 shows the structure of the MIMIC model. Structural equations were linear and the measurement equations were of Ordered Logit type (see Greene & Hensher 2010). Indicators were associated with the level of agreement with the phrases presented on Table 1, under the Attitude category, while demographic and consuming-behaviour data were used as explanatory variables in the structural equations. Social Cohesion and Sophistication achieved a Cronbach's alpha of 0.45 and 0.46, respectively, rather low for psychometric standards but nevertheless relevant given that each comprises only three and four indicators, respectively. All links among latent variables and indicators are significant. The model's RMSEA was 0.047 (less than 0.05 at 69% confidence level). Table 2-4 shows the structural equations parameters: the utility thresholds of the Ordered Logit models are not reported here.

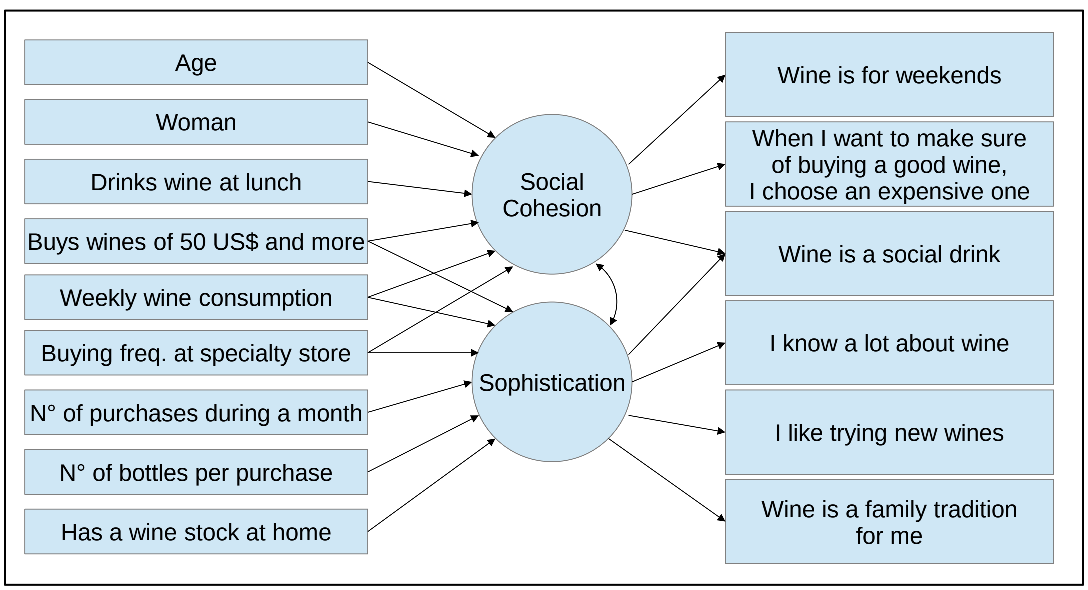


Figure 2-2 - Structure of the MIMIC model

Table 2-4 - Estimated parameters of the structural equation in MIMIC model

	Social Cohesion		Sophistication	
	Coeff.	**t-test**	**Coeff.**	**t-test**
Age	0.010	2.84		
Woman	-0.259	-2.61		
Drinks wine at lunch	-0.382	-2.37		
Buys wines of 12.5 US$ and more	-0.296	-2.38	0.338	2.66
Weekly wine consumption	-0.172	-6.35	0.113	3.99
Buying frequency at specialty store	-0.128	-2.29	0.169	2.94
N° of purchases during a month			0.182	4.60
N° of bottles per purchase			0.011	1.92
Has a wine stock at home			0.744	4.53
Link to "Wine is for weekends"	0.717	10.12		
Link to "Wine is a social drink"	0.379	7.92	0.392	9.65
Link to "When I want to make sure…"	0.337	8.51		
Link to "I know a lot about wine"			0.554	13.59
Link to "I like trying new wines"			0.425	9.83
Link to "Wine is a family tradition for me"			0.488	12.7
Cronbach's alpha	0.45		0.46	

Social cohesion attempts to measure how much a given respondent perceives wine-drinking as social activity, therefore associating wine-drinking with non-work-related social gatherings, and qualifying wine for its social appraisal more than its organoleptic characteristics. *Sophistication*, on the other hand, seeks to capture an interest for wine itself more than the consuming occasion, therefore focusing on self-reported knowledge, variety seeking and a perception of tradition in wine-drinking. Despite their differences, though, both attitudes look at wine as a social drink (i.e. one to share with friends or family). The estimated correlation between both construct is not significant (0.08, with a t-test of 1.19).

Table 2-5 - Estimated coefficients of the HDC model (choice part)

Attribute	Level	Coeff.	t-test
Grape variety	Cabernet Sauvignon	0	
	Merlot	-0.710	-4.70
	Merlot x Social Cohesion	-0.741	-3.66
	Carménère	-0.223	-2.42
	Carménère x Social Cohesion	-0.540	-3.64
	Syrah	-1.640	-6.02
	Syrah x Sophistication	0.755	6.69
Label design	Delicate	0	
	Contrast	0	
	Natural	0	
	Natural x Social Cohesion	-0.172	-3.20
Advice (Recommendation)	None	0	
	Salesman	0	
	Salesman x Sophistication	0.121	3.47
	Critic	0.223	2.86
	Critic x Sophistication	0.189	4.27
	Friend	0.175	2.25
	Friend x Sophistication	0.140	3.17
Alcohol content	° G.L.	-0.115	-2.61
	° G.L. x Social Cohesion	0.061	2.90
	° G.L. x Sophistication	0.117	5.60
Price	Price (US$)	0.060	1.49
	Price x Social Cohesion	-0.028	-2.06
	Price x Sophistication	-0.044	-2.59
Type of wine	Varietal	0	
	Reserve	0.671	1.93
	Grand Reserve	0.811	2.08
	Grand Reserve x Social Cohesion	-2.470	-4.19
Discount	10%	0.769	3.48
	10% x Sophistication	-0.209	-2.04
	20%	1.090	4.04
	20% x Sophistication	-0.309	-2.60
Others	Constant	1.090	4.09
	Panel effect std. Deviation	0.256	4.43
	Scale factor: Advice	5.070	3.34
	Scale factor: Grape variety	1.755	2.93
Number of observations (number of individuals)		3382 (274)	
Number of estimated parameters		29	
Log likelihood		-4175.06	
AIC		8408.1	
Corrected Rho square		0.187	

Once the MIMIC model was estimated, the two latent variables were calculated for each individual and included in the DCM, with a normally distributed error component to account for the estimation error in the previous stage. Simple interactions between latent variables and wine attributes were considered (i.e. multiplying the attribute and the latent variable). Coefficients, t-tests and goodness-of-fit of the HDC model (choice part) are shown in Table 2-5. The parameters for main effects and interactions are shown on different rows (values of 0 represent the base level or a level statistically equivalent to the base). Unlike the ML model, the HDC model does not consider any coefficient as random.

Table 2-6 - Most relevant percentiles of expected reduced form effects in the HDC model

		Parameter value in sample (percentile)			Coeff.<0
Attribute	**Level**	**25%**	**50%**	**75%**	
Grape Variety	Merlot	-0.603	-0.427	-0.211	89%
	Carménère	-0.145	-0.017	0.141	53%
	Syrah	-0.347	-0.095	0.144	61%
Label design	Natural	0.025	0.066	0.116	14%
Advice	Friend	0.415	0.461	0.506	0%
	Critic	0.547	0.610	0.670	0%
Alcohol	°G.L.	0.068	0.103	0.128	5%
Price	Price (US$)	-0.029	-0.019	-0.008	88%
Discount	10%	0.275	0.341	0.411	0%
	20%	0.360	0.458	0.561	1%
Type of wine	Reserve	0.671	0.671	0.671	0%
	Grand Reserve	1.169	1.754	2.476	1%

The reduced form effect of an attribute for each individual is an addition of normally distributed normal variables, i.e. $\beta_{kn} + \sum_l \gamma_{lk}(z_{ln} + \omega_l)$. *Therefore the reduced form effect is itself normally distributed (except in the case of Reserve, which is does not interact with any latent variable, and therefore has a fixed effect). However, for building this table only the expected value of each individual's effect was considered, i.e.* $\beta_{kn} + \sum_l \gamma_{lk} z_{ln}$.

Table 2-6 presents the value of each parameter for relevant percentiles of the sample, as well as the percentage of the sample with a negative parameter. Variability produces some potentially counterintuitive results. For example, 1% of respondents dislike high discounts (20%), and a similar result is obtained for the Grand Reserve type of wine. Notwithstanding, as only few consumers are in this situation, it could even be assumed that for those respondents these levels are not relevant. More importantly, 12% of consumers seem to like higher prices, however, this percentage is significantly smaller than that implied by the ML model (46%).

2.4 Discussion

While the ML model assumes that preferences are homogenous among participants (except for their sensitivity to price), the HDC model considers systematic preference variations through two latent variables: *Social cohesion* and *Sophistication*.

In the food and beverages literature, O'Neill et al. (2014) studied consumers' preferences for home-cooked meal options using latent variables to explain preference heterogeneity. Using a more complex approach, they estimate attitudes toward each level of their attributes (i.e. they estimate one latent variable for each level of each attribute). Even though their approach can be more informative than ours, it is also more difficult to apply on a larger scale, and harder to interpret for continuous variables. Instead, our approach is more straightforward, and successfully explains the heterogeneity in preferences. Other approaches to unravelling preference heterogeneity include the use of latent classes (Mueller et al. 2010) and different choice strategies (Adamowicz & Swait 2013). While these are promising methodologies, they are also somewhat more difficult to implement than latent variables.

In our case study, the HDC model provides a better fit to the data (a lower AIC index; 8408 against 8839). Perhaps more relevant, the HDC model captures heterogeneity in a more detailed way than the ML model, explaining it through latent variables, in turn related to observable socio-demographics. While the ML model shows that Cabernet Sauvignon and Carménère grape varieties were equally preferred by consumers, the HDC model revealed that actually only half of the sample (47%) liked Carménère. Both models agreed on the natural design being the most preferred one, as well as on salesman's advice being irrelevant and critic's advice being preferred over friend's, as well as the decreasing effect of discounts (20% discount is not twice as good as 10%). But the HDC model was able to differentiate between the effects of Reserve and Grand Reserve. Finally, consumer's rejection for low alcohol levels was captured by both models: consumers associate low alcohol content with low quality wine not aged on barrels.

The size reduction of the (estimated) part of the population with positive price coefficient is another important benefit of the HDC model. While 46% of the population has a positive parameter in the ML model, only 12% does in the HDC model. This is likely due to the higher flexibility of the HDC model, allowing it to reproduce the real distribution of the price coefficient more precisely.

In this study, choice models including respondents' latent traits (HDC) outperform flexible and powerful discrete choice models (such as a ML model with random coefficients), thanks to their ability to explain preference heterogeneity. Despite the more complex experimental set-up and analysis, HDC model results are significantly more informative and easier to extrapolate to the population level. Also, our application shows that introducing consumers' latent characteristics in the utility function through interactions is a simple and effective way to explain heterogeneity. However, much of the HDC advantages depend on using an effective questionnaire to measure consumers' latent characteristics. The best practice would be to use standard and validated questionnaires, but these are usually too long for a SC survey (Brunner & Siegrist 2011's has 97 items). The development of shorter instruments could be most valuable.

3 MODELLING CHOICE WHEN PRICE IS A CUE FOR QUALITY: A CASE STUDY WITH CHINESE WINE CONSUMERS

3.1 Introduction

Price is a key attribute in choice experiments. It is not only relevant for consumers and producers, but from a modelling perspective it is also used to calculate willingness to pay (WTP) estimates and the price elasticity of demand. However, the effect of price on consumers can be twofold. The quality of certain products, such as new foods and beverages, is uncertain before purchase because it cannot be fully evaluated until after consumption (Nelson, 1970; Grisolía et al., 2012). Other products, such as jewellery and some medicines have uncertain qualities even after purchase, as consumer do not have the means or knowledge to determine them. In these cases, consumers resort to extrinsic cues to determine product quality (i.e. they construct an *expected quality*). Any attribute that can be perceived before purchase, such as packaging, publicity, health claims, store advertising, etc., can constitute an extrinsic cue for quality. Among these, price may become a highly relevant cue for quality (Leavitt, 1954), as consumers tend to assume that higher prices are associated with higher quality in the case of many products. In these cases, price has a double effect: a positive one due to its role as a cue for quality, and a negative effect due to the consumers' budget constraints.

In discrete choice models, the double effect of price can generate an endogeneity problem, causing coefficient estimates to be biased. This happens because as modellers, we do not observe consumers' expected quality for the product, and as this variable correlates with price, omitting it from the utility function makes price endogenous.

Even though the literature offers several approaches to deal with endogeneity in discrete choice models (Guevara 2015), the latent variable approach is particularly suitable for cases where quality is uncertain to the consumer at the time of purchase. It also provides a reliable framework both from a methodological and a behavioural perspective, as it employs a tested econometric approach (hybrid choice models, Walker & Ben-Akiva 2002) and a well-developed behavioural theory (signalling mechanisms, Milgrom & Roberts 1986).

The approach consists in modelling expected quality as a latent variable, explained by the product's observable attributes (including price), while the actual purchase choice is explained by the trade-off between price and expected quality. This easily fits the frame of a stated preference (SP) experiment, where besides recording participants' choices, only an additional indicator of quality is required. Under the appropriate structure, the modeller can correct for endogeneity and measure both the positive and negative effects of price, while keeping the analysis in line with behavioural theory and not overwhelming respondents with excessive additional tasks.

In this paper, we use wine as a case study to test the latent variable approach to correct for endogeneity, in accordance to behavioural theory. To this end, we use a computer-based stated choice experiment which was responded by a particular sample of Chinese wine consumers, experts and students. We find that the method provides promising results, and propose further topics for future research.

The rest of the paper is structured as follows. Section 2 presents a brief literature review about the double effect of price and endogeneity in discrete choice models and their treatment in the foods and beverages literature. Section 3 provides details of the survey, the sample of participants and the models used. Results are presented on section 4 and discussed in section 5.

3.2 Literature review

3.2.1 Double effect of price

In traditional economic theory, price is expected to have a negative effect on the purchase probability due to consumers' budget constraints; however, under some circumstances a positive effect may also exist. Scitovsky (1945) proposes that higher prices can be attractive if consumers assume price to be a cue for quality (i.e. they assume that price and quality are positively correlated), a rational assumption in perfect markets. Leavitt (1954) did one of the first experimental measurements of this phenomenon, discovering a tendency to choose the most expensive product when there were no other cues for quality, especially on product categories with heterogeneous levels of quality (i.e. vertical differentiation).

Later studies confirmed the association between price and quality, and therefore the positive effect of higher prices on choice probability. Rao & Monroe (1989) showed that the price-quality association grew stronger as the price difference between alternatives increased, through a meta-analysis. Caves & Greene (1996) found a positive correlation between price and expert's quality ratings in 200 products, while controlling for other variables. They also found that the magnitude of the price-quality correlation depended on the product category and its vertical differentiation. Dodds *et al.* (1991) proposed and estimated a model where price positively influences perceived quality, and negatively influences willingness to buy, while controlling for brand and store information in the case of calculators and stereo headset players.

Another possible explanation for the positive effect of price on purchase probability is what Lichtenstein *et al.* (1993) call *prestige sensitivity*, i.e., a "favourable perception of the price cue based on feelings of prominence and status that higher prices signal to other people about the purchaser". This concept has been employed mainly in the area of fashion, and found to be strongly related with brand perception (Deeter-Schmelz *et al.* 2000), as *other people* see brands, not prices. This phenomenon is also known as *Veblen Effect* (Veblen 1899/1994), and is directly related with the status provided by the consumption, and only indirectly related with price. Bagwell & Bernheim (1996) claim that "… in a theory of conspicuous consumption that is faithful to Veblen's analysis, utility should be defined over consumption and status, rather than over consumption and prices". Therefore, this effect could be controlled for, to a reasonable degree, by including brand in the analysis.

The positive effect of price on perceived quality has also been studied in the case of wines. Plassman *et al.* (2007) showed that higher prices can positively influence markers of pleasure in the brain activity, even though the wine itself remains unchanged. Aqueveque (2006; 2008) found a negative effect of price on perceived risk and a positive effect on perceived quality, though this effect tended to disappear when experts' ratings were present and the consumption occasion did not involve other people. Lewis & Zalan (2014) showed that higher prices increased both reported enjoyment and willingness to pay among wine consumers.

3.2.2 Endogeneity in discrete choice models and some ways to deal with it

From an econometric perspective, the double effect of price generates an endogeneity problem. In this sub-section we present a simple framework to understand how endogeneity is caused by the price – quality association, and review some alternatives to deal with. The different approaches to deal with endogeneity are discussed and evaluated based on their applicability to the problem at hand, that is, a stated choice experiment where the main source of endogeneity is the price – quality association.

Endogeneity occurs when an explanatory variable is correlated with the error term of the model. This can be due to many reasons: omission of an explanatory variable correlated with an included variable, measurement errors in explanatory variables, simultaneous determination of both the dependent and one or more of the explanatory variables, self-selection bias, among others (Guevara 2015). Endogeneity is a serious problem as it renders the estimated parameters inconsistent (see Wooldridge 2002, section 15.7.2 for a proof on binary choice models).

One important source of endogeneity in the case of price's double effect is the omission of perceived quality as an explanatory variable. The omission of other unobservable attributes correlated with price can also play a role in the endogeneity problem (Guevara & Ben-Akiva 2012); however, if these attributes are relevant, they should also be correlated with perceived quality. Simultaneous determination is likely not a severe problem at the microscopic scale, because price is exogenous for each individual, as s/he does not influence price.

More formally, consider the following true model for the utility U of individual n, for alternative j on choice scenario t.

$$U_{njt} = X_{njt}\beta_X + Y_{njt}\beta_Y + \varepsilon_{njt} \tag{3.1}$$

where X_{njt} and Y_{njt} are attributes of alternative j, ε_{njt} is an independent identically distributed error among alternatives, scenarios and individuals, and β_X and β_Y are parameters to be estimated. Now suppose the modeller does not observe Y_{njt}, therefore she estimates the following model.

$$U_{njt} = X_{njt}\beta_X + \eta_{njt} \tag{3.2}$$

where $\eta_{njt} = Y_{njt}\beta_Y + \varepsilon_{njt}$. If X and Y are correlated, then so are η_{njt} and X, introducing endogeneity in the model and therefore rendering the estimated $\hat{\beta}_X$ inconsistent. In our particular case, if we consider X to be a vector of attributes including price, and Y to be perceived quality, then the price-quality association would induce correlation between X and Y, generating an endogeneity problem. Then we would say that the explanatory variable price is endogenous.

For discrete choice models, the most popular five ways to correct for endogeneity are the BLP method proposed by Berry *et al.* (1995), the use of proxies, the control function approach (CFA), the multiple indicators solution (MIS) and the use of latent variables (Guevara 2015).

The BLP method requires market level data in the form of market shares for several different markets. This data is used to capture the endogeneity in constants for each market. This data requirement makes the method unsuitable for models estimated only with consumer-level information, such as our case study.

The Proxy approach consists in including proxies of the unobserved variable in the utility function. A proxy must satisfy two requirements: (i) it must be independent of the choice model's error term and (ii) the difference between the proxy and the unobserved variable should be independent of all other explanatory variables. Both requirements can be fulfilled if the proxy is exogenous to the choice, it is measured with no error, and it is the cause of the unobserved variable (i.e. it is both exogenous to the unobserved variable and it correlates with it). Therefore, the main difficulty of this method is to find an appropriate proxy. For example, a proper proxy for the comfort experienced by a new passenger on a train is the density of passengers in the train before s/he boards.

A proxy for perceived quality should be able to explain it while not being correlated with price. An objective measurement of quality should be a good proxy for perceived quality only if the objective quality does not correlate with price; however, it is not clear that such a measurement exists. In the case of wine, expert ratings may not be appropriate either as their ability to measure objective quality has been seriously questioned (Lawless 1984, Hodgson 2009), as well as their relationship with consumer's quality perception (Lattey *et al.* 2009, Gokcekus & Nottebaum 2011, D'Alessandro & Pecotich 2013, Hopfer & Heymann 2014). And even though consumers do use experts' ratings as a proxy for quality when available in hypothetical situations (Aqueveque 2006, Mastrobuoni *et al.* 2014), several studies have indicated that consumers are not really aware of them in real conditions (Channey 2000, Johnson & Bruwer 2004, Atkin & Thach 2012). Furthermore, it is likely that if an objective measurement of quality exists, it would correlate with price due to production costs.

As our experiment used fictional wines, no real experts' quality ratings were available, neither did we include fictional ratings as an extra attribute because Chinese consumers do not seem to consider experts' ratings (at least in the form of prizes or written recommendation) among the most relevant cues for quality (Goodman 2009).

In the particular case of wine, the weather during growth and harvest could be used as a proxy for quality, as wine quality is expected to depend largely on them. But the weather only influences the sensory (or intrinsic) quality of wine, and therefore it would not reflect the expected quality before purchase, when the consumer has not tasted the wine yet. Also, the weather is not available for fictional wines in a SP context.

Another method to correct for endogeneity is the Control Function (CF) approach (Villas-Boas & Winer 1999, Petrin & Train 2010), which is analogous to the Instrumental Variables approach on linear models (Wooldridge 2002, chapter 5). The CF approach requires the modeller to identify instrumental variables for the endogenous explanatory variable (in our case: price). The instrumental variables must fulfil two requirements: (i) correlate with the endogenous explanatory variable and (ii) be independent of the error terms. The estimation procedure has two stages: first, the endogenous variables are regressed on the instrumental and other exogenous explanatory variables, and then the residuals of this regression are included in the choice utility along with the endogenous and exogenous explanatory variables. This way the new extended model is consistently estimated. Estimation can also be performed in a single step using Full Information Maximum Likelihood (Villas-Boas & Winer 1999, Train 2009 section 13.5, Guevara 2015). The main difficulty with this procedure is finding adequate instrumental variables.

Production costs are useful instruments (Villas-Boas & Winer 1999), but they are hardly available for real products, and do not exist in the case of fictional ones. The price of similar alternatives can also be used (Guevara & Ben-Akiva 2006), but once again, they do not exist in the context of hypothetical choices. And even though it is possible to design a stated choice experiment where the weather, the price of similar alternatives or other instruments are fictionally developed, its implementation would be convoluted and probably unrealistic. In summary, CF is hardly applicable on stated choice datasets, such as ours.

A Multiple Indicator Solution (MIS) is yet another way to correct for endogeneity in discrete choice models (Guevara & Polanco 2016). This approach is a mixture between the use of a proxy and a control function. The method requires two indicators of the omitted variable. Indicators are only required to correlate with the unobserved variable, and not to be exogenous to the choice. The idea is to include the first indicator in the utility function, using it as a proxy for the omitted variable and therefore transferring the endogeneity from the original endogenous variable to the indicator. Then, the second indicator serves as an instrument to correct the endogeneity of the first indicator, using the CF approach. The second indicator is a valid instrument for the first indicator, as both are correlated because both are explained by the omitted variable; it is also uncorrelated with both the original error tem of the utility function and the first indicator's error term, under the assumption that both indicators are redundant in the structural equation of utility if the omitted variable is included (Guevara 2015).

In the case of price-quality associations, one would only require two indicators of quality to apply the MIS approach. Unlike proxies, indicators can be noisy and they do not need to have a causal relation with the omitted variable, but quite the contrary, it is the omitted variable that causes and explains both indicators. Therefore, simple quality ratings from the consumers or experts could be used. The former would be preferable though, as they measure expected quality directly. When applied to solve the endogeneity problem due to the price-quality association, the MIS approach could effectively provide consistent estimates for both the positive and negative effects of price, through the first indicator and price coefficients, respectively. However, two reliable and independent (given the omitted variable) indicators must be available. As we only had a single quality indicator in our dataset we could not apply the MIS approach.

Finally, the Latent Variable approach to endogeneity correction consists in explicitly modelling the omitted variable as a latent variable. To do this, two pieces of information are required: (i) at least one indicator of the omitted (latent) variable, and (ii) one or more exogenous explanatory variables for the omitted variable. This method requires strong distributional assumptions, as the structural relation between the omitted variable, its explanatory variables, and the choice is explicitly (and parametrically) formulated. However, its data requirements (at least in the context of this study) are easier to fulfil, as it does not require hard-to-find proxies or instrumental variables, and it only requires one quality indicator.

The Latent Variable approach is the only method that provides a consistent behavioural model in the context of price-quality associations. Therefore, it allows to clearly separate the positive and negative effects of price, and to separately model the perception of quality, and the willingness to buy. This is particularly useful when consumers cannot perceive the quality of a product and therefore must infer it from observable attributes.

3.2.3 Endogeneity in the foods and beverages literature

In the foods and beverage choice literature, endogeneity has been considered mainly in the context of price's simultaneous determination due to supply and demand equilibration. Using a panel of scanner data at the household level and discrete choice models, Villas-Boas & Winer (1999) applied the CF approach to test and control for endogeneity in the yoghurt and ketchup market. They found evidence of endogeneity, which they explained on the simultaneous determination of price. Also using household data, but analysing it through a discrete-continuous model, Richard & Padilla (2009) analysed the impact of promotions in fast food consumption. They also found evidence of price endogeneity using a CF approach, which they again explained on the simultaneous determination of price. O'Neill *et al.* (2014) recognized that their analysis of food choices could be affected by endogeneity, but did not explicitly control for it.

In the wine choice literature, endogeneity has been explicitly controlled for mostly in the context of aggregate demand models. Cuellar & Huffman (2008) used aggregate data to estimate the price elasticity using linear models with grape prices as instrumental variables to correct for endogeneity. Stasi *et al.* (2011) used Italian market aggregate data and simultaneous equation modelling to measure the impact of geographical indicators, while correcting for endogeneity using several instrumental variables, such as lagged prices and seasonal dummies. Michis & Markidou (2013) used aggregate data from Cyprus and a system of simultaneous equations to identify the determinants of wine price, and took market concentration and competitors' prices as instrumental variables to correct for price endogeneity.

To the best of our knowledge, only two papers deal with the endogeneity problem when modelling wine demand at the individual level using stated choice experiments. In particular, although Appleby *et al.* (2012) do not mention endogeneity explicitly, their approach can be seen as using Wine Spectator's ratings as a proxy for quality, yielding reasonable results. However, as discussed in the previous sub-section, the use of experts' ratings as proxies for quality is highly questionable.

Mastrobuoni *et al.* (2014) used a two-stage process (somewhat similar to our approach) to separate the positive and negative effects of price in a SP experiment. However, they mixed the Proxy and Latent Variable approaches to correct for endogeneity. Their experiment appears to yield reasonable results, but the method is not applicable to situations without tasting, it resorts to experts' ratings as a proxy for quality and uses a sequential estimation process, which could lead to new endogeneity problems as the deterministic part of the first stage logit's utility is a noisy (and therefore endogenous) proxy for quality.

In this paper, we use the latent variable approach to correct for endogeneity. Our particular application is a stated wine choice experiment where consumers provided a single quality indicator per alternative, additionally to their choices. Due to the way our data was collected, we are not able to offer any comparison of the Latent variable approach with other methods. BLP requires market-level data, which does not exist in a SP experiment. The proxy method in a SP setting implies providing an expert ranking for consumers to use as a proxy for quality, but as consumers do not seek this information in real settings we did not include it the experiment. The CFA is not applicable as there are no available instruments in a SP setting, and we only have one quality indicator in our dataset, therefore the MIS approach cannot be applied either (as it requires at least two indicators).

3.3 Materials and methods

3.3.1 Survey design

In association with a private Chilean Vineyard, we designed a computer-based Stated Choice (SC) experiment (Rose & Bliemer 2009; Rose *et al.* 2008; Ortúzar & Willumsen 2011, section 3.4) that was applied to a sample of Chinese wine consumers, including experts, students and regular consumers. Respondents were presented with six choice scenarios (also called choice exercises) with three alternatives each (Caussade *et al.* 2005), plus a non-purchase alternative if they rather wished to opt out.

We considered four attributes in the SC experiment (Table 3-1): label design (6 levels), grape variety (3 levels), name and "story" of the brand (3 levels) and price (3 pivoted levels). In addition, in every choice scenario we also stated one out of two consuming occasions (formal and informal). Attributes were selected after a literature review (see, for example Lockshin & Corsi 2012), focus groups, previous experience with Chilean consumers (Palma *et al.* 2013), and advice from experts on the Chinese wine market. The "story" attribute, in particular, was proposed by these experts, and included both the name of the wine and a short statement describing its origin (the name and the statement were not shuffled, instead they were always paired in the same way). The objective was to provide a narrative for the product, for example, one story presented the wine as an old family tradition, while another presented it as the last innovation of a young entrepreneur.

Table 3-1 - Attributes and their levels (levels' order have been altered)

	Label	Grape variety	Story	Price	Consuming occasion
0	Label 0	Red Blend	Hacienda	Informal low	Informal: "an informal dinner with friends"
1	Label 1	Shiraz	Don Juan	Informal mean	
2	Label 2	Cabernet Sauvignon	Union	Informal high	
3	Label 3			Formal low	Formal: "a formal dinner"
4	Label 4			Formal mean	
5	Label 5			Formal high	

Before facing the SC scenarios, participants provided the minimum and maximum amounts of money they would be willing to pay for a bottle of wine on a formal and on an informal occasion. The phrasing of the question was: "Imagine that you need to buy a wine for the following occasions. How much would you be willing to spend? Please indicate a minimum and a maximum amount of money you would be prepared to pay for each occasion". Price levels of the SC experiment were pivoted based on these values at the individual level, i.e. each participant saw prices based on his/her own reported buying range for each occasion. This allowed us to make sure that participants did not see alternatives with prices outside their regular buying range, therefore avoiding them ruling out alternatives considered either too cheap or too expensive.

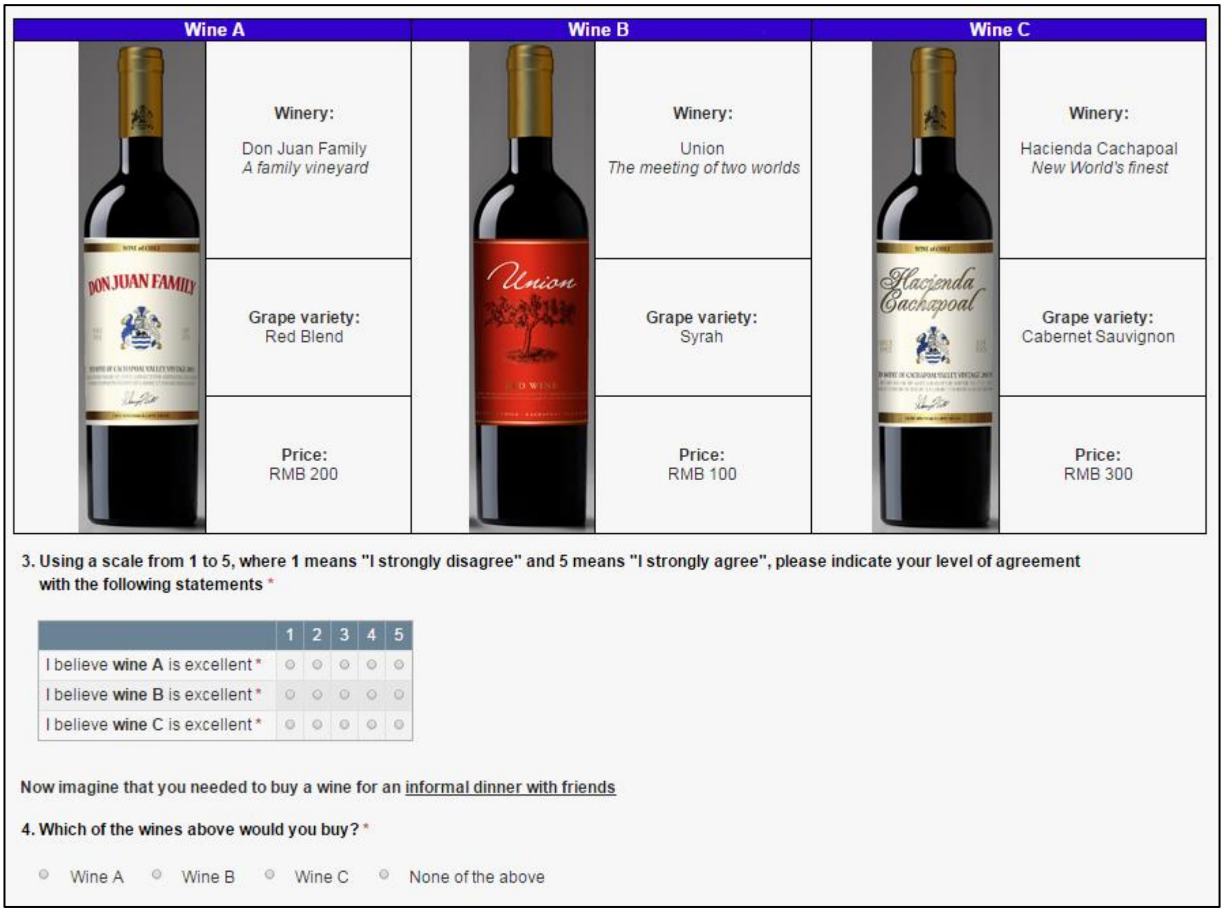
As participants provided different buying ranges for formal and informal occasions, six different price levels were calculated for each participant: informal low (the minimum price the participant would pay for a wine to drink at an informal occasion), informal high (the maximum price in the same case as above), informal mean (the midpoint between the previous two) and three more levels analogous to the previous ones, but for formal occasions. The occasion associated with each scenario determined which set of prices (formal or informal prices) were used.

Consuming occasion only varied between scenarios. Introducing more than one consuming occasion per scenario would have made the experiment unrealistic, as individuals seem to choose differently based on the consuming occasion (Dubow 1992; Quester & Smart 1998; Martínez-Carrasco *et al*. 2006, Jaeger & Rose 2008).

We generated a D-efficient balanced design assuming a simple MNL model using N-gene (http://choice-metrics.com/). We used null *priors* for the experts' design, who answered the experiment first, and then used the experts' results as *priors* for the design for the rest of participants. The experimental design had twelve choice scenarios divided into two blocks of six choice scenarios each, to which respondents were assigned randomly. The presentation orders of both scenarios and alternatives were randomized.

Before choosing the wine they would buy in each scenario, respondents had to provide their level of agreement with the phrase "I believe this wine is excellent" for each alternative presented, using a 5-point Likert scale. This information was used as an indicator of quality for each alternative. Then, respondents were told about the consuming occasion, and asked to make their choices (including the opt-out option). Figure 3-1 shows an example of a choice scenario.

Before facing the choice scenarios, participants also had to rate each of the considered grape varieties using a 5-points Likert scale. Based on these ratings, we built a grape variety ranking for each participant excluding ties; that is, when a participant gave the same rating for two or three grape varieties, we excluded them from the ranking.



Wine A	Wine B	Wine C
Winery: Don Juan Family *A family vineyard*	**Winery:** Union *The meeting of two worlds*	**Winery:** Hacienda Cachapoal *New World's finest*
Grape variety: Red Blend	**Grape variety:** Syrah	**Grape variety:** Cabernet Sauvignon
Price: RMB 200	**Price:** RMB 100	**Price:** RMB 300

3. Using a scale from 1 to 5, where 1 means "I strongly disagree" and 5 means "I strongly agree", please indicate your level of agreement with the following statements *

	1	2	3	4	5
I believe **wine A** is excellent *	○	○	○	○	○
I believe **wine B** is excellent *	○	○	○	○	○
I believe **wine C** is excellent *	○	○	○	○	○

Now imagine that you needed to buy a wine for an informal dinner with friends

4. Which of the wines above would you buy? *

○ Wine A ○ Wine B ○ Wine C ○ None of the above

Figure 3-1 - Example of choice scenario with quality indicator for each alternative (labels have been altered)

These rankings were exploded (Chapman & Staelin 1982, Ortúzar & Willumsen, 2011, section 8.7.2.3), creating “grape variety choices” in our dataset generating up to two new observations per respondent. As an example, let us consider a participant whose ranking was: (1st) Cabernet Sauvignon, (2nd) Shiraz and (3rd) Red Blend. In this case, the first "grape variety choice" would be between three wines with the same attributes, except grape variety: wine A would be a Cabernet Sauvignon, wine B would be a Shiraz and wine C would be a Red Blend, and the participant would choose wine A. The second "grape variety choice" would be between wines B and C only (wine A would not be available), and the participant would choose wine B. For participants whose rankings where shorter (due to ties), only one or none "grape variety choice" were generated. When modelling, we multiplied the utility of the "grape variety choices" by a scale factor, so we could control for differences in variance among the traditional choices and the "grape variety choices". However, this scale parameter turned out to be not significant in the ML model, so we removed it from its reported version.

3.3.2 Sample

A total of 180 participants answered the survey; however, after data cleaning only 168 responses were considered valid. The main reason to eliminate respondents was unreasonable price ranges, that were either too low (maximum was less than 1.6 USD) or too high (minimum was more than 10% of their monthly income).

The sample was divided into three groups: experts (21), regular consumers (81) and students (66). We introduced this classification, as it would help the private vineyard developing a more detailed strategy aimed at connoisseurs (experts), regular consumers and millennials (students). Most experts worked in the wine industry, mainly in marketing or trade departments, while others were wine critics. Regular consumers were mostly professionals and office clerks from different industries, including some scholars. All students were enrolled in some of the wine-related courses taught at the College of Horticulture at CAU.

We used a convenience sample; therefore, there is no guarantee that it represents the average Chinese wine consumer, nor any particular segment of the Chinese wine market. Experts and consumers received a small monetary incentive for their participation and performed the experiment in a laboratory, in a controlled environment. Students, on the other hand, were invited to participate in the experiment during classes, and answered the survey later using their own computers in an uncontrolled environment. Most students (97%) were under 30 years old; more details about the sample are shown in Table 3-2.

Given the age and profile of the students, their answers for the formal occasion were removed from the analysis, as their self-reported price ranges tended to be unreasonable. On average, students set a minimum price of 17% of their income and a maximum of 129% for formal occasions; instead, experts and consumers set an average price range for the same occasions between 4% and 11% of their income. Therefore, at the end 810 wine choices were collected (126 from experts, 486 from consumers and 198 from students).

Table 3-2 - Sample description

	Experts	Consumers	Students	Total
Respondents	21	81	66	168
Gender				
Female	11	43	50	104
Male	10	38	16	64
Age				
18 - 24	0	0	1	1
25 - 30	2	20	64	86
31 - 35	10	35	1	46
36 - 40	4	10	0	14
41 - 50	2	8	0	10
51 - 60	2	7	0	9
>60	1	1	0	2
Maximum level of education attained				
12th grade or less	1	2	0	3
Graduated high school	1	2	2	5
Some college, no degree	0	7	62	69
Associate degree	3	2	0	5
Bachelor's degree	3	61	2	66
Post-graduate degree	13	7	0	20
People in household				
Unknown	1	0	0	1
1	1	3	0	4
2	4	12	1	17
3	9	37	48	94
>3	6	29	17	52
Household monthly income (USD)				
<800	0	5	12	17
<1600	4	23	30	57
<2400	3	19	15	37
<3200	6	6	4	16
<4000	3	13	4	20
<4800	2	1	0	3
<5600	1	6	0	7
>5600	2	8	1	11
Average buying price range (USD)				
Min Informal	24	18	17	19
Max Informal	83	69	64	69
Min Formal	51	79	57	66
Max Formal	159	221	345	262

All participants rated the three grape varieties included in the experiment, giving rise to a personal ranking, which was exploded providing up to two additional choices per participant (as mentioned above, ties were excluded). Experts provided 27, consumers 111 and students 59 of these choices. Considering all choices (both wine and grape variety choices), 1007 observations were used for estimation.

3.3.3 Modelling

Two models were estimated with the available data: a traditional Mixed Logit (ML) model with random coefficients without considering an endogeneity correction (McFadden & Train 2000, Train 2009, chapter 6) and a Hybrid Choice (HC) model using random coefficients and the latent variable approach to correct for endogeneity (Ortúzar and Willumsen, 2011, section 8.4.3; Bolduc & Alvarez-Daziano 2010, Guevara 2015). Comparing both models allows determining how effective the latter is in dealing with endogeneity.

In the ML model, all attributes explain choice by entering the utility function directly (Figure 3-2).

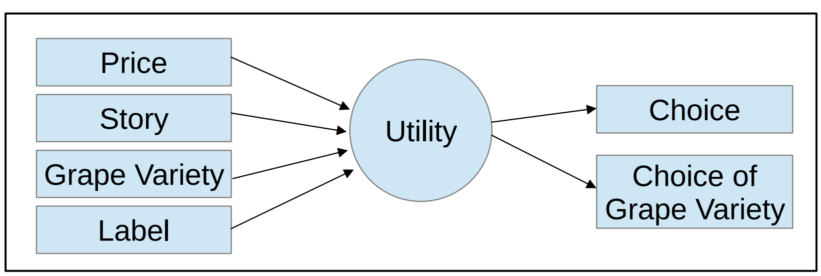


Figure 3-2 - ML model structure (for each alternative)

The deterministic utilities of the alternatives, their full utilities and the model's likelihood for one individual are shown in equations (3.3), (3.4) and (3.5), respectively.

$$V_{jtn} = X_{jtn}\beta_{Xn} + \left(\beta_{price} + \beta_{price}^{expert} expert_n + \beta_{price}^{student} student_n\right) price_{jtn} \quad (3.3)$$

$$U_{jnt} = V_{jnt} + \epsilon 1_{jnt} \quad (3.4)$$

$$L(\vec{\imath}_n) = \int \left(\prod_t \frac{e^{V_{itn}}}{\sum_j e^{V_{jtn}}} \right) \prod_g \frac{e^{V_{ign}}}{\sum_j e^{V_{jgn}}} \varphi\left(\beta_{Xn} \middle| \mu_{\beta_X}, \Sigma_{\beta_X}\right) d\beta_{Xn} \quad (3.5)$$

where V_{jtn} is the deterministic part of the choice utility for alternative *j* in scenario *t* for respondent *n*. X_{jtn} is a vector of alternative *j*'s attributes (excluding price); β_{Xn} is a vector of random parameters representing respondent *n*'s preferences for attributes other than price; $expert_n$ and $student_n$ are dummies which take the value 1 if respondent *n* is an expert or a student respectively, and 0 otherwise; $price_{jtn}$ is the alternative's price. U_{jnt} is the alternative's full random utility, and $\epsilon 1_{jnt}$ is an iid Extreme Value type 1 random error that gives the choice probability its logit form. $\vec{\imath}_n$ is the vector of choices made by respondent *n*; V_{itn} is the deterministic part of the utility of the chosen alternative *i* in choice scenario *t* by respondent *n*; V_{ign} is the deterministic part of the utility of the chosen alternative *i*, by respondent *n*, on the "grape variety choice" g; $\varphi\left(\beta_{Xn} \middle| \mu_{\beta_X}, \Sigma_{\beta_X}\right)$ is the multivariate normal density function of all random coefficients included in the β_{Xn} parameter vector, with vector μ_{β_X} as mean and the diagonal matrix Σ_{β_X} as variance. Finally, μ_{β_X},Σ_{β_X}, β_{price}, β_{price}^{expert} and $\beta_{price}^{student}$ are parameters to be estimated.

No additional error components were included to model the pseudo-panel effect. We did test a specification with error components, as proposed by Daly & Hess (2010), but the error components' standard deviations were non-significant, so we removed them from the final specification. However, as the randomness in β_{Xn} is between and not within participants (Revelt & Train 1998), correlation between the observations of each respondent is present, even though some confounding effects could occur (Daly & Hess 2010). Consuming occasion was not considered in the final specifications either, as we tested several ways to interact it with the different attributes, and none was significant.

Unlike the ML model, the HC model explains the choices made by consumers as a trade-off between an alternative's expected quality and its price. Each alternative's expected quality is modelled as a latent variable, which is explained by its attributes including price (Figure 3-3).

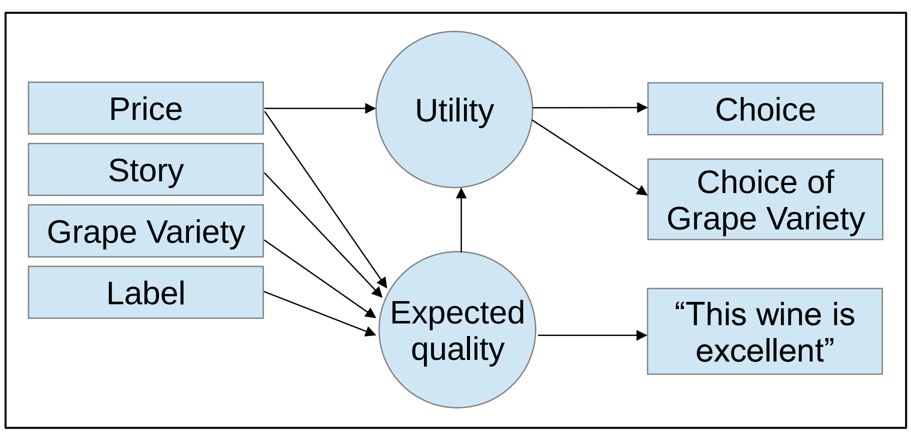


Figure 3-3 - HC model structure (for each alternative)

Price was included as an explanatory variable both in the structural equation of the expected quality and in the choice utility in the HC model. Its first coefficient was expected to capture the positive effect of price as a cue for quality, while the second intended to measure the negative effect of price owing to the consumers' budget restrictions. Therefore, the price coefficient was expected to be positive in the expected quality's structural equation and negative in the choice utility function.

In the HC specification, a Multinomial Logit (MNL) model was used to link the utility with choices, and an ordered logit model (Greene & Hensher 2010) to link expected quality and level of agreement with the phrase "This wine is excellent". The expected quality's structural equation (3.6), its measurement equation (3.7), the ordered logit probability function (3.8) and the deterministic part of the choice utility (3.9) are as follows:

$$EQ_{jtn} = X_{jtn}\alpha_{Xn} + \left(\alpha_{price} + \alpha_{price}^{expert} expert_n + \alpha_{price}^{student} student_n\right)price_{jtk} + \eta_{jn} + \omega_t \quad (3.6)$$

$$measurement_{jnt} = \lambda EQ_{jtn} + \epsilon 2_{jnt} \quad (3.7)$$

$$P\left(IQ_{jtn} = l\right) = \frac{1}{1 + e^{\lambda EQ_{jtn} - \delta_l}} - \frac{1}{1 + e^{\lambda EQ_{jtn} - \delta_{l-1}}} \quad (3.8)$$

$$\begin{gathered} V_{jtn} = \beta_{EQ} EQ_{jtn} + \left(\beta_{price} + \beta_{price}^{expert} expert_n + \beta_{price}^{student} student_n\right)price_{jtn} \\ W_{jnt} = \left(1 + \mu_{student} student_n\right)\left(1 + \mu_{gVarRnk}\, gVarRnk_{jtn}\right)V_{jtn} \end{gathered} \quad (3.9)$$

where EQ_{jtn} is participant n's expected quality of alternative j in scenario t; X_{jtn} is a vector of alternative's attributes (except for price); α_{Xn} is a vector of normally distributed random parameters representing participant n's preferences, with vectors

μ_{α_X} as mean and the diagonal matrix Σ_{α_X} as variance. η_{jn} is a normally distributed error component with mean 0 and standard deviation fixed to 1 (this is a requirement for identification in the structural equation model). These error components capture the expected quality's determinants that are not observed by the modeller (Bahamonde-Birke *et al.*, 2015) and correlate observations of the same respondent by being invariant across choice scenarios (Daly & Hess 2010). ω_t is an iid normal error component with mean zero and variance σ_ω^2 to be estimated, correlating the expected quality of all wines observed on the same choice situation. $measurement_{jnt}$ is the ordered logit's latent variable depending on expected quality and its iid Extreme Value type 1 error component $\epsilon 2_{jnt}$, that gives the measurement its ordered logit form. $P\left(IQ_{jtn} = l\right)$, is the ordered logit probability of quality indicator IQ_{jtn} (level of agreement with the phrase "This wine is excellent") being equal to l and V_{jtn} is the deterministic part of the choice utility. The dummies $expert_n$ and $student_n$ take the value 1 if respondent *n* is an expert or student, respectively, and 0 otherwise. W_{jnt} is the deterministic part of the utility scaled by factors $\mu_{student}$ and $\mu_{gVarRnk}$ when the observation belongs to a student or is a "grape variety choice". The dummy variable *gVarRnk* takes the value 1 if the observation is a "grape variety observation" and 0 otherwise. Scale factors analogous to these ones were tested in the ML model, but were not significant, therefore they were removed from the final model. Finally, μ_{α_X}, Σ_{α_X}, α_{price}, α_{price}^{expert}, $\alpha_{price}^{consumer}$,λ, δ_l, β_{EQ}, β_{price}, β_{price}^{expert}, $\beta_{price}^{student}$, $\mu_{student}$ and $\mu_{gVarRnk}$ are parameters to be estimated. Note that $\delta_0$and δ_5 were set to $-\infty$ and $+\infty$, respectively, for identification purposes. Finally, just as in the ML model, no error components or interactions with consuming occasion were included in the utility, as they both were not significant.

The likelihood function of the HC model is presented in equation (3.10).

$$L(\vec{\imath},\vec{l}) = \int\limits_{\overrightarrow{\eta_n},\omega_t,\alpha_{Xn}} \left[\prod_t \left(\prod_j P\left(IQ_{jnt} = l_{jtn}\right)\right) \frac{e^{W_{itn}}}{\sum_j e^{W_{jtn}}}\right] \prod_g \frac{e^{W_{ign}}}{\sum_j e^{W_{jgn}}} \varphi(\overrightarrow{\eta_n}|0,1)\varphi(\omega_t|0,\sigma_\omega^2)\varphi\left(\alpha_{Xn}|\mu_{\alpha_X},\Sigma_{\alpha_X}\right) d\overrightarrow{\eta_n} d\omega_t d\alpha_{Xn} \quad (3.10)$$

where $\vec{\imath}$ represent the vector of choices and $\vec{l}$ the vector of quality indicators; W_{itn} is the deterministic part of the utility of the chosen alternative *i* in choice scenario *t* by respondent *n*; W_{ign} is the deterministic part of the utility of the chosen alternative *i* by respondent *n* on the "grape variety choice" g; $\overrightarrow{\eta_n}$ is the vector containing all three η_{jn} associated with the expected quality of each of the three alternatives; $\varphi(\overrightarrow{\eta_n}|0,I)$ is the multivariate normal density function for the vector of error components associated with expected quality, with mean a vector of zeros and a 3x3 identity matrix for variance. $\varphi(\omega_t|0,\sigma_\omega^2)$ is the normal density function with mean 0 and variance σ_ω^2. Finally, $\varphi\left(\alpha_{Xn}\middle|\mu_{\alpha_X},\Sigma_{\alpha_X}\right)$ is the multivariate normal density function with the vector μ_{α_X} as mean, and the diagonal matrix Σ_{α_X} as variance.

Both models were estimated using the Python version of Biogeme (Bierlaire 2003). Monte Carlo techniques were used to estimate the integrals on the likelihood functions, as these do not have a closed analytical form. This consists in randomly drawing a large number of points from $\varphi\left(\beta_{Xn}\middle|\mu_{\beta_X},\Sigma_{\beta_X}\right)$ (in the case of the ML model) or $\varphi(\overrightarrow{\eta_n})$, $\varphi(\omega_t|0,\sigma_\omega^2)$ and $\varphi\left(\alpha_{Xn}\middle|\mu_{\alpha_X},\Sigma_{\alpha_X}\right)$ (in the case of the HC model) and then evaluating $\left(\prod_t \frac{e^{V_{itn}}}{\sum_j e^{V_{jtn}}}\right)\prod_g \frac{e^{V_{ign}}}{\sum_j e^{V_{jgn}}}$ (in the case of the ML) or $\Big[\prod_t \Big(\prod_j P\Big(IQ_{jnt} =$

$\left.\left. l_{jtn}\right)\right)\frac{e^{W_{itn}}}{\sum_j e^{W_{jtn}}}\Big]\prod_g \frac{e^{W_{ign}}}{\sum_j e^{W_{jgn}}}$ (in the case of the HC), for each of these points. The average of all these evaluations is a consistent estimator of the integral's value (Train 2009, chapters 9 and 10).

3.4 Results

Table 3-3 and 4-4 show the ML and HC models' estimated coefficients as well as their goodness of fit measures. All reported t-test are robust (i.e. they were calculated using the "sandwich estimator" clustering by respondent). Both models were estimated using 1000 Modified Latin Hypercube Sampling draws (Hess *et al*. 2006). Even though we tested interactions with consuming occasion in both models, none turned out to be significant, so we removed these from the final specifications reported in this document. The same holds true for the scale factors for "grape variety choices" and participant classes (experts and students) in the ML model. In the HC model, instead, the "grape variety choices" and the students' scale factors were significant and, therefore, kept in the model. We also kept the non-significant main effects of attributes to facilitate comparisons between models, and to avoid endogeneity problems due to the omission of relevant attributes.

Results indicate that the HC model works as expected. The *Price* coefficients have the expected signs (i.e. positive in the expected quality's structural equation and negative in the choice utility). The quality indicator (level of agreement with the phrase "this wine is excellent") strongly correlates with expected quality, as reflected by a positive and significant parameter λ. Finally, expected quality has a positive and significant effect on choice utility.

Table 3-3 - Coefficients and goodness of fit measures for the ML model (robust t-test)

		Main effect		Standard deviation	
		Value	t-ratio	Value	t-ratio
Choice utility	Grape variety 1	0.000	0.00	0.646	4.71
	Grape variety 2	-0.036	-0.32	0.445	2.52
	Label 1	-0.338	-2.20	0.586	2.40
	Label 2	-0.008	-0.06	0.069	0.84
	Label 3	0.045	0.23	0.990	3.09
	Label 4	-0.375	-2.05	0.892	2.38
	Label 5	0.171	1.06	0.317	0.73
	Story 1	-0.101	-0.85	0.338	1.40
	Story 2	0.145	1.32	0.267	0.90
	Price	-0.001	-0.97		
	Price x experts	-0.005	-1.57		
	Price x students	-0.006	-1.36		
	Center position	0.212	2.41		
	No purchase	-2.010	-7.30		
Goodness of fit indicators	Number of parameters				23
	Number of observations (respondents)				1007 (168)
	Loglikelihood				-1170.974
	ρ^2				0.114
	Adjusted ρ^2				0.096
	Corrected ρ^2				0.022
	First Preference Recovery (FPR)				0.331
	FPR Expected value				0.387
	Chance recovery (CR)				0.274

In the ML model none of the *Price* parameters are significant, but exhibit the expected negative sign. Instead, in the HC model all *Price* parameters have also the expected sign: positive in the expected quality structural equation, and negative in the choice utility; but most of them are significant using a one-tail t-test (t-test critical value of 1.645 at 95% significance). In particular, only students exhibit a significant use of price as a cue for quality, while all classes exhibit a significantly negative effect of price, though with different intensities: students are the most sensitive, followed by experts and regular consumers.

Table 3-4 - Coefficients and goodness of fit measures of the HC model (robust t-tests)

		Main effect			Standard deviation	
		Value	t-ratio		Value	t-ratio
Expected quality	Grape variety 1	0.086	0.420		1.250	2.520
	Grape variety 2	-0.247	-1.140		1.190	3.620
	Label 1	-0.617	-2.600		0.946	3.560
	Label 2	-0.070	-0.370		0.771	2.710
	Label 3	-0.411	-1.400		1.450	2.700
	Label 4	-0.863	-2.480		1.410	4.280
	Label 5	0.069	0.350		0.914	2.880
	Story 1	-0.548	-2.710		0.794	2.030
	Story 2	-0.188	-1.290		1.050	3.780
	Price	0.001	1.540			
	Price x experts	0.003	1.360			
	Price x students	0.009	2.130			
	σ_{ω}	0.621	2.020			
	λ	0.854	4.240			
	Threshold 1	-5.100	-15.100			
	Threshold 2	-3.000	-11.770			
	Threshold 3	-0.263	-1.130			
	Threshold 4	2.320	8.800			
Choice utility	Expected quality	0.640	4.490			
	Price	-0.002	-2.460			
	Price x experts	-0.008	-2.650			
	Price x students	-0.028	-1.880			
	Center position	0.308	2.880			
	No purchase	-2.400	-9.410			
	$\mu_{gVarRnk}$	-0.684	-3.670			
	$\mu_{student}$	-0.608	-4.680			
Goodness of fit indicators	Number of parameters					35
	Number of observations (respondents)					1007 (168)
			with indicators		without indicators	
	Loglikelihood		-4184.61			-1181.42
	ρ^2		0.714			0.106
	Adjusted ρ^2		0.712			0.079
	Corrected ρ^2		0.072			0.013
	AIC		4254.6			1251.4
	BIC		8611.2			2604.9
	First Preference Recovery (FPR)		0.331			
	FPR Expected value		0.385			
	Chance recovery (CR)		0.274			

Concerning attributes other than price, even though there are similarities between both models, results are not always consistent between them. The main effects of *Grape variety* are zero in both models, meaning that –on average- there is no particular grape variety preferred over others. However, as all standard deviations of *Grape variety* are statistically significant, preferences for grape varieties are highly heterogeneous among participants. Both models agree on labels 1 and 4 being –on average- less preferred than the base label, though with significant variability in the population. Both models also agree on labels 2, 3 and 5 to be –on average- equivalent to the base label. But both models disagree on how preferences for labels 2, 3 and 5 distribute among the population, with the ML model implying that only preferences for label 3 have significant variability, while the HC model suggests that the preferences for all three labels do. Finally, the effect of *Story* is also different in both models: while the ML results imply that all stories are equivalent, the HC model recognizes story 1 to be the least preferred on average, and preferences for story 1 and 2 have significant variability among the population.

Both models indicate that most of the main effects are statistically equivalent to zero. This is probably due to preferences being highly heterogeneous among consumers, cancelling out on average. As there is no single grape variety, label or story clearly superior to the others, preferences are only a matter of taste. This reflects on the relatively high values of the standard deviations estimated for most parameters, a phenomenon better captured by the HC model than by the ML model. This variability seems to be inherent to all consumers, and not an artefact arising from mixing different classes of them (i.e. experts, regular consumers and students). We tested removing students -probably the most eccentric class- and found no evidence of a decrease in preference variability, nor an increase of t-tests on their average effects.

We tested the effect of alternatives' position on choice by including constants for the left and central alternative (see Figure 3-1) in the utility function. Results were consistent in both models, with only the central position achieving significance. We therefore kept a constant for the central alternative in the final model, effectively controlling for presentation order bias.

The goodness of fit indices of both models must be compared with care, as their structures are different: while the ML model takes into consideration only the consumers' choices, the HC model also includes the expected quality indicators. Therefore, only the choice part of the HC model must be taken into account when comparing fit indices (Table 3-4 presents goodness of fit indices differentiated for the whole HC model and its choice component). As expected, the ML model fits choices better, as all its parameters are exclusively dedicated to fit them, unlike the HC model, where the grape variety, label and story parameters must reproduce the respondents' answers for both the expected quality indicators and the choices. This extra restriction implies a difference of 10 points between their log-likelihoods, and a global loss of fit as the ρ^2, adjusted ρ^2, corrected ρ^2, and Akaike and Bayesian information criterions (AIC and BIC) point out. This loss of fit is significant ($p<0.01$) according to Horowitz (1983)'s test for non-nested models.

In principle the prediction capacity of both models could be tested in-sample and out-of-sample. The First Preference Recovery (FPR) or “percent correctly predicted” is an index of prediction accuracy, which assumes that the alternative with the highest predicted probability is chosen, and then it compares this prediction with the actual choices to determine how many times the prediction was “accurate”. However, this is a poor index, as Train (2009, page 69) explains: “The researcher has only enough information to state the probability that the decision maker will choose each alternative. (…). This is quite different from saying that the alternative with the highest probability will be chosen each time.” Following Gunn & Bates (1982) we present the actual FPR its expected value and the value of Chance Recovery (CR), i.e. the prediction by chance. Results are as expected, because market shares in unlabelled experiments tend to be similar to chance recovery, as otherwise the experiment would be unbalanced towards a particular alternative. Furthermore, the FPR is expected to improve significantly if we used individual level parameters for prediction (Train 2009, chapter 11).

3.5 Discussion

In this study, the HC model using expected quality as a latent variable allowed us to successfully reduce price endogeneity. This reflects on the increased t-test of the price coefficients in the choice utility for all groups of participants. This improvement is caused by the separation of the positive effect of price due to its role as a cue for quality, and its negative effect due to the participants’ budget restrictions. While the positive effect is captured in the structural equation of perceived quality, the negative effect is captured in the choice utility.

Comparing our results with other wine studies can only be done in general terms. Most comparable studies were not performed on the same market as ours, so price sensitivities are expected to change. However, it is possible to analyse the general behaviour of price coefficients estimated in studies both with and without endogeneity correction.

Among the studies that do not correct for endogeneity, most tend to find non-linear effects of price, such that mid-range prices provide higher utilities than lower and higher prices. This is likely due to the double effect of price: people may think that wines below some price are of low quality, therefore utility increases with price for a certain interval, but after overcoming a given price threshold, the budget restriction outweighs the price-quality association and the choice utility decreases again. Lockshin *et al*. (2006) do not report the coefficients of their estimated model, but plot simulations showing how market shares first increase with price, reach a peak at about US$ 11 and then decrease again after that point. Similarly, Mtimet & Albisú (2006) used a quadratic form for price finding a similar concave shape, with the peak utility at about US$ 7. Using dummies for price levels and latent classes, Remaud *et al*. (2008) and Mueller *et al.* (2010) found that some classes had this same concave behaviour.

Unlike other studies, Barreiro-Hurlé *et al.* (2008) and Stasi *et al*. (2014) obtained monotonic decreasing effects for price in the choice utility without correcting for endogeneity. Stasi *et al.* (2014) used customised (pivoted) prices for alternatives, varying among 90% and 140% of the average wine price in the area and obtained a negative and significant price coefficient; but it is possible that their strategy for determining alternatives' prices only allowed them to capture the decreasing part of the price-utility curve (i.e. where the budget effect overweighs the price-quality association). Something similar might had happened in the work of Barreiro-Hurlé *et al*. (2008), who found a negative and highly significant price coefficient using four price levels: 3, 7, 10 and 14 Euros (about 4, 9.5, 14 and 19 US$). According to Mtimet & Albisú (2006), who also studied the Spanish market, three of these levels would fall into the part of the price-utility curve where the budget effect overweighs the price-quality association. Palma *et al*. (2013) also found a negative coefficient for price; however, they explicitly pivoted the alternatives' prices above the participants self-reported willingness to pay for the considered occasion (from 100% to 160%).

Papers that do correct for endogeneity at an individual level yield results as expected. Appleby *et al.* (2012) used experts' ratings as a proxy for quality when modelling a stated purchase decision. They found a negative and significant effect of price, as well as a positive effect for the experts' ratings. However, as very few attributes were included in their study, it is possible that participants relied on the experts' ratings more than they would under more realistic conditions (Channey 2000, Johnson & Bruwer 2004, Goodman 2009, Atkin & Thach 2012).

Mastrobuoni *et al.* (2014) estimated both the positive and negative effects of price separately, finding a negative and significant coefficient for the budget effect of price, and a positive and significant effect for the price-quality association, though only up to 5 Euro (about US$ 7). Their approach to endogeneity correction could be considered as mixed, as they explicitly separated the modelling of both effects (i.e. used a latent variable approach), but also included experts' ratings as a proxy for quality. In their experiment, consumers tasted a set of wines, then chose their preferred alternative, and finally chose the one they would buy. With the first answer the authors modelled the perceived quality using experts' ratings (which consumers do not see) as a proxy for sensory quality. Then, they explained the (hypothetical) purchase decision as a trade-off between price and perceived quality. This approach has three main limitations. First, as it includes tasting, the method is not suitable for situations where the consumer has not tasted the wine (e.g. a first buy). Secondly, and as mentioned before, the use of experts' ratings as a proxy for sensory quality has been questioned (Hodgson 2009). Finally, the proposed estimation method neglects the inherent noise of perceived quality, therefore introducing endogeneity (the measurement of perceived quality becomes a noisy proxy). We tested an analogous procedure to Mastrobuoni *et al.* (2014) with our dataset, yielding only positive coefficients for price.

In our application, significant coefficients were obtained for the positive effect of price on students, and on all classes of participants for the negative effects of price. Consumers and experts' positive effect of price had the expected sign, with (one-sided) t-tests' p values of 0.06 and 0.09. Results seem to be robust to the particular model structure, as we estimated models without random parameters and with random price parameters, and results remained analogous (i.e. the sign of the price coefficients remained the same, and their t-tests did not decrease significantly).

Students were found to be the most price-sensitive group, but also those who more strongly associated price with quality, as it would be expected for lower-income and less knowledgeable consumers. Experts appeared to be more price sensitive than regular consumers, probably because they purchased wine more frequently than regular consumers and therefore looked for cheaper alternatives. This, however, is only a possible explanation, as we did not record consuming nor purchase frequency in our questionnaire. We tested for other possible explanations, such as income effect and non-linearity in the effect of price, but none of them turned out to be significant.

The positive effect of price could be overstated because of our experimental design. By asking participants at the beginning of the experiment what their minimum and maximum willingness to pay for wine were, and then using these values throughout the SC scenarios, we might have reinforced the use of price as cue for quality. Let us consider the following situation. Participants, when asked for their minimum WTP, think of the lowest quality wine they would be willing to buy and state their WTP for it. Then, when asked for their maximum WTP, respondents think of the best quality wine they have tried, and state their WTP for it. This would lead them to associate immediately the low price with low quality and the high price with high quality during the SC experiment. This, however, could not happen if participants did not use price as a cue for quality. If that was the case, then their willingness to pay range would be completely determined by their budget constraint and the lowest wine price in the market. Therefore, even though our experimental design might have artificially increased the positive effect of price to some degree, it could not have artificially induced it. To avoid this potential problem, we recommend using predetermined price ranges when applying the endogeneity correction method.

Concerning attributes other than price, results of the model with (HC) and without (ML) endogeneity correction are generally aligned, but the HC model seems to provide more information. Preferences for grape variety are highly heterogeneous in the sample, making it impossible to declare a single variety as preferred –on average- over the others. Labels 1 and 4 are less preferred than the base label, while preferences for labels 2, 3 and 5 seem to be equivalent to the base label on average, though both models disagree on the (significant) level variability of these preferences. Finally, while the ML model makes no difference among preferences for stories, the HC model suggests that story 1 is –on average- significantly less preferred than the base story. These results indicate that price endogeneity might not only affect the price parameters, but also other attributes' parameters, though to a lesser degree.

Contrary to some published literature (Quester & Smart 1998, Hall 2003, Martinez-Carrasco *et al.* 2006), we found that the effect of consuming occasion was non-significant for Chinese respondents. Several factors may have influenced this result. First, we may have described the consuming occasion without enough detail, making it difficult for participants to picture themselves in it. Secondly, it may be that simply stating the consuming occasion is not enough to evoke such a context in the mind of

Chinese consumers; therefore, more compelling methods should be tried in the future. Finally, in formal occasions the Veblen (or snob) effect is more likely to play an important role, but this effect usually manifests itself through brand value. As brands were fictional in our experiment, this effect was probably absent, therefore diminishing the effect of consuming occasion.

Several simple improvements could be applied to the method employed in this paper in order to correct for endogeneity. First, more than one expected quality indicator could be used, though it remains to be determined what indicators would be best. Secondly, it is not necessary to collect an expected quality indicator for each alternative, as it would be possible to separate the survey into two parts: one where only quality indicators are collected (i.e. a series of wines the expected quality of which had to be assessed), and another one where only choices are required (i.e. as in a traditional SC experiment). This could allow optimizing the data collection method by using different efficient designs for each stage, but might decrease the correlation between expected quality and choice.

Despite the limitations of this particular application, modelling quality as a latent variable seems to be a promising approach to deal with endogeneity while being consistent with commonly accepted behavioural frameworks and not demanding excessive extra effort from respondents. Additionally, this method does not require difficult–to-find proxies or instruments. Finally, the method seems to be fairly robust, as it worked on a relatively small and very heterogeneous sample, using a single quality indicator, and on a choice experiment that was not incentive compatible, where choosing an expensive wine had no actual consequence on participants.

Even though this particular case study was concerned with wine choice, the modelling structure can be applied to any product the quality of which is uncertain to the consumer, even after considering observable attributes. Most food and beverage products fit this description, but also many leisure activities do too (e.g. selecting a travel company, choosing a show or a play, etc.), as well as some sparsely bought products or services the quality of which is hard to determine by the consumer even after purchase (e.g. jewellery, some medicine, broadband providers, etc.).

Our approach should not be of much use in cases were the main source of endogeneity is the Veblen effect. In such cases, endogeneity is caused by the unobserved social benefits of conspicuous consumption, which are correlated with price, but are not related with perceived or expected quality. Therefore, modelling quality as a latent variable in such cases would not provide any new information; even more, it might lead to the wrong conclusion that price itself has a positive effect on consumers, when in reality it is conspicuous consumption that provides utility to the consumer. In these situations, including the brand of the product or a measure of its social appreciation might be more useful.

It is very likely that consumers present both the Veblen effect and the use of price as a cue for quality at the same time. This is probably more problematic in a revealed preference context, were brand and quality uncertainty go hand in hand, but less so in a SP experiment with fictional brands, such as the one analysed in this paper. As no real brands were presented, there is no benefit to be obtained from conspicuous consumption, beside that provided by the observable attributes (*e.g.* a particular label looking more luxurious than another).

Modelling quality as a latent variable assumes that prices are exogenous. Therefore, this method corrects endogeneity only due to the use of price as cue for quality, but does not correct endogeneity due to price’s simultaneous determination (i.e. supply and demand equilibrium). If this later effect is to be considered, then an additional endogeneity correction method especially suited for it should be used.

The latent variable approach for endogeneity correction shows highly promising results, but its real performance should be measured against a revealed preference study, an area we are currently working on. The method should also be compared with other available approaches to correct for endogeneity, notably the Control Function Approach (CFA) and Multiple Indicator Solution (MIS).

4 WHEN EXPENSIVE IS GOOD: MODELLING THE WINE REPURCHASE DECISION

4.1 Introduction

Foods and beverages are experience goods (Nelson 1970, Grisolía *et al.* 2011). Some of their relevant attributes, such as taste and aroma, cannot be perceived before purchase. This makes it impossible for consumers to fully determine product quality before buying them. In this context, and throughout this paper, we understand quality as a subjective appreciation of the product by the consumer.

According to Grunert (2005), as quality is not available to guide the purchase decision the first time consumers face a new product, they construct an *expectation of quality* to guide their choice. Expected quality is determined only by the product's attributes that can be perceived before purchase and consumption. These attributes are called *extrinsic*, and typical examples are the packaging, price, health claims, in-store displays, etc. The effect of extrinsic attributes on the expected quality, however, is mediated by each consumer's perception of them.

Only after purchase and consumption, can consumers perceive the physical (or sensory) attributes of a product and determine the *experienced quality* of it (i.e. the product's actual quality, as experienced by the consumer). These attributes are called *intrinsic*, and are those that can only be perceived during consumption, with taste and aroma being the most prominent ones. Preferences for intrinsic attributes are also heterogeneous among consumers. Furthermore, *experienced quality* is not determined only by the intrinsic attributes, but also by the *expected quality* and the consumption context.

After tasting the product, the next time consumers face a purchase opportunity their decision should also be influenced by their recalling of its *experienced quality*. This means that even if consumers liked a given product, they may not buy it again if they cannot remember its name or how much they liked it. The whole process is summarized in Figure 4-1.

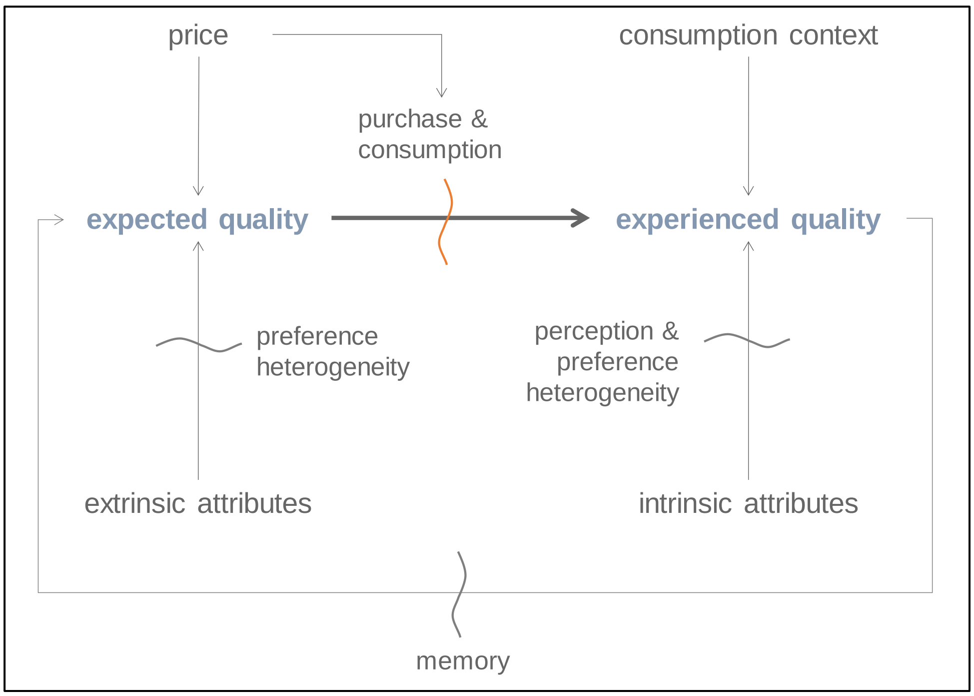


Figure 4-1 - Hypothetical process of purchase and consumption of food and beverage products

Note that price can have a double and opposite effect on the purchase decision process. A higher price can positively influence the probability of choosing a product if consumers use price as a cue for quality, that is, consumers believe that more expensive products are of higher quality (Palma *et al.* 2016). On the other hand, a higher price can diminish the probability of choosing a product given consumers' budget restrictions and, consequently, their tendency to avoid higher prices. These two effects fit within the theoretical framework exposed in Figure 4-1. If we consider the purchase decision as driven by a trade-off between *expected quality* and *price*, then the positive effect of price can be captured as a positive influence of *price* on *expected quality*, and its negative effect by a direct (negative) influence of price in the purchase decision.

The positive effect of price is documented both in the consumer behaviour literature in general, as well as in the wine literature in particular. Scitovsky (1945) proposes than assuming a positive correlation between price and quality is reasonable in perfect markets, where consumers have perfect information. Leavitt (1954) takes an empirical approach, showing (for a relatively small sample) that consumers do choose the more expensive products when they do not have enough information to determine the product's quality, especially in the case of products with high vertical differentiation (i.e. relevant quality differences within the product category). Dodds *et al.* (1991) take a more formal approach, proposing a structural equations model (SEM) where price positively influences perceived quality, and negatively influences willingness to buy. Their data -which also controls for brand and store- confirms the model.

In the wine literature, Plassman *et al.* (2007) showed that consumers associate price and quality at a neuronal level. They gave a set of subjects the same wine twice, without telling them: first with a low price tag, then with a high price tag. Consumers were subject to functional Magnetic Resonance Imaging (MRI) while tasting the wines, and they were also asked to rank their level of pleasantness for each wine. Both measurements indicated a higher level of pleasantness when the price was higher. Inspired by these results, Lewis & Zalan (2014) investigated the relationship between price, enjoyment, and willingness to pay of wine consumers, finding that higher prices positively influenced the other two measurements. Finally, Aqueveque (2006; 2008) found that higher prices reduced perceived risk, while increasing perceived quality on wine consumers. But they also found that the positive effect of price on perceived quality weakened when a proxy for the product's quality (such as expert's rating) was provided.

Failure to consider the double effect of price when modelling choice can lead to specification issues. Under a traditional approach, interpretation of the price coefficient can be confusing. As a proof of concept, consider the following true form of consumer n's willingness to buy product j (WTB_{nj}).

$$WTB_{nj} = \beta_p price_j + \beta_q Q_{nj} + \varepsilon_{nj} \tag{4.1}$$

$$Q_{nj} = \gamma_p price_j + X_j \gamma_x + \omega_{nj} \tag{4.2}$$

where $price_j$ is the price of product j; Q_{nj} is product j's quality, as perceived by consumer n; X_j is a vector of product's attributes; ε_{nj} and ω_{nj} are uncorrelated and independent random error components specific to each product and individual; and β_p, β_q and γ_x are parameters to be estimated. This model is a simple operationalization of the first purchase of a product under the behavioural model on

Figure 4-1. WTB_{nj} is determined by a trade-off between perceived quality and price (as $\beta_p < 0$ and $\beta_q > 0$ are assumed), and *perceived quality* is determined by the product's observable attributes *price* and *X*. Now suppose the modeller does not observe Q_{nj} (as it happens in reality), and that she estimates the following model:

$$WTB_{nj} = \alpha_p price_j + X_j \alpha_x + \eta_{nj} \tag{4.3}$$

where $\alpha_p = \beta_q + \beta_q \gamma_p$, $\alpha_x = \beta_q \gamma_x$ and $\eta_{nj} = \beta_q \omega_{nj} + \varepsilon_{nj}$. The estimated parameters are still consistent, as η_{nj} does not correlate with *price* or *X*, but the estimation and interpretation of α_p can be complicated. α_p measures price's net effect on *WTB*, which is a mix of two potentially very different effects. The positive effect is most probably associated with marginally decreasing quality improvements (i.e. quality gains diminish for higher prices). To see this clearer, consider three wines costing 3.90, 8.90 and 13.90 US$. The quality difference between the first and the second is probably larger than the difference between the second and the third. On the other hand, the negative effect of price probably depends on each consumer's income, as richer consumers may be more willing to spend higher amounts. But quality perception may not be influenced by consumers' income.

Measuring both effects of price using a single indicator of choice is not possible, as choice alone only provides information about the net effect of price. Therefore, if we are to measure both effects separately, we require at least one additional indicator that measures a single effect of price (positive or negative). Once we control for one of the two effects of price, then the choice indicator can provide us with the size of the other effect.

Modelling the positive and negative effect of price separately poses several benefits. First, it allows giving different functional forms to each effect of price. This can help reduce the estimation error of price's effects, but can also provide more complex forms for demand as a function of price. For example, if price is a cue for quality, then demand may be upward for a certain range of prices. A second benefit is the extra information provided by the second indicator. This additional information can help not only reduce the estimation error of price, but also of other attributes.

In this paper, we focus on modelling the repurchase decision, that is, when consumers know both the extrinsic and intrinsic attributes of the product- while considering the double effect of price and measuring the relative weight of intrinsic and extrinsic attributes on the purchase decision. Even though the double effect of price has been empirically demonstrated on wine consumers (Plassman *et al.* 2007, Lewis & Zalan 2014, Aqueveque 2006; 2008) and theoretically modelled (Grunert 2005), few have proposed a method to measure both effects in choice experiments (Lockshin *et al*. 2006, Mastrobuoni *et al*. 2014) and –to our knowledge- none in a statistical and economics consistent way.

We test two different approaches to control for price's double effect. The first is based on the Multiple Indicator Solution (MIS, Guevara & Polanco 2016) method, and consists in using two indicators of quality: one as an endogenous proxy for the omitted variable, and another as an instrumental variable for the first one. The second is based on the Latent Variable (LV) approach (Guevara, 2015) and consists in explicitly modelling *perceived quality* as a latent variable. In this model the positive effect of price is captured by the *perceived quality*'s structural equation, and its negative effect is captured as a direct influence on the purchase decision. We compare these two models with a third one where we do not control for price's double effect. We therefore estimate, report and discuss three models: one without separating price's double effect (called Traditional model), one using the MIS method (called Simplified model), and finally one using the LV approach to separate both effects (called Full model).

As a case study, we use data from a stated choice (SC) experiment on wine preferences, answered by 122 Chilean consumers. The experiment included tasting and aimed to measure preferences for three attributes: price, label and four synthetically added flavours to the wine.

The rest of the paper is organized as follows. The next section presents a short review about measuring the double effect of price in the wine literature. On section 3 we present the sample's main characteristics, review the design of the experiment and describe the statistical models used to analyse the data. We present results for each of the three estimated models in section 4, and in section 5 we discuss them and propose further topics of research.

4.2 Literature review

The importance of price as a cue for quality has been noted before in the wine preference literature. Lockshin *et al*. (2006) discovered that the effect of price on (hypothetical) purchase probability had an inverse U shape, i.e. an increase in price raised purchase probability for cheap wines, but diminished it for expensive wines. They found the optimum price (i.e. the top of the inversed U curve) to be around 12 AUD (~16 US$) for the Australian Market, though this value changed depending on consumers' characteristics and wine's attributes. Using a similar approach, Mtimet & Albisú (2006) used a quadratic function to model the effect of price on consumers' utility. They also found the inverse U shape, with a maximum utility around 7 US$. Remaud *et al*. (2008) and Mueller *at al.* (2010b) used dummy variables to measure the effect of price, as well as latent classes to consider preference heterogeneity among consumers. They found the same inversed U shape for the effect of price, but only for some classes of consumers.

This can be explained by considering the double effect of price: when a bottle of wine is cheap, its quality is assumed to be poor, therefore an increase in its price generates an important increase in its perceived quality (as now the wine is not cheap), but a relatively small decrease on the consumer's perceived utility due to price (as the wine is still not expensive). On the other hand, when an expensive wine increases its price, the net effect is negative, as gains in quality are not that important (because the wine was already perceived as good) but the budget restrain gets more binding when it was already quite strained. Both the inverse U shape, as well as its explanation based on the behavioural model, indicate a non-linear behaviour of either the positive or negative effects of price (or both).

Other studies have ignored the double effect of price, but still found a decreasing and monotonic effect on purchase probability, just as economic theory proposes. However, most of these findings may be due to their experimental designs including prices only in the decreasing part of the inverse U-shaped curved described by Lockshin *et al*. (2006). Stasi *et al.* (2014) used prices between 90% and 140% of the average wine price on the area of their study. Similarly, Palma *et al*. (2013) used prices between 100% and 160% of each consumer's self-reported maximum willingness to pay for a bottle of wine. Barreiro-Hurlé *et al*. (2008) used four price levels (3, 7, 10 and 14 euros), three of which fell on the decreasing part of the inverse U-shaped curve described by Mtimet & Albisú (2006), who also studied the Spanish market.

Measuring the positive and negative effects of price has also been done before in the wine preference literature, but not without some issues. Costanigro *et al*. (2014) used Wine Spectator's grading system (i.e. an expert's rating) as a proxy for *expected quality*. Proxies must be exogenous to both the choice utility and the omitted variable, while correlating with the second. In this particular case, these requirements translate into the expert's rating having to satisfy two requirements: (i) it should be an objective measure of quality, so it is exogenous to the choice and the *expected quality*; and (ii) it must correlate with consumers' own *expected quality*. Both of these requirements have been seriously questioned (Lawless 1984, Hodgson 2009, Lattey *et al.* 2009, Gokcekus & Nottebaum 2011, D'Alessandro & Pecotich 2013, Hopfer & Heymann 2014). Also, including experts' ratings as an extra attribute does not seem appropriate as most consumers do not seem to be aware of them (Channey 2000, Johnson & Bruwer 2004, Atkin & Thach 2012).

Mastrobuoni *et al.* (2014) explicitly measured the positive and negative effects of price using a two-stage approach. In a SC experiment with tasting, they asked participants to choose both the best wine among the alternatives, and the one they would actually buy. Then, they first used a logit model to explain the choice of the best wine using all observable attributes -including price- and an expert rating as a proxy for participants liking of each wine (participants tasted the wine, but they did not observe the experts' rating). Then, they used another logit model to explain the (hypothetical) purchase decision of participants, as a trade-off between price and *experienced quality*, using the first logit's deterministic part of the utility as a proxy for perceived quality.

Though Mastrobuoni *et al.* (2014)'s approach effectively distinguishes both effects of price, it has some problems. First of all, and as mentioned before, it is not clear that an expert's rating is a good proxy for consumers' *expected* nor *experienced quality*. Secondly, the sequential estimation process used by Mastrobuoni *et al.* (2014) might introduce bias in the estimation of the second logit model. For the deterministic utility of the first logit to be an adequate proxy of experienced quality in the second logit, it must satisfy three conditions: (i) be exogenous to the second logit's utility; (ii) be exogenous to *experienced quality*; and (iii) correlate with *experienced quality* (see Guevara 2015 for more details). The second condition is not met, as the first logit's deterministic part of the utility is a noisy measure of the real *experienced quality*, and therefore is not exogenous to it. Authors should have included the whole utility from the first logit model (including its error term) to avoid problems. But their omission introduces a new source of endogeneity in the model, therefore making the estimates of the second logit model potentially biased.

4.3 Material and Methods

4.3.1 Sample

Table 4-1 - Participant's main socio-demographic characteristics

		Male	Female	Total
Number of participants		72	50	122
Age (in years)	18 - 20	15	8	23
	21 - 23	31	21	52
	24 - 25	13	10	23
	26 - 28	4	6	10
	29 - 30	3	0	3
	31 - 35	4	4	8
	36 - 40	2	1	3
Maximum level of Education of the head of the Household	Incomplete primary school	1	0	1
	Complete primary school	2	0	2
	Incomplete high school	1	1	2
	Complete high school	6	3	9
	Incomplete technical	2	1	3
	Complete technical	7	3	10
	Incomplete university	5	4	9
	Complete university	30	26	56
	Incomplete postgrad	2	2	4
	Complete postgrad	16	10	26
Monthly Household Income (in USD)	0 to 214	1	1	2
	215 to 429	4	2	6
	430 to 643	4	4	8
	644 to 964	3	6	9
	965 to 1393	9	3	12
	1394 to 2250	14	9	23
	2251 to 3107	11	6	17
	3108 and more	23	18	41
Consuming Frequency	Never or almost never	11	6	17
	1 to 3 times a month	19	18	37
	Once a week	27	12	39
	2 to 5 times a week	12	13	25
	Everyday or almost everyday	1	1	2
	Without information	2	0	2

Participants were mainly graduate and postgraduate students, but also young professionals working at the university campus where the experiment took place. No incentive was offered for participating in the experiment. Data collection was performed in three waves, reaching a total of 122 participants. Their main socio-demographic characteristics are summarized in Table 4-1. Most participants were 25 years old or younger, from highly educated households (most household's heads held university-level degrees) as well as fairly affluent (approximately half of the sample's households belonged to the country's richer 20%, according to Ministerio de Desarrollo Social 2014). Even though the sample size is small, the sample composition successfully represents the high-income millennial target market.

4.3.2 Survey design

Our study aimed to model the repurchase decision of wine consumers, i.e. the decision to buy or not to buy a bottle of wine after having tasted the product. To this end, we used a stated choice (SC) experiment, a technique that has several advantages over using revealed preference (RP) data (i.e. actual purchase records).

In particular, it is easy to collect several choices from each respondent, it allows using alternatives that are not available in the market, and makes use of experimental designs that increase the quality of the collected data. The main disadvantage of SC data is that individual responses may suffer from some biases (Ortúzar and Willumsen, Chapter 3), although the mental processes that drive decision making are thought to be the same (Louviere *et al.* 2000).

The experiment was divided in five stages (Figure 4-2). In the first, participants were asked to blindly taste five different wines, one at a time in a randomized order, each of them uniquely identified by a three-digit number (e.g. “642”). For each wine, participants were asked to provide two indicators of their sensory liking: their level of overall liking on a 9-point Likert scale (from “I dislike this wine very much” to “I like this wine very much”) and a qualitative proxy for their willingness to pay for that wine using a 5-point Likert scale (“I would pay much less that I am used to” to “I would pay much more than I am used to”).

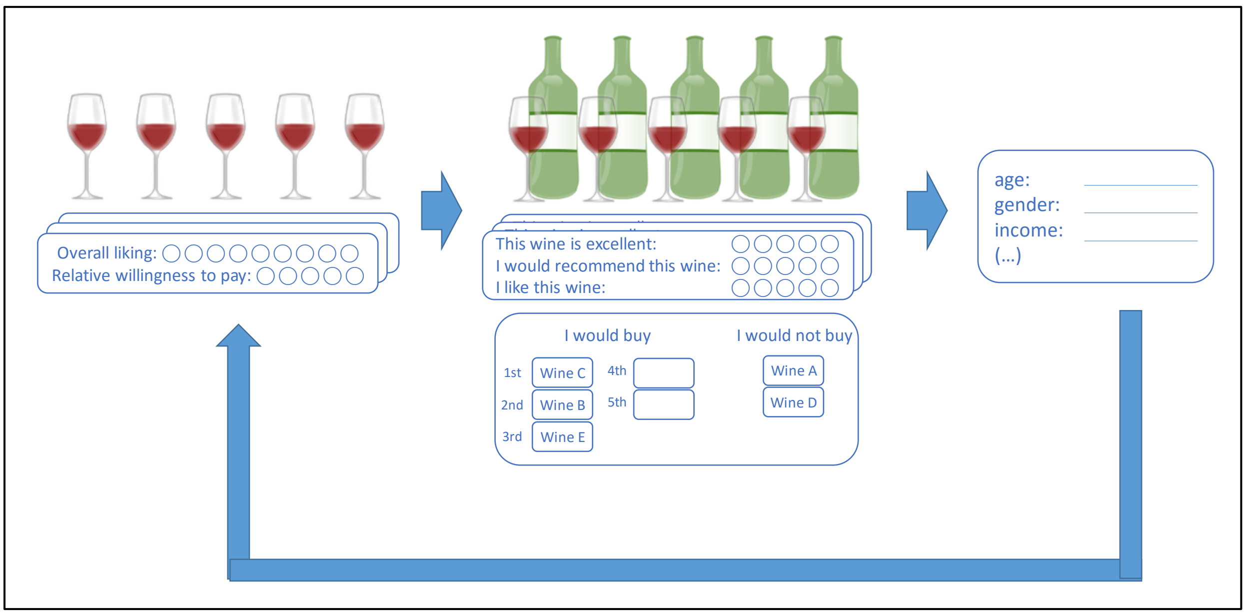


Figure 4-2 – Illustration of the experiment's procedure. Glass of wine by rinze, Wine bottle by Solarisphere (openclipart.org) (cc)

The SC exercise was the second stage of the experiment (Figure 4-3). We simultaneously presented five alternative wines to participants on a screen, each of them described by three attributes: label, price, and taste. An image showed each wine's label, and under it we showed its price in Chilean pesos (Ch$). We described the taste attribute by referencing the wines consumers had just tasted on the previous stage (e.g. the first wine was said to be sample "642"). As the tasted samples were still on each participant's table, they were able to try the wine again if they wanted to.

Before answering the SC exercise, consumers had to provide three indicators of quality for each of the five alternative wines using a 5-point Likert scale, presented as an unlabelled 5-star rating. The indicators of quality were the participant's level of agreement with the following statements applied to each alternative: "This wine is excellent", "I would recommend this wine to my friends" and "I like this wine".

The SC exercise considered a single scenario, where participants had to rank the alternatives they would actually buy. Participants could leave alternatives out of the ranking if they would not buy them in a real purchase situation. We asked consumers to consider a party at a friend's house as the wine's consuming occasion. We decided to ask for a ranking instead of a single choice to increase the amount of information collected.

Once the SC exercise was completed, participants moved to the third stage, where they had to answer a short questionnaire about their socio-demographic characteristics and consuming habits. The third stage also served as a forced pause during the experiment, as stages fourth and fifth were repetitions of the first and second stages. In these last two stages we used a different three-digit code for the tasted samples, therefore inducing consumers to think they were evaluating new wines. The repetition of the experiment was mainly to ensure a large enough number of observations.

While the blind tastings of the first and fourth stages aimed to measure the sensory appreciation of each wine, the SC exercise during the second and fifth stages aimed to measure the effect of both sensory appreciation and extrinsic attributes on the purchase decision. The tasting was made in blind conditions so it would be easier to measure the effect of intrinsic (i.e. sensory) attributes on participants' sensory appreciation. However, in the SC exercises, the effect of both intrinsic and extrinsic attributes on the purchase decision was measured at the same time, just as it happens when consumers make actual repurchase decisions.

The SC exercise followed a D-efficient design (Rose & Bliemer 2009; Rose *et al.* 2008; Ortúzar & Willumsen 2011, section 3.4), assuming a multinomial logit model, with 20 choice scenarios divided in 10 blocks of two scenarios each. At the beginning of the survey, participants were randomly assigned to one block, then they would answer the first scenario of the block at the second stage, and the second one on the fifth stage of the experiment. Both scenario and alternative orders were randomized to avoid position bias. *Priors* for the efficient design were first assumed to be zero, and then updated after each data collection session.

Figure 4-3 – Example of a SC exercise (labels are not the real ones used on the experiment)

Three attributes were considered in the SC exercise: *price*, *label* and added *flavour*. These attributes, as well as their respective levels, were determined in association with the private vineyard who partially financed the study. The relevance of label and price is well stablished on the wine marketing literature (Lockshin & Hall 2003; Lockshin & Corsi 2012), and even though the relevance of overall sensory liking is also acknowledged (Siegrist & Cousin 2009, Goodman 2009, Combris *et al.* 2009, Mueller *et al.* 2010a), the relevance of particular flavours and aromas is still hard to measure. Instead, measuring the impact of intrinsic attributes has been mostly limited to simple ones, such as sugar level and fat content (Enneking *et al.* 2007, Hopper *et al.* 2012).

Five different wines were considered in the experiment, all of which were the same base wine, except that four of them had synthetically added flavours. As the base wine was the same in all cases, a single attribute describing the added flavour was sufficient to fully describe the differences in sensory characteristics between the wines (as in traditional choice models only differences between alternatives matter), making a full description by a trained panel unnecessary. The considered attributes and their levels are shown in Table 4-2, though most levels are not reported by request of the private vineyard who funded part of the study. We restrained the experiment to only three attributes to avoid requiring an excessively large sample of respondents.

Table 4-2 – Attributes and levels used in the SC exercise

Level	**Label**	**Price (US$)**	**Added flavour**
Base	Base label	2.17	None
1	Label 1	3.15	Flavour 1
2	Label 2	4.98	Flavour 2
3	Label 3	7.98	Flavour 3
4	Label 4	11.65	Flavour 4

4.3.3 Modelling

All estimated models are based on the behavioural framework described in Figure 4-1. However, as our focus was on the repurchase decision, we reformulated the behavioural model from that perspective into the structure displayed in Figure 4-4.

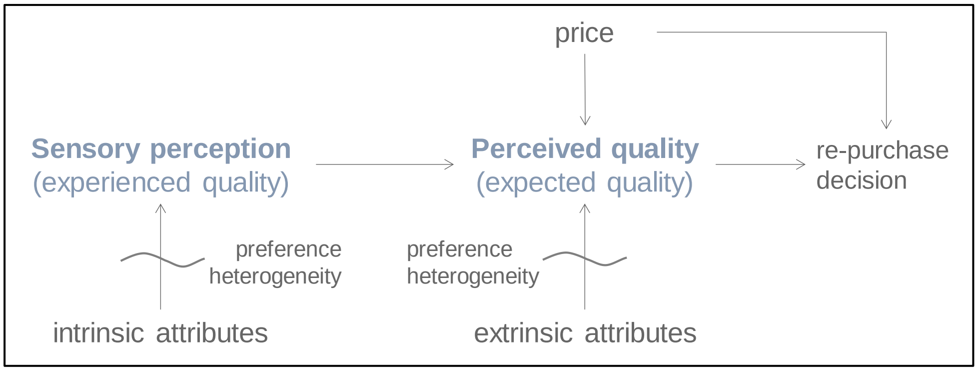


Figure 4-4 - Behavioural model for repurchase

Several simplifications had to be considered to keep modelling tractable:

(i) We assumed that the *experienced quality* associated with the first time consumers tasted the product was independent of *expected quality*, that is, *experienced quality* is only influenced by intrinsic attributes. This could be the case if consumption happened a long time ago and the consumer forgot the extrinsic attributes of the product, or if the first tasting was done in new-blind conditions (e.g. at a friend's house where only a glass is served). We made this assumption to facilitate measuring the effect of intrinsic attributes on liking, and this led us to rename *experienced quality* as *sensory perception*.

(ii) The effect of time and memory cannot be modelled, as the hypothetical re-purchase decision in our experiment was done almost immediately after tasting the wine.

(iii) The effect of consumption context was ignored, as we performed the experiment in a single controlled environment (a mobile laboratory), making it impossible to measure the influence of a real consuming occasion on the re-purchase decision.

(iv) As the experiment focused on re-purchase instead of the first purchase of a product, what we called "expected quality" in Figure 4-1 is now more accurately called *perceived quality*, as it is also influenced by the wine's tasting.

We estimated three different models based on Figure 4-4's repurchase behavioural model; albeit with different simplifications, all of them are Hybrid Choice Models (HCM, Ortúzar and Willumsen, 2011, section 8.4.3; Bolduc & Alvarez-Daziano 2010). The first one (Traditional model) ignores the double effect of price (i.e. price as a cue for quality and a strain to the budget restriction), and it represents the traditional way of using HCMs to model wine purchase decisions. The second (Simplified model) is a sequential implementation of the full behavioural model. It considers the double effect of price and is based on the MIS approach (Guevara & Polanco 2016). Finally, the third one (Full model) it is a full implementation of the repurchase behavioural model of Figure 4-4, therefore effectively considering the double effect of price. The Traditional and Full models were estimated using the

Python version of Biogeme (Bierlaire 2003), while the Simplified model was estimated in R (R Core Team 2015) using the maxLik package (Henningsen *et al.* 2011). In the remaining of this section, each model is presented and described in detail.

4.3.3.1 Traditional model

This model does not consider the double effect of price (i.e. its effect as a cue for quality and as a strain to consumers' budget restrictions). The model is based on a simplification of the repurchase behavioural model (Figure 4-4): *perceived quality* is ignored and *price, sensory perception* and the extrinsic attributes are assumed to influence the re-purchase decision directly. This represents the traditional approach of choice models applied to wine preferences, which includes all attributes directly in the utility function, assuming that their effects on the purchase decision are not mediated by any construct other than utility itself. The only exception are the intrinsic attributes, which are mediated by the variable *sensory perception*. This structure is represented in Figure 4-5.

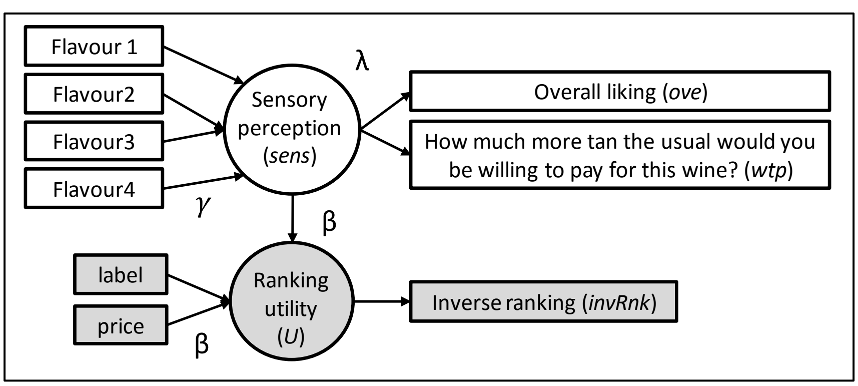


Figure 4-5 - Structure of traditional model

In Figure 4-5, the observable attributes and indicators are drawn in squares, while the non-observable (latent) variables are depicted in circles. The first LV is *sensory perception* (*sens*), representing how much the participant liked the wine during the blind tasting. This LV depends only on the wine's intrinsic attributes, and reflects the liking indicators provided by participants during the blind tasting sessions. The second latent variable is (inverse) *ranking utility*, which is determined by the alternative's *sensory perception,* its *label* and its *price*. Even though we cannot directly observe *ranking utility*, we observe the *inverse ranking* of each alternative, which depends on its *ranking utility*. The structure of this model is not able to separate the double effect of price, as only its net effect (i.e. the mix of both the positive and negative effects of price) is captured in the *ranking utility*'s structural equation.

The white part of Figure 4-5 is a structural equation model (SEM) with linear structural equations and ordered logit models (Train 2009, chapter 7.3) as links between the latent variables and their indicators. The grey part of Figure 4-5 is an ordered logit model with random coefficients and error components. Both parts can be estimated sequentially, but we estimated them simultaneously to achieve higher efficiency.

The white part of Figure 4-5 is implemented by equations (4.4) to (4.10). Equation (4.4) is the structural equation of *sensory perception* (*sens*). Equations (4.5) and (4.6) are the measurement equations of *sensory perception*, where l_{int}^{ove} and l_{int}^{wtp} are latent variables explaining participants' indicators of *sensory perception*. The probability of observing each indicator with a value *s* are described in equations (4.7) and (4.8) in their conditional form, and in equations (4.9) and (4.10) in their unconditional form.

$$sens_{int} = (1 + \kappa\, I_{t=2})(F_{it}\gamma_{Fn} + \omega_n) \tag{4.4}$$

$$l_{int}^{ove} = \lambda_{ove} sens_{int} + \epsilon_{int}^{ove} \tag{4.5}$$

$$l_{int}^{wtp} = \lambda_{wtp} sens_{int} + \epsilon_{int}^{wtp} \tag{4.6}$$

$$P(ove_{int} = s|sens_{int}) = P(\tau_{s-1}^{ove} < l_{int}^{ove} < \tau_s^{ove}|sens_{int}) = \frac{1}{1 + e^{\lambda_{ove} sens_{int} - \tau_s^{ove}}} - \frac{1}{1 + e^{\lambda_{ove} sens_{int} - \tau_{s-1}^{ove}}} \tag{4.7}$$

$$P(wtp_{int} = s|sens_{int}) = P\left(\tau_{s-1}^{wtp} < l_{int}^{ove} < \tau_s^{wtp}\middle|sens_{int}\right) = \frac{1}{1 + e^{\lambda_{wtp} sens_{int} - \tau_s^{wtp}}} - \frac{1}{1 + e^{\lambda_{wtp} sens_{int} - \tau_{s-1}^{wtp}}} \tag{4.8}$$

$$P(ove_{int} = s) = \int\limits_{\omega_n,\gamma_{Fn}} P(ove_{int} = s|sens_{int})\phi(\omega_n|0,1)\phi(\gamma_{Fn}|\mu_F,\Sigma_F)d\omega_n d\gamma_{Fn} \tag{4.9}$$

$$P(wtp_{int} = s) = \int\limits_{\omega_n,\gamma_{Fn}} P(wtp_{int} = s|sens_{int})\phi(\omega_n|0,1)\phi(\gamma_{Fn}|\mu_F,\Sigma_F)d\omega_n d\gamma_{Fn} \tag{4.10}$$

where $sens_{int}$ is the value of *sensory perception* for alternative *i* in SC exercise *t*, for individual *n*; $I_{t=2}$ is a dummy variable taking the value 1 if *t*=2 (i.e. the alternative belongs to the second set of wines tasted by the participant); F_{it} is a vector of dummy variables indicating the added flavour to the alternative; ω_n is an independent identically distributed (iid) normal random error with mean 0 and standard deviation equal to 1, capturing the unobserved (by the modeller) determinants of *sensory perception* and correlating observations from the same individual (pseudo panel effect). The scalar κ and the vector γ_{Fn} are parameters to be estimated. κ is a scale factor to allow for differences in scale between the first and second group of tasted wines (Ortúzar & Willumsen 2011, Chapter 8). γ_{Fn} represents participant *n*'s preferences for the added flavours, and is assumed to follow a multivariate normal distribution on the population with the vector μ_F as mean and the diagonal matrix Σ_F as variance, both to be estimated.

l_{int}^{ove} and l_{int}^{wtp} are the latent variables of the ordered logit models explaining the observed levels of indicators *ove* (overall liking) and *wtp* (the answer to "How much would you be willing to pay for this wine"). Both variables depend only on $sens_{int}$, which is scaled by λ_{ove} and λ_{wtp}, which are parameters to be estimated. The error components ϵ_{int}^{ove} and ϵ_{int}^{wtp} are iid Extreme Value type 1, and give the probability of observing some particular level of the indicators an ordered logit form.

$P(ove_{int} = s|sens_{int})$ and $P(wtp_{int} = s|sens_{int})$ are the conditional probabilities of observing level s of indicator *ove* and *wtp*, respectively. These are conditional probabilities given a known value of $sens_{int}$. According to the definition of the ordered logit model (Train 2009, chapter 7.3), these probabilities are equal to l_{int}^{ove} and l_{int}^{wtp} belonging to an interval the limits of which ($\tau_0^{ove}, \ldots, \tau_9^{ove}$ and $\tau_0^{wtp}, \ldots, \tau_5^{wtp}$) must be estimated, except for $\tau_0^{ove}, \tau_9^{ove}, \tau_0^{wtp}, \tau_5^{wtp}$, which are fixed to $-\infty$ and $+\infty$ in the case of lower and upper bounds respectively (a restriction necessary for identification).

Finally, $P(ove_{int} = s)$ and $P(wtp_{int} = s)$ are the unconditional probabilities of observing level s of indicator *ove* and *wtp*, respectively. The integrals are necessary to integrate both random components of $sens_{int}$ (i.e. ω_{int} and γ_{Fn}). $\phi(.|\mu, \sigma^2)$

represents the normal distribution probability density function with mean μ and variance σ^2.

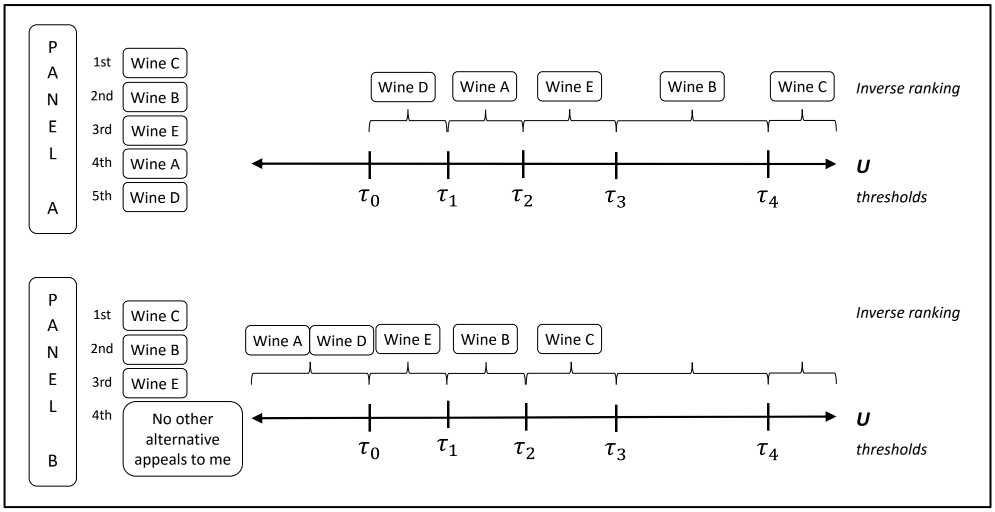


Figure 4-6 - Structure for modelling rankings made by SC respondents. Rankings constructed by consumers are on the left, and their corresponding “inverse rankings” on the right.

The grey part of Figure 4-5 models the ranking constructed by consumers in each SC exercise using an ordered logit model (Train 2009, chapter 7.4) with random coefficients (McFadden & Train 2000; Train 2009, chapter 6). Instead of working directly with the ranking provided by respondents, we used an inverse ranking: as the inverse ranking position increases, so does the preference for the alternative and its *ranking utility*. If consumers would not buy a particular alternative (maybe because they did not like it) they could leave it out of the ranking. We assigned the inverse ranking position 0 to those excluded alternatives, then position 1 to the last alternative of the ranking, position 2 to the next one from bottom-up, and so on until the top of each participant's ranking. Figure 4-6 presents two examples of inverse ranking building: one where the participant ranked all five alternatives (panel A), and another one where the participant excluded two alternatives from the ranking (panel B).

Equations (4.11) to (4.14) describe the implementation of the ranking. Equation (4.11) is the deterministic part of the *ranking utility*, while equation (4.12) presents the full *ranking utility* with its random error component. Equation (4.13) defines the conditional probability of observing alternative *i* in the inverse ranking position *r*, given that the deterministic utility is known. This probability has the traditional ordered logit form. Finally, equation (4.14) presents the unconditional form of the same probability.

$$V_{int} = L_{it}\beta_{Ln} + \beta_{sens}sens_{int} + \beta_p p_{it} \tag{4.11}$$

$$U_{int} = V_{int} + \varepsilon_{int} \tag{4.12}$$

$$P(invRnk_{int} = r|V_{int}) = P\left(\tau_{r-1}^{invRnk} < U_{int} < \tau_r^{invRnk}|V_{int}\right) = \left(\frac{1}{1 + e^{V_{int}-\tau_r^{invRnk}}} - \frac{1}{1 + e^{V_{int}-\tau_{r-1}^{invRnk}}}\right) \tag{4.13}$$

$$P(invRnk_{int} = r) = \int\limits_{\omega_n,\gamma_{Fn},\beta_{Ln}} P(invRnk_{int} = r|V_{int})\phi(\beta_{Ln}|\mu_L,\Sigma_L)\phi(\gamma_{Fn}|\mu_F,\Sigma_F)\phi(\omega_n|0,1)d\beta_{Ln}d\gamma_{Fn}d\omega_n \tag{4.14}$$

V_{int} is the deterministic part of the utility of alternative *i* in SC situation *t*, for respondent *n*; L_{it} is a vector of dummy variables indicating the label of the alternative; $sens_{int}$ is the value of *sensory perception*; p_{it} is the price of the alternative. β_{Ln}, β_{sens} and β_p are parameters representing respondents' preferences for *label*, *sensory perception* and *price*, respectively. In particular, β_{Ln} is a respondent-specific vector of coefficients with one element associated with each particular label. It is assumed to be randomly distributed on the population, following an independent multivariate Normal distribution with vector μ_L as mean, and a diagonal matrix Σ_L as variance. β_p and β_{sens} are assumed to have a single value common to all the sample; therefore, μ_L, Σ_L, β_p and β_{sens} must be estimated.

U_{int} represents the full random utility, where ε_{int} is an iid extreme value type 1 random error component, representing the extra information observable by consumers but not by the modeller, and giving the probability of observing the alternative in a given position *r* its ordered logit form.

$P(invRnk_{int} = r|V_{int})$ is the probability of individual *n* positioning alternative *i* in SC situation *t* at an inverse ranking position *r*, assuming V_{int} given. This is defined as the probability of observing $\tau_{r-1}^{invRnk} < V_{int} < \tau_r^{invRnk}$, where $\tau_{-1}^{invRnk}, \dots, \tau_5^{invRnk}$ are thresholds to be estimated, except the extremes $\tau_{-1}^{invRnk} = -\infty$ and $\tau_5^{invRnk} = +\infty$ which must be fixed to allow model identification. $P(invRnk_{int} = r)$ is the unconditional form of the same probability. The integral is necessary to remove the conditioning on V_{int} and its three random components (β_{Ln}, γ_{Fn} and ω_n). $\phi(.|\mu,\sigma^2)$ represents the normal distribution probability density function with mean μ and variance σ^2.

As all the previous integrals do not have a closed form, they were solved using Monte Carlo methods (Train 2009, chapters 9 and 10). First, a big number (*K*) of random points from $\phi(\mu_L,\Sigma_L)$, $\phi(\mu_F,\Sigma_F)$ and $\phi(0,1)$ are drawn, each called β_{Ln}^k,

γ_{Fn}^{k} and ω_{n}^{k}, respectively. Then $P(ove_{int} = s|sens_{int})$, $P(wtp_{int} = s|sens_{int})$ and $P(invRnk_{int} = r|V_{int})$ are calculated for each point k. The averages of all those evaluations are consistent estimators of the integrals.

The model was estimated using Simulated Full Information Maximum Likelihood, as integrals had to be estimated using Monte Carlo methods. Equation (4.15) shows the full likelihood function of the Traditional model.

$$P_{int}(\beta_{Ln}, \gamma_{Fn}, \omega_n) = P(ove_{int} = s|sens_{int})P(wtp_{int} = s|sens_{int})P(invRnk_{int} = r|V_{int})$$

$$L_{Traditional} = \prod_{n} \int_{\omega_n, \gamma_{Fn}, \beta_{Ln}} \prod_{t} \prod_{i} P_{int}(\beta_{Ln}, \gamma_{Fn}, \omega_n)\, \phi(\beta_{Ln}|\mu_L, \Sigma_L)\phi(\gamma_{Fn}|\mu_F, \Sigma_F)\phi(\omega_n|0,1)d\beta_{Ln}d\gamma_{Fn}d\omega_n \tag{4.15}$$

4.3.3.2 The Simplified model

The Simplified model considers both effects of price, that is, the positive effect due to its use as a cue for quality and its negative effect considering the budget constrain. Conceptually, the Simplified model is a sequential implementation of the repurchase behavioural model (Figure 4-4). Its implementation is based on linear regressions, two stage least squares (2SLS) and the MIS (Wooldridge 2010 chapter 5, Guevara & Polanco 2016), and it comprises three independent stages, each guaranteeing consistent estimators for all parameters.

Figure 4-7 shows the structure of the Simplified model. The first stage measures the impact of intrinsic attributes on consumers' liking (not on the purchase decision), allowing identification of consumers' preferred added flavour and their relative effect on liking. The second stage measures the impact of price and all extrinsic attributes on consumers' perception of quality (not on the purchase decision), allowing identification of the preferred label and measuring the effect of price on consumers' quality perception. The third stage measures the negative effect of price on the consumers' purchase decision (i.e. the effect of price due to consumers' budget constraints, allowing estimation of a price elasticity of demand.

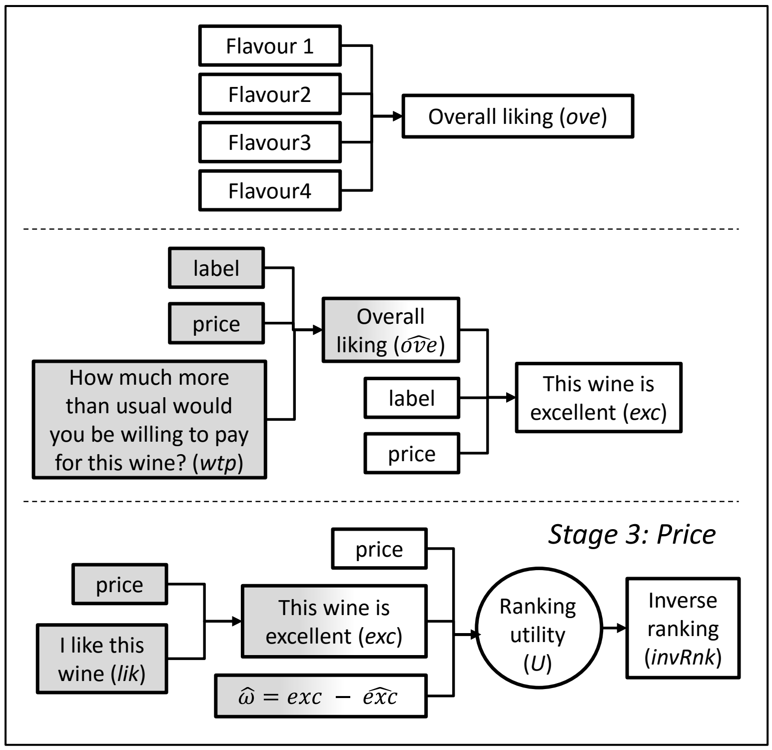


Figure 4-7 - Structure of the Simplified model

The first stage measures the impact of *added flavours* on that participant's liking of wines, and is described by the linear model in equation (4.16):

$$ove_{int} = \alpha_0 + F_{it}\alpha_F + \epsilon^1_{int} \tag{4.16}$$

ove_{int} is consumer n's level of liking of alternative i in SC exercise t; F_{it} is a vector of dummy variables indicating the added flavour of alternative i; ϵ^1_{int} is an iid normal error with mean zero and standard deviation $\sigma_{\epsilon 1}$; and α_0, α_L and $\sigma_{\epsilon 1}$ are parameters to be estimated. Unlike the Traditional and Full models, preferences for added flavours are considered to be common among respondents (i.e. α_L is a vector of fixed parameters).

Equation (13) can be regarded as a simplification of a linear MIMIC model, similar to the one used to measure *sensory perception* in the Traditional model. Assume that the structural equation of such MIMIC model is $sens_{int} = F_{it}\gamma_F + \omega_{int}$, where ω_{int} is an iid normal random error component with mean zero and unknown variance, and γ_F (the vector of taste parameters) is common across the sample. Let us further assume a single indicator (*ove*) with the linear measurement equation: $ove_{int} = \alpha_0 + \alpha_{sens} sens_{int} + \upsilon_{int}$, where υ_{int} is another iid normal error component with mean zero, variance to be estimated, and uncorrelated with ω_{int}. Replacing the structural equation in the measurement equation, we arrive at equation (13), with $\alpha_F = \alpha_{sens}\gamma_F$ and $\epsilon_{int}^1 = (\alpha_{sens}\omega_{int} + \upsilon_{int})$, which is nothing else than a traditional linear model. The estimated parameters will be consistent as long as ω_{int} and υ_{int} are uncorrelated to each other and with F_{it}. The first condition is satisfied, as any correlation between ω_{int} and υ_{int} is through F_{it}, which is included in the equation and therefore its influence is controlled for. The second condition is also met as F_{it}, like all wine attributes, is exogenous to both the reported and true level of liking (i.e. *ove* and *sens*, respectively), and therefore also to ω_{int} and υ_{int} as the wine attributes are not determined in any way by the participants. Another –probably simpler- way to see the correctness of equation (13) is to conceive *added flavour* as a proper proxy for *sensory perception*, because *added flavour* determines the level of *sensory perception* (see Guevara 2015 for a deeper discussion of this issue).

The second stage measures the effect of extrinsic attributes (*label*), *price* and *sensory perception* (*sens*) in the participants' *perceived quality* (*qual*). Only the positive effect of price is expected be captured at this stage. Conceptually, we want to estimate the following model:

$$qual_{int} = \beta_0 + L_{it}\beta_L + p_{it}\beta_p + \beta_{sens}sens_{int} + \varepsilon_{int} \tag{4.17}$$

but as *perceived quality* and *sensory perception* are not observed, they must be replaced by their indicators *exc* (level of agreement with the phrase "this wine is excellent") and *ove* (level of overall liking), respectively. Replacing *perceived quality* (the dependent variable) by its indicator *exc* is not a problem, as it is analogous to the process of replacing *sensory perception* by its indicator at the first stage. However, replacing *sensory perception* (an explanatory variable) by its indicator *ove* is problematic, as *ove* becomes an endogenous variable (see Guevara 2015). Therefore, we must use two stages least square (2SLS, Wooldridge 2010 chapter 5.2) to obtain consistent estimates. Equations (4.18) and (4.19) describe the first and second stages of 2SLS, depicted in grey and white, respectively, in Figure 4-7. Equation (4.18) is an auxiliary linear model using *sensory perception*'s second indicator *wtp* as instrument. Equation (4.19) is the main regression we are interested in, and it uses the predicted values of the first stage regression as an exogenous version of *ove*.

$$ove_{int} = \beta_0^{aux} + L_{it}\beta_L^{aux} + \beta_p^{aux} p_{it} + \beta_{wtp}^{aux} wtp_{int} + \delta_{int} \tag{4.18}$$

$$exc_{int} = \beta_0 + L_{it}\beta_L + \beta_{ove}\widehat{ove}_{int} + \beta_p p_{it} + \epsilon_{int}^2 \tag{4.19}$$

ove_{int} is participant *n*'s level of overall liking of alternative *i* in SC exercise *t*; L_{it} is a vector of dummy variables indicating the alternative's *label*; p_{it} is the *price* of the alternative; wtp_{int} is participant *n*'s answer to the question "How much more than usual would you be willing to spend on this wine"; δ_{int} is an iid normal variable with mean zero and variance σ_δ to be estimated. $\beta_0^{aux}, \beta_L^{aux}, \beta_p^{aux}, \beta_{wtp}^{aux}$ and σ_δ are parameters to be estimated, all of which are considered to have a single value for all the population (i.e. they are fixed parameters).

exc_{int} is participant *n*'s level of agreement with the phrase "This wine is excellent" concerning alternative *i* from SC exercise *t*; $\widehat{ove}_{int}$ is the predicted value of *ove* from equation (14); and ϵ_{int}^2 is an iid normal variable with mean zero and variance $\sigma_{\epsilon 2}$ to be estimated. $\beta_0, \beta_L, \beta_{ove}, \beta_p$ and $\sigma_{\epsilon 2}$ are fixed parameters to be estimated (i.e. they are not random). β_p is expected to be positive as it measures the use of price as a cue for quality. The standard errors in equation (15)'s must be corrected to take into account the estimation errors from the first stage (Wooldridge 2010, chapter 5.2.2). The role of *ove* and *wtp* in this stage are interchangeable.

The third stage measures the negative impact of price on the purchase probability, due to price being a strain in the consumers' budget constraint. The "purchase decision" is actually modelled as an inverse ranking of preferences, just as described for the Traditional model, but with a different utility function. We would like to define the deterministic part of utility as a trade-off between *price* and *perceived* *quality* (*qual*), that is: $V_{int} = \gamma_p p_{it} + \gamma_{qual} qual_{int}$. But as *qual* is not observed, we must replace it by its indicator *exc*, and this is an endogenous variable in the utility. To correct *exc*'s endogeneity we must use the Control Function approach (CF, Guevara & Polanco 2016), which requires us to estimate the model in two steps. First, we must estimate an auxiliary linear model instrumenting *exc* with at least one other indicator of *perceived quality*. For this we have two additional indicators: (i) level of agreement with the phrase "I would recommend this wine to my friends" (*rec*), and (ii) "I like this wine" (*lik*), but we only used the second one as it performed better. Second, we estimated the logit model including the estimated residuals of the auxiliary regression as an additional explanatory factor in the utility function.

Equations (4.20) to (4.24) summarize the third stage of the Simplified model. Equation (4.20) depicts the auxiliary linear model and equations (4.21) to (4.24) present the "main" ordinal logit model. Equation (4.21) presents the deterministic part of utility of the ordered logit model for the inverse ranking; equation (4.22) shows the full utility with its random error term; equation (4.23) presents the probability of observing a particular inverse ranking *r* for an alternative; and equation (4.24) presents the log-likelihood of the ordered logit model.

$$exc_{int} = \gamma_0^{aux} + \gamma_p^{aux} p_{it} + \gamma_{lik}^{aux} lik_{int} + \omega_{int} \tag{4.20}$$

$$V_{int} = \gamma_p p_{it} + \gamma_{exc} exc_{int} + \gamma_\delta \widehat{\omega}_{int} \tag{4.21}$$

$$U_{int} = V_{int} + \varepsilon_{int} \tag{4.22}$$

$$P(invRnk_{int} = r) = P\left(\tau_{r-1}^{invRnk} < U_{int} < \tau_r^{invRnk}\right) = \frac{1}{1 + e^{V_{int} - \tau_r^{invRnk}}} - \frac{1}{1 + e^{V_{int} - \tau_r^{invRnk}}} \tag{4.23}$$

$$\prod_n \prod_t \prod_i P(invRnk_{int} = r) \tag{4.24}$$

exc_{int} is participant *n*'s level of agreement with the phrase "This wine is excellent" concerning alternative *i* from SC exercise *t*; p_{it} is a scalar indicating the alternative's *price*; lik_{int} is participant *n*'s level of agreement with the phrases "I like this wine" (we left out the *rec* indicator as it performed poorly); ω_{int} is an iid normally distributed error with mean zero and variance σ_ω^2 to be estimated; U_{int} is the ordered logit's full utility; $\hat{\omega}_{int} = exc_{int} - \widehat{exc}_{int}$ is the estimated residual from the auxiliary linear regression; ε_{int} is an iid Extreme Value type 1 error component, giving the probability of observing a given ranking its ordered logit form. All γ's are parameters to be estimated, all of which are considered to be fixed (i.e. there are no random parameters). The roles of *exc* and *lik* are interchangeable within this stage.

$P(invRnk_{int} = r)$ is the conditional probability of observing alternative *i* from SC exercise *t* in position *r* of the inverse ranking for participant *n*; $\tau_{-1}^{invRnk}, \ldots, \tau_5^{invRnk}$ are thresholds to be estimated, except for $\tau_{-1}^{invRnk} = -\infty$ and $\tau_5^{invRnk} = +\infty$, which are fixed (a requirement for identification of the model). The estimated parameters of the ordered logit model, as well as their standard errors, are biased as they do not consider the estimation error of $\hat{\omega}_{int}$, therefore, they must be calculated using bootstrap (Guevara & Polanco 2016).

From a methodological perspective (Guevara & Polanco, 2016), the omitted variable indicators (*exc*, *lik*) must depend only on the omitted variable (*perceived quality*), and not on other attributes of the product. In our model, even though the indicators correlate with *price*, they depend on it only through the *perceived quality* latent variable itself, therefore avoiding any direct effect of *price* on the indicator. This assumption is aligned with the repurchase behavioural model (Figure 4-4).

4.3.3.3 The Full model

The Full model is a full implementation of the repurchase behavioural model in Figure 4-4, and is designed to capture the double effect of price. Its implementation is based on three interrelated latent variables: *sensory perception* (*sens*), *perceived quality* (*qual*) and *ranking utility* (*U*), as shown in Figure 4-8. The first represents how much the participant liked the wine in the blind tasting; *qual* represents the wine's subjective quality, which should be positively influenced by price if used as a cue for quality, and *U* represents the trade-off between *perceived quality* and *price*, capturing the negative effect of price due to the budget constraint.

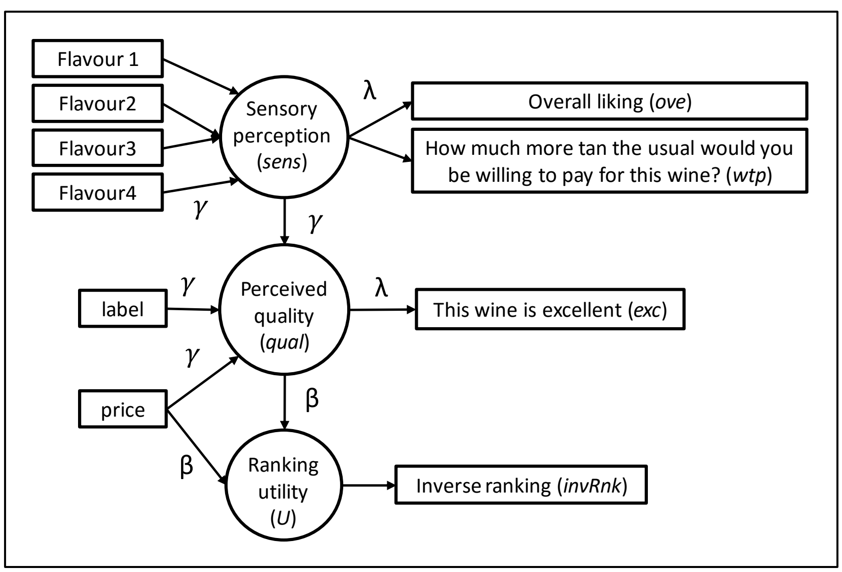


Figure 4-8 - Structure of the HC model

Sensory perception's structural and measurement equations are modelled in exactly the same way as in the Traditional model (equations 4.11 to 4.14).

Perceived quality's structural and measurement equations are depicted on equations (4.25) to (4.28). Equation (4.25) is the *perceived quality*'s structural equation, which depends on the wine extrinsic attributes (*labels*), its *sensory perception* and *price*. Only the positive effect of *price* is theorized to influence *perceived quality*. Equation (4.26) is *perceived quality*'s measurement equation, where l_{int}^{exc} is a latent variable relating *perceived quality* to its indicator *exc* (level of agreement with the phrase "I think this wine is excellent"). Equation (4.27) presents the conditional probability of observing the indicator *exc* with a value *q*. Equation (4.28) presents the same probability but in its unconditional form (i.e. integrating out all random components of qual).

$$qual_{int} = L_{it}\gamma_{Ln} + \gamma_{pA}(1 - e^{\gamma_{pB}p_{it}}) + \gamma_{sens}sens_{int} + \psi_n \tag{4.25}$$

$$l_{int}^{exc} = \lambda_{exc}qual_{int} + \epsilon_{int}^{exc} \tag{4.26}$$

$$P(exc_{int} = q|qual_{int}) = P(\tau_{q-1}^{exc} < l_{int}^{exc} < \tau_q^{exc}|qual_{int}) = \frac{1}{1 + e^{\lambda_{exc}qual_{int} - \tau_q^{exc}}} - \frac{1}{1 + e^{\lambda_{exc}qual_{int} - \tau_{q-1}^{exc}}} \tag{4.27}$$

$$P(exc_{int} = q) = \int_{\omega_{int},\gamma_{Fn},\psi_{int},\gamma_{Ln}} P(exc_{int} = q|qual_{int})\phi(\omega_{int}|0,1)\phi(\gamma_{Fn}|\mu_F,\Sigma_F)\phi(\psi_n|0,1)\phi(\gamma_{Ln}|\mu_L,\Sigma_L)d\omega_{int}d\gamma_{Fn}d\psi_{int}d\gamma_{Ln} \tag{4.28}$$

$qual_{int}$ is the value of the latent variable *perceived quality* for alternative *i* of SC exercise *t* for participant *n*; L_{it} is a vector of dummy variables indicating the *label* of the alternative; $\gamma_{pA}(1 - e^{\gamma_{pB}p_{it}})$ measures the positive effect of price, where p_{it} is the alternative's *price*; $sens_{int}$ is the level of the latent variable *sensory perception*; ψ_n is an iid normal random error with mean zero and unit variance (a restriction necessary for identification), it captures the unobserved (by the modeller) determinants of *qual* and it allows correlating observations from the same individual; γ_{Ln}, γ_{pA}, γ_{pB} and γ_{sens} are parameters to be estimated. γ_{Ln} is a vector of parameters for each individual, assumed to follow a multivariate normal distribution on the population with vector μ_L as mean and diagonal matrix Σ_L as variance. Scalars γ_{pA}, γ_{pB} and γ_{sens} are considered to be common among the sample (i.e. they are fixed parameters).

The positive effect of price is theorized to have diminishing quality improvements. We tested different functional forms to model this behaviour: dummy variables, log transformation of *price* and an asymptotic form. We kept the last one as it fitted better than the log transformation and had the benefit of being continuous, unlike the dummy formulation. The asymptotic formulation is defined by two parameters: γ_{pA} determining the maximum amount of quality that price can possibly inspire on consumers, and γ_{pB} determining how fast that amount is achieved. We believe these parameters to be highly dependent on the price niches under study.

l_{int}^{exc} is a latent variable relating *qual* to its indicator *exc*; λ_{exc} is a parameter to be estimated; if it is positive and significant, it means that *exc* is a suitable indicator of *perceived quality*; ϵ_{int}^{exc} is a iid Extreme Value type 1 random error, that gives $P(exc_{int} = q|qual_{int})$ its ordered logit form.

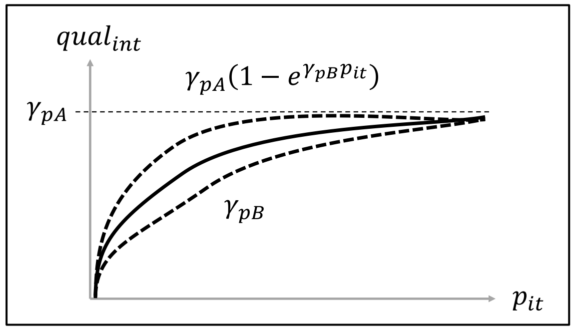


Figure 4-9 - Amount of perceived quality inspired by price. The benefits in perceived quality of higher prices are marginally decreasing.

exc$_{int}$ is the level of agreement with the phrase "I think this wine is excellent", and $P(exc_{int} = q|qual_{int})$ is the conditional probability of observing a level of agreement q with that phrase, given a known value of *qual*. We did not use the additional perceived quality indicators *rec* and *lik* as they performed poorly in the Full model. $\tau_0^{exc}, \dots, \tau_5^{exc}$ are thresholds to be estimated, except the extremes which are fixed to $-\infty$ and $+\infty$ for identification purposes. $P(exc_{int} = q)$ is the unconditional probability of observing level q of indicator *exc*, and it requires to integrate all random parameters in *qual*: $\omega_{int}, \gamma_{Fn}, \psi_n$ and γ_{Ln}.

Equations (4.29) to (4.32) present the inverse ranking implementation in the Full model. This implementation is analogous to the one in the Traditional model, except for having a different utility function (equations 4.29 and 4.30). Equation (4.31) presents the conditional probability of observing an alternative in the inverse ranking position *r*, given a known value of *qual*, while equation (4.32) presents the same probability in its unconditional form.

$$V_{int} = \beta_p p_{it} + \beta_q qual_{int} \tag{4.29}$$

$$U_{int} = V_{int} + \varepsilon_{int} \tag{4.30}$$

$$P(invRnk_{int} = r|qual_{int}) = P\left(\tau_{r-1}^{invRnk} < U_{int} < \tau_r^{invRnk}\middle|qual_{int}\right) = \frac{1}{1 + e^{V_{int} - \tau_r^{invRnk}}} - \frac{1}{1 + e^{V_{int} - \tau_{r-1}^{invRnk}}} \tag{4.31}$$

$$\begin{aligned} &P(invRnk_{int} = r) \\ &= \int_{\omega_{int}, \gamma_{Fn}, \psi_{int}, \gamma_{Ln}} P(invRnk_{int} = r|qual_{int}) \phi(\omega_{int}|0,1) \phi(\gamma_{Fn}|\mu_F, \Sigma_F) \phi(\psi_{int}|0,1) \phi(\gamma_{Ln}|\mu_L, \Sigma_L) d\omega_{int} d\gamma_{Fn} d\psi_{int} d\gamma_{Ln} \end{aligned} \tag{4.32}$$

V_{int} is the deterministic part of the *ranking utility* for alternative *i* of SC exercise *t*, as perceived by participant *n*; p_{it} is the *price* of the alternative (as a scalar); $qual_{int}$ is the value of the *perceived quality* of the alternative; and β_p and β_q are parameters to be estimated, both of them are considered fixed in the population. U_{int} is the full *ranking utility*, including its iid Extreme Value type 1 error component ε_{int}, that gives the ranking probability $P(invRnk_{int} = r|qual_{int})$ its ordered logit form.

The formulation of U_{int} represents a trade-off between price and perceived quality when making the purchase decision. While β_q is expected to be positive, β_p is expected to be negative, that is, *perceived quality* is a desirable attribute of the alternative, while *price* is an undesirable attribute of the alternative. Therefore, β_p aims to capture the negative effect of price as a strain to the budget restriction.

$P(invRnk_{int} = r|qual_{int})$ is the conditional probability of alternative *i* from SC exercise *t* being ranked on position *r* in the inverse ranking by participant *n*, given a known value of $qual_{int}$. $\tau_{-1}^{invRnk}, \ldots, \tau_5^{invRnk}$ are thresholds to be estimated, except for the extremes, which are fixed to $-\infty$ and $+\infty$ for identification purposes. As the (inverse) *ranking utility* has no random components other than those from *perceived* *quality* and its ordered logit error component, the unconditional probability only requires to integrate $\omega_{int}, \gamma_{Fn}, \psi_n$ and γ_{Ln} (the same as in equation 24), i.e. the random components in perceived quality.

Equation (4.33) presents the likelihood of the full model. We estimated the model simultaneously, using (simulated) Full Information Maximum Likelihood. As the integrals do not have a closed form, we used Monte Carlo methods to estimate them (Train 2009, chapters 9 and 10).

$$P_{int}(\beta_{Ln}, \gamma_{Fn}, \omega_n) = P(ove_{int} = s|sens_{int})P(wtp_{int} = s|sens_{int})P(exc_{int} = q|qual_{int})P(invRnk_{int} = r|qual_{int})$$

$$L_{Full} = \prod_n \int_{\omega_{int},\gamma_{Fn},\psi_{int},\gamma_{Ln}} \prod_t \prod_i P_{int}(\beta_{Ln}, \gamma_{Fn}, \omega_n)\phi(\omega_{int}|0,1)\phi(\gamma_{Fn}|\mu_F, \Sigma_F)\phi(\psi_n|0,1)\phi(\gamma_{Ln}|\mu_L, \Sigma_L)d\omega_{int}d\gamma_{Fn}d\psi_{int}d\gamma_{Ln} \tag{4.33}$$

4.4 Results

Results from each of the three estimated models are presented in this section. The Traditional, Simplified and Full models differ in the way the double effect of price is considered (i.e. its positive effect as a cue for quality and its negative effect as a strain to the budget constrain). The Traditional model ignores the double effect of price, using the traditional formulation of discrete choice models; the Simplified model estimates each stage of the repurchase behavioural model (Figure 4-4) separately; and the Full model fully implements the repurchase behavioural model and estimates it jointly.

Table 4-3 - Coefficients and goodness of fit indicators for the traditional model

(the thresholds of the ordered logit are not reported)

	Sensory model				Ranking model	
	Coeff.	**t-test**			**Coeff.**	**t-test**
Flavour 1 mean	-0.629	-1.62		Sensory perception	0.545	3.18
Flavour 1 s.d	1.700	3.79		Price	0.028	1.68
Flavour 2 mean	-0.231	-1.37		Label 1 mean	0.154	1.13
Flavour 2 s.d.	1.130	4.28		Label 1 s.d.	1.030	4.96
Flavour 3 mean	-0.634	-3.45		Label 2 mean	0.037	0.22
Flavour 3 s.d.	1.530	4.97		Label 2 s.d.	1.100	4.07
Flavour 4 mean	-1.220	-2.41		Label 3 mean	0.097	0.60
Flavour 4 s.d.	1.670	5.15		Label 3 s.d.	1.120	6.25
λ ove	1.230	3.59		Label 4 mean	-0.174	-1.05
λ wtp	1.360	3.68		Label 4 s.d.	0.985	4.31
scale factor	0.189	2.13				
Number of parameters						38
Number of observations (n° respondents)						1220 (122)
Number of draws (MLHS)						10000
Log-likelihood of the whole model						-5878.9
Log-likelihood of the ranking component						-2027.4
ρ^2 of the ranking component						0.002
Corrected ρ^2 of the ranking component						-0.006
Adjusted ρ^2 of the ranking component						-0.008

The estimated parameters and goodness of fit indicators for each model are presented in Table 4-3 to 5. The log-likelihood values of the whole models are not directly comparable, as their structures are different. However, the log-likelihood of each model ranking component can be compared directly with the others. For this reason, we report the regular, corrected and adjusted Rho squares based only on the ranking component log-likelihood values. All models were estimated using 10000 draws from the Modified Latin Hypercube sampling algorithm (Hess *et al.* 2006).

Table 4-4 - Coefficients and goodness of fit indicators for the Simplified model

(1st stage's t-test might be biased, the thresholds of the ordered logit are not reported)

		Auxiliary regression			**Main regression**	
		Coeff.	**t-test**		**Coeff.**	**t-test / C.I.**
Stage 1	(Intercept)				5.861	42.79
	Flavour 1				-0.611	-3.15
	Flavour 2				-0.279	-1.44
	Flavour 3				-0.619	-3.20
	Flavour 4				-1.455	-7.51
	R^2					0.05
	corrected R^2					0.05
Stage 2	(Intercept)	0.928	7.91		0.672	5.65
	lab1	-0.020	-0.20		0.042	0.46
	lab2	0.066	0.67		-0.045	-0.48
	lab3	-0.011	-0.11		0.065	0.70
	lab4	-0.086	-0.86		-0.046	-0.50
	pri	0.009	0.99		0.029	3.38
	s1.tongo				0.306	19.66
	s2	1.644	60.16			
	R^2		0.75			0.24
	corrected R^2		0.75			0.24
Stage 3	(Intercept)	0.419	8.01			
	pri	0.017	3.06		-0.017	[-inf , -0.002]
	exc				1.163	[1.001 , +inf]
	lik	0.752	51.69			
	residual				-0.785	[-inf , -0.547]
	Number of parameters					3
	Number of observations (n° respondents)					1220 (122)
	Number of bootstrap draws					5000
	Log-likelihood					-1912.10
	R^2 / ρ^2		0.69			0.140
	Corrected R^2 / ρ^2		0.69			0.137
	Adjusted ρ^2					0.116

The Full and Simplified models successfully separate and measure the positive and negative effects of price, as all price parameters have the expected sign and are significant. The Traditional model only captures an insignificant (p-value=0.11) positive net effect of price.

Table 4-5 - Coefficients and goodness of fit indicators for the Full model

(the thresholds of the ordered logit are not reported)

	Sensory perception				**Perceived quality**	
	Coeff.	**t-test**			**Coeff.**	**t-test**
Flavour 1 mean	-0.464	-2.21		Label 1 mean	0.109	0.66
Flavour 1 s.d	1.790	7.12		Label 1 s.d.	0.367	1.23
Flavour 2 mean	-0.226	-1.56		Label 2 mean	0.027	0.14
Flavour 2 s.d.	1.220	7.52		Label 2 s.d.	0.896	4.74
Flavour 3 mean	-0.656	-3.59		Label 3 mean	0.165	0.92
Flavour 3 s.d.	1.760	12.00		Label 3 s.d.	0.731	2.64
Flavour 4 mean	-1.380	-7.91		Label 4 mean	-0.220	-1.13
Flavour 4 s.d.	1.580	8.94		Label 4 s.d.	0.471	1.62
λ ove	1.250	8.85		Price 1	1.440	3.02
λ wtp	1.370	10.09		Price 2	-0.205	-1.70
scale factor	0.201	2.56		Sensory perception	0.861	6.64
				λ exc	0.693	7.92
	Ranking Utility					
Perceived quality	0.756	4.68				
Price	-0.041	-2.62				
Number of parameters						46
Number of observations (n° respondents)						1220 (122)
Number of draws (MLHS)						10000
Log-likelihood of the whole model						-7429.2
Log-likelihood of the ranking component						-1988.4
ρ^2 of the ranking component						0.017
Corrected ρ^2 of the ranking component						0.013
Adjusted ρ^2 of the ranking component						0.011

Preferences for added flavours and labels are fairly aligned between models. Table 4-6 presents the percentages of the population -as extrapolated from the sample- that like each added flavour and label more than the base level. These percentages are calculated disregarding t-tests, as available values are the best estimates available, though small differences should be considered with care. All models agree that -on average- there is no added flavour superior to the base wine (i.e. the wine without any artificially added flavour). Flavour 2 is the most preferred added flavour (after the base) according to all models, and Flavour 4 is the least preferred one. Preferences for labels seem to be very heterogeneous, though Labels 1 and 3 could have an advantage on more than for half of the population, while Label 4 is the least preferred.

Table 4-6 - Percentage of the population with a positive coefficient for each attribute level.

Attribute	Traditional	Simplified	Full
Flavour 1	36%	0%	40%
Flavour 2	42%	0%	43%
Flavour 3	34%	0%	35%
Flavour 4	23%	0%	19%
Label 1	56%	100%	62%
Label 2	51%	0%	51%
Label 3	53%	100%	59%
Label 4	43%	0%	32%

From a marketing perspective, it is interesting to analyse how a wine's market share changes as a function of its price. Figure 4-10 shows the effect of price on the probability of positioning a wine on the top of the ranking. As participants were asked to rank the wine they would like to buy the most in the first place, this probability is a proxy for its market share. Figure 4-10 displays a simulation of the probability of being on top of the ranking, considering a 5 US$ wine with an initial probability of 20% of being on the top of the ranking (i.e. the point where all curves intercept in the figure). Each curve shows each model's prediction of this probability when price changes to the value in the horizontal axis. This simulation considers that price is the only attribute that changes, and it was calculated based on the marginal effect of price on the probability:

$$\frac{\partial P(invRnk = 5)}{\partial p} = \frac{\partial V}{\partial p} P(invRnk = 5) \left(1 - P(invRnk = 5)\right) \tag{4.34}$$

This analysis is only an approximation to the real probabilities in the sample, as a rigorous approach would require selecting a set of levels to analyse, calculate each participant's probability for every point in the plot and then aggregate them to obtain the demand curve. Notwithstanding, it effectively shows the behaviour of each model's prediction when price changes.

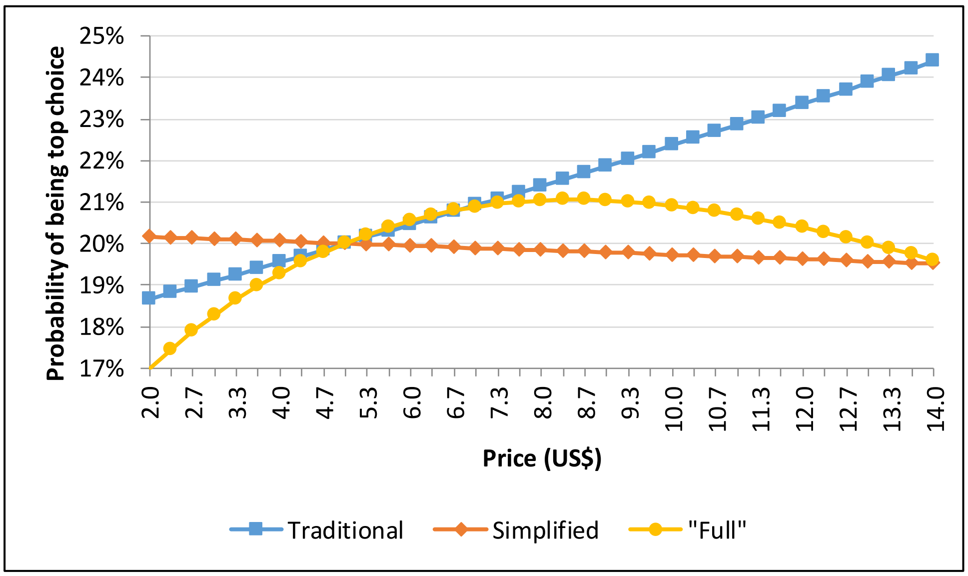


Figure 4-10 – Probability of being on top of the ranking as a function of price (ceteris paribus).

Demand predictions as price change are quite different among models: the Traditional model predicts demand to increase with price, the Simplified model predicts demand to slowly drop as price rises, and the Full model predicts an inverse U shape for demand. To build these predictions, we ignored that the Traditional model's price coefficient is not significant at 95% confidence.

4.5 Discussion

Our results show that ignoring price's double effect (i.e. as a cue for quality and as a strain to the budget constrain) when modelling demand can lead to inaccurate results, especially when dealing with products the quality of which is hard to determine. In our particular application to wine demand, ignoring price's double effect led to a model where consumers appear to be either insensitive to price or to always prefer higher prices, depending on how rigorous we are with significance levels. In any case, ignoring the effect of price leads to clearly unrealistic result.

All three models are fairly consistent on their results regarding attributes other than price: the base wine without any added flavour and the wine with added flavour n°2 are the most preferred ones, and there is no clearly superior label, at least at an average level. Despite this similarity, the Traditional and Full models have an advantage over the Simplified model, as both consider preference heterogeneity by means of random parameters. This allows estimating individual-level parameters, something impossible when fixed parameters are assumed, as in the Simplified model. This limitation, however, could be easily overcome by using more flexible (but complex) models in the first and second stage of the Simplified model, for example, full MIMIC models with random parameters. We did not do this in our application, as we wanted to provide the simplest possible alternative to estimating both effects of price separately.

When it comes to price, all models predict very different demand behaviour. As Figure 4-10 shows, the Traditional model predicts demand to increase as price rises.

This is unrealistic, as wine is not a Giffen good. Instead its upward demand when price increases is due to price's double effect, not to a lack of substitutes and low income of consumers, as in the case of Giffen goods. Furthermore, this effect is not significant with 95% confidence under a two-tailed t-test.

The Simplified model predicts a monotonically decreasing demand as price increases. The decreasing speed is determined by the interaction between price's effect in the third stage's auxiliary regression and the main ordered logit model's utility. To see this more clearly, consider price's marginal effect on the probability of placing a wine on top of the ranking, when price enters both equations with forms $f(p_{it})$ and $f^{aux}(p_{it})$ respectively (in our application, $f(p_{it}) = \gamma_p p_{it}$ and $f^{aux}(p_{it}) = \gamma_p^{aux} pri_{it}$.)

$$\begin{aligned}
\frac{\partial V_{simplified}}{\partial p} &= \frac{\partial}{\partial p}\left(f(p_{it}) + \gamma_{exc} exc_{int} + \gamma_\delta (exc_{int} - \widehat{exc}_{int})\right) \\
&= \frac{\partial f(p_{it})}{\partial p} - \gamma_\delta \frac{\partial \widehat{exc}_{int}}{\partial p} \\
&= \frac{\partial f(p_{it})}{\partial p} - \gamma_\delta \frac{\partial}{\partial p}(\gamma_0^{aux} + f(p_{it}) + \gamma_{lik}^{aux} lik_{int}) \\
&= \frac{\partial f(p_{it})}{\partial p} - \gamma_\delta \frac{\partial f^{aux}(p_{it})}{\partial p}
\end{aligned} \quad (4.35)$$

The problem is that f and f^{aux} must have the same form for the method to work, even though they may have different parameters. For example, if price is included in a linear way in the utility function of the main "regression" (in our case the ordered logit), then it must be linear too in the auxiliary regression. This is a requirement of the MIS approach, as the auxiliary regression must control for the same explanatory variables than the main regression, plus the instruments. This restriction becomes a problem because the main regression measures the negative effect of price and the auxiliary regression measures its positive effect, and both most likely do not have the same form. While the benefits of higher prices are expected to be marginally decreasing, there is no reason to believe the same for the negative effect of price. So, for example, while a log transformation of price may seem reasonable for the positive effect of price, it may not be so for the negative effect. Furthermore, this restriction makes it impossible to achieve the inverse U-shape effect of price on choice probability detected by Lockshin *et al.* (2006) for any continuous transformation of price. The only chance to reproduce that effect would be to turn price into a discrete variable (e.g. create dummies for price ranges). We tested this approach with our data, but we lost significance of the price parameters associated with the dummy variables.

The Full model successfully reproduces the inverse U-shape of demand when price changes. The initial increase of demand as price rises is driven by the positive effect of price (i.e. consumers perceive the product's quality as higher, and therefore are more willing to pay for it). However, after approximately 8.5 US$, quality ceases to increase due to price, and then the negative effect of price dominates, decreasing demand for higher prices. The exact point of maximum demand depends on the functional form of price both in the *perceived quality*'s structural equation (positive effect) and in the *rank utility* (negative effect), although the maximum seems to be fairly robust to different specifications. For example, we tested a model with a log transformation of price in the *perceived quality*'s structural equation which translated the maximum demand closer to 9 US$, representing a change of just 6% from our reported model.

The inverse U-shape demand could also be obtained introducing price in a quadratic form in the utility of the Traditional model, but with several disadvantages. The first is that estimation errors increase dramatically. We tested this approach in our dataset and obtained very low t-tests (< 0.50) for both the linear and quadratic coefficients of price in the *ranking utility*. The Full model minimizes this error because it uses different indicators for the positive and negative effects of price, that is the *perceived quality* indicators and the ranking position, respectively (*exc* and *invRnk* in our data). A second disadvantage is that the inverse U-shape is bound to be symmetric, an artificial restriction that is not necessary in the Full model. A third disadvantage is that the upward slope of the effect of price in the utility does not have a clear economic interpretation, something that does not happen in the Full model, as the direct effect of price on utility is always negative. Finally, a quadratic form for price in the utility makes estimating willingness to pay (WTP) for attributes a difficult task. As WTP would have the form $\frac{\partial V/\partial x}{\partial V/\partial p} = \frac{\partial V/\partial x}{\beta_p + 2\beta_{p^2} p_{it}}\ with\ \beta_p > 0\ and\ \beta_{p^2} < 0$, its sign changes depending on price and is undefined for the point of maximum demand of the product: $p = -\beta_p/2\beta_{p^2}$. In the Full model, instead, WTP can always be estimated using only the negative coefficient of price (as long as its form is not quadratic). The positive effect of price should not be considered for WTP calculations as in that case the measurement would stop being *ceteris paribus*, due to price –and quality- changing.

Despite the Full model's superiority, the Simplified model achieves significantly better results than the Traditional approach. Furthermore, if more flexible forms are used on each of the Simplified model stages (e.g. random parameters) its performance could be significantly boosted. So even though the Simplified model's structural limitations should not be forgotten, it seems to be an interesting tool to obtain quick and consistent results from a dataset.

The Full model appears to be the most robust and flexible approach when modelling price's double effect. And even though its implementation and estimation can be cumbersome, its benefits outweigh its complications. It allows for different functional forms for price's positive and negative effects, it increases the precision of the estimated parameters by avoiding estimation of the net effect and including additional sources of information (i.e. quality indicators), it allows for WTP estimation at any point of the demand curve, it allows the demand curve to take more flexible and realistic forms, and provides a consistent modelling framework from a behavioural, economic and statistical perspective.

None of the estimated models considers the *Veblen Effect* (Veblen 1899/1994). This effect relates to conspicuous consumption, that is, when consumers obtain utility not only from the product itself, but from the social recognition associated with the consumption of exclusive or expensive products. This effect is different to using price as a cue for quality, as in the case of the Veblen effect, utility is not due to the price itself, but due to the product's social recognition. Therefore, none of the estimated models consider the Veblen Effect. Including brands in the experiment should control for the Veblen effect to a reasonable extent, but we did not do it in our experiment as we were focusing in new products, with an unknown brand to consumers.

Further research is necessary to determine the best indicators for *perceived quality*. Even though we collected three of them (the level of agreement with the phrases “This wine is excellent”, “I would recommend this wine to my friends”, and “I like this wine”) and they were highly correlated (Cronbach Alpha of 0.93), only the first one behaved as expected. From a qualitative perspective, it is easy to see that even though the three indicators are related, they do not take into account the same factors. While the first one (“This wine is excellent”) captures only the consumer’s subjective evaluation of quality, the second one (“I would recommend this wine to my friends”) brings into play the consumer friends’ preferences. Finally, the third indicator (“I like this wine”) may have been interpreted as focusing more on the sensory attributes of wine, instead of on the product as a whole. We believe that other ways of measuring perceived quality that depend less on interpretation may render better results, such as ranking alternatives based on their quality.

In summary, our results suggest that ignoring the double effect of price as a cue for quality and as a strain in the budget constraint may lead to specification issues and higher estimation error. These problems are especially relevant in the case of products the quality of which is hard to asses by consumers. This is generally the case for new foods and beverage products (though with different intensity), but also for other product categories, such as jewellery, art and some medications. Our proposed solution consists in using quality indicators to account for subjective quality (Simplified model) or to explicitly implement the full behavioural model (Full model). Both alternatives provide good results, with the second alternative being more flexible and robust, but also more cumbersome to estimate. More research on simplifications to the Full model estimation (e.g. sequential estimation) and more flexible forms for the Simplified model are required, as both approaches could lead to easier yet informative ways of measuring price's double effects. We are also working on testing these approaches on revealed preference data, i.e. actual purchases, and using more complex intrinsic attributes.

5 CONCLUSIONS

Consumer preferences are always difficult to measure. But preferences for food and beverages present a set of particularities that set them apart: they are highly variable and may change depending on the amount of information available to consumers. Despite the awareness of these particularities on the theoretical literature (Grunert 2005), the first applications of discrete choice models to food and beverage products in the applied literature mostly ignored them.

The main objective of this thesis was to show that discrete choice models are a useful and adequate tool to understand and measure consumers' perception and preferences for food and beverage products. To achieve that, three specific objectives had to be fulfilled:

i. Consider preference heterogeneity in a way that eases extrapolation of results to the general population
ii. Separately measure the positive effect of price as a cue for quality and its negative effect as a strain to the consumer's budget restriction.
iii. Measure the effect of both intrinsic and extrinsic attributes in a way that allows comparing their relative importance

While the first objective was only partially achieved, the second and third were fully accomplished. Concerning the first objective, even though we managed to explain heterogeneity based on consumers' latent attitudes, it is hard to extrapolate these attitudes to the population level. Therefore, more research is required in finding what observable characteristics of consumers are useful in explaining their preferences.

This also relates to the first hypothesis we set to test: “The level of involvement correlates with consumers’ preferences and can be explained, to a reasonable extent, by consumers’ socio-demographic characteristics and consuming habits.” We found this hypothesis to be only partially true, as level of involvement correlates mainly with consumers’ purchase and drinking habits, but only weakly with their socio-demographic characteristics (Chapter 2).

The second objective was achieved by harmonizing the discrete choice model structure with the structure of the behavioural model. By explicitly modelling consumers’ perceived quality, we were able to capture the positive effect of price. And by explaining choice as a trade-off between quality and price we could measure the negative effect of price (Chapters 3 and 4). We also offered a simpler –but less flexible- way to measure both effects of price using the MIS approach to endogeneity correction (Chapter 4).

This relates to the second hypothesis: “Consumers use price as a cue for quality, leading to price having a double effect in utility. Measuring only the net effect of price can lead to inconsistent estimates.”. We proved true the first part of this hypothesis in Chapters 3 and 4, where both effects of price were found to be significant. We discussed the potential endogeneity of price and showed how to solve it in Chapter 3.

The third objective was achieved once again by harmonizing the discrete choice and the behavioural model structures. We assumed that the effect of the intrinsic attributes was mediated by consumers’ sensory likings (or sensory perceptions) of each wine. Then, the interaction of sensory liking and extrinsic attributes determine the wine’s experienced (or perceived) quality. This modelling approach allowed us to compare the effect of intrinsic and extrinsic attributes almost directly under a single joint modelling approach (Chapter 4).

The approach used to achieve the third objective also allowed us to prove the third hypothesis: “The impact of intrinsic attributes on the consumer’s choice is mediated by their liking of the product.”

A more detailed discussion of each objective is presented in the next three sections. Section 5.1.4 summarizes this thesis’ main contributions and section 5.1.5 closes the chapter with future research proposals.

5.1.1 About preference heterogeneity

Across the three experiments we tested different ways of considering preference heterogeneity. The first implemented latent variables measuring consumers’ attitudes towards wine, more specifically, measuring consumers’ level of involvement with wine at an individual and social level. Agreeing with results from other researches, our latent variables proved to be useful explaining preference heterogeneity.

However, it proved hard to relate consumers’ attitudes to their observable socio-demographic characteristics. Instead, their attitudes were more correlated with their consuming behaviour. This poses a difficulty from a marketing perspective, as consuming habits are not as widely available for the population as their socio-demographic characteristics.

We also found a limited performance by our original instrument to measure attitudes. We recommend –instead– using well stablished and tested instruments, such as those proposed by Brunner & Siegrist (2011) or Ogbeide & Bruwer (2013). The second one is particularly interesting, as it is only 13 items long. However, the problem of relating these attitudes to observable –and available– consumers' characteristics is not solved.

In later experiments we used random coefficients to capture preference heterogeneity. We used random coefficients for both extrinsic and intrinsic attributes, except for price. Using random coefficients for price makes it harder to estimate willingness to pay and price elasticity. Furthermore, from a theoretical perspective we believe price sensitivity –on average– should be more dependent on income than on other non-observable characteristics of consumers.

While it is true than random coefficients do not explain preference heterogeneity in the detailed way than latent variables do, they do provide a broad perspective on how preferences distribute in the population. This allows, for example, to estimate the percentage of consumers who like a particular grape variety (chapter 3) or added flavour (chapter 4).

Using random coefficients to capture preference heterogeneity also offers the possibility to estimate individual level parameters. This was not done in any paper, as their objectives were not to draw specific marketing recommendations. But this technique could provide a much more precise perspective on consumers' preferences and, if the sample is representative enough, it would be possible to extrapolate results to the population level.

A third possible way to model preference heterogeneity is using latent classes. This consists in defining *a priori* a number of different classes with different unknown preferences, and letting the model find them. In this modelling approach, each respondent has a probability of belonging to each class, determined by a classification function that can depend on consumers' observable characteristics. This approach has two main problems, though. First, it is hard to determine the number of classes, as the model's fit always improves as the number of classes increases. Secondly, it is not clear what explanatory variables should be used in the classification function, leading to the same problem of attitudinal variables: observable consumers' characteristics such as their socio-demographic characteristics do not seem to correlate with their tastes.

More research is necessary to determine the most appropriate way to explain preference heterogeneity. Socio-demographic characteristics do not seem to be helpful in explaining consumers' choices, and while attitudes are, they do not correlate with any of the consumers' observable characteristics either. Until more informative results are obtained, using random coefficients and a posterior estimation of individual level parameters could provide the most flexible approach to measuring preference heterogeneity (though not explain it).

5.1.2 About the double effect of price

The double effect of price, as a cue for quality and as a strain to the budget constraint, is well known in the literature but has not being consistently incorporated in discrete choice models. Several authors have reported and controlled for the double effect of price in wine choice models using quadratic forms for price in the utility function. This approach requires the linear effect of price to have a positive coefficient and the quadratic effect to have a negative coefficient, therefore describing a concave parabola (i.e. an inverse U-shape).

While the use of a quadratic form of price in the utility may be useful for prediction, it poses at least three main limitations. First, using a quadratic form for price in the utility is not economically consistent, and it complicates the interpretation and estimation of willingness to pay (WTP). WTP becomes a function of price (as described in chapter 4), making its interpretation at best difficult for the increasing part of the price's parabola, and its calculation impossible for its maximum. Second, there is no guarantee that the effect of price is symmetrical in its positive and negative dimensions, and the quadratic form implies symmetry. This could bias predictions.

Lastly, the quadratic form of price in the utility may render estimated parameters inconsistent due to endogeneity if at least one of its effects is not quadratic. As discussed in chapter 4, the double effect of price can render price endogenous in the choice utility function if its effects' forms are poorly defined.

We propose an economical and behaviourally consistent way for considering the double effect of price. Modelling *expected quality* as a latent variable and capturing the effect of extrinsic attributes in it allows isolating the positive effect of price. Then, choice is posed as a trade-off between *perceived quality* and price, capturing the negative effect of price. This allows modelling each effect separately, avoiding endogeneity problems and increasing the estimation procedure efficiency by including additional indicators for *expected quality*.

The proposed methodology worked well in both the Chinese (Chapter 3) and Chilean (Chapter 4) datasets, and using only extrinsic (Chapter 3) and a mix of extrinsic and intrinsic attributes (Chapter 4).

The proposed approach is based on the latent variable approach for endogeneity correction (Guevara 2015) and is applicable not only to food and beverages, but to any product the quality of which is hard to determine by consumers. Examples of such products are jewellery, medication and art.

5.1.3 About measuring the effect of intrinsic and extrinsic attributes simultaneously

Few articles in the literature study the joint effect of extrinsic and intrinsic attributes on food and beverages choices. The main difficulty is associated with two factors: perception of intrinsic attributes may vary greatly among consumers, and it is hard to associate liking ratings (such as the traditional hedonic Likert scales) to purchase decisions.

We propose a method based on the use of latent variables providing an easy and straightforward way to link hedonic liking and purchase decision. We use a latent variable representing consumers' *sensory perception*, with hedonic ratings as indicators and intrinsic attributes as explanatory variables in the structural equations. This allows for separating the modelling of sensory perception from *expected quality*, as the behavioural model proposes. This is relevant, as according to the behavioural model the influence of the intrinsic attributes on the re-purchase decision is mediated by *experienced quality*, and do not influence the repurchase decision directly. In other words, it is *how much* the consumer liked the wine that influences the re-purchase, not its particular flavour or aroma. This structure led us to an important corollary: intrinsic attributes should probably not interact directly with extrinsic attributes, but *experienced quality* should. This, of course, could change depending on the particular product under study.

We tested our approach using a set of simple intrinsic attributes: added flavours. The methodology needs to be tested under more demanding conditions. The most relevant situation is where the product's sensory profiles are not determined by a single variable, as in paper III, but through a set of them. That is the case of products described by a sensory profile, such as those provided by trained sensory panels.

5.1.4 Main contributions of this thesis

There are three main contributions of this thesis to the current state of the art.

1. We proposed a novel and useful way to measure and control for the double effect of price. Price's double effect has been widely recognized in literature, but it has not been considered in choice models in a consistent and efficient way. Our proposal is consistent with a generally accepted behavioural model, is correct from both an economics and statistics perspective, and does not increase respondents' burden significantly.

2. We proposed a novel way to jointly consider the effect of intrinsic and extrinsic attributes in consumers' purchase decisions. Our approach is rooted on a generally accepted behavioural model, is correct from a statistical perspective and it easily allows measuring the effect of both types of

attributes under realistic conditions. Also, it does not require any extra data that is not already collected in most preference studies with tasting.

3. We confirmed that attitudes are useful explaining preference heterogeneity, but we also confirmed their lack of correlation with consumers' observable socio-demographic characteristics.

In general, we proposed harmonizing the structure of choice and behavioural models, which not only allows reproducing consumers' choices or measuring their preferences, but leads to a better understanding of their decision processes. This understanding, in turn, makes extrapolation of results easier and potentially broader. It is not enough to simply predict choices in the immediate future, but to be able to recognize trends and model their impact –hopefully– in the medium term.

5.1.5 Future research

Many areas are still in need of development to allow for more reliable predictions and measures of preferences using discrete choice models.

Probably the most urgent next step for research is validating results with revealed preference data, i.e. actual purchases. All the results presented in this thesis were obtained using hypothetical choices under non-incentive compatible experiments. But we also performed an incentive-compatible experiment where consumers could actually buy the wine they selected. Partial results were presented by Benjamín Lizana in his Master Thesis (Lizana 2016). They indicate a strong difference on price sensitivity between occasions when consumers believe their answers will not have real implications, and when they are facing actual purchase situations. Despite this, the double effect of price was confirmed in both the (supposedly) hypothetical decision and the actual purchase.

The inclusion of more complex intrinsic attributes in our modelling approach is also a pending task to be completed using the data collected at our larger incentive-compatible study.

More research is needed also in identifying the best method to incorporate preference heterogeneity. More particularly, it would be helpful to identify under what circumstances the latent classes approach and the random coefficients approach with individual parameters perform better.

Even though the explicit modelling of price's double effect has proven to be useful, it is still not clear what the best way to collect information on a product's perceived quality is. We measured quality (as perceived by consumers) using consumers' levels of agreement with several different phrases, but only one performed as expected ("I believe this wine is excellent"). Other possible quality indicators should be tested, such as a ranking of quality, selecting the alternative with highest quality, or maybe even using consumers' online ratings of wines. A deeper study on how confident the consumer is in his appreciation of perceived quality might also be of interest. This could be modelled using heteroskedastic MIMIC models (Bollen 1989).

Finally, choice strategies other than random utility should be tested, as there have been claims that consumers incur in risk reducing strategies when choosing wine (Johnson & Bruwer 2004). In that sense, random regret minimization could be an interesting strategy to test.

6 REFERENCES

Adamowicz, W. and Swait, J.D. (2013). Are food choices really habitual? Integrating habits, variety-seeking, and compensatory choice in a utility maximizing framework. *American Journal of Agricultural Economics* **95**, 17 - 41.

Angulo, A. and Gil, J. (2007). Risk perception and consumer willingness to pay for certified beef in Spain. *Food Quality and Preference* **18**, 1106 – 1117.

Angulo, M., Gil, M., Gracia, A. and Sánchez, M. (2000). Hedonic prices for Spanish red quality wine. *British Food Journal* **102**, 481-493.

Appleby, C., Costanigro, M., Thilmany, D. and Menke, S. (2012). Measuring consumer willingness to pay for low-sulfite wine: a conjoint analysis. Working Paper No. 117, American Association of Wine Economists.

Aqueveque, C. (2006) Extrinsic cues and perceived risk: the influence of consumption situation. *Journal of Consumer Marketing* **23**, 237 – 247.

Aqueveque, C. (2008) The effect of extrinsic cues on imported wine evaluations: an experimental approach. *Journal of Food Products Marketing* **14**, 33-47.

Ares, G., Varela, P., Rado, G. and Giménez, A. (2011). Identifying ideal products using different consumer profiling methodologies. Comparison with external preference mapping. *Food Quality and Preference* **22**, 581 – 591.

Atkin, T. and Thach, L. (2012) Millennial wine consumers: risk perception and information search. *Wine Economics and Policy* **1**, 54 – 62.

Bagwell, L.S. and Bernheim, D. (1996) Veblen effects in a theory of conspicuous consumption. *The American Economic Review* **86**, 349 – 373.

Bahamonde-Birke, F.J., Kunert, U., Link, H. and Ortúzar, J. de D. (2015) About attitudes and perceptions – finding the proper way to consider latent variables in discrete choice models. *Transportation*, 1-19. doi: 10.1007/s11116-015-9663-5.

Barreiro-Hurlé, J., Colombo, S. and Cantos-Villar, E. (2008). Is there a market for functional wines? Consumer preferences and willingness to pay for resveratrol-enriched red wine. *Food Quality and Preference* **19**, 360 – 371.

Barreiro-Hurlé, J., Gracia, A. and de-Magistris, T. (2010). Does nutrition information on food products lead to healthier food choices? *Food Policy* **35**, 221 – 229.

Bastian, S., Collins, C. and Johnson, T. (2010). Understanding consumer preferences for Shiraz wine and Cheddar cheese pairings. *Food Quality and Preference* **21**, 668 – 678.

Berry, S., Levinsohn, J. and Pakes, A. (1995) Automobile prices in market equilibrium. *Econometrica* **63**, 841 – 890.

Bhat, C.R. and Pinjari, A. (2014). Multiple discrete-continuous choice models: a reflective analysis and a prospective view. In S. Hess and A. Daly (eds.), *Handbook of Choice Modelling*, Edward Elgar, Cheltenham.

Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland.

Bolduc, D. and Alvarez-Daziano, R. (2010). On estimation of hybrid choice models. In S. Hess and A. Daly (eds,), *Choice Modelling: The State-of-the-Art and the State-of-Practice.* Emerald Group Publishing, Bingley.

Bollen, K.A. (1989). *Structural Equations with Latent Variables*. John Wiley and Sons, Chichester.

Brunner, T. and Siegrist, M. (2011) A consumer-oriented segmentation study in the Swiss wine market. *British Food Journal* **113**, 353 – 373.

Castro, M., Bhat, C.R., Pendyala, R.M. and Jara-Díaz, S.R. (2012). Accommodating multiple constraints in the multiple discrete-continuous extreme value (MDCEV) choice model. *Transportation Research* **46B**, 729 – 743.

Caussade, S., Ortúzar, J. de D., Rizzi, L.I. and Hensher, D.A. (2005) Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research* **39B**, 621-640.

Caves, R.E. and Greene, D.P. (1996) Brand's quality levels, prices, and advertising outlays: empirical evidence on signals and information costs. *International Journal of Industrial Organization* **14**, 29 – 52.

Cerda, A., Torres, M.J. and García, L. (2010). Preferencias y disposición a pagar por vinos ecológicos de parte de los consumidores de la Región del Maule, Chile. *Panorama Socioeconómico* **40**, 60 – 71 (in Spanish).

Chamberlin, E.H. (1933). *The Theory of Monopolistic Competition*. Harvard University Press, Cambridge, MA.

Chaney, I. (2000) External search effort for wine. *International Journal of Wine Marketing* **12**, 5 – 21.

Chapman, R.G. and Staelin, R. (1982) Exploiting rank ordered choice set data within the stochastic utility model. *Journal of Marketing Research* **19**, 288 – 301.

Charters, S. and Pettigrew, S. (2003). I like it but how do I know if it's any good? Quality and preference in wine consumption. *Journal of Research for Consumers*, **5** 1021-1027.

Charters, S. and Pettigrew, S. (2006). Product involvement and the evaluation of wine quality. *Qualitative Market Research* **9**(2), 181-193.

Cohen, E. and d'Hauteville, F. (2008). A cross-cultural comparison of choice criteria for wine in restaurants. *4th International Conference of the Academy of Wine Business Research*, Siena.

Combris, P., Bazoche, P., Giraud-Héraud, E. and Issanchou, S. (2009). Food choices: what do we learn from combining sensory and economic experiments? *Food Quality and Preference* **20**, 550 – 557.

Cortez, P., Cerdeira, A., Almedida, F., Matos, T. and Reis, J. (2009). Modelling wine preferences by data mining from physicochemical properties. *Decision Support Systems* **47**, 547-553.

Costanigro, M., Appleby, C. and Menke, S.D. (2014). The wine headache: consumer perceptions of sulphites and willingness to pay for non-sulphite wines. *Food Quality and Preference* **31**, 81 – 89.

Cuellar, S.S. and Huffman, R. (2008) Estimating the demand for wine using instrumental variable techniques. *Journal of Wine Economics* **3**, 172 – 184.

D'Alessandro, S. and Pecotich, A. (2013) Evaluation of wine by expert and novice consumers in the presence of variations in quality, brand and country of origin cues. *Food Quality and Preference* **28**, 287 – 303.

Daly, A. and Hess, S. (2010) Simple approaches for random utility modelling with panel data. In *European Transport Conference*, October 2010.

Deeter-Schmelz, D.R., Moore, J.N. and Goebel, D.J. (2000) Prestige clothing shopping by consumers: a confirmatory assessment and refinement of the Precon scale with managerial implications. *Journal of Marketing Theory and Practice* **8**, 43 – 58.

Dodds, W.B., Monroe, K.B. and Dhruv, G. (1991) Effects of price, brand, and store information on buyers' product evaluation. *Journal of Marketing Research* **28**, 307 – 319.

Dubow, J.S. (1992). Occasion-based vs. user-based benefit segmentation: a case study. *Journal of Advertising Research* **25**, 72 – 82.

Enneking, U., Neumann, C. and Henneberg, S. (2007) How important intrinsic and extrinsic product attributes affect purchase decision. *Food Quality and Preference* **18**, 133 – 138.

Ferreira, V., Escudero, A., Campo, E. and Cacho, J. (2008). The chemical foundations of wine aroma: a role game aiming at wine quality, personality and varietal expression. In *Proceedings of the Thirteenth Australian Wine Industry Technical Conference,* Adelaide.

Gokcekus, O. and Nottebaum, D. (2011) The buyer's dilemma – whose rating should a wine drinker pay attention to? Working Paper N° 91, American Association of Wine Economists.

Goodman, S. (2009). An international comparison of retail consumer wine choice. *International Journal of Wine Business Research* **21**, 41 – 49.

Gracia, A. and de 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. and Hensher, D.A. (2010). *Modelling Ordered Choices: A Primer*. Cambridge University Press, Cambridge.

Grisolía, J.M., López, F. and Ortúzar, J. de D. (2012) Sea Urchin: from plague to market opportunity. *Food Quality and Preference* **25**, 46-56.

Grunert, K. (2005). Food quality and safety: consumer perception and demand. *European Review of Agricultural Economics* **32**, 369–391.

Guadagni, P. and Little, J. (1983). A logit model of brand choice calibrated on scanner data. *Marketing Science* **2**, 203 – 236.

Guevara, C. A. and Polanco, D. (2016). Correcting for endogeneity due to omitted attributes in discrete-choice models: the multiple indicator solution (MIS). *Transportmetrica A* (in press).

Guevara, C.A. (2015) Critical assessment of five methods to correct for endogeneity in discrete-choice models. *Transportation Research* **82A**, 240 – 254.

Guevara, C.A. and Ben-Akiva, M. (2006) Endogeneity in residential location choice models. *Transportation Research Record* **1977**, 60 – 66.

Guevara, C.A. and Ben-Akiva, M. (2012) Change of scale and forecasting with the control-function method in logit models. *Transportation Science* **46**, 425 – 437.

Guinard, J.X., Uotani, B. and Schlich, P. (2001). Internal and external mapping of preferences for commercial lager beers: comparison of hedonic ratings by consumers blind versus with knowledge of brand and price. *Food Quality and Preference* **12**, 243 – 255.

Hall, E.J. (2003). *The Influence of Occasion on Consumer Choice: An Occasion Based, Value Oriented Investigation of Wine Purchase, Using Means-End Analysis*. Ph.D. Thesis, School of Agriculture and Wine, University of Adelaide.

Henningsen, Arne and Toomet, Ott (2011). maxLik: A package for maximum likelihood estimation in R. Computational Statistics 26(3), 443-458. DOI 10.1007/s00180-010-0217-1.

Hess, S., Train, K. and Polak, J. (2006) On the use of a modified Latin hypercube sampling (MLHS) method in the estimation of a mixed logit model for vehicle choice. *Transportation Research* **40B**, 147 – 163.

Hodgson, R.T. (2009) How expert are “expert” wine judges. *Journal of Wine Economics* **4**, 233 – 241.

Hopfer, H. and Heymann, H. (2014) Judging wine quality: do we need experts, consumers or trained panelists? *Food Quality and Preference* **32**, 221 – 233.

Hopper, K., Mai, R., Zahn, S., Hoffmann, S. and Rohm, H. (2012) Integrating sensory evaluation in adaptive conjoint analysis to elaborate the conflicting influence of intrinsic and extrinsic attributes on food choice. *Appetite* **59**, 949 – 955.

Horowitz, J. (1983) Statistical comparison of non-nested probabilistic choice models. *Transportation Science* **17**, 319 – 350.

Jaeger, S. and Rose, J.M. (2008). Stated choice experimentation, contextual influences and food choice: a case study. *Food Quality and Preference* **19**, 539–564.

Jarvis, W., Mueller, S. and Chiong, K. (2010). A latent analysis of images and words in wine choice. *Australasian Marketing Journal* **18**, 138 – 144.

Jiménez, F., Marshall, B., Ortega, J. and Foster, W. (2006). Factores que intervienen en la frecuencia de consumo de vino en el sector oriente de Santiago, Chile. *Economía Agraria* **10**, 37-52 (in Spanish).

Johnson, T. and Bruwer, J. (2004) Generic consumer risk-reduction strategies (RRS) in wine-related lifestyle segments of the Australian wine market. *International Journal of Wine Marketing* **16**, 5 – 35.

King, E., Kievit, R., Curtin, C., Swieger, J., Pretorius, I., Bastian, S. and Francis, L. (2010). The effect of multiple yeasts co-inoculations on Sauvignon Blanc wine aroma composition, sensory properties and consumer preference. *Food Chemistry* **122**, 618 – 626.

Kotler, P. and Keller, K.L. (2012). *Marketing Management*. Prentice Hall, Boston, MA.

Krystallis, A., Linardakis, M. and Mamalis, S. (2010). Usefulness of the discrete choice methodology for marketing decision-making in new product development: an example from the European functional foods market. *Agribusiness* **26**, 100 – 121.

Lattey, K.A., Bramley, B.R. and Francis, I.L. (2009) Consumer acceptability, sensory properties and expert quality judgments of Australian Cabernet Sauvignon and Shiraz wines. *Australian Journal of Grape and Wine Research* **16**, 189 – 202.

Lawless, H.T. (1980). A comparison of different methods for assessing sensitivity to the taste of phenylthiocarbamide PTC. *Chemical Senses* **5**, 247 – 256.

Lawless, H.T. (1984) Flavor description of white wine by "expert" and non-expert wine consumers. *Journal of Food Science* **49**, 120 – 123.

Leavitt, H.J. (1954) A note on some experimental findings about the meaning of price. *The Journal of Business* **27**, 205 – 210.

Lewis, G. and Zalan, T. (2014). Strategic implications of the relationship between price and willingness to pay: evidence from wine-tasting experiment. *Journal of Wine Economics* **9**, 115-134.

Lichtenstein, D.R., Ridgway, N.M. and Netemeyer, R.G. (1993) Price perceptions and consumer shopping behavior: a field study. *Journal of Marketing Research* **30**, 234 – 245.

Lizana, B. (2016). *Modelación de Preferencias de Consumidores de Vino en el Mercado Chileno: Análisis de Preferencias Reveladas.* Tesis de Magister en Ciencias de la Ingeniería, Pontificia Universidad Católica de Chile (in Spanish).

Lockshin, L. and Corsi, A.M. (2012). Consumer behaviour for wine 2.0: a review since 2003 and future directions. *Wine Economics and Policy* **1**, 2 – 23.

Lockshin, L. and Hall, J. (2003) Consumer purchasing behaviour for wine: what we know and where we are going. *International Wine Marketing Colloquium*, Adelaide.

Lockshin, L., Jarvis, W., d'Hauteville, F. and Perrouty, J.P. (2006). Using simulations from discrete choice experiments to measure consumer sensitivity to brand, region, price, and awards in wine choice. *Food Quality and Preference* **17**, 166 – 178.

Lockshin, L., Spawton, A. and Macintosh, G. (1997). Using product, brand and purchasing involvement for retail segmentation. *Journal of Retailing and Consumer Services* **4**, 171-183.

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: analysis and applications*. Cambridge: Cambridge University Press.

Louviere, J.J. and Islam, T. (2008). A comparison of importance weights and willingness-to-pay measures derived from choice-based conjoint, constant sum scales and best-worst scaling. *Journal of Business Research* **61**, 903-911.

Martínez-Carrasco, L., Brugarolas, M., Del Campo, F.J. and Martínez, A. (2006). Influence of purchase place and consumption frequency over quality wine preferences. *Food Quality and Preference* **17**, 315–327.

Mastrobuoni, G., Peracchi, F. and Tetenov A. (2014) Price as a signal of product quality: some experimental evidence. *Journal of Wine Economics* **9**, 135 – 152.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (ed.), *Frontiers in Econometrics*. Academic Press, New York.

McFadden, D. and Train, K.E. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics* **15**, 447-440.

Michis, A.A. and Markidou, A.G. (2013) Determinants of retail wine prices: evidence from Cyprus. *Empirical Economics* **45**, 267 – 280.

Milgrom, P. and Roberts, J. (1986) Price and advertising signals of product quality. *Journal of Political Economy* **94**, 796 – 821.

Ministerio de Desarrollo Social (2012). Límites mínimos y máximos del ingreso autónomo per-cápita del hogar (Encuesta CASEN 2011). Observatorio Social. Retrieved March 22, 2013, from: http://observatorio.ministeriodesarrollosocial.gob.cl/layout/doc/casen/ingresos2011cuadro3.xlsx (in Spanish).

Ministerio de Desarrollo Social (2014) CASEN 2013: Ingresos: Cuadro 10: Promedio, mínimoc y ,áximos de los deciles de ingreso autónomo per cápita del hogar. Retrieved on May 30th 2016. Available at: http://observatorio.ministeriodesarrollosocial.gob.cl/casen/casen-documentos.php?c=85&m=1&a=2013

Mora, M., Magner, N. and Silva, R. (2010). Segmentación de mercado de acuerdo a estilos de vida de consumidores de vino orgánico Región Metropolitana de Chile. *IDESIA* **28**, 25-33 (in Spanish).

Moskowitz, H. and Silcher, M. (2006). The applications of conjoint analysis and their possible uses in Sensometrics. *Food Quality and Preference* **17**, 146 – 165.

Mtimet, N. and Albisu, L.M. (2006). Spanish wine consumer behaviour: a choice experiment approach. *Agribusiness* **22**, 343 – 362.

Mueller, S., Lockshin, L. and Louviere, J.J. (2010a). What you see may not be what you get: asking consumers what matters may not reflect what they choose. *Marketing Letters* **21**, 335-350.

Mueller, S., Lockshin, L., Saltman, Y. and Blanford, J. (2010b). Message on a bottle: the relative influence of wine back label information on wine choice. *Food Quality and Preference* **21**, 22-32.

Mueller, S., Osidacz, P., Francis, I.L. and Lockshin, L. (2010c). Combining discrete choice and informed sensory testing in a two-stage process: can it predict wine market share? *Food Quality and Preference* **21**, 741-754.

Nelson, P. (1970) Information and consumer behaviour. *Journal of Political Economy* **78**, 311 – 329.

O'Neill, V., Hess, S. and Campbell, D. (2013). A question of taste: recognising the role of latent preferences and attitudes in analysing food choices. *Food Quality and Preference* **32**, 299 – 310.

ODEPA (2012). Boletín de vinos y pisco: producción, precios y comercio exterior. Oficina de Estudios y Políticas Agrarias. Retrieved March 24, 2013, from: odepaweb/servicios-informacion/Boletines/BVinos_0113.pdf (in Spanish).

ODEPA (2016). Exportaciones de vino con denominación de origen (total) al mundo: Código SACH 22042110. Retrieved on May 23th 2016. Available on: http://www.odepa.cl/series-anuales-por-producto-de-exportaciones-importaciones/ (in Spanish).

Ogbeide, O.A. and Bruwer, J. (2013). Enduring involvement with wine: predictive model and measurement. *Journal of Wine Research* **24**, 210-226.

Orth, U. and Malkewitz, K. (2008) Holistic package design and consumer brand impressions. *Journal of Marketing* **72**, 64-81.

Ortúzar, J. de D. (2010). Estimating individual preferences with flexible discrete-choice-models. *Food Quality and Preference* **21**, 262–269.

Ortúzar, J. de D. and Willumsen, L.G. (2011). *Modelling Transport*. John Wiley and Sons, Chichester.

Pages, J. (2005) Collection and analysis of perceived product inter-distances using multiple factor analysis Application to the study of 10 white wines from the Loire Valley. *Food Quality and Preference* **16**, 642 – 649.

Palma, D., Ortúzar, J. de D., Casaubon, G., Rizzi, L.I. and Agosin, E. (2013). Measuring consumer preferences using hybrid discrete choice models. *Working Paper N° 137*, American Association of Wine Economists, New York.

Palma, D., Ortúzar, J. de D., Rizzi, L.I., Guevara, C.A., Casaubon, G. and Ma, H. (2016). Modelling choice when price is a cue for quality. *Journal of Choice Modelling* (in press).

Petrin, A. and Train, K. (2010) A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research* **47**, 3 – 13.

Plassmann, H., O'Doherty, J., Shiv, B. and Rangel, A. (2008) Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1050 – 1054.

Quester, P. and Smart, J. (1998) The influence of consumption situation and product involvement over consumers' use of product attribute. *Journal of Consumer Marketing* **15**, 220 – 238.

R Core Team (2014) R: a language and environment for statistical computing [computer software]. R Foundation for Statistical Computing. Available from http://www.R-project.org.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rao, A.R. and Monroe, K.B. (1989) The effect of price, brand name, and store name on buyers' perceptions of product quality: an integrative review. *Journal of Marketing Research* **26**, 351 – 357.

Raveau, S., Alvarez-Daziano, R., Yáñez, M.F., Bolduc, D. and Ortúzar, J. de D. (2010). Sequential and simultaneous estimation of hybrid discrete choice models: some new findings. *Transportation Research Record* **2156**, 131-139.

Remaud, H., Mueller, S., Chvyl, P. and Lockshin, L. (2008, July) Do Australian wine consumers value organic wine? *4th International Conference of the Academy of Wine Business Research*, Siena.

Revelt, D. and Train, K. (1998) Mixed logit with repeated choices: household's choices of appliance efficiency level. *The Review of Economics and Statistics* **80**, 647 – 657.

Richard, T.J. and Padilla, L. (2009) Promotion and fast food demand. *American Journal of Agricultural Economics* **91**, 168 – 183.

Rose, J.M. and Bliemer, M.C.J. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews* **29**, 1-31.

Rose, J.M., Bliemer, M.C.J., Hensher, D.A. and Collins, A. (2008). Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research* **42B**, 395-406.

Rosseel, Y. (2012). Lavaan: an R package for structural equation modelling. *Journal of Statistical Software* **48**, 1-36.

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.

Schnettler, B. and Rivera, A. (2003). Características del proceso de decisión de compra de vino en la IX Región de la Araucanía, Chile. *Ciencia e Investigación Agraria* **30**, 1-14 (in Spanish).

Scitovsky, T. (1945) Some consequences of the habit of judging quality by price. *The Review of Economic Studies* **12**, 100 – 105.

Siegrist, M. and Cousin, M-E. (2009) Expectations influence sensory experience in a wine tasting. *Appetite* **52**, 762 – 765.

Sillano, M. and Ortúzar, J. de D. (2005). Willingness-to-pay estimation with mixed logit models: some new evidence. *Environment and Planning* **37A**, 525-550.

Spawton, T. (1990). Marketing planning for wine. *International Journal of Wine Marketing* **2**, 2-49.

Stanislawski, D. (1975). Dionysus westward: early religion and the economic geography of wine. *Geographical Review* **65**, 427-444.

Stasi, A., Bimbo, F., Viscecchia, R. and Seccia, A. (2014) Italian consumers' preferences regarding dealcoholized wine, information and price. *Wine Economics and Policy* **3**, 54 – 61.

Stasi, A., Nardone, G. and Viscecchia, R. (2011) Italian wine demand and differentiation effect of geographical indications. *International Journal of Wine Business Research* **23**, 49 – 61.

Train, K. (2009). *Discrete Choice Models with Simulation*. Cambridge University Press, Cambridge.

Veblen, T. (1899/1994) *The Theory of the Leisure Class: An Economic Study of Institutions*. Unwin Books, London (1899); reprinted by Dover Publications, New York (1994).

Villas-Boas, J.M. and Winer, R.S. (1999) Endogeneity in brand choice models. *Management Science* **45**, 1324 – 1338

Walker, J. and Ben-Akiva, M. (2002). Generalized random utility model. *Mathematical Social Sciences* **43**, 303–343.

Williamson, P.O., Lockshin, L., Francis, I.L. and Mueller, S. (2016). Influencing consumer choice: short and medium term effect of country of origin information on wine choice. *Food Quality and Preference* **51**, 89 – 99.

Wooldridge, J. (2002) *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Mass.

Yañez, M.F., Raveau, S. and Ortúzar, J. de D. (2010). Inclusion of latent variables in mixed logit models: modelling and forecasting. *Transportation Research* **44A**, 744 – 753.

7 APPENDIX I: TELL ME WHY YOU LIKE TO DRINK WINE: DRINKING MOTIVES AS A BASIS FOR MARKET SEGMENTATION

7.1 Introduction

Wine is a complex product from a consumer's point of view. To start, it is an experience product, *i.e.* it can only be fully appreciated after consumption because it is not possible to smell or taste it before a bottle is opened (Grunert 2005). Secondly, wine presents a vast and depth sensory variety (Ferreira *et al.* 2007), which is often overwhelming to the new consumer (Charters & Pettigrew 2003). Furthermore, wine is deeply rooted in the history of Western culture, bearing complex social (Mouret *et al.* 2013), cultural and religious symbolisms (Stanislawski 1975).

Wine complexity induces complex behaviour on consumers, making it harder for the wine industry to understand the market. Preferences for wine seem to be not only heterogeneous among the population (Blackman *et al.* 2010), but also variable within individuals (Mueller & Szolnoki 2010) and strongly context-dependant (Ritchie 2007). Since understanding consumers is the first step to effective marketing, a method to untangle this complexity is required.

Segmentation is one of the most traditional techniques used to understand consumers (Smith 1956), and as such, it has been applied to the wine market (Lockshin & Hall 2003, Lockshin & Corsi 2012). Traditionally, segmentation has been based on four factors or variables (Arnould *et al.* 2002): demographics, geographic location, behaviour, and psychological characteristics of the consumers. On the wine industry, the latter has been the most explored (Spawton 1990, Lockshin *et al.* 1997, Bruwer *et* *al.* 2002), but some have proposed occasion based segmentations (Dubow 1992) and even mixed approaches (Quester & Smart 1998).

Within psychological segmentation, several approaches exist with no one standing out as clearly superior. Level of involvement and lifestyle has been two of the most studied segmentation variables, yet neither of them considers the influence of drinking context, such as the consumption occasion. Consumption occasion is a relevant factor in wine liking and purchasing, as empirically shown by Martínez-Carrasco *et al.* (2006), Hersleth *et al.* (2003) and Hall (2003). A third alternative, also within the psychological segmentation approach, is to segment the market based on the motives behind drinking. According to the theory of planned behaviour, “intentions are assumed to capture the motivational factors that influence a behaviour” (Ajzen 1991). Therefore, consumers’ motives for drinking wine would be linked to their behaviour, making motives a useful segmentation variable.

Despite all the benefits of motive-based segmentation, this approach is not common on wine market research. Meanwhile, the need for an efficient segmentation of the wine market remains unsatisfied.

This study explores the most relevant motives for drinking among Chilean premium wine consumers, and compares them with motives discovered on other populations. These motives are complemented with a simple proposal for a consumer segmentation based on wine knowledge. This is the first exploratory stage in stablishing a motive-based segmentation for the Chilean wine market. As we use only qualitative methods the discovered motives will be tested quantitatively in future studies. Nevertheless, our results present important similarities with other findings reported in the literature, making us confident about the relevance of our results.

Chilean premium wine consumers are defined as those who regularly buy and drink wine costing more than US$ 10 per 75cl bottle. Even though wine has been consumed in Chile since the arrival of the Spanish conquerors, the premium niche has experienced a significant growth since the nineties, in similar fashion to other emerging economies where premium wine consumption is a relatively new trend too.

7.2 Literature Review

Most wine market segmentation studies are based on psychological variables. Spawton (1990) was one of the first studies that, based on research by McKinna (1987), identified four consumer segments: *Connoisseurs* (knowledgeable segment, they consume often and they like to experiment with new wines), *Aspirational drinkers* (concerned with self-image projected, buy fashionable brands, seek advice), *Beverage wine consumers* (consume often but do not experiment) and *New wine drinkers* (new to wine, drink mostly on premises). A decade later Hall & Winchester (2000) empirically confirmed three of Spawton's (1990) four segments, changing the *New wine drinkers* for an *Enjoyment-oriented* segment.

After Spawton (1990), other researchers have explored different segmentation variables but mainly within the psychological spectrum. Level of involvement, wine-related lifestyles and occasion have been the most studied segmentation variables, even though there has also been some work on motive-based segmenting (Dubow

1992), as well as other attempts based on consumer behaviour (Mueller & Lockshin 2013; Goodman *et al.* 2002). This section discusses some relevant work on each segment.

7.2.1 Segmentation based on level of involvement

Involvement is a measure of how relevant the consumption, purchase, and brand of a product are for a consumer. Lockshin *et al.* (1997) proposed segmenting based on these three types of involvement identifying five segments: the *Choosy buyer*, the *Brand conscious*, the *Uninvolved* consumer, the *Interested* consumer and the *Lazy involved* consumer.

Aurifeille *et al.* (2002) segmented Australian and French consumers using level of involvement, discovering five clusters too but different from the ones proposed by Lockshin *et al.* (1997). The authors did not name their segments and only one of them appeared to be stronger among French than among Australian consumers.

Charters & Pettigrew (2006) discovered that the perception of quality was affected by the consumers' level of involvement. Unlike other studies, these authors did not differentiate between product, brand and purchase involvement (even though they acknowledged the taxonomy), using instead a single "enduring involvement" measure, which is more related to product involvement than to the two other types of involvement. Through a qualitative study, they found that highly involved consumers experienced a more complex response to wine, being this not only a sensorial effect, but also cognitive and affective. Consumers with a lower level of involvement, instead, only experienced sensorial responses. This difference would allow highly involved consumers to consider more factors when evaluating wine quality. Finally, they observed a tendency among consumers with low involvement to consider quality as a subjective judgement, while highly involved consumers thought about quality on a more objective way.

Johnson & Bastian (2007) proposed segmenting consumers based on their level of wine knowledge. Using an Australian consumer sample, they discovered that objective knowledge correlated poorly with acute sensory perception, and that more knowledgeable consumers tend to drink more often and spend more on wine, than less knowledgeable consumers.

Bruwer & Hang (2013) designed a 24-item instrument (drawn from several sources) to measure wine involvement. Their instrument is based on a general model by Kapferer and Laurent (1985) that considers five factors: *interest*, *behaviour, ritual, pleasure*, and *risk*. The model was validated, even though risk presented a slightly low reliability. The authors found that the instrument was useful in explaining Bring Your Own Bottle (BYOB) behaviour on a South Australian sample of wine consumers.

Ogbeide & Bruwer (2013) proposed both a theoretical model and a measurement instrument for *enduring involvement* (as opposed to the short-lived *situational involvement*). They suggested that *enduring involvement* “relies upon the consumer, based on the ability of the product to express the inner needs”, and therefore it derived from consumer’s lifestyle self-image, interest, ego, and value. In the author’s model, *enduring involvement* reflects on the consumer’s follow-up behaviour, *i.e.*

consumer's cognitive and behavioural reactions to wine. But the influence of *enduring involvement* is moderated by contextual factors such as opportunities, knowledge and consumer's characteristics, all of which influence the persistency, direction, and intensity of *enduring involvement*. This theoretical model was elaborated from discussions in the literature and focus groups. After presenting their theoretical model, Ogbeide & Bruwer (2013) proposed a 13-item instrument to measure *enduring involvement*, which was composed of three dimensions: self-image/sign value, pleasure/interest and lifestyle/enjoyment. The instrument was validated on a sample of 140 Australian consumers.

7.2.2 Segmentation based on wine-related lifestyle

Lifestyle is a psychographic segmentation variable that aims to capture information about consumers and their long term purchasing habits through linking products with values (Grunert *et al.* 1993). Bruwer *et al.* (2002) introduced this type of segmentation strategy onto the wine market, even though the work of McKinna (1986) could be classified in this category too.

Bruwer et al. (2002), Johnson & Bruwer (2003) and Bruwer & Li (2007), can be seen as a serial proposition of a refined segmentation for the Australian wine market. The first paper was exploratory, and found five segments of wine drinkers, the second one empirically confirmed the segmentation (even though with some amends), and the third aimed to check if the segmentation remained stable on time. Only three wine drinkers' segments remain unchanged along the three studies: (i) *Conservative, knowledgeable* (also called *Purposeful inconspicuous*); (ii) *Basic* and (iii)

Enjoyment-oriented, *social*. The other two segments changed from work to work, but they always seem to be subdivisions of the *Connoisseur* segment of Spawton (1990). They concluded that the segmentation should be updated from time to time.

The work of Mora *et al.* (2010) is also framed on the wine-related lifestyle tradition, and was performed on Chilean consumers. Based on the ideas of Kassarjian (1971) and Laroche *et al.* (2001), the study focused on consumer's attitudes toward organic wine, identifying three segments in relation with the consumption of organic wine: *Social and indifferent* towards it, *Positive attitude* towards it, and *Consumer of organic products*.

7.2.3 Segmentation based on consuming occasion

Although segmentation based on consuming occasion is probably one of the least popular strategies in the literature, the approach is very promising. As Ritchie (2007) points out, wine satisfies different functions on consumer lives, and those functions are highly dependent on context. Consumers do not behave the same way (*i.e.* do not buy the same wine) when buying a bottle for themselves, buying it as a gift, taking it to a dinner party, or sharing it with someone at a restaurant.

Hall & Lockshin (2000) used means-end chain analysis to study the perceptual differences between consuming occasions. Means-end chain analysis aims to link personal values to product characteristics. It assumes that the product is the mean used by consumers to achieve their end (the end being values that are important to consumers). Given that values act as motives for consumers (Vinson *et al.* 1997), the work of Hall & Lockshin (2000) is able to relate occasions with motives, and thus constitutes a mixed approach to consumer behaviour analysis. The authors worked with nine predefined values (motives), taken from Kamakura & Novak (1992), and eight consuming occasions set *a priori*: intimate dinner, meal with friends, meal with family, business related, outdoor BBQ/picnic, party/celebration, self, and with friends.

Finally, Quester & Smart (1998) –unlike the previous work discussed– advocated for an approach that considers both the consumption occasion and the level of product involvement of the consumer. Using conjoint analysis, they demonstrated that the behaviour of consumers changed based both on consuming occasion and level of product involvement. In their experiment, the authors only considered buying wine for three fixed occasions: meal at home during the week, dinner party at a friend's house, and a gift for the employer.

7.2.4 Segmentation based on motives

Few marketing studies rely on motives for segmenting. The most relevant ones are probably Dubow (1992) and Brunner & Siegrist (2011). The former compares two segmenting approaches based on drinking motives[1]. The first (called consumer-based) classifies users based on their general wine drinking motives (i.e. their motives without specifying a particular consuming occasion). The second approach (called occasion-based) classifies consuming occasions based on the particular motives experienced by consumers on each occasion. This requires consumers not

[1] Dubow (1992) measured drinking motives using a 33-item questionnaire. This questionnaire, however, is not validated as a psychometric instrument, as no factor analysis is performed.

only reporting all their consuming occasions, but also their motives for each. The resulting occasion segments –called “need states”- can be interpreted as prototypical consuming occasions, just as any user-based segment can be understood as a prototypical consumer. Dubow (1992) discovered five segments with each approach, but only three were common to both. He concludes that occasion-based segmentation is more informative than user-based segmentation, at least in the case of products where consumers are not loyal to a brand. The five segments associated with the occasion-based approach are: *Social* (drinks to share with others), *Introspective* (drinks to improve mood), *Semi-temperate* (drinks light wine), *Food enhancement* (drinks to enhance food), and *Oenophilic* (drinks for the aroma, enjoys choosing wine).

Dubow (1992)’s work was based on the hypothesis that different consuming occasions differ due to the motive they entangle. This hypothesis is supported by the finding that occasion-based segmentation is preferable to user-based segmentation. Therefore, motives should not be invariably associated with particular consumers, as they may be influenced by different motives on different occasions. This inherent variability of consumers is an often overlooked aspect on most segmentation studies, probably because most use cross-sectional data.

In an effort to combine different segmentation approaches, Brunner & Siegrist (2011) proposed a single 81-item instrument to measure consumer involvement, motives/lifestyle and purchase behaviour. Using principal component analysis, they found 17 dimensions, with two of them associated with involvement, ten with motives/lifestyle, and five with purchase behaviour. Their instrument, comprising items both new and taken from previous literature, shows good reliability (Cronbach's α ranging from 0.52 to 0.92) on a large (n = 929) sample of German-speaking Swiss wine consumers. Later, consumers were segmented based on their scores on the different dimensions, using hierarchical cluster analysis. Six segments were found: *Price conscious, Involved*, *Image-oriented*, *Indifferent*, *Basic,* and *Enjoyment-oriented*. The importance of motives becomes evident on the work of Brunner & Siegriest (2011), as motives –combined with lifestyle- concentrate the biggest number of dimensions.

Even though not a segmentation study, Charters & Pettigrew (2008) performed an extensive qualitative study on motives for wine drinking. They identified two dimensions on the experience of drinking wine: experiential and symbolic. The experiential dimension entails three aspects: personal enjoyment of wine, no matter if it is sensorial or cognitive; the feelings when consuming wine with a meal or with others; and the relaxing effect. The symbolic dimension also considers three aspects: the drinking ritual and its different meanings; links between wine and consumers' personal history; and the way consumers want others to see them. Charters & Pettigrew (2008) concluded that the experiential dimension is the most relevant for the majority of consumers, even though it interacts with the symbolic one. The authors also note that –for some consumers– wine is something that "gives meaning to life". This relates to Charters & Pettigrew (2005) perception of wine as a quasiaesthetic product, *i.e.* a product that is not only considered a beverage, but that can also be appreciated in a sensorial and cognitive way, just like art.

Despite not being related to market segmentation, it is interesting to look at the field of health psychology, where a significant amount of work on drinking motivations has been done. A key work in this field is the one by Cooper *et al.* (1992), where an instrument to measure drinking motivations, called Drinking Motives Questionnaire (DMQ), was developed and validated. The instrument measures three drinking motives: *social* (drinking because it is customary on certain social occasions), *coping* (drinking to control negative emotions), and *enhancement* (drinking to enhance positive emotions). The authors also show that these motives can (partially) predict drinking volume and frequency.

In Cooper (1994) and Cooper *et al.* (1995), a revised form of the original DMQ - called DMQ–R - was developed and validated. This new instrument categorizes drinking motives as internal/external and positive/negative. From the crossing of these categories, four drinking motives arise: *enhancement* (internal positive), *coping* (internal negative), *social* (external positive, *i.e.* drinking to have fun with others), and *conformity* (external negative, *i.e.* drinking to fit in the social group). The DMQ-R was also validated by Kuntsche *et al.* (2008) on Swiss, Canadian and American samples of adolescents.

More recently, there have been several studies that expand or question the DMQ-R and its underlying model. Kuntsche & Stewart (2009) prove that the motives of adolescent's peer groups influence the individual's drinking motives, therefore proving that the drinking motives of peers indirectly influence the individual's drinking habits. Crutzen & Kuntsche (2013) validated the DMQ-R model on adults.

Grant *et al.* (2007) proposed dividing the coping motive on two: *coping-anxiety* and *coping-depression*, as anxiety and depression may have a different effect on alcohol use. They validated their modified DMQ-R on college students, even though they report some reliability problems with the social motive.

Crutzen *et al.* (2010) argue that the use of alcohol to cope with stress is not really proven. They study a sample of Dutch students (19-26 years old), and show that stress (both daily as well as due to life events) does not have a significant effect on alcohol consumption, neither directly nor through the coping motive. But they do confirm a link between the coping motive, and alcohol-related problems.

Finally, Moran & Saliba (2012) propose that *Taste* (*i.e.* drinking alcohol because the user likes its taste) and Health (*i.e.* drinking because it is healthy) are also relevant motives to drink alcohol. And even though they do not use the full DMQ-R instrument, they find that the *Taste* motive is the most endorsed by their South Australian sample, while almost a quarter of it also endorsed the *Health* motive.

Even though both the marketing and health psychology literature on drinking motives focus on the link between motives and consuming behaviour, the health psychology literature is only interested in risky consumption. This means that the DMQ-R and all of its variations aim only to identify potentially risky drinkers, and there is no guarantee that the purchase behaviour of individuals with different DMQ-R motives may be different, or may change in a predictable way. Therefore, even though the DMQ-R is a validated instrument for measuring drinking motives, there is no evidence that it may be useful for marketing applications.

7.2.5 Segmentation based on behaviour

Instead of classifying consumers according to their psychological characteristics, some authors have attempted to segment consumers based on their self-reported purchase behaviour. Some interesting studies employing this approach are those by Mueller & Lockshin (2013) and Goodman *et al.* (2002), both using best-worst scaling methods.

7.3 Methodology

A sample of premium wine consumers were first recruited to participate in in-depth interviews. Two recruiting methods were used. The first was an open invitation published in the Facebook page of the *Centro de Aromas y Sabores* in Santiago, Chile. All interested individuals had to go through a filter in order to be accepted; the requirements were having bought and drunk premium wine at least once during the last month. The second recruiting process was an invitation sent by e-mail to a subset of consumers affiliated to a popular Wine Club. While the first process aimed to recruit occasional consumers, the second sought to capture more experienced users (*i.e.* individuals who buy and consume wine regularly). An heterogeneous sample of consumers (based on their socio-demographic characteristics) was finally selected to participate.

Fourteen consumers were interviewed in depth, four of them being members of the Wine Club. Half of the sample was female, and six individuals were forty years or older. All interviews were performed on places selected by the consumers, so they would feel comfortable (*i.e.* mainly at their homes or workplaces). Theoretical saturation was reached after coding and thoroughly analysing seven interviews. The rest of the interviews were listened and evaluated, but not coded, since they did not provide significant additional information. Table 7-1 shows all respondent's characteristics.

The in-depth interviews were episodic (Flick, 2004), i.e. one-on-one interviews, where the respondent was asked to narrate a number of events related to the research subject, allowing him/her to elaborate those aspects that s/he considered relevant. In this case, they had to describe the last time they bought wine, the last time they drank wine and the last time they gave wine as a present. The analysis was performed following the thematic coding proposed by Flick (2004), which considers four stages: coding, categorization, category description, and category harmonization. The methodology aims to extract basic underlying categories from the interviewee's discourse, and then – through iteration both within and between interviews - reach global categories (*i.e.* motives). Based on results of the last stage, a consensus map (Zaltman 2003, Novak & Cañas 2008) was also built.

Table 7-1 - Main characteristics of the seven in-depth interview respondents

Gender	Age	Occupation	Wine club member	Was interview coded'
Female	24	Secretary/Student	No	No
Male	31	Engineer	No	No
Male	30	Engineer	No	No
Male	29	Engineer	No	No
Male	62	Engineer	Yes	No
Male	41	Engineer	Yes	No
Male	36	Engineer	Yes	No
Female	42	Commercial assistant	No	Yes
Female	29	Technician	No	Yes
Female	43	Secretary	No	Yes
Male	43	PE teacher	No	Yes
Female	41	Secretary	No	Yes
Female	32	Journalist	Yes	Yes
Female	22	Student	No	Yes

Subsequently, a trained specialist conducted six focus groups. Each session lasted about one and a half hours. Each focus group was classified based on gender, age, consuming frequency and the participants' wine drinking experience (in years). None of the participants were affiliated to a Wine Club. Table 2 shows the main characteristics of the participants of each focus group. Even though no income restriction was included in the recruiting process for the focus groups, only medium and high income individuals participated. This probably happened because all participants had to have bought and drunk at least one bottle of premium wine recently.

Table 7-2 - Focus group filters

Focus Group	Gender of participants	Level of involvement with wine	Age
1	Males	High	30 to 60
2	Females	High	30 to 60
3	Males & Females	Medium (novice drinker)	25 to 35
4	Males & Females	Medium (novice drinker)	36 to 60
5	Males & Females	Low	25 to 35
6	Males & Females	Low	36 to 60

All sessions were structured around seven topics: purchasing process, purchase experience, consuming occasion, product choice, relation between wine attributes and consuming occasion, sensorial experience, and information sources. For closure, a list of the most relevant attributes of wine was constructed. The motives for drinking wine were not explicitly included in the script of the focus groups, as they were intended to show up on their own during the discussion. Audio recordings of all the focus groups sessions were made, for later analysis.

7.4 Results

Figure 7-1 presents the consensus map (Novak & Cañas, 2008) resulting from the in-depth episodic interviews and focus groups analyses. Four main motives were identified: *Social cohesion*, *Sophistication*, *Self-indulgence* and *Tradition*. Furthermore, a basic dichotomous classification of consumers was also proposed: *Occasional* and *Connoisseur*. All of them are developed below.

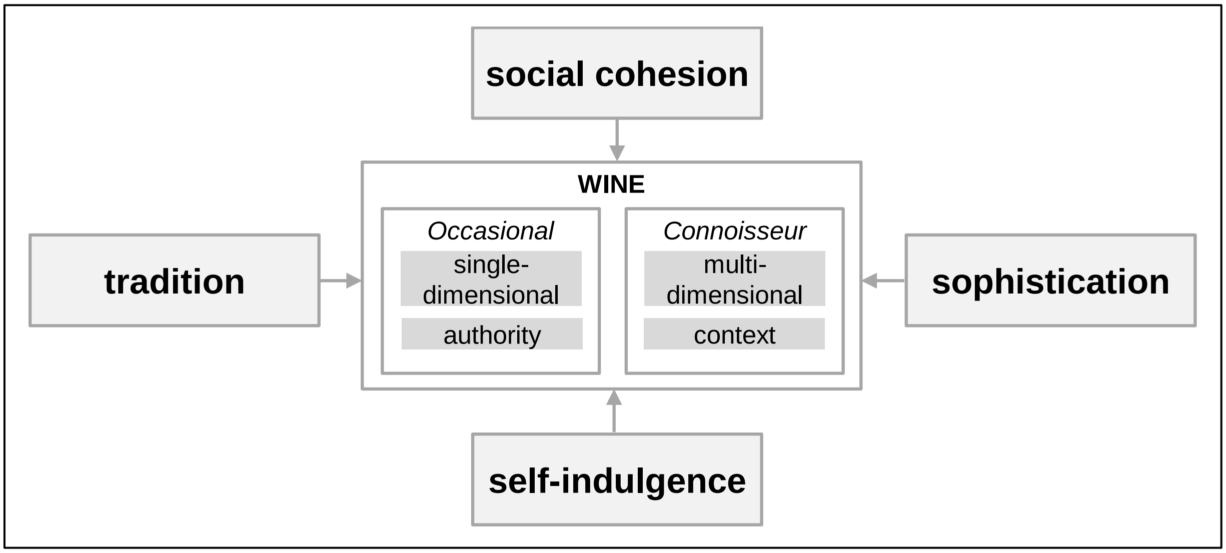


Figure 7-1 - Consensus map with four motives for drinking wine, and two classes of consumers

7.4.1 Motives for drinking wine

Social cohesion, *Sophistication*, *Self-indulgence* and *Tradition* are four motives to drink wine. These motives determine the way consumers relate with wine by giving it a particular role in their life, as well as in their relation to others. More than one motive can be present on a given consumer, as s/he faces different consuming occasions.

Social Cohesion, relates to the social environment in which wine drinking often happens, an environment that validates wine consumption. In Chile, as in many other societies, wine is the preferred drink with meals, and it is also used in social gatherings and celebrations, even religious ones, as in the case of the Catholic mass. On social occasions, wine helps people to relax, easing social relations. This is in direct connection with wine's alcoholic nature, as several participants expressed. And even though drunkenness is often condemned, moderate wine consumption is not negatively perceived.

Wine can also ease social relations by providing a conversation subject. When a good wine is on the table, talking about the wine itself can work as a conversation starter. Even more, trying new wines can be the very reason to get together with friends.

Even though *Social cohesion* explains wine drinking based on social interactions, wine is actually considered a drink for intimate occasions by consumers. This means that wine is perceived as a drink to share with close friends or family, not to be drunk at massive parties. This perception relates to wine being a relaxing drink, not a drink to “get tipsy”, as other mass consumption products might be.

Sophistication is born from the desire of being unique. A sophisticated individual wants to be distinguished from others, and being an expert on wine allows him to differentiate. The fact that a particular consumer is moved by *Sophistication* does not necessarily imply that the individual's interest in wine is purely utilitarian. *Sophistication* may entail a genuine passion for wine as an object (*i.e.* the individual is honestly attracted by wine and its world). The sophisticated individual will seek to learn more about wine just as music-lovers do not only listen to music, but educate themselves on it. This motive can make wine a hobby or even a passion.

When someone is passionate about wine, an intellectual interest on the product is observed. This means that consumers moved by *Sophistication* will seek on wine a device for distinction rather than a product for sensorial experience. Their interest can range from learning about the history of wine, to the particular wine maker. Again, this does not mean that they necessarily disregard its sensorial qualities, as many enjoy the exercise of looking for the particular aromas mentioned on the back label of the wines they drink.

This motive, in its passionate form, is usually remarkable on *Connoisseur* consumers (i.e. consumers with very good knowledge of wine). This often makes these consumers a reference, enabling them to assume an authority role for other, less informed, wine drinkers.

Self-indulgence represents the tendency to gratify oneself, in the sense of attempting to add pleasure to everyday life. This does not necessarily correspond to an hedonist way of life, but rather the sense of reward for one's efforts, or an “I deserve to be happy and enjoy myself” way of thinking. This particular motive was more commonly observed on less juvenile consumers (over forty years old, approx.).

The motive to drink by *Self-indulgence* requires a strong association between wine and pleasure. This pleasure can arise from two main sources. First, as was explicitly mentioned by consumers, it can arise from the sensory properties of wine (*i.e.* consumers enjoy wine’s aroma and flavour). Secondly, and not explicitly mentioned by consumers, the reward may come from drinking something exclusive (*i.e.* to buy and drink an expensive wine makes the consumer feel rewarded). Both sources of pleasure are not mutually exclusive and often present themselves together.

Even though *Self-indulgence* seems to be a rather individualistic motive to drink wine, it can also include others (close individuals). An example of this is a romantic dinner with a partner, where wine is carefully picked up to meet the expectations and allowing to share a nice experience.

Tradition represents the socio-cultural influence on wine consuming behaviour. This comes from many sources, being family and closer relationships the most remarkable. The image of the father (or grandfather) at the head of the table, having wine with the meal, helps legitimizing wine consumption, and gives it a halo of filial and fraternal bond. This may motivate wine consumption as an attempt to live once again those warm feelings.

Tradition also adds a strongly masculine aspect to wine representation. This was particularly noticeable on middle-aged women, but several males also fondly remembered the first time they drank wine with their fathers. This association may also help explaining why red wine is preferred over white among Chilean male consumers: as wine is considered male, it should be strong, a characteristic more suitable for red than white wine.

The consumption of wine on annual festivities, such as the national holydays, Christmas, or New Year’s Eve, is also linked to *Tradition* as a motive. In Chile, these festivities are often celebrated with a family meal, where wine is most commonly the preferred drink. This argument may be extended to also include other occasional celebrations, such as marriages.

7.4.2 Consumer classes

Besides identifying four motives for drinking wine, two basic classes of wine consumers were detected: *Occasional* and *Connoisseur* consumers. More than being the result of a strict classification exercise they represent archetypical consumers, so each individual is expected to present characteristics of both classes but with one dominating the other. The main (observable) difference among the two classes is their level of confidence on their own wine-related decisions.

The **Occasional consumer** tends to perceive wine as a single-dimensional product, with quality being the main concern. But quality is a vague concept for *Occasional* consumers. Although they often perceive it as an objective characteristic, they believe they lack enough knowledge about wine to make a correct judgement about quality. They believe themselves to be incapable - at least to some extent - to differentiate a good wine from a bad one. Therefore, they resort to figures of authority when choosing wine. If no authorized source of information is available, the *Occasional* consumer heavily relies on price, assuming that more expensive wines provide higher quality.

Occasional consumer's evaluations of wine tend to be static and independent of context. This is a consequence of considering external opinions as the only valid parameter to judge a wine, and the tendency of these consumers to strip those opinions of their context and generalized them as absolute rules. This leads to occasional consumers having a list of "good wines" they buy and consume, disregarding the particulars of the occasion they are drink at.

***Connoisseur* consumers** perceive wine as a multi-dimensional product, meaning that a particular wine is not good or bad *per se*, but it has certain characteristics which makes it good or bad depending on the drinker's personal tastes and drinking occasion. This perception is only achievable if consumers have a certain level of experience or knowledge about wine. However, the specific depth of this knowledge is largely irrelevant, as only a small amount is necessary. The most relevant thing is that these consumers are confident enough on their knowledge, so the *Connoisseur* can judge wines based on it. And trusting one's own judgement has more to do with confidence than with actual knowledge.

Even though *Connoisseur* consumers perceive wine in a multidimensional way, they still consider quality in their evaluation, but their perception of quality can depend on a series of contextual factors. Among these factors, price can play an important role. Therefore, *Connoisseur* consumers may associate higher prices with higher quality, but not in a necessarily strict manner. Most consider that as price grows, it becomes less probable to find defects on wine.

Both classes of consumers interact with each other. The *Connoisseur* often plays the role of authority for the *Occasional* consumer. At the same time, *Occasional* consumers often see themselves on a learning path to become a *Connoisseur*. Most *Occasional* consumers in the sample did manifest explicitly their will to learn more about wine. This, nonetheless, could be a particularity of our sample, because of a self-selection bias (as many participants postulated to participate on the study, they may had been more interested in wine than the average consumer).

The way to become a *Connoisseur* consumer seems to be experience. When *Connoisseurs* were asked about how they became confident about their own knowledge, most of them answered that it was through experience (i.e. having tested many different wines puts them in a better position to judge new ones). This perception is shared by most occasional consumers, who referred to their figures of authority as someone who had tried many different wines. But the access to a large diversity of wines is mediated by each consumer's budget. This is perceived as a limitation by consumers with lower income, making it harder for them to eventually become *Connoisseurs*.

To be a *Connoisseur* consumer does not necessarily imply a higher volume of wine consumption, but it does seem to involve a higher degree of diversity on purchase behaviour. *Occasional* consumers often buy a limited set of products, and tend not to try new wines unless an “authority” recommends them (a *Connoisseur* friend or a good review they might have read). *Connoisseurs*, instead, tend to buy new wines, either by taking a chance or by searching for information themselves (looking up products on the internet or periodically checking blogs or wine clubs). This does not mean that *Connoisseurs* do not rely on habit when buying wine, but to a far lesser extent than occasional consumers.

7.5 Discussion

The four motives discovered can be related to previously reported motives in the literature; in particular on Brunner & Siegrist (2011)[2], which itself draws from several previous works. In that study, the authors presented a *Sociability* dimension within their instrument, which can be directly related to the *Social Cohesion* motive found among Chilean consumers. Our *Tradition* motive, on the other hand, can be assimilated to Brunner & Siegrists (2011)'s *Tradition and Food* dimensions on the same instrument. *Self-indulgence*, as we described it, has elements from four different dimensions in Brunner & Siegrist (2011): *Fun*, *Pleasure*, *Health* and *Recreation*. Finally, our *Sophistication* motive can be associated with five dimensions on Brunner & Siegrist (2011)'s instrument, *i.e. Knowledge*, *Events*, *Self-expression*, *Style* and *Intellectual Challenge*. The *Knowledge* dimension can also be equated with consumer's confidence, that we also argue is very relevant for Chilean consumers.

While Brunner & Siegrist (2011) developed their instrument in an attempt to summarize all wine segmentation "traditions" (involvement, motives/lifestyle, and purchase), we used an exploratory approach. Both approaches present a great level of agreement, supporting the idea that Chilean premium wine consumers may be moved by similar motives than other populations.

The four discovered motives can also be related to some DMQ-R drinking motives. The *Social* motive of the DMQ-R could include both *Social Cohesion* and *Tradition*

---

[2] We are grateful to an unknown referee for having made us notice this.

motives, while *Enhancement* could comprise *Sophistication* and *Self-Indulgence*. No negative motives for wine drinking were discovered in our study, probably due to the fact that most participants were wine enthusiast.

Our results also agree – at least to some degree - with the theoretical model proposed by Ogbeide & Bruwer (2013). In that paper, the authors considered *enduring involvement* to rely on the matching between consumer's inner needs and the product's ability to fulfil them. If we consider motives to be a reflection of consumer's inner needs, then *enduring involvement's* dimensions should closely relate to consumer motives. We can observe this relation by loosely associating Self-image/sign value to *Sophistication*, Pleasure/interest to *Self-Indulgence*, and Lifestyle/enjoyment to *Tradition* and *Social Cohesion*.

A more evident relation between our results and those reported by Ogbeide & Bruwer (2013) is the finding of confidence as a moderator of motives and enduring involvement, respectively; they report Wine Knowledge as a moderator of enduring involvement, mentioning that it … "gives a feeling of security and consequently an indifference attitude toward risk".

We also found evidence of Dubow (1992)'s hypothesis on motivations being dependent on consuming occasion. Because although some consumers might tend more to one motive than to another, they all showed variability when narrating different consuming occasions. This means that different motives can act upon the same consumer at different times. This result reinforces the conclusion of Ritchie (2007), Martínez-Carrasco *et al.* (2006), Hersleth *et al.* (2003), Hall (2003) and Hall

& Lockshin (2000), all of whom postulate that wine purchase behaviour is highly contextual. A segmentation based on motives, therefore, should not associate consumers to motives on a fixed manner, regardless of consuming occasion. It is the consumer's motives on each particular occasion that drive his/her behaviour.

Summarising, we believe that Brunner & Siegrist (2011)'s instrument is a promising one to segment Chilean premium wine consumers. Nevertheless, we also consider that it lacks the ability to dynamically assign different motives to different consuming occasions. One possible solution would be to apply the *involvement* and *purchase behaviour* part of the instrument only once to the consumer, and to apply several times the motives/lifestyle part, under the assumption of different consuming occasions. However, this could be exhausting for the consumer, as the motives/lifestyle part of the instruments is easily the longest (46 items).

Identification of the main psychological factors that influence consumer behaviour could also help improving some statistical model's prediction of consumer's choice. These factors could be included on discrete choice models, as done in O'Neill et al. (2014), where the use of latent variables for modelling taste and attitudes significantly improved the model's predictions.

7.6 Conclusions

Four motivations for drinking wine were discovered among Chilean premium wine consumers: *Social cohesion, Sophistication, Self-indulgence* and *Tradition*. Also, two classes of consumers were identified: *Occasional* and C*onnoisseurs*. While the latter tend to trust their own judgements about wine, *Occasional* consumers do not. These results, though only exploratory, are in line with previous findings, notably Brunner & Siegrist (2011) and Dubow (1992). Future work will include designing and empirically validating a segmenting instrument for Chilean consumers, considering that motives can vary between consuming occasions.

8 APPENDIX II: MODELLING WINE CONSUMER CHOICES IN AN INCENTIVE-COMPATIBLE EXPERIMENT

8.1 Introduction

As proposed by Grunert (2005) in his *Total Food Quality Model*, the purchase process of food and beverage products has three stages: the first time a product is purchase, the tasting (consumption) of the product, and the decision to purchase or not the product again (re-purchase). During the first purchase only extrinsic attributes are available (i.e. attributes that can be perceived before tasting the product), therefore the decision must be made based on an expectation of quality rather than the product's "true" quality. During the tasting stage consumers can finally perceive the products intrinsic attributes (i.e. taste and aroma, mainly) and therefore perceive the "real" or experienced quality. During future purchase occasions consumers will be able to recall the experienced quality as well as being influenced again by the extrinsic attributes if they have changed or have being forgotten.

Among food and beverages products, wine is a particularly interesting example. Wine's cultural and sensory complexity (Stanislawski 1975, Ferreira et al. 2007) is very demanding for consumers, to the point of overwhelming them (Charters & Pettigrew 2003). This complexity makes it difficult to determine wine's quality, therefore forcing consumers to strongly rely on extrinsic cues. At the same time, the incredible diversity in the wine market makes it very likely that consumers are constantly facing alternatives that they had not tried (maybe not even heard of) before. This makes of wine a remarkable product to study consumers' purchase behaviour, within the food and beverages category.

The whole purchase process of wine has seldom been studied as a whole on the food and beverages preference literature. A notable exception is Mueller et al (2010c), where the purchase, taste and repurchase stages were simulated and modelled. However, that study faces three main limitations: (i) it does not measure the influence of extrinsic attributes individually, measuring their combined effect instead; (ii) even though the modelling of the three stages is interrelated, its estimation is sequential and therefore results are difficult to compare between stages; and (iii) data comes from simulated decisions in a setting where choices do not have any real consequence (i.e. it is not an incentive-compatible experiment).

We designed and executed an incentive-compatible experiment to study the three stages of wine purchase. Data from our experiment can be conceived as revealed preference (i.e. actual purchase decisions) collected in a controlled environment. This mixes the benefits of traditional revealed preferences, i.e. their grounding in actual behaviour, with the benefits of declared preferences, i.e. its flexibility and possibility to optimize the data collection.

We analysed the data from our experiment using a single model combining observations from each stage of the experiment. This allows us to compare results between stages and providing a consistent interpretation of the whole process. Our model is capable of measuring the effect of individual attributes -both extrinsic and intrinsic- and compare their influence on the purchase decision. Our model is firmly rooted on the Total Food Quality model, therefore enjoying a strong behavioural grounding.

Section 8.2 describes the experiment, summarizes the sample's main characteristics and presents the model used to analyze the data. Section 8.3 presents some preliminary results of the experiment, and section 8.4 discusses them.

8.2 Materials and methods

In this section, the experiment's design is presented (section 8.2.1), followed by a summary of the sample's main characteristics (section 8.2.2) and a description of the model used to analyze the data (section 8.2.3).

8.2.1 Design of the experiment

The experiment had four stages: registration, choice on shelf, tasting and re-purchase. As soon as participants arrived to the experiment location we gave them a tablet PC, where they would input all information of the experiment.

The first stage was **registration**. This could be completed at each participant's place (if they had sign up for the experiment ahead of the data collection day), or at the experiment's site (for those who hadn't signed up previously to their arrival). During this stage, participants had to provide some socio-demographic information as well as some details about their wine purchasing and consuming habits.

Then, participants faced a screen with four wines, and they had to provide their level of agreement with three phrases for each of the four wines, using a 7-point Likert scale. The phrases were "This is wine is distinguished", "This is a quality wine" and "This is an attractive wine". Consumers' level of agreement with those three phrases were intended as indicators of their expected quality, i.e. their expectation of quality based solely on extrinsic attributes, as they could not taste the wines. Figure 8-1 shows an example of the screen where participants had to provide their expected quality indicators.

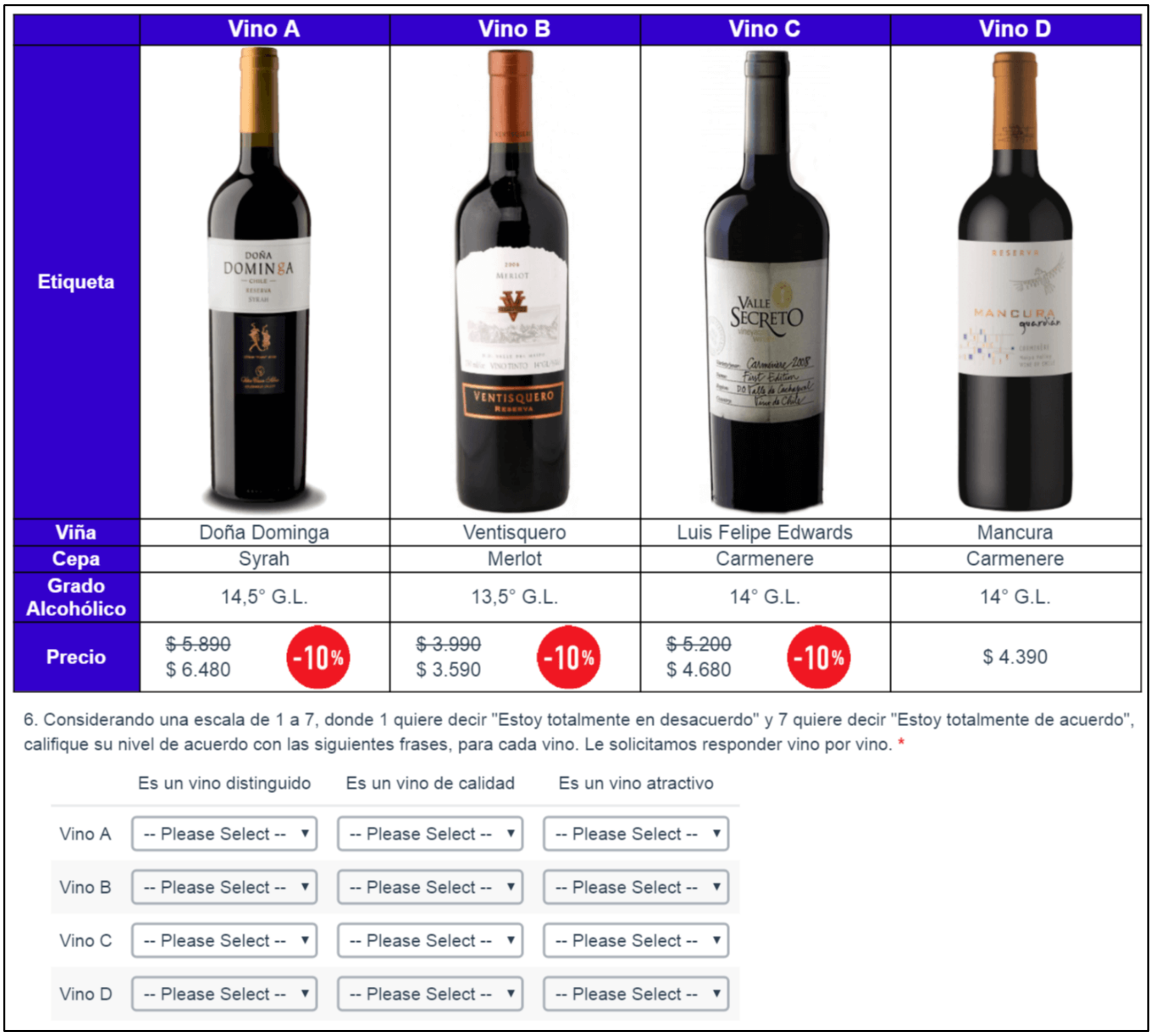


Figure 8-1 - Example of quality indicators collection during the registering stage

Each wine was described by six attributes: type of label, wine maker, grape variety, alcohol content, price and discount. Type of label had three levels: delicate, natural and contrast, according to the definitions of Orth & Malkewitz (2008). There were

19 different wine makers among 24 possible wines to show. Grape variety had four possible levels: Cabernet Sauvignon, Merlot, Carménère and Shiraz. Alcohol content had three possible levels: 13.5, 14.0 and 14.5 percent of alcohol. Price (before discount) was determined by the wine's market price. Discount had three levels: 0, 10 and 20 percent. The design of this stage followed an orthogonal design based only on label design, grape variety and alcohol content, assuming only two levels for alcohol content: low (13.5), and high (14.0 and 14.5).

After completing the registration, participants went on to the **choice on shelf** stage. Each participant was randomly assigned to one of six shelfs, each of them with 24 wines (see Figure 8-2). There, they had to build a ranking of the six wines they would like to buy the most. In the first place, they had to place the wine they would buy if they could only buy one bottle. In the second place they had to place the wine they would buy if their first preference was not available, and so on until the sixth position. Participants could make shorter rankings if they did not find six wines they would like to buy.

The purpose of the "choice on shelf" stage was to capture participants' preferences for wine when only extrinsic attributes were available, i.e. to capture their behaviour during a first buy. We attempted to use mostly unknown brands so participants were likely not familiar with the available wines, but we could not explicitly ask their level of familiarity with the wines as this would have greatly extended the duration of the experiment (we tested it during a pilot study, and most participants complained).

We asked participants to build a ranking of their favourite wines, instead of a single choice, to maximize the number of observations per individual. However, the ranking limited each participant to observing only a single shelf.

Figure 8-2 - Participants during the choice on shelf stage

Even though participants could observe all extrinsic attributes of the alternatives (they could examine the bottles for as long as they wanted to), we restricted the analysis to seven attributes: type of label (see Orth & Malkewitz 2008), grape variety, alcohol content, price, discount, horizontal position and vertical position. Levels are presented on Table 8-1.

Table 8-1 - Attributes and their levels

Type of label	**Grape variety**	**Alcohol content**	**Price**	**Discount**	**Horizontal position**	**Vertical position**
Delicate	Cabernet Sauvignon	13.5	(market price, approx. from 7 to 18 USD)	0%	Left	Down
Contrast	Merlot	14.0		10%	Centre	Centre
Natural	Carménère	14.5		20%	Right	Up
	Shiraz					

We used a D-efficient experimental design for this stage of the experiment. Only the type of label, the grape variety, the alcohol content, the discount and the position of the alternatives were considered for the design. For design purposes, alcohol content was considered as only having two levels: low (13.5) and high (14.0 and 14.5). We did not considered price in the design, as we used the products' market price in the experiment. We tried letting the design determine the wines' prices during a pilot study, but then consumers would always choose the wines with the lowest prices with respect to their market price, contaminating the experiment. We used priors from a pilot study for the efficient design, and assumed a simple MNL model. The full design had a single block with 24 choice situations. And even though we set up the 24 shelves, each respondent only faced one of them. We did four sessions of data collection, each of them with six different shelves. As mentioned before, each consumer was randomly assigned to one of the six available shelves.

After completing the shelf choice, participants would go on to the **tasting stage**, in a different room. Based on each consumer's ranking from the previous stage, a set of fives wines was automatically assigned for each participant to taste. Three of these five wines were drawn from the top three positions of each participant's ranking. The other two wines were selected based on each participant's first choice during the previous stage, in such a way that they would maximize the coverage of the sensory space of all 24 wines. This "maximization" was done by visual inspection, as shown in Figure 8-3. In the graph, each dot represents a wine in the sensory space (as described by two principal components). To select the additional two wines of the sample, the participant's first choice was found in the graph, and then a triangle was built with the chosen wine as one of its vertex, in such a way that the triangle would cover the maximum possible area of the graph.

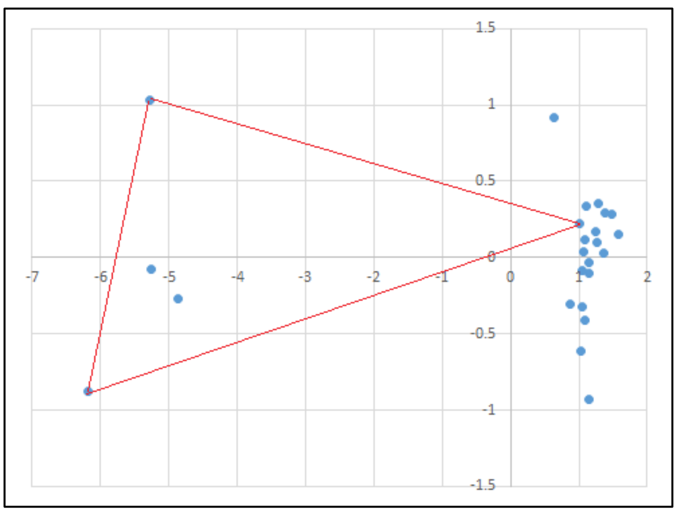


Figure 8-3 - Sensory space of all wines, as described by two principal components. The triangle maximizes the coverage for any given vertex (wine)

The sensory space was build based on a sensory profile for each of the 24 wines provided by a trained panel. The trained panel performed quantitative descriptive analysis (QDA) of the 24 wines, based on 27 different attributes (including colour). Then, we performed a principal component analysis (PCA) on the profiles of the 24 wines, keeping a solution of only two principal components explaining approximately 40% of the data set variation. We only kept two components because we needed to plot the wines in the space of the principal components.

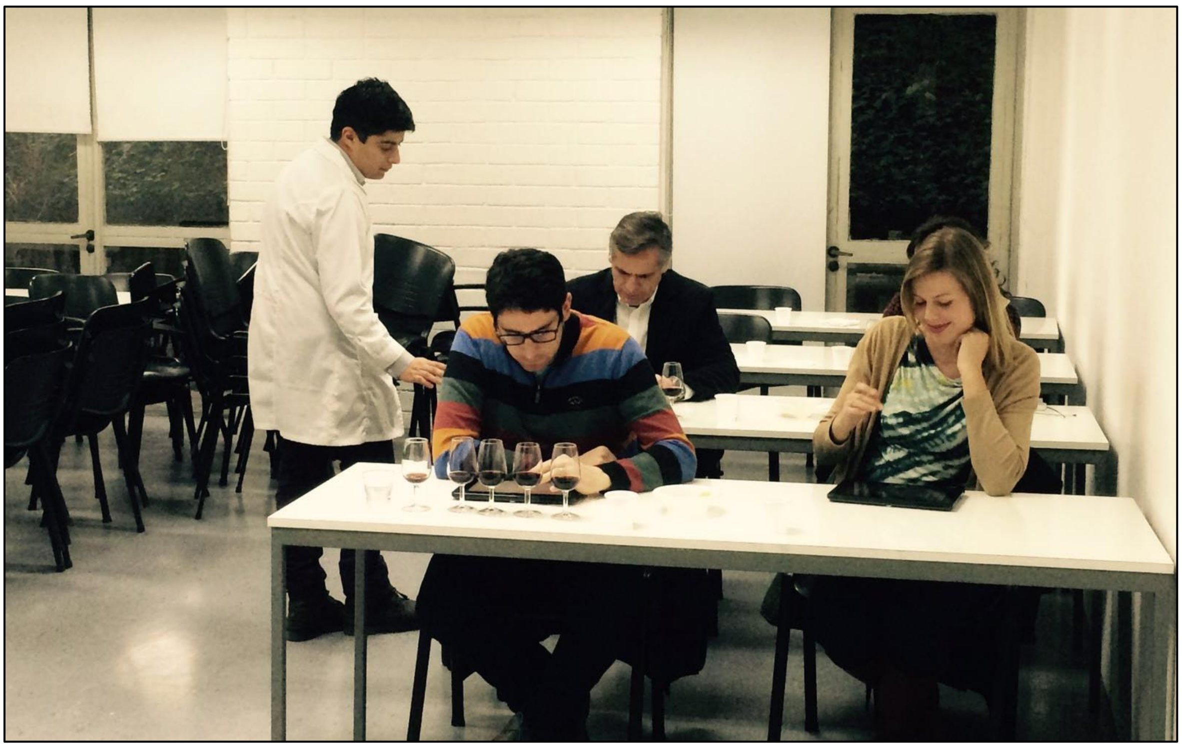


Figure 8-4 - Participants during the tasting stage

Participants then blindly tasted the set of five wines in a randomized order (Figure 8-4). Tasting was done blindly, i.e. participants could not see any extrinsic attribute of the wines they were tasting, to more easily measure the effect of intrinsic attributes. After tasting each wine, participants had to provide their level of liking on a 7-point Likert scale. They also had to check from a predefined set of adjectives those who -in their opinion- described the wine they were tasting. These adjectives were non-technical wine descriptors collected during a focus groups conducted previous to this experiment. Participants could select all or none of the attributes for each wine if they so preferred. The adjectives were: sweet, fruity, tastes like red berry, wood, chocolate, bitter, sour, dry, tastes like leather, tastes like alcohol, tastes flat, thick.

After completing the tasting stage, participants could proceed to the final stage of the experiment: the **repurchase**. During this stage, participants saw in their screen the full information of the five wines they had just tasted, i.e. all of their extrinsic attributes as well as the level of liking they reported (Figure 8-5).

In this stage, participants had the possibility of actually purchasing up to three bottles of any of the five wines they had previously tasted, in any combination. For example, they could buy three bottles of the same wine, or three different bottles, or just two bottles, or none at all. This was an actual purchase decision, as participants had to pay for the wines they wanted to buy. Prices and discounts were the same that respondents faced on the "shelf choice" stage of the experiment.

Participants could not choose any wine other that the five wines they had tasted, at least three of which they had also selected from the shelf among the top three of their ranking. As this stage wanted to measure the re-purchase decision, i.e. when consumers have both seen the product's extrinsic attributes as well as tasted its intrinsic attributes, it was of the greatest importance that consumers had full information on the alternatives. Therefore, we could not allow them to choose any wine they had not tasted in the previous stage.

This stage did not had an experimental design as its alternatives were fully determined by the previous stages.

Vino A	Vino B	Vino C	Vino D	Vino E
Odjfell	ChateauLosBoldos	Apaltagua	ValleSecreto	ViuManent
Cabernet Sauvignon	Syrah	Cabernet Sauvignon	Carmenere	Cabernet Sauvignon
13.5° G.L.	14° G.L.	14° G.L.	14.5° G.L.	13.5° G.L.
Nota degustación: 1	Nota degustación: 2	Nota degustación: 3	Nota degustación: 4	Nota degustación: 5
~~$ 6.750~~ $ 5.400 -20%	~~$ 4.790~~ $ 4.310 -10%	~~$ 5.200~~ $ 4.160 -20%	~~$ 8.150~~ $ 6.520 -20%	~~$ 4.590~~ $ 4.130 -10%

6. De los vinos mostrados en la tabla, ¿Qué vino(s) le gustaría comprar?
Puede comprar hasta tres botellas. Usted puede llevar más de una botella de cada vino.
Si no desea comprar, elija la alternativa "No compraré". *

1er vino a comprar -- Please Select --

2do vino a comprar -- Please Select --

3er vino a comprar -- Please Select --

Figure 8-5 - Example of choice during the re-purchase stage

8.2.2 Sample

128 people participated on the experiment. Most of them (57%) were students aged between 18 and 25 years old, representing the Millennial group. The overrepresentation of this group was due to the experiment being performed in a university campus, so many students participated spontaneously.

Figure 8-6 - A happy participant takes home his wines after purchasing them in the last stage of the experiment

Invitations to participate were send to member of a Chilean wine social network (wineCLR), including a link to the registration stage of the experiment and offering an extra 10% discount to those who registered in advance. We also announced a lottery among those who participated, with six prizes of 12 bottles of wine each.

Table 8-2 - Main characteristics of the sample

		Male	Female	Total
Number of people		77	51	128
Consuming	Almost never	0	6	6
Frequency	Few times a year	8	8	16
	Once a month	12	5	17
	Two times a month	23	10	33
	Once a week	11	11	22
	Two times a week	14	8	22
	Almost everyday	7	3	10
	Everyday	0	0	0
	More than one time a day	2	0	2
Average number of bottles per purchase		1.92	1.75	1.83
Always or	Supermarket	54	33	87
almost always	Specialty store	6	5	11
buys there	Liquor store	15	8	23
	Internet	2	1	3
	Catalog	3	0	3
Maximum	Less than 5	5	3	8
buying	Between 5 and 7	30	25	55
range	Between 8 and 14	34	17	51
(USD)	Between 15 and 21	4	3	7
	Between 22 and 29	2	2	4
	30 or more	2	1	3
Mantains a cellar at home		38	23	61
Age	18 to 25	44	29	73
	26 to 35	23	12	35
	36 to 45	4	5	9
	46 to 55	2	2	4
	56 to 65	3	3	6
	66 or more	1	0	1
Education	Incomplete university	37	26	63
	Complete university	22	13	35
	Postgrad	11	9	20
Household	Number of people	4.03	3.65	3.84
	Number of adults	3.53	3.10	3.32
Household	Less than 1429	16	14	30
income	Between 1430 and 4285	29	18	47
(USD)	More than 4286	32	19	51

8.2.3 Modelling

Each stage of the experiment was modelled separately. However, results from the registering and choice-in-shelf stage were very similar, so only the results of the choice-in-shelf stage are reported. Originally, these two stages aimed to measure two related but different things: the registering stage was supposed to capture the *expected quality* of wine bottles (i.e. the subjective quality based solely on extrinsic attributes), while the choice-in-shelf stage was supposed to capture a trade-off between *expected quality* and price. However, participants seem to ignore the negative effect of price on the choice-in-shelf stage. In other words, during the choice-in-shelf stage, participants seem to build a ranking of wines based on their perceived quality (i.e. *expected quality*) and not their real purchase intent (which would consider the negative effect of price). Therefore, results from the two stages are very similar, and only results from the choice-in-shelf stage are presented.

The **choice-in-shelf stage** was modelled as an ordered logit. The dependant variable was the inverted position in the ranking assigned by participants to each wine. The inverted ranking position was calculated by inverting the ranking made by participants. For example, if participant n had the following ranking: 1st wine A, 2nd wine B, …, 6th wine F; then wine A would have an inverted ranking value of 6, wine B of 5, wine C of 4, …, and wine F of 1. All wines left out of the ranking would have an inverted ranking value of 0. This coding method provided 24 observations per participants, i.e. one observation for each wine in the shelf per participant.

The utility of each wine was explained by its attributes (see Table 8-1) and their interactions with some characteristics of the participants. Specifically, only two characteristics of the participants resulted significant: their age included as a dummy, *millennial*, taking the value 1 if the participant was 25 years or younger; and their income, also included as a dummy, *high income*, taking the value 1 if the participant's household had a monthly per capita income equal or bigger than one million Chilean pesos (approximately 1428 USD). Utility functions were linear in parameters.

The **tasting stage** was also modelled as an ordered logit, with the overall rating as dependant variable and the sensory profile of the wines as explanatory variables. As the sensory profile provided by the trained panel had too many descriptors (27), we reduced its dimensional applying Principal Component Analysis.

The **repurchase stage** was modelled using Bhat (2008)'s multiple discrete continuous extreme value (MDCEV) model. This model combines a discrete choice (what to buy) with a continuous one (how much to buy). It also allows choosing more than one product, e.g. buy two units of product A and one of product B. Therefore, this model fits precisely the nature of the repurchase stage.

The MDCEV model has two formulations: one where an outside good is considered, and another one without an outside good. The outside good is analogous to the opt-out or no-purchase alternative. At the same time, including the outside good allows introducing a budget restrain in the model: all budget must be spent, with the outside good capturing all budget not spent on the products under study. We only use the version with the outside good, as using the version without it would require us to throw away all observations where no purchase was made, because the model without outside goof does not allow a non-purchase alternative.

In the full MDCEV not all parameters are estimable, due to empirical identification problems. Therefore, some kind of normalization must be implemented. Bhat (2008) proposes three different normalizations: (i) the alpha profile, (ii) the gamma profile, and (iii) the alpha-gamma profile. In this particular application we use the second, as it produces the simpler likelihood function.

The utility function of the MDCEV model under the gamma profile is as follows (individual subscript is dropped for ease of presentation).

$$U(x) = \sum_{k=1}^{K} \gamma_k \psi_k ln\left(\frac{x_k}{\gamma_k} + 1\right) \tag{8.1}$$

Where x is the vector of consumed units, γ_k and ψ_k are parameters associated with product k (K is the total number of products available), and x_k is the k^{th} element of vector x, i.e. the amount of product k purchased. This utility is maximized by the individual subject to the following budget constrain.

$$\sum_{k=1}^{K} p_k x_k = B \tag{8.2}$$

Where p_k is the unit price of product k, and B is the total budget of the individual. If no outside good is considered, then B becomes the amount of money spent in all purchased goods by the individual. In our case, B is the maximum amount of money consumers could have spent, i.e. three times the price of the more expensive alternative. Therefore, in our formulation each participant spends $3 \max_k p_k - p^T x$ in the outside good, where p is the vector of unit prices and x in the vector of purchases.

The γ_k parameter represents the satiation effect associated to product k. As γ_k becomes bigger, the less satiated the individual gets by product k, meaning he or she would tend to buy more. The ψ_k parameter represents the attractiveness of product k, and as it grows, the product becomes more likely to be purchased. Both parameters must be positive.

A simpler way of understanding the function of the ψ_k and γ_k parameters is that ψ_k determines the discrete choice (i.e. what to buy) and γ_k determines the continuous choice (i.e. how much to buy). This interpretation, however, is not completely accurate, as the decisions of what and how much to buy are taken jointly, so both parameters interact and play a part on both decisions. Nevertheless, interpreting ψ_k as the attractiveness of the product and γ_k as its satiation parameter is a useful conceptualization and we will employ it in the rest of the paper.

To make the model more informative, we make ψ_k depend on the product's attribute and γ_k on the individual's characteristics. Therefore, we define each parameter as follows.

$$\psi_k = e^{z_k \beta + \varepsilon_k} \tag{8.3}$$

$$\gamma_k = e^{c\gamma} \tag{8.4}$$

Where z_k is a vector of attributes of product k, c is a vector of characteristics of the individual, and β and γ (without sub-indices) are vectors of parameters to be estimated. Finally, ε_k is a iid extreme value type I random error component, specific to each product for each individual. Its scale value σ is identifiable and can therefore be estimated.

Among the attributes of the product (i.e. on the z_k vector) price is not included, as this enters the maximization process through the budget restriction. Therefore, unlike traditional choice models, there is no single parameter for price in the utility or attractiveness (ψ_k) of a product.

All models were estimated using R (R Core Team 2015) and its maxLik package (Henningsen *et al.* 2011).

8.3 Results

Results of the **choice-in-shelf stage** are presented in Table 8-3.

This stage reveals that the perception of a wine's quality varies significantly between younger consumers (millennials) and older ones. While most consumers have a strong positive perception of Carmenere grape variety, this is not so strong among millennials and higher income individuals. Unlike older consumer, millennials also dislike the Shiraz grape variety. And while older consumers prefer delicate (i.e. more traditional) labels, millennials seem more open to more modern (contrast) and figurative (natural) designs, though these trends are not significant at the 95% confidence level.

Table 8-3 – Estimates of the choice-in-shelf modelling stage. Parameters not significant at the 95% confidence level are presented in grey.

		Estimate	Std. error	t value
Grape Variety	Merlot	0.006	0.115	0.05
	Carmenere	0.438	0.167	2.62
	x Millennial	-0.372	0.182	-2.04
	x High income	-0.382	0.167	-2.28
	Shiraz	0.034	0.155	0.22
	x Millennial	-0.415	0.187	-2.21
Label design	Contrast	-0.278	0.134	-2.08
	x Millennial	0.237	0.158	1.50
	Natural	-0.213	0.140	-1.52
	x Millennial	0.208	0.165	1.26
Alcohol content		0.368	0.109	3.37
Price*	log(USD)	1.131	0.168	6.73
Discount	10%	0.048	0.103	0.47
	20%	0.135	0.102	1.32
Fit indices	Loglikelihood			-3079.95
	Number of parameters			20
	Number of observations			3120
	Number of individuals			130
	ρ^2			0.526
* before discount				

All consumers agree on higher alcohol content as a sign of higher quality, and with higher prices being a cue of quality. Quality perception, however, does not increase linearly with price, as the best fit is obtained using a logarithmic transform of price. This means that price increases provide decreasing quality perception improvements (e.g. a 10 USD wine is not perceived as twice as good as a 5 USD wine). Finally, as was to be expected, price discounts do not increase the quality perception, though higher discounts may make a wine more attractive to purchase.

Concerning the **tasting stage**, we did not observe significant correlation between the participants' liking and the wines' sensory description. Despite trying different numbers of principal components solution (from two to eight) and different rotations as explanatory variables, we did not obtain significant parameters. We also tested different model structures (random parameters, error components, systematic taste variations) and additional explanatory variables (grape variety, price and wine cluster based on their sensory profile), but could not obtain significant parameters.

The only variables useful in explaining participants' liking are the participant-provided adjectives, but these are endogenous. Participants seem to have divided the adjectives in two groups: positive (sweet, fruity, tastes like red berry, wood, chocolate) and negative (bitter, sour, dry, tastes like leather, tastes like alcohol, tastes flat, thick). Consistently with this classification, participants assigned positive adjectives to the wines they liked, and negative adjectives to the wines they did not liked. This means that adjectives are not reflecting the wines' sensory attributes, but instead the level of liking of consumers. This hypothesis is supported by the high correlation between overall liking and adjectives, and the lack of correlation between wines and adjectives across participants, i.e. the adjectives assigned to wines are not consistent between participants.

Therefore, we could not find models capable of explaining participants' level of liking of wines.

When modelling the **repurchase** stage, we explain the discrete choice part of the purchase (i.e. which wines to buy, the ψ_k parameter) using each wine's *expected quality* and *sensory perception*. As the first is derived only from extrinsic attributes, and the second only from intrinsic attributes, this allows us to compare the influence of both type of attributes in the purchase decision. Price also plays a negative role, due to the budget constrain considered in the MDCEV model. The positive effect of price -i.e. as a cue for quality- is already captured in the *expected quality*.

But instead of predicting each individual's *expected quality* and *sensory perception* for each wine, which would include a significant amount of error, we use the participant-provided *inverted ranking* and *overall liking* as indicators of *expected quality* and *sensory perception*, respectively. But as these two indicators are endogenous, we must use the Control Function approach (Petrin & Train 2010) to correct for this problem.

The Control Function approach requires us to perform two auxiliary linear regressions. The first one explains the *inverted ranking* based on the wine's observable attributes (i.e. we use the extrinsic attributes as instruments). The second linear regression explains *overall liking* based on the adjectives provide by consumers, i.e., we use the adjectives as instruments for *overall liking*. Finally, we must include the residual of both auxiliary regressions in the attractiveness (ψ_k) of each alternative. Therefore, we define the attractiveness of each bottle of wine as follows.

$$\psi_k = e^{\beta_0+\beta_q invRnk_k+\beta_s ove_k+\beta_{qr} resInv_k+\beta_{sr} resOve_k+\beta_I \ln(\frac{income}{nPeople})_k} \tag{8.5}$$

Where $invRnk_k$ is the *inverse ranking* position of wine k, ove_k is its *overall liking* (both expressed on a 1 to 7 Likert scale); $resInv_k$ and $resOve_k$ are the residuals from the *inverse ranking* and *overall liking* auxiliary regressions, respectively; and all β's are scalar parameters to be estimated. Among the parameters, β_q and β_s are the most relevant ones, as they are proxies for the impact of extrinsic and intrinsic attributes in the purchase decision. β_{qr} and β_{sr} are only necessary for the endogeneity correction and are of no particular interest, except for them having to be negative, as required by the Control Function approach. β_I represents how much the individual's household income influences the attractiveness of wine. This term is the same for every alternative (as it depends on the individual, not the alternative), but unlike in traditional choice models, this is not a problem in the MDCEV model. As each alternative's attractiveness is multiplied by a decreasing factor representing satiation, the common elements in the attractiveness do not cancel out, as long as the consumed amount varies among alternatives.

We tested several different definitions of the satiation parameter γ_k, but a single constant value provided the best fit and significance level. We tested definitions where satiation depended on income, age and other participant's characteristics, but to no avail. We also tested definitions where γ_k depended on alternative's attributes, but they were also ineffective. Therefore, we kept a single constant. This means that the satiation effect is the same for all participants and all products, suggesting a preference for variety instead of large amounts.

Parameter estimates, their standard errors or 95% confidence intervals, and fit indices of the MDCEV model are presented in Table 8-4. The only parameters not significant at the 95% confidence level are the residuals from the auxiliary linear regressions, but they exhibit the expected sign.

As expected, both the *inverse ranking* and the *overall liking* have a positive influence on the wines' attractiveness. The value of their parameters are proxies of the weight of the wine's *expected quality* and *sensory perception*. Interestingly, *sensory perception* seems to be twice as important as *expected quality*, implying that intrinsic attributes are twice as important as extrinsic attributes when repurchasing, at least when little time has passed since tasting the wine.

The significance of income in the attractiveness of alternatives and not in the satiation effect indicates that participants tended to favoured variety over bigger amounts of a single product. In other words, instead of buying several bottles of a single wine, participants were more inclined to buy several different bottles, each one of a different wine.

The outside good's α parameter, though significantly different from 1, is closer to 1 than 0. This goes against several other applications of the MDCEV model (e.g. Bhat 2008), where α tended to zero. However, the higher value of α has economic sense, as it implies low satiation of the outside good. This is in line with approximately 45% of the sample not buying any wine.

The low number of observations did not allow testing more sophisticated forms of the attractiveness and the γ_k parameter.

Table 8-4 – Estimates of the repurchase stage

		Estimate	Std. error	t value
Attractiveness $(\boldsymbol{\psi_k})$	Constant	-3.504	1.044	-3.36
	Inverse ranking	0.219	0.091	2.40
	Overall liking	0.424	0.084	5.05
	Residual Inv. Rnk.	-0.153	0.094	-1.63
	Residual Ove. Lik.	-0.098	0.068	-1.44
	ln(income/nPeople)	0.176	0.089	1.99
Satiation	γ	0.713	0.193	3.70
	γ_k	2.040	C.I. = [1.49;	2.8]
Outside good	α'	1.164	0.356	3.27
	α	0.762	C.I. = [0.64;	0.85]
Scale	σ'	-0.708	0.146	-4.86
	σ	0.493	C.I. = [0.39;	0.63]
Fit indices	LogLikelihood			-395.24
	Number of parameters			9
	Number of observations			130
	Number of individuals			130
	ρ^2			0.172
$\gamma_k = e^{\gamma}$; $\alpha = \frac{1}{1+e^{-\alpha'}}$; $\sigma = e^{\sigma'}$				

8.4 Discussion

While the proposed experimental design effectively captured the expected quality, the sensory perception and the trade-off between both of them and price, there are still several aspects than could be improved. First of all, participants did not perceive the difference between the registration and the choice-in-shelf stages. For future applications, it is recommended to (i) drop the ranking on the shelf stage in favour of a multi-purchase, i.e. participants can actually buy as many bottles as they want in front of the shelf; and (ii) make participants pay for their purchase in the shelf immediately to assure the compatibility of incentives. After the transaction, participants can be invited to the blind tasting.

Given the difficulty in finding correlation between wines' sensory profile, as developed by a trained panel (QDA), we encourage researchers to use different methods to profile a wine. Particular attention should be given to methods where the same consumers are the ones who develop the profile, instead of experts or trained panels. Also, we recommend using methods with no explicit enumeration of attributes is made. In our case, we used a loose implementation of the Check All That Apply (CATA) method, proposing to consumers a series of adjectives drawn from focus groups. However, the interpretation of these adjectives is so ambiguous that a stable profile could not be constructed.

A promising profiling method using regular consumers is Projective Mapping or Napping (Pagès 2005). In it, consumers are given a blank sheet, a set of products, and are asked to place the products in the sheet in a way that similar products are close, and dissimilar products are far away. This methods' main limitations are that it can be quite demanding for consumers; it requires several samples to be evaluated simultaneously (which can make it difficult to use with alcoholic products). Furthermore, unlike QDA, Napping does not provide a clear cut set of attributes, making it difficult to employ its output on econometric models such as discrete choice. More research on these are is much needed.

Despite these difficulties, some valuable results were be obtained from the data gathered. Concerning the formation of expected quality on the mind of the consumers, we were able to identify systematic differences between young (millennials) and older consumers, especially on their perception of grape varieties. Once again, we confirm the use of price as a cue for quality from part of the consumers, and we also find that discounts do not seem to influence consumers' perception of quality.

On the repurchase stage we discovered that the sensory liking is twice as relevant as the expected quality, at least shortly after tasting the product. This relevance could decrease as time passes and consumers may forget how much they liked a particular wine. However, this implies that a tasting on location can greatly incentivize sales if the product is liked by consumers. This result, however, can be hard to operationalize, as it is still hard to identify what makes a wine more likeable (for reasons already discussed).

The use of the MDCEV model allowed discovering the importance of variance seeking by consumers. They seem to appreciate diversity in their purchase over repeating purchase of a single wine. This is not a strategy to reduce risk, as in the experiments consumers had already tasted the wine, effectively removing the risk of choosing a wine they dislike. Therefore, this results points to a variety seeking behaviour on the studied consumers.

A limitation of our application was that we limited the number of potential purchases to three bottles. This restriction was not considered in the modelling, but it does not seem to be of much relevance giving the variety seeking behaviour of consumers. Nevertheless, we recommend not to impose this restriction on similar studies, even though there are extensions of the MDCEV model that can incorporate such restrictions (Castro *et al.* 2012).

In conclusion, the experimental design and the analysis that comes with it shows promising results, despite the difficulties in this particular application. The MDCEV proved to be a powerful and flexible tool, capable of revealing insight not accessible through more traditional model.