

When expensive is good: Addressing price-quality association in choice modelling.

David Palma*

Choice Modelling Centre and Institute for Transport Studies
University of Leeds
d.palma@leeds.ac.uk

Stephane Hess

Choice Modelling Centre and Institute for Transport Studies
University of Leeds

Cristian Angelo Guevara

Department of Transport, Faculty of Engineering
Universidad de Chile

Luis Ignacio Rizzi, Juan de Dios Ortúzar

Department of Transport and Logistics
Pontificia Universidad Católica de Chile

Gerard Casaubon

Center for Research and Innovation
Viña Concha y Toro

Acknowledgments

We are grateful to the Chilean Fund for the Development of Scientific and Technological Research (FONDECYT) through Projects 1121058 and 1150590. Thanks are also due to the Millennium Institute in Complex Engineering Systems (ICM: P05-004F; FONDECYT: FB016) and the European Research Council through the consolidator grant 615596-DECISIONS, for having partially financed this work.

* Corresponding author

When expensive is good: Addressing price-quality association in choice modelling

Abstract

Evidence in the literature points to consumers using price as a cue for quality when judging products. This leads to price having a double effect on the purchase probability: a positive effect due to its use as a cue for quality, and a negative effect due to consumers' budget constraints. This can induce endogeneity issues when modelling demand.

Based on existing behavioural models and econometric techniques, we propose and test three different approaches to correct for price endogeneity while maintaining behavioural interpretability of estimated parameters in choice models. The first and second approaches fully implement a behavioural model explaining the double effect of price, using Latent Variables and a Control Function, respectively. The third approach is based on the Multiple Indicator Solution and focuses only on the negative effect of price. We compare these three approaches to a traditional one estimating only the net effect of price.

Using data from a wine choice experiment ($n=122$), we conclude that estimating only the net effect of price can lead to biased results concerning consumers' price sensibility. We recommend using the Control Function approach when both the positive and negative effect of price are relevant, and the Multiple Indicator Solution when only the negative effect of price is of interest.

Keywords: price-quality, choice, endogeneity, wine, experience good, perceived quality, choice modelling

1 Introduction

1.1 The role of price in purchase decisions

From the perspective of consumers, the purchase decision can be conceived as a trade-off between *perceived cost* and *perceived quality* (Scitovsky 1945, Zeithaml 1988, Dodds et al. 1991, Grunert 2005), both of which are subjective concepts. Perceived quality can be defined as the superiority or excellence of a product (Zeithaml 1988), when compared against others. Consumers' perception of quality are heterogeneous among themselves, and may differ from those of manufacturers or any other "objective" definition of quality, such as product ratings. Perceived cost, on the other hand, includes the price of the product, as well as any other sacrifice (monetary or not) the consumer may have to incur in due to buying and consuming the product (e.g. travelling to the store, cooking the product, etc.). As our focus in this study is to understand the role of price in the purchase decision, from now on we will assume that perceived cost is fully determined by price alone. While perceived quality has a positive effect on the likelihood of buying a product, cost has a negative effect; yet price can influence both constructs.

Perceived quality is constructed based on product attributes (or cues), which are processed by consumers according to their own preferences and perception biases. Attributes can be classified in two groups: *extrinsic* and *intrinsic* (Olson & Jacoby 1972, Zeithmal 1988, Grunert 2005). Intrinsic attributes are those which cannot be changed without changing the product, as they are usually related to the product chemical and physical composition. Taste, aroma and colour are some of the most relevant intrinsic attributes of a product. Extrinsic attributes are those surrounding the product, and can be changed without altering the product itself, e.g.: price, packaging, advertising, etc. What specific attributes are relevant for determining the perceived

quality of a product will depend on the nature of the product (Zeithmal, 1988), still, a large body of literature shows that price is often an important determinant of perceived quality.

Price can have a positive effect on perceived quality, as many consumers believe that more expensive products are of higher quality. A large body of literature supports this finding. Scitovsky (1945) proposes that assuming a positive correlation between price and quality is reasonable under perfect information, but not in its absence. Leavitt (1954) shows (in a relatively small sample) that consumers do choose the more expensive products when they have trouble determining the quality of products. Rao & Monrow (1989) confirms the existence of price-quality association through meta-analysis. Dodds *et al.* (1991) use empirical data to prove that price positively influences perceived quality, and negatively influences willingness to buy. Yan & Sengupta (2011) show that price-quality association strengthens when consumers are unfamiliar with the product they are purchasing, are buying it for someone else, or lack information about it. Langhe *et al.* (2014) show that the strength of the price-quality association depends on how heterogeneous quality is perceived to be within price niches.

1.2 Price-quality association and price endogeneity

The use of price as a cue for quality, when not dealt with appropriately, can lead to specification issues. In the best case, it can lead to efficiency loss, and in the worst can lead to biased estimates. To see this more clearly, consider the following simple representation of intention to buy as a trade-off between perceived cost and quality.

$$ITB_{nj} = \beta_p h(\text{price}_j) + \beta_q q_{nj} + \varepsilon_{nj} \quad (1)$$

$$q_{nj} = \gamma_p g(\text{price}_j) + X_j \gamma_x + \omega_{nj} \quad (2)$$

where ITB_{nj} is consumer n 's intention to buy product j , $h(price_j)$ is a transformation of the price of product j , representing perceived cost; q_{nj} is product j 's quality, as perceived by consumer n ; $g(price_j)$ is a transformation of price; X_j is a vector of product's attributes (extrinsic and intrinsic); 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. We assume $\beta_p < 0$ and $\beta_q > 0$. Now suppose the modeller does not observe q_{nj} (as it happens in reality), and therefore estimates the following model:

$$ITB_{nj} = \alpha_p h(price_j) + X_j \alpha_x + \eta_{nj} \quad (3)$$

If $h = g$, then $\alpha_p = \beta_p(1 + \beta_q \gamma_p)$, $\alpha_x = \beta_q \gamma_x$ and $\eta_{nj} = \beta_q \omega_{nj} + \varepsilon_{nj}$, making the estimated α parameters consistent, as η_{nj} does not correlate with $price$ or X . But if $h \neq g$, then $\alpha_p = \beta_p$, $\alpha_x = \beta_q \gamma_x$ and $\eta_{nj} = \beta_q \gamma_p g(price_j) + \beta_q \omega_{nj} + \varepsilon_{nj}$, introducing correlation between $price$ and the error term η_{nj} , therefore causing an endogeneity problem leading to inconsistent estimators of the α parameters. In summary, endogeneity due to price-quality association will be present unless price influences perceived quality and cost in the same way. Unfortunately, this is often not the case.

The effect of price on perceived quality is most probably associated with marginally decreasing quality improvements (i.e. diminishing quality gains per unit price increases). 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, despite the price difference between them being the same. 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. All

these considerations make highly unlikely for both effects of price to have the same functional form.

1.3 Measuring the double effect of price

One way to solve the problem of price endogeneity is to separately model the positive and negative effect of price, allowing each effect to have a different functional form. However, measuring the two effects is not possible when we only observe consumers' purchase decision, as it is often assumed in most models of demand. To estimate both effects we need information about the trade-off between quality and price (provided by the choice), but also about the perception of quality alone, to isolate the effect of price as a cue for quality.

In this study, we propose the combined use of indicators of quality and choice data to measure both the positive and negative effect of price on hypothetical purchase decisions of wine. Even though the double effect of price is well documented, to the best of our knowledge there have been few attempts to consider this effect when modelling demand. Furthermore, in the particular case of wine, all proposed methods have important shortcomings (Lockshin *et al.* 2006, Mastrobuoni *et al.* 2014, discussed in section 2).

We test three different approaches to control for price's double effect. The first approach uses Latent Variables (LV, Guevara 2015) to explicitly model perceived quality. The second approach uses the indicators of quality provided by consumers and the Control Function Technique (Petrin & Train 2010; Villas-Boas & Winer 1999) to separate both effects. The third approach only aims at capturing the negative effect of price using the Multiple Indicator Solution method (MIS, Guevara & Polanco 2016), a special form of the Control Function technique. We compare these three models with a base model where we only measure the net effect of price.

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 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 report the sample's main characteristics, present the design of the experiment, and describe the statistical models used to analyse the data. We present results for all estimated models in section 4, and in section 5 we discuss them and propose further topics of research.

2 Literature review

The importance of price as a cue for quality has been noted before in the wine preference 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, showing higher activity in zones associated with pleasure when tasting wines tagged as more expensive. Lewis & Zalan (2014) found higher prices to increase both self-reported enjoyment and willingness to pay. Aqueveque (2006; 2008) found that higher prices reduced perceived risk and increased perceived quality among wine consumers, but that this relationship weakened in presence of expert ratings. Kardes *et al.* (2004) also found price-quality association to decrease as consumers have more information available, and are not in a hurry to choose.

From a modelling perspective, Lockshin *et al.* (2006) discovered that the effect of price on the (hypothetical) purchase probability had an inverted U shape, i.e. an increase in price raised

purchase probability for cheap wines, but diminished it for expensive wines. They found the inflexion point 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, finding the inflexion point at 7 US\$. Remaud *et al.* (2008) and Mueller *et al.* (2010b) used dummy variables to measure the effect of price, also finding the inverted U shape for the effect of price for at least some of the consumers in their samples.

The inverted U shaped effect of price on the likelihood to buy can be explained by considering the double effect of price. When a product is cheap, the net effect of a price increase on purchase likelihood is positive: a small increase in price can be cue for an important quality improvement (as the product is no longer classified as "cheap"), while the price increase is not big enough to discourage purchase. On the other hand, when an expensive product increases its price, the net effect is negative, as gains in quality are marginal (because the wine was already perceived as good) but the budget constraint gets even more binding. This explanation reinforces the hypothesis that at least one of the effects of price is non-linear.

Other studies have ignored the double effect of price, but still found a decreasing and monotonic effect on purchase probability, just as traditional economic theory proposes. However, most of these findings may be due to their experimental designs including prices only in the decreasing part of the inverted 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.* (2017) 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 fall on the decreasing part of the inverted U-shaped curve described by Mtimet & Albisú (2006), who also studied the Spanish market.

Separately 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 experts' ratings (Wine Spectator's scores) as a proxy for perceived quality. However, perceived quality is a subjective assessment, and there is ample literature showing that experts' ratings do not correlate with consumers' perception, nor do they measure an underlying objective quality (Lawless 1984, Hodgson 2009, Lattey *et al.* 2009, Gokcekus & Nottebaum 2011, D'Alessandro & Pecotich 2013, Hopfer & Heymann 2014).

Mastrobuoni *et al.* (2014) explicitly measured the positive and negative effects of price using a two-stage approach. In a Stated Choice experiment with tasting, they asked participants to choose both the best wine among the alternatives (stage 1), and the one they would actually buy (stage 2). They modelled perceived quality using responses from the first stage, and explained the purchase decision as a trade-off between price and the estimated perceived quality. As the experiment included tasting but participants did not provide a tasting evaluation, authors used experts' ratings as a proxy for participants sensory liking which, as already mentioned, is a questionable approach.

Palma *et al.* (2016) used a similar but more robust approach than Mastrobuoni *et al.* (2014), by using a Likert Scale type of indicator for perceived quality instead of single choice. The Latent variable approach presented in this paper expands the work done by Palma *et al.* (2016) by also considering the effect of intrinsic variables.

3 Material and Methods

3.1 Sample

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 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, its composition successfully represents the high-income millennial target market of the wines under examination.

3.2 Survey design

Our study aimed to model 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 data, i.e. actual purchase records. In particular, its ease of collection of multiple choices from each respondent, 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 hypothetical bias (Ortúzar and Willumsen, Chapter 3), although the mental processes that drive decision making are thought to be the same (Louviere *et al.* 2000).

Table 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
Monthly Household Income (in USD)	Incomplete postgrad	2	2	4
	Complete postgrad	16	10	26
	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
Consuming Frequency	2251 to 3107	11	6	17
	3108 and more	23	18	41
	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
Every day or almost every day	1	1	2	
Without information	2	0	2	

The experiment was divided in five stages (Figure 1). In the first, participants were asked to blindly taste five different wine samples, one at a time in a randomized order, each of them uniquely identified by a three-digit number (e.g. “642”). Participants were asked to provide two indicators of their sensory liking for each sample: 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

assessment of their willingness to pay for a bottle of that wine sample using a 5-point Likert scale (from “I would pay much less than I am used to” to “I would pay much more than I am used to”).

Participants then faced a SC exercise during the second stage of the experiment (Figure 2). In a screen, we presented to participants the bottle label and price (in Chilean pesos, Ch\$) of each of the tasted samples. Participants could easily associate each sample with its price and label through its code, e.g. in Figure 2 the first label corresponds to sample “642”. As the samples were still on each participant’s table, they were able to try them again if they wanted to.

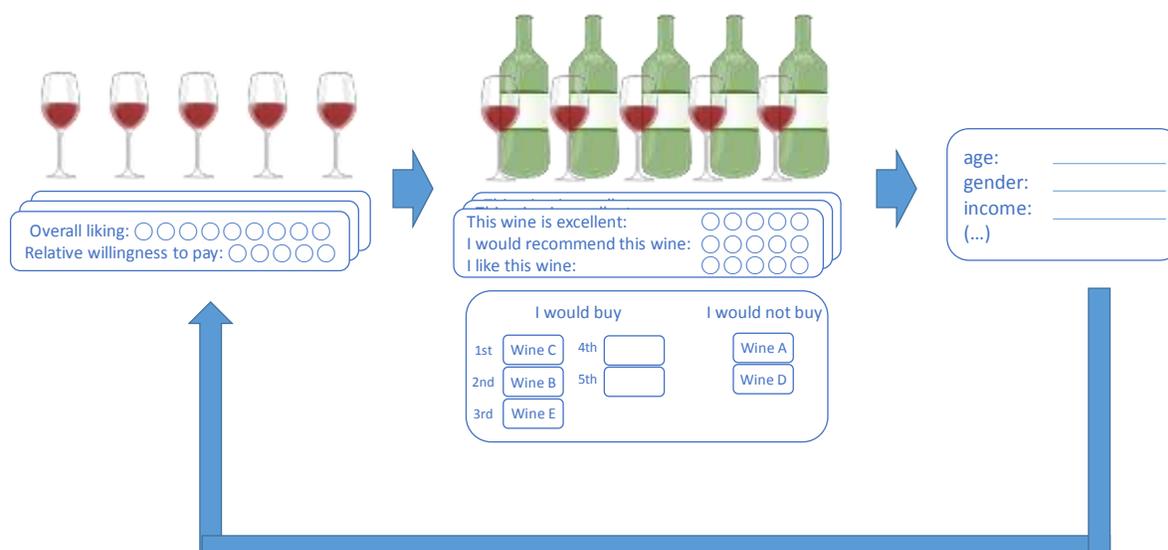


Figure 1 – Stages of the experiment. The first and second stage were repeated.
Glass of wine by rinze, Wine bottle by Solarisphere (openclipart.org) (cc)

The SC exercise had two parts. First, participants had to provide three indicators of quality for each sample, by reporting their level of agreement with the statements “This wine is excellent”, “I would recommend this wine to my friends”, and “I like this wine”. They did so using 5-point Likert scale from “I disagree completely” to “I agree completely”.

	Wine A	Wine B	Wine C	Wine D	Wine E
Wine	642	267	761	558	359
Bottle					
Price	\$ 6.990	\$ 1.300	\$ 2.990	\$ 2.990	\$ 2.990

12. Using a scale from 1 to 5 stars, where 1 star means "I disagree completely" and 5 stars mean "I agree completely", report your level of agreement with the following phrases (We recommend you answering by wine). *

	This wine is excellent	I would recommend this wine to my friends	I like this wine
Wine A	✘ ★ ★ ★ ★ ★	✘ ★ ★ ★ ★ ★	✘ ★ ★ ★ ★ ★
Wine B	✘ ★ ★ ★ ★ ★	✘ ★ ★ ★ ★ ★	✘ ★ ★ ★ ★ ★
Wine C	✘ ★ ★ ★ ★ ★	✘ ★ ★ ★ ★ ★	✘ ★ ★ ★ ★ ★
Wine D	✘ ★ ★ ★ ★ ★	✘ ★ ★ ★ ★ ★	✘ ★ ★ ★ ★ ★
Wine E	✘ ★ ★ ★ ★ ★	✘ ★ ★ ★ ★ ★	✘ ★ ★ ★ ★ ★

13. If you went to a party at a friend's house, which wine would you buy? Please make a ranking of the wines above considering all you know about them. When you are done making your ranking, choose the alternative "I am not attracted by other alternatives". *
Drag items from the left-hand list into the right-hand list to order them.

Wine A	➡	1. Wine C	⬇
Wine D	➡	2. Wine B	⬇
4. I am not attracted by other alternatives	➡	3. Wine E	⬇

Figure 2 – Example of a SC exercise (labels are not the real ones used on the experiment)

After providing the three indicators of quality for each sample, 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. This also served as a forced pause during the experiment, as stages fourth and fifth were repetitions of the first and second stages, but with samples, labels and prices scrambled. As different sample identification codes were used in the repetition of stages one and two, participants were induced to think they were tasting different samples. The repetition of these stages was to increase the number of observations. In this paper we do not investigate the consistency of preferences across repetitions of the SC exercise.

The SC exercise followed a D-efficient design (Rose & Bliemer 2009; 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. The first scenario was used on the second stage of the experiment, and the second scenario in the fifth stage. 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.

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 winemaker who partially financed the study. The relevance of label and price is well established in the wine marketing literature (see, for example, the review by Lockshin & Corsi 2012). The relevance of flavours is less clear. While the relevance of overall sensory liking is well acknowledged (Siegrist & Cousin 2009, Goodman 2009, Combris *et al.* 2009, Mueller *et al.* 2010a), measuring the relevance of individual flavours and aromas is difficult. A sign of this difficulty is that literature has mainly focused on the effect of simple flavours, such as sugar and fat (Enneking *et al.* 2007,

Hopper *et al.* 2012), but not on other more complex flavours and aromas, such as wood, butter or vanilla.

Five different wine samples 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 samples (as in logit models only differences in alternatives attributes matter). This made a full sensory description of wines by a trained panel unnecessary. The considered attributes and their levels are shown in Table 2, though most levels are not reported by request of the private winemaker who funded part of the study. We restrained the experiment to only three attributes to avoid requiring an excessively large sample of respondents.

Table 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

3.3 Modelling

Figure 3 presents the behavioural model guiding our work. Sensory quality represents how much the consumer likes the taste and aroma of the wine. Perceived quality represents the subjective quality of the product according to the consumer, and Intention to buy represents the trade-off between perceived quality and price that will determine the consumer's purchase decision, it is assumed to be negatively affected by the direct effect of price and positively affected by perceived quality. Following Grunert (2005), we assume the influence of intrinsic attributes on

perceived quality to be mediated by sensory quality. Sensory quality, perceived quality and intention to buy are mental constructs of the consumer, and therefore they are not observed.

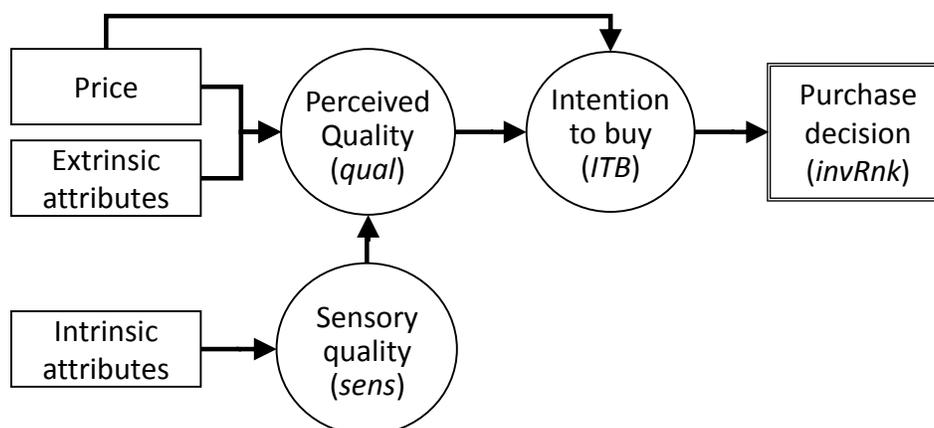


Figure 3 - Behavioural model.

Figure 3 does not represent the first time consumers buy a product, as they do not perceive the product intrinsic attributes then. Instead, it represents the second and subsequent purchases of a product, with the restriction of assuming sensory quality to be independent of extrinsic attributes. Even though this restriction is most likely false (Asioli *et al.* 2017), it fits the design of our experiment, and it could be easily relaxed if other sources of data are used.

We estimated four different models based on Figure 3. The first three models consider the double effect of price in increasingly simpler ways, while the fourth is a base model that only measures the net effect of price. All models were estimated using R (R Core Team 2015) and its maxLik package (Henningsen *et al.* 2011). In the remaining of this section, each model is presented and described in detail.

3.3.1 The Latent Variable (LV) model

The LV model is a full implementation of the behavioural model in Figure 3, through a Hybrid Choice Model (Ortúzar and Willumsen, 2011, section 8.4.3; Bolduc & Alvarez-Daziano 2010).

This model is designed to capture the double effect of price, and uses latent variables to model all unobservable elements: *sensory quality* (*sens*), *perceived quality* (*qual*) and *intention to buy* (*ITB*).

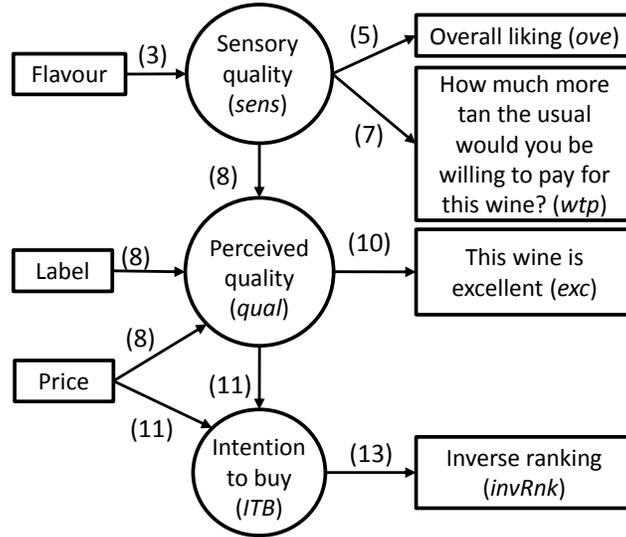


Figure 4 - Structure of the HC model. Number in parenthesis refer to equations.

Figure 4 shows the implementation of Figure 3's behavioural model as a structural equation model (SEM). We used linear structural equations and ordered logit models (Train 2009, chapter 7.3) as links between the latent variables and their indicators.

Equations (3) to (9) describe the modelling of *sensory quality* as a latent variable. Equation (3) is *sens*' structural equation, while equations (4) to (7) are its measurement equations. As the link between *sens* and its indicators *ove* and *wtp* are ordered logits, we write their conditional probabilities in the form of equations (4) and (6). We obtain their unconditional probabilities (equations 5 and 7) by integrating over *sens*' error component.

$$sens_{int} = (1 + \kappa_{sens} I_{t=2})(F_{it}\gamma_F + F_{it}Cl_n\gamma_{FCl} + \omega_n) \quad (3)$$

$$P(ove_{int} = s | sens_{int}) = \frac{1}{1 + e^{\lambda_{ove} sens_{int} - \tau_s^{ove}}} - \frac{1}{1 + e^{\lambda_{ove} sens_{int} - \tau_{s-1}^{ove}}} \quad (4)$$

$$P(ove_{int} = s) = \int_{\omega_n} P(ove_{int} = s | sens_{int}) \phi(\omega_n | 0, 1) d\omega_n \quad (5)$$

$$P(wtp_{int} = s | sens_{int}) = \frac{1}{1 + e^{\lambda wtp_{int} sens_{int} - \tau_s^{wtp}}} - \frac{1}{1 + e^{\lambda wtp_{int} sens_{int} - \tau_{s-1}^{wtp}}} \quad (6)$$

$$P(wtp_{int} = s) = \int_{\omega_n} P(wtp_{int} = s | sens_{int}) \phi(\omega_n | 0, 1) d\omega_n \quad (7)$$

In equation (5), $sens_{int}$ is the value of sensory quality 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; Cl_n is a dummy variable taking value 1 if participant n belongs to “cluster 2”, a subset of participants identified to have similar sensory preferences; ω_n is an independent identically distributed (iid) normal random error with mean 0 and standard deviation equal to 1. This normalization is necessary to identify the model. ω_n captures the unobserved (by the modeller) determinants of *sensory perception*, and correlates observations from the same individual (pseudo panel effect, as proposed by Daly & Hess 2010). The scalar κ_{sens} and the vector γ_F are parameters to be estimated. κ_{sens} 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_F represent participants average preferences for the added flavours.

Cl_n is an auxiliary variable used exclusively to improve the fit of the sensory quality model. Using the k-means technique, we clustered participants based on their answers to the sensory indicators (*ove* and *wtp*) for the first set of wines. We identified two clusters. The variable Cl_n takes the value 1 for those participants classified in the second cluster. We used this auxiliary variable because sensory preferences are extremely heterogenous among participants, to the point

of making this part of the model futile if F_{it} were the only explanatory variables. We tested using participants' sociodemographic characteristics as additional factors to improve the fit of the model, but to no avail. We used the Cl_n variable, despite it being endogenous, because it was the simpler way to improve fit, and while it may bias the parameters of *flavour*, we are not interested in them in this analysis. A more correct approach would require doing a latent class modelling for sensory quality. Such an approach would provide similar fit at the cost of only one more parameter, but it can significantly increase computation time and makes for a much more complex model exposition.

Equations (4) and (6) describe the conditional probabilities of observing a level s for *ove* (overall liking) and *wtp* (the answer to “How much would you be willing to pay for this wine”) when the value of $sens_{int}$ (including ω_n) is given. λ_{ove} and λ_{wtp} are parameters to be estimated. They measure the strength and direction of the relationship between the latent variable $sens$ and its indicators. If these parameters are positive and significant, it means that *ove* and *wtp* are suitable indicators for $sens$. Conditional probabilities depend on thresholds $\{\tau_0^{ove}, \dots, \tau_9^{ove}$ and $\tau_0^{wtp}, \dots, \tau_5^{wtp}\}$, which 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). Equations (5) and (7) are the unconditional probabilities of observing level s of indicator *ove* and *wtp*, respectively. The integrals are necessary to average out the random component ω_n in $sens_{int}$. $\phi(\cdot | \mu, \sigma^2)$ represents the normal distribution probability density function with mean μ and variance σ^2 .

Equation (8) is the structural equation of perceived quality, and equations (9) and (10) are their measurement equations.

$$qual_{int} = (1 + \kappa_{qual} I_{t=2})(L_{it}\gamma_L + \gamma_p \ln(p_{it}) + \gamma_{sens} sens_{int} + \psi_n) \quad (8)$$

$$P(exc_{int} = q | qual_{int}) = \frac{1}{1 + e^{\lambda_{exc} qual_{int} - \tau_q^{exc}}} - \frac{1}{1 + e^{\lambda_{exc} qual_{int} - \tau_{q-1}^{exc}}} \quad (9)$$

$$P(exc_{int} = q) = \int_{\omega_n, \psi_n} P(exc_{int} = q | qual_{int}) \phi(\omega_{int} | 0, 1) \phi(\psi_n | 0, 1) d\omega_n d\psi_n \quad (10)$$

In equation (8), $qual_{int}$ is the value of the perceived quality latent variable for alternative i of SC exercise t for participant n ; κ_{qual} is a scale factor to be estimated, L_{it} is a vector of dummy variables indicating the *label* of the alternative; p_{it} is the price of the alternative, and $\ln(p_{it})$ is its logarithmic transformation to capture the positive effect of price (we tested several other transformations, but the logarithm fitted the data better); ψ_n is an iid normal random error with mean zero and unit variance (a restriction necessary for identification). γ_L , γ_p and γ_{sens} are parameters to be estimated, with γ_L being a vector of parameters, one for each label.

Equation (9) shows the conditional probability of observing a level of agreement q with the phrase “I think this wine is excellent” (exc_{int}), given a known value of $qual$. We did not use any other available indicators for perceived quality (levels of agreement with “I like this wine” and “I would recommend this wine to my friends”), as they performed poorly. λ_{exc} is a parameter to be estimated, so are thresholds $\tau_1^{exc}, \dots, \tau_4^{exc}$. However, $\tau_0^{exc} = -\infty$ and $\tau_5^{exc} = +\infty$ are fixed for identification purposes. Equation (10) shows the unconditional form of the same probability, which requires integrating over both ω_n and ψ_n .

Concerning *intention to buy*, we did not use participants’ purchase ranking as an indicator, but instead we used a transformation of it that we called *inverted ranking* ($invRnk$). The transformation consisted in assigning an $invRnk=0$ to all alternatives excluded from a participant’s purchase ranking. Then, we assigned $invRnk=1$ to the last alternative in a participant

ranking, $invRnk=2$ to the second one from bottom-to-top, $invRnk=3$ to the third alternative from bottom-to-top, and so forth until there were no alternatives left. Figure 5 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). This transformation has the benefit of associating a higher intention to buy to a higher inverted ranking, making all parameters in the model easier to interpret.

Equations (11) to (13) describe the modelling of intention to buy (ITB).

$$ITB_{int} = (1 + \kappa_{itb} I_{t=2})(\beta_q qual_{int} + \beta_p p_{it}) \quad (11)$$

$$P(invRnk_{int} = r | ITB_{int}) = \left(\frac{1}{1 + e^{ITB_{int} - \tau_r^{invRnk}}} - \frac{1}{1 + e^{ITB_{int} - \tau_{r-1}^{invRnk}}} \right) \quad (12)$$

$$P(invRnk_{int} = r) = \int_{\omega_n, \psi_n} P(invRnk_{int} = r | ITB_{int}) \phi(\omega_n | 0, 1) \phi(\psi_n | 0, 1) d\omega_n d\psi_n \quad (13)$$

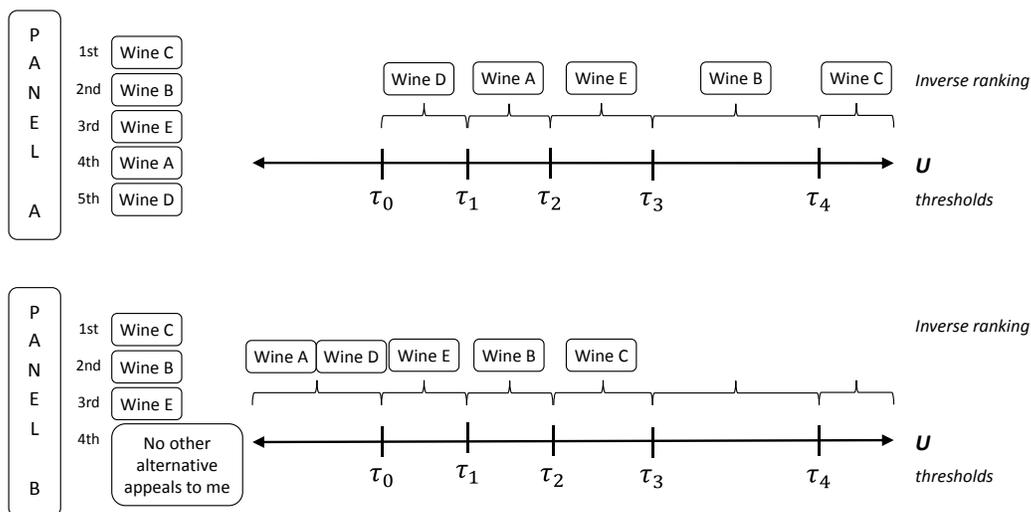


Figure 5 - Structure for modelling rankings made by SC respondents. Rankings constructed by consumers are on the left, and their corresponding "inverse rankings" on the right.

Equation (11) presents intention to buy structural equation, ITB_{int} is participant n 's intention to buy alternative i in SC situation t ; κ_{itb} is a scale factor to be estimated. β_q and β_p are parameters to be estimated, theory leads us to expect $\beta_q > 0$ and $\beta_p < 0$. We used a linear trade-off between price and quality, as advised by Hagerty (1978) and Levin & Johnson (1984).

Equation (12) is the conditional probability of participant n setting alternative i of SC situation t in the *inverted ranking* position r , given ITB_{int} . Thresholds $\tau_0^{invRnk}, \dots, \tau_4^{invRnk}$ must be estimated, while $\tau_{-1}^{invRnk} = -\infty$ and $\tau_5^{invRnk} = +\infty$ are fixed to allow model identification. Equation (13) presents the unconditional form of the same probability. It requires integrating both random components in ITB_{int} : ω_n and ψ_n .

All integrals in equations (5), (7), (10) and (13) do not have a closed analytical form, therefore they were solved using Monte Carlo methods (Train 2009, chapters 9 and 10). To do this, a big number (K) of random points from $\phi(0,1)$ are drawn, called ω_n^k and ψ_n^k , respectively. Then the conditional probabilities (equations 4, 6, 9 and 12) are calculated for each point k . The averages of all those evaluations are consistent estimators of the integrals. We use the Modified Latin Hypercube Sampling method (Hess et al. 2006) to construct the random draws k .

Equation (14) presents the likelihood of the LV model. We estimated the model simultaneously, using Simulated Full Information Maximum Likelihood.

$$P_{int}(\omega_n, \psi_n) = P(ove_{int} | sens_{int})P(wtp_{int} | sens_{int})P(exc_{int} | qual_{int})P(invRnk_{int} | qual_{int})$$

$$L_{LV} = \prod_n \int_{\omega_n, \psi_n} \prod_t \prod_i P_{int}(\omega_n, \psi_n) \phi(\omega_n | 0,1) \phi(\psi_n | 0,1) d\omega_n d\psi_n \quad (14)$$

3.3.2 Control Function (CF) model

This model -just like the LV model- fully reproduces Figure 3's behavioural model, and therefore captures both the positive and negative effect of price. But unlike the LV model, which uses latent variables to represent the unobserved constructs, the CF model replaces them by its indicators. Indicators, however, are endogenous (Guevara 2015). To see this, consider the following simple equation system linking an unobserved variable q , an indicator of it IND , and an outcome of interest OUT .

$$IND = \lambda_0 + \lambda_q q + \varepsilon_{IND} \quad (15)$$

$$OUT = \beta_0 + \beta_q q + \beta_p p + \varepsilon_{OUT} \quad (16)$$

where λ_0 , λ_q , β_0 , β_q , and β_p are parameter; q is an unobserved explanatory variable, and p is an observed explanatory variable. ε_I and ε_{ITB} are error terms. As the modeller does not observe q , he instead estimates $OUT = \alpha_0 + \alpha_q IND + \alpha_p p + u_{OUT}$, with u_{OUT} the new error term. Then, $\alpha_0 = \beta_0$; $\alpha_q = \frac{\beta_q}{\lambda_q}$; $\alpha_p = \beta_p$; and $u_{OUT} = \varepsilon_{OUT} - \frac{\beta_q}{\lambda_q}(\lambda_0 + \varepsilon_{IND})$. As ε_{IND} correlates with IND , IND is endogenous in the estimated model.

To control for indicators endogeneity, we use the Control Function Approach (Petrin & Train 2010). This approach has two stages: (i) regress the endogenous variable on a set of instruments, and (ii) introduce the residual of the first stage in the regression of interest. Instruments used in the first stage must be independent of the error term of the second stage regression, but at the same time they must correlate with the endogenous variable. The idea behind this approach is that the first stage residuals capture the endogenous part of the endogenous variable, therefore including them in the second stage regression allows us to literally control for the source of endogeneity.

The CF model replaces both sensory quality and perceived quality by one of their indicators, and must therefore apply the Control Function Approach twice. Figure 6 presents the structure of the model.

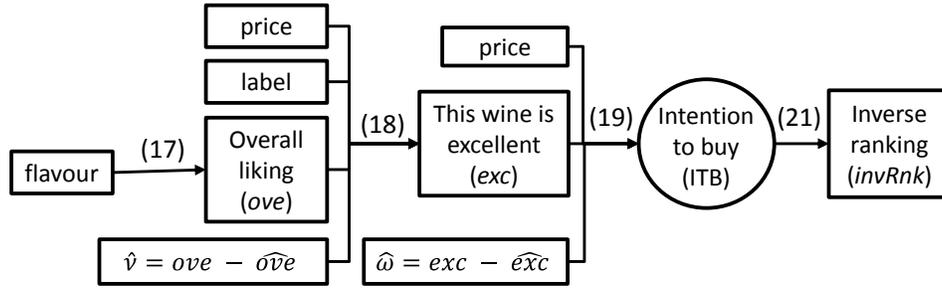


Figure 6 – CF model structure. Number is parenthesis refer to equations.

In practical terms, the CF model is divided in three stages. First we estimate a linear regression with *ove* as the dependent variable and *flavours* as the explanatory variables (equation 17). The second stage is another linear regression, this time with *exc* as the dependent variable, and *price*, *label*, *ove* and the residual of the first stage as explanatory variables (equation 18). The third stage is an ordered logit with intention to buy explaining inverse ranking, just as in the LV model. Intention to buy is a linear function of *price*, *exc* and the residual of the second stage (equation 19), its measurement are presented in equation (20) and (21). While the second stage captures the positive effect of price, the third stage captures its negative effect.

$$ove_{int} = \alpha_0 + F_{it}\alpha_F + F_{it}\alpha_{F2}Cl_n + v_{int} \quad (17)$$

$$exc_{int} = \gamma_0 + \gamma_p \ln(p_{it}) + L_{it}\gamma_L + \gamma_{ove}ove_{int} + \gamma_v \hat{v}_{int} + \omega_{int} \quad (18)$$

$$ITB_{int} = (1 + \kappa_{ITB}I_{t=2})(\beta_p p_{it} + \beta_{exc}exc_{int} + \beta_\omega \hat{\omega}_{int} + \psi_n) \quad (19)$$

$$P(invRnk_{int} = r | ITB_{int}) = \left(\frac{1}{1 + e^{ITB_{int} - \tau_r}} - \frac{1}{1 + e^{ITB_{int} - \tau_{r-1}}} \right) \quad (20)$$

$$P(invRnk_{int} = r) = \int_{\psi_n} P(invRnk_{int} = r | ITB_{int}) \phi(\psi_n | 0, \sigma_{ITB}) d\psi_n \quad (21)$$

Equation (17) is a linear regression instrumenting *ove*. v_{int} is an iid normal error component with mean 0 and standard deviation σ_{ove} to be estimated; α_0 , α_F , and α_{F2} are parameters to be estimated. All other variables are described in section 3.3.1. The CF approach, as described by Petrin & Train (2010) requires the instrumenting regression to include all controls used in the main regression. In our case, this would imply including all explanatory variables of equation (18) in equation (17). However, explaining participants' sensory perception based on labels consumer had not even seeing at the time is not reasonable. Therefore, we only include extra controls if they are reasonable from a practical and theoretical perspective.

Equation (18) is a linear regression instrumenting *exc*. We attempt to capture the positive effect of price (p_{it}) in this equation. \hat{v}_{int} is the residual of equation (17), and ω_{int} is an iid normal error component with mean 0 and standard deviation σ_{exc} to be estimated. All γ parameters are to be estimated.

Equation (19) describes consumer n 's intention to buy alternative i in SC situation t (ITB_{int}). κ_{ITB} is a scale factor for the second SC situation; $\hat{\omega}_{int}$ is the residual of equation (18); ψ_n is a normal error component with mean 0 and standard deviation σ_{ITB} to be estimated, used to introduce correlation between all observations of the same participant (Daly & Hess 2010). All β parameters are to be estimated.

Equation (20) is the conditional probability of observing alternative i in SC situation t , ranked r in the *inverted ranking* by participant n , given ITB_{int} . τ_0 to τ_4 are thresholds to be estimated, while τ_{-1} and τ_5 are fixed to $-\infty$ and $+\infty$ respectively for identification purposes. Equation (21) is the unconditional form of the same probability, where ψ_n must be integrated over. As this integral does not have a closed analytical form, we used MonteCarlo Integration to estimate it.

Even though this model can be estimated sequentially, i.e. one stage at a time, we estimated it simultaneously using Simulated Full Information Maximum Likelihood. This technique provides accurate standard errors. Equation (22) shows the CF model likelihood function.

$$L_{CF} = \prod_n \int \prod_t \prod_i \phi\left(\frac{\hat{v}_n}{\sigma_{ove}} \middle| 0,1\right) \phi\left(\frac{\hat{\omega}_n}{\sigma_{exc}} \middle| 0,1\right) P(invRnk_{int} | ITB_{int}) \phi(\psi_n | 0, \sigma_{ITB}) d\psi_n \quad (22)$$

3.3.3 The MIS model

The MIS model reproduces only part of Figure 3's behavioural model. It focuses in capturing the negative effect of price, therefore modelling only the intention to buy. Even though this model does not capture the positive effect of price, it does capture its negative effect consistently. In a similar way to the last stage of the CF model, the MIS model replaces the construct perceived quality by one of its indicators (*exc*). As the indicator is endogenous, it uses the Multiple Indicator Solution approach (MIS, Guevara & Polanco 2016) to control for endogeneity.

The MIS approach is a special implementation of the CF approach for indicator variables. It simply states that when using the CF approach, a good instrument for an indicator is a second indicator of the same unobserved variable. A second indicator is correlated with the first indicator because both are caused by the same unobserved variable, but it is uncorrelated to the error term of the first indicator as long as both indicators were collected independently. This makes the second indicator a good instrument for the first one.

Figure 7 present the MIS model structure, and equations (23) to (27) describe the model.

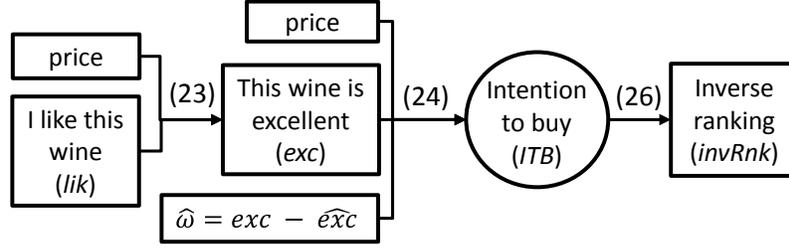


Figure 7 - MIS model structure. Numbers in parenthesis refer to equations.

$$exc_{int} = \gamma_0 + \gamma_p p_{it} + \gamma_{lik} lik_{int} + \omega_{int} \quad (23)$$

$$ITB_{int} = (1 + \kappa_{ITB} I_{t=2}) (\beta_p p_{it} + \beta_{exc} exc_{int} + \beta_\omega \hat{\omega}_{int} + \psi_n) \quad (24)$$

$$P(invRnk_{int} = r | ITB_{int}) = \left(\frac{1}{1 + e^{ITB_{int} - \tau_r}} - \frac{1}{1 + e^{ITB_{int} - \tau_{r-1}}} \right) \quad (25)$$

$$P(invRnk_{int} = r) = \int_{\psi_n} P(invRnk_{int} = r | ITB_{int}) \phi(\psi_n | 0, \sigma_{ITB}) d\psi_n \quad (26)$$

$$L_{MIS} = \prod_n \int_{\psi_n} \prod_t \prod_i \phi\left(\frac{\hat{\omega}_n}{\sigma_{exc}} \middle| 0, 1\right) P(invRnk_{int} | ITB_{int}) \phi(\psi_n | 0, \sigma_{ITB}) d\psi_n \quad (27)$$

Equation (24) is a linear regression and is the first stage of the MIS approach. lik_{int} is participant n 's level of agreement with the phrase “I like this wine” concerning alternative i in SC situation t , i.e. it is an indicator of perceived quality, just as exc ; ω_{int} is a iid normal error component with mean 0 and standard deviation σ_{exc} to be estimated. All parameters γ are to be estimated. Equations (24) to (26) are analogous to equations (19) to (21) of the CF model. Equation (27) is the MIS model likelihood function for all participants.

The integrals in equations (26) and (27) were calculated using Monte Carlo methods. Even though this model can be estimated sequentially, we used the simultaneous approach as the latter provides accurate standard errors.

3.3.4 The OL model

This model does not attempt to reproduce Figure 3's behavioural model, instead adopting a more traditional approach where all attributes (*extrinsic*, *intrinsic* and *price*) influence *intention to buy* directly. This model attempts to only estimate the net effect of price. Figure 8 present this model's structure.

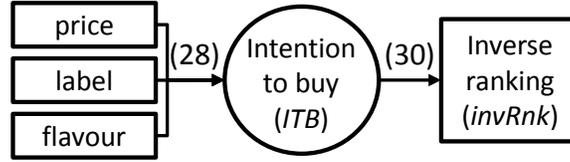


Figure 8 - OL model structure

Equation (29) present *intention to buy*'s functional form. κ_{ITB} is a scale parameter for the second SC situation; ψ_n is a normal random error component with mean 0 and standard deviation σ_{ITB} to be estimated, whose objective is to correlate observations by the same respondent (Daly & Hess 2010). Equations (29) and (30) are the conditional and unconditional probabilities of observing alternative i ranked r in the *inversed ranking* of respondent n in SC situation t . Equation (32) is the likelihood of the OL model for all participants. We estimated the OL model using Simulated Maximum Likelihood (i.e. we used Monte Carlo methods to calculate integrals without closed form).

$$ITB_{int} = (1 + \kappa_{ITB}I_{t=2})(\beta_p p_{it} + L_{it}\beta_L + F_{it}\beta_F + F_{it}\beta_{F2}Cl_n + \psi_n) \quad (28)$$

$$P(invRnk_{int} = r | ITB_{int}) = \left(\frac{1}{1 + e^{ITB_{int} - \tau_r}} - \frac{1}{1 + e^{ITB_{int} - \tau_{r-1}}} \right) \quad (29)$$

$$P(invRnk_{int} = r) = \int_{\psi_n} P(invRnk_{int} = r | ITB_{int}) \phi(\psi_n | 0, \sigma_{ITB}) d\psi_n \quad (30)$$

$$L_{OL} = \prod_n \int_{\psi_n} \prod_t \prod_i P(invRnk_{int} | ITB_{int}) \phi(\psi_n | 0, \sigma_{ITB}) d\psi_n \quad (31)$$

4 Results

Results from each of the four estimated models are presented in this section. The first two models (LV and CF) measure both the positive and negative effects of price, the third model (MIS) measures only the negative effect, and the fourth (OL) captures only its net effect, acting as a base model for comparison purposes.

Estimated parameters and goodness of fit indicators for each model are presented in Table 3 to 6. 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 (i.e. the part of the model concerned with *ITB* alone) can be compared directly with the others.

Preferences for added flavours and labels are fairly aligned between models. The first cluster of participants is insensitive to flavour changes, except for flavour 3, which they dislike. The second cluster dislikes all added flavour, always preferring the unaltered wine. Concerning preferences for labels, no model captures any significant effect. This is probably due to high preference heterogeneity as an (unreported) model with random coefficients showed no significant means but significant standard deviations for the effect of labels.

The LV, CF and MIS models work as expected. Both the LV and CF models successfully capture the positive and negative effect of price, with a positive and significant coefficient for price on perceived quality and a negative and significant one in intention to buy. The MIS model effectively captures price negative effect, with a negative and significant coefficient in intention to buy. The OL model, instead, does not achieve significance for its only price coefficient, even after attempting several transformations of it.

Table 3 – LV model coefficients and goodness of fit indicators (thresholds not reported)

		Coefficient	Robust t-test
Sensory quality	Flavour 1 Cluster 1	0.328	1.55
	Cluster 2	-2.098	-4.32
	Flavour 2 Cluster 1	0.004	0.02
	Cluster 2	-1.003	-2.74
	Flavour 3 Cluster 1	-0.958	-3.20
	Cluster 2	-2.116	-4.54
	Flavour 4 Cluster 1	0.127	0.57
	Cluster 2	-1.674	-3.71
Quality perception	Price (log transform)	0.684	3.16
	Sensory perception	0.713	3.67
	Label 1	0.096	0.52
	Label 2	0.037	0.19
	Label 3	0.109	0.61
	Label 4	-0.278	-1.44
Intention to buy	Price (linear)	-0.092	-2.97
	Quality perception	0.976	6.86
Scale factor	Sensory quality	0.110	1.19
	Quality perception	-0.127	-1.35
	Intention to buy	0.176	0.94
Measurement equations parameters	λ sensory 1	0.903	7.55
	λ sensory 2	1.018	7.11
	λ quality	0.506	6.02
Number of observations (individuals)		1220 (122)	
Number of parameters (draws)		43 (500)	
		Full model	Ranking
Loglikelihood		-7775.9	-1971.8
Rho square		0.105	0.099
Adjusted Rho square		0.100	0.095

Table 4 – CF model coefficients and goodness of fit indicators (thresholds not reported)

		Coefficient	Robust t-test
Sensory quality (ove)	Intercept Cluster 1	6.034	35.47
	Cluster 2	5.590	17.07
	Flavour 1 Cluster 1	0.186	0.83
	Cluster 2	-2.100	-5.01
	Flavour 2 Cluster 1	-0.004	-0.02
	Cluster 2	-0.925	-2.32
	Flavour 3 Cluster 1	-1.082	-3.99
	Cluster 2	-2.055	-4.25
	Flavour 4 Cluster 1	-0.022	-0.09
	Cluster 2	-1.683	-3.68
Quality perception (exc)	Intercept	0.266	1.08
	Price (log transform)	0.164	2.76
	Sensory indicator (ove)	0.364	9.47
	Residual of sensory ind.	-0.099	-2.31
	Label 1	0.055	0.66
	Label 2	-0.006	-0.07
	Label 3	0.074	0.88
	Label 4	-0.066	-0.75
Intentio to buy	Price (linear)	-0.046	-2.75
	Quality indicator (exc)	2.370	13.05
	Residual of quality ind.	-0.715	-5.23
Scale factor	Intention to buy	-0.006	-0.26
Error components s.d.	Sensory (σ_{ove})	1.959	45.22
	Quality (σ_{exc})	1.021	37.91
	Ranking (σ_{ITB})	2.265	12.19
Number of observations (individuals)		1220 (122)	
Number of parameters (draws)		30 (500)	
		Full model	Ranking
Loglikelihood		-5917.8	-1609.4
Rho square		0.794	0.214
Adjusted Rho square		0.793	0.209

Table 5 – MIS model coefficients and goodness of fit indicators (thresholds not reported)

		Coefficient	Robust t-test
Quality	Intercept	0.419	5.47
Perception (exc)	Quality indicator (<i>lik</i>)	0.752	23.24
	Price (linear)	0.017	2.90
Intention to buy	Price	-0.040	-2.56
	Quality indicator (exc)	2.286	14.89
	Residual of quality ind.	-1.296	-5.54
Scale factor	Intention to buy	-0.004	-0.18
Error components s.d.	Quality (σ_{exc})	0.671	12.10
	Ranking (σ_{ITB})	2.337	12.81
Number of observations (individuals)			1220 (122)
Number of parameters (draws)			14 (500)
		Full model	Ranking
Loglikelihood		-2810.1	-1565.0
Rho square		0.637	0.236
Adjusted Rho square		0.636	0.231

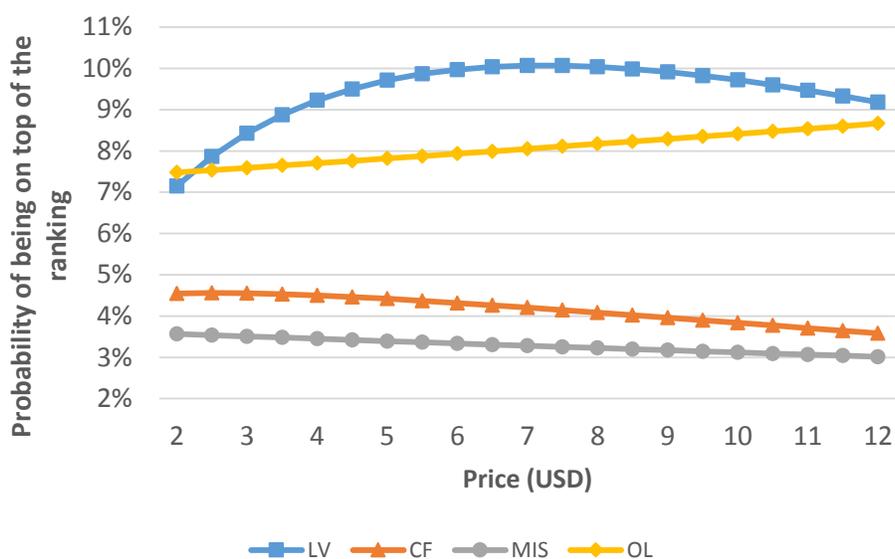
Figure 9 – Probability of being on top of the ranking as a function of price (*ceteris paribus*).

Figure 9 shows results from a simulation of the probability of a participant from the first cluster ranking an unaltered wine with the base label as his top choice. In the case of the CF and MIS models, average values of *ove*, *exc* and *lik* for such kind of wines were used. Even though the OL's price coefficient is not significant, it is still considered in the simulation, therefore predicting an increasing probability of being on top of the ranking as *price* grows.

Table 6 – OL model coefficients and goodness of fit indicators (thresholds not reported)

		Coefficient	Robust t-test
Intention to buy	Price (linear)	0.016	1.00
	Label 1	0.112	0.66
	Label 2	0.073	0.43
	Label 3	0.039	0.23
	Label 4	-0.231	-1.33
	Flavour 1 Cluster 1	0.126	0.60
	Cluster 2	-1.506	-3.57
	Flavour 2 Cluster 1	-0.185	-0.82
	Cluster 2	-0.680	-1.73
	Flavour 3 Cluster 1	-0.708	-3.05
	Cluster 2	-1.078	-2.56
	Flavour 4 Cluster 1	0.061	0.26
	Cluster 2	-1.019	-2.26
	Other parameters	Scale factor	0.099
	Error component s.d. (σ_{ITB})	1.270	11.61
Number of observations (individuals)		1220 (122)	
Number of parameters		20	
Number of draws		500	
Loglikelihood		-1967.6	
Rho square		0.039	
Adjusted Rho square		0.030	

5 Discussion

Price can have a double effect on the probability of choosing a wine: (i) a positive effect if consumers associate price and quality, and (ii) a negative effect due to consumers' budget

constrain. This double effect is well documented in a large amount of products (Leavitt 1954, Dodds *et al.* 1991, Yan & Sengupta 2011), as well as wine in particular (Plassman *et al.* 2007, Lewis & Zalan 2014, Kardes *et al.* 2004, Palma *et al.* 2016).

Based on ideas by Dodds *et al.* (1991) and Grunert (2005), we proposed a behavioural framework (Figure 3) incorporating the double effect of price. We then proposed three econometric implementations of this model with varying degrees of detail. The LV and CF models are full implementations of the behavioural model using latent variables and the Control Function approach, respectively. The MIS model focuses exclusively in the negative effect of price using the Multiple Indicator Solution. For comparison purposes, we also estimated a base model estimating only the net effect of price.

Results from all models are consistent regarding attributes other than price, but differ significantly when it comes to the effect of price. Therefore, our results indicate that ignoring the double effect of price may bias the estimation of price coefficients, but not of other attributes. Hence, we highly recommend correcting for price endogeneity when estimating price sensitivity for products whose quality might be associated to price by consumers.

5.1 Comparison of fit

In terms of fit, from best to worse, models can be ordered as: (i) MIS, (ii) CF, (iii) OL, and (iv) LV. This ranking was constructed based on the log-likelihood of only the comparable part of all models, i.e. their modelling of intention to buy.

It is not surprising that the LV model fits the data the worst, due to the way it uses the additional information available to it. We have three sources of information: the alternatives ranking (*invRnk*, which all models use in the same way); the indicators of sensory quality (*ove* and *wtp*); and the indicators of perceived quality (*exc* and *lik*). The LV model uses the indicators to train

sub-models that predict the level of sensory quality and perceived quality. The CF and MIS models, instead, use the indicators as proxies for sensory quality and perceived quality, therefore achieving higher accuracy compared to the LV model.

In our particular case study, the sensory and perceived quality sub-models within the LV model are particularly poor due to the high degree of preference heterogeneity in our sample, making the difference in fit between the LV model and the rest more evident. However, even in perfectly homogenous samples, the fit of the LV model will never surpass that of the MIS and CF models, simply because estimating a value will always be less accurate than using the value itself.

Furthermore, the OL model will always have a better fit than the LV model, as the OL model its parameters must only fit the intention to buy, while the LV model must try to fit sensory quality and perceived quality as well (see Vij & Walker 2016).

The reason why the MIS model fits the data better than the CF model is similar to why the CF model fits better than the LV model. As the MIS model relies more heavily on the use of indicators as proxies for the latent variables, it captures each participant perception more accurately, therefore achieving a higher fit.

Considering the fit of each model, we discourage the use of the LV model. The CF model can provide similar insight into how sensory and perceived quality are constructed in the mind of the consumer, while providing much higher fit.

5.2 Comparison of predictions

Concerning predictions, Figure 9 summarizes the differences between all models. The OL model predicts an increasing probability of preferring a wine as its price grows. This is due to the OL model failing to correctly capture price's net effect. The net effect of price is most probably not linear (see discussion the end of section 1.2), but as the OL model only uses a linear effect, it

only captures a non-significant positive effect, leading to a probability of choice that increases with price. While it is possible to use a more complex form for the effect of price in the OL model (such as a quadratic function or a mix of a linear and log transformations) this does not significantly improve the fit of the model, nor the significance of its price parameters. Though not reported, we tested more complex forms of including price in the OL model, always finding non-significant parameters for price. This is probably due to the OL model not having enough information to disentangle the true effect of price, highlighting the importance of mimicking the behavioural model of Figure 3 when modelling.

The LV model predicts a curve with a maximum demand near 7 USD, but its prediction is far away from the one of the best fitting models. The LV model is the only one whose prediction has a shape consistent with the one proposed by Lockshin *et al.* (2006). The best-fitting models, instead, show a much flatter prediction. This could be indicative that excessively low prices do not decrease demand, but only keep it from increasing.

Predictions from the CF and MIS models are relatively similar. The main difference between them is that the CF model predicts an almost flat demand until 4 USD, then it starts decaying faster as the price grows; the MIS model –instead– predicts a linear decrease of demand as price increases. We tested using a log transform of price in the auxiliary (instrumental) regression of the MIS model, but fit and prediction do not change significantly.

While in the MIS model is tempting to interpret price's coefficient in the auxiliary regression as the positive effect of price, this is not correct. The MIS model's auxiliary regression has two explanatory variables: price and a second indicator of perceived quality (*lik*). As both are highly correlated, their coefficients may be confounded to some degree, and therefore cannot be interpreted in a reliable way. Nevertheless, the coefficient of price in the auxiliary regression must be considered when predicting.

Given the similarities between predictions of the CF and MIS models, we recommend using the CF model only when measuring the positive effect of price is paramount, e.g. when developing a new product's pricing strategy. When dealing with an already existing product, we recommend using the simpler MIS approach to do market share prediction or calculating price-demand elasticity.

5.3 General considerations

The main value of measuring the positive effect of price lays in identifying price ranges where the producer does not want to position its product. For example, based on the curve described by the CF model, pricing the product anywhere below 4 USD would lead to missing revenue, as price can be risen without decreasing demand. Optimum pricing, however, cannot be determined from Figure 9 alone, as marginal production cost would have to be factored in. Identifying these areas of sub-optimal prices is more relevant for new products, as pricing strategies of already existing products cannot be easily changed.

Our work has some relevant limitations. Most importantly, when we discuss the positive effect of price we do not consider 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. 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 a brand unknown to consumers.

As our objective was studying the double effect of price, we did not focus on improving model fit by capturing preference heterogeneity. Instead, we only used an ad-hoc clustering approach to

improve fit of sensory liking. Palma *et al.* (2017) discusses several methods to consider preference heterogeneity in discrete choice models in a similar setting.

The indicators of *perceived quality* used in our study could be improved. Even though we collected three indicators with a high Cronbach alpha ($\alpha=0.93$), one of them (*rec*) did not work and another worked only partially (*lik*), as they hindered fit when included in the modelling. The three indicators were participants' level of agreement with the phrases "this wine is excellent" (*exc*), "I would recommend this wine to my friends" (*rec*) and "I like this wine" (*lik*). Identifying better indicators for perceived quality is an important area of future research.

In conclusion, our research shows that neglecting the double effect of price can lead to seriously biased estimation of price sensitivity. However, the bias does not seem to affect other attributes. Measuring the positive effect of price can be particularly useful when devising pricing strategies for new products, as it helps avoiding sub-optimal price ranges. We recommend using a model based on the Control Function approach to measure the positive and negative effects of price simultaneously, and a model based on Multiple Indicator Solution when only the negative effect of price is of interest.

6 References

- 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.
- Asioli, D., Varela, P., Hersleth, M., Almli, V.L., Olsen, N.V and Næs, T. (2017) A discussion of recent methodologies for combining sensory and extrinsic product properties in consumer studies. *Food Quality and Preference* **56**, 266 – 273.

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.

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.

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.

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

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.

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.

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.

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.

- Grunert, K. (2005). Food quality and safety: consumer perception and demand. *European Review of Agricultural Economics* **32**, 369–391.
- 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.
- Hagerty, M. (1978) Model testing techniques and price-quality tradeoffs. *Journal of Consumer Research* **5**, 194 – 205.
- 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.
- 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.

- Kardes, F., Cronley, M., Kellaris, J. and Posavac, S. (2004) The role of selective information processing in price-quality inference. *Journal of Consumer Research* **31**, 368 – 374.
- Langhe, B., van Osselaer, S., Puntoni, S. and McGill, A. (2014) Fooled by heterocedastic randomness: local consistency breeds extremity in price-based quality inferences. *Journal of Consumer Research* **41**, 978 – 994.
- 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. (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.
- Levin, I. and Johnson, R. (1984) Estimating price-quality tradeoffs using comparative judgments. *Journal of Consumer Research* **11**, 593 – 600.
- 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.
- 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., 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.

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

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.

Ministerio de Desarrollo Social (2014). CASEN 2013: Ingresos: Cuadro 10: Promedio, mínimo y máximos 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>

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

Mueller, S., Lockshin, L., Saltman, Y. and Blanford, J. (2010a). 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. (2010b). 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.

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

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: A case study with Chinese consumers. *Journal of Choice Modelling* **19**, 24 – 39.

Palma, D., Ortúzar, J. de D., Rizzi, L.I. and Casaubon, G. (2017). Modelling consumers' heterogenous preferences: a case study with Chilean wine consumers. *Australian Journal of Grape and Wine Research*, 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.

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/>.

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.

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

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.

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.

Train, K. (2009). *Discrete Choice Models with Simulation* (2nd edition) 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.

Yan, D. and Sengupta, J. (2011) Effects of construal level on the price-quality relationship. *Journal of Consumer Research* **38**, 376 – 389.

Zeithaml, V.A. (1988) Consumer perceptions of price, quality and value: A means-end model and synthesis of evidence. *Journal of Marketing* **52**, 2 – 22.