Estimation of Consumers' Demand Function

Evgeniya Goryacheva

SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

Economics Discipline Group, Business School
University of Technology Sydney

February 2019

Abstract

In this thesis, we test the existence of the behavioral component in the consumer's decision-making process that captures the direct influence of other available products and their characteristics on the consumer's utility. We introduce this behavioral component to the empirical demand model and show that it plays an important role in the widely used approach of employing rival products' characteristics as instruments to overcome the price endogeneity problem in demand estimation. To do so, we use a dataset on individuals' choices of the red wines from an experiment. The obtained results show that the exclusion condition is not satisfied for some of the rival products' characteristics, but is satisfied for other rival products' characteristics.

We extend the choice model by allowing the subjective evaluations of the products' quality in the consumer's utility function. We exploit the unique survey design of the discrete choice experiment on wine choice with random prices to estimate the coefficients of consumers' demand function for wines. The consumers form their subjective evaluations of the quality of the new wine from the bottle design and label information. Consumers' subjective evaluations of the wine's quality may be correlated with unobserved product characteristics. To solve the endogeneity problem of the subjective evaluations, we use characteristics of other wines from the randomly formed choice set as instruments. The existence of the individuals' behavioral bias allows us to use other product characteristics as instruments. Additionally, we study how the purpose of consumption affects individuals' choices of the wines.

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Evgeniya Goryacheva, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Business School at the University of Technology Sydney. This thesis is wholly my own work unless otherwise indicated in the references or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution. This research was supported by the Australian Government Research Training Program.

Signature: Production Note: Signature removed prior to publication.

Date: 27 February 2019

ACKNOWLEDGEMENTS

This thesis used the data from the Web survey conducted by the GMO Research Company which was kindly provided by Dr. Kazuko Nakata (Setsunan University).

I would like to thank my supervisor Professor Susumu Imai for his great guidance in developing me as a researcher and for the fruitful discussions on research ideas. I am grateful to my supervisors, Professor Mikhail Anufriev and Associate Professor Shiko Maruyama, for their support during my PhD and for the helpful advice on improving my thesis. I thank the current and former Directors of Graduate Studies, Peter Siminski, David Goldbaum, and Joshua Chan, for their patience and taking care of EDG PhD students.

I am thankful to the participants of research seminars at the University of Technology Sydney, Hokkaido University, ESAM 2018 at Auckland University of Technology and CMES 2018 at Fudan University for their valuable comments.

Capstone Editing provided copyediting and proofreading services, according to the guidelines laid out in the university-endorsed national “Guidelines for Editing Research Theses”.

PREFACE

This is a conventional thesis structured as a series of chapters. This thesis includes an introduction to the research study, a review of the literature, three chapters and a conclusion. The first chapter describes the experimental design and the dataset that was used in this thesis. The second chapter introduces a demand model that describes the direct influence of the characteristics of other products in a choice set on consumer's utility and the estimation results. This chapter also tests the instrumental validity of rival products' characteristics. The third chapter is devoted to the demand models that capture the role of quality's subjective evaluations and the purpose of consumption in consumer's choice.

Contents

List of Figures

List of Tables

1 Introduction

The discrete choice models are a workhorse in Empirical IO used to estimate consumers' demand. When the consumer chooses the product, she also takes into account other alternatives available to her. In discrete choice models, the choice of a particular product depends on the utility of this product and the utilities of other products in the choice set. The connection between different alternatives is captured by the functional form of the demand. However, these models do not consider that the characteristics of other products may directly affect the utility. The utility of the product reflects the satisfaction level that the consumer expects to obtain by choosing the product. This satisfaction level is relative to the other available products in the choice set. For example, a consumer's choice set includes the wine with her favorite grape variety. In this case, the consumer may give high evaluation to this product and lower evaluations to other products in the choice set. The structure of the choice set itself affects consumers' preferences. Consumer's reaction to the particular product's characteristic may also depend on how often the characteristic occurs in the choice set.

In this thesis, we introduce a behavioral component of the consumer's decision-making process that captures the direct influence of other available products and their characteristics on the consumer's utility to the empirical demand model. Moreover, we show that this behavioral component plays an important role in the widely used approach of employing other products' characteristics as instruments for the market prices in demand function estimation.

For our analysis, we use a dataset on individuals' choices of the red wines produced in different countries obtained from the experiment conducted by Kazuko Nakata (Setsunan University), Susumu Imai (Hokkaido University), and Yuka Ohno (Hokkaido University). In the experiment, the participants were also asked to provide their evaluations of each of the wine in their choice sets. To show the presence of the behavioral component in a consumer's decision-making process, we estimated whether the characteristics of other products in the choice set affect the individual's subjective evaluation of the wine. To further consider the issue, we also estimated a multinomial logit model and a random coefficient logit model. We argue that the result obtained by using subjective evaluation of the product as a proxy for its utility is more robust to model misspecification than the discrete choice model estimation where only choices are observable.

Several theories in behavioral economics are related to the influence of the other products' characteristics on consumer's choice considered in this thesis. One of them is the model of context-dependent preferences that considers the structure of choice set as a context. According to this model, people get extra value if the product is better in some aspects than other available products. Other alternatives and their characteristics may affect a consumer's perception of the product through different behavioral effects. The results obtained in this research show that at least one of them occurs in the consumer's evaluation of the product and its utility. This evidence is enough to question the validity of rival products' characteristics as instruments in demand estimation.

Rival products' characteristics are widely used instruments to solve the price endogeneity problem caused by a potential positive correlation between the price and unobserved product's characteristics. More specifically, the unobserved product's characteristics include quality. A monopoly or oligopoly firm with a high-quality product optimally raises its price to take advantage of consumers' appreciation of the product. Further, to produce high-quality products, firms need to use high-quality inputs that are more expensive. These high inputs' costs lead to high product price. The positive correlation between the price and the error term leads to the positive bias in the price coefficient estimates.

The rival product characteristics can be considered as proper instruments if both relevance and exclusion conditions are satisfied. The relevance condition is satisfied if endogenous variable and instrumental variable are correlated. This condition can be easily tested by regressing the endogenous variable on instrumental variable and checking the F-statistic. The instrumental variable satisfies the exclusion condition if it is not correlated with the error term in the equation of interest. This condition requires the instrumental variable to affect the dependent variable only through the endogenous variable but not directly. It is impossible to test whether the exclusion condition is satisfied by using market data because the endogenous variable is correlated with the error terms. The true error term can only be recovered when econometrician knows the true parameter values, which are unknown. We test whether the exclusion condition is satisfied for the characteristics of rival products by using a unique experimental design. First, prices are set randomly to eliminate their possible endogeneity problem. Second, in our experiment, individuals are provided with random choice sets.

The obtained results show that the exclusion condition is not satisfied for some of the rival products' characteristics but is satisfied for other rival products' characteristics. If researchers use rival products' characteristics as an instrument for the prices and they directly affect consumer's utility, it leads to the biased estimated coefficients for the prices. In this case, we cannot recover the true price elasticities of the products. Recovering price elasticity is important for predicting market changes and formulating taxation policy.

We show that characteristics of other products in the choice set directly enter consumer's utility function. Not including them in the utility may result in omitted variable bias in demand function estimation. For this reason, we propose a demand model that takes into account this behavioral component of consumers' decision-making process.

In this thesis, we also introduce demand model specification in which individuals' subjective quality evaluations are part of the consumer's utility function. We estimate the parameters of this demand model by using the same dataset on individuals' choices of the red wines obtained from the experiment. A distinctive feature of our experimental design is a collection of individuals' subjective evaluations of the wines' qualities. These evaluations help us to show the role that the wine's quality plays in consumers' preferences by estimating quality's coefficient in the consumer's demand function.

In our research, we consider the situation when the consumers choose the wine that they have not tried before, and internal product characteristics are unknown to them. The only information they may use to predict the quality of the product is package design and the label of the product. By observing external product characteristics, consumers form their evaluations of the products' quality. The evaluation of the product reflects the level of the quality that the consumer expects to obtain from this product. These evaluations of the quality affect the utility level that the consumer expects to get if she chooses the product.

When we introduce subjective evaluations of quality into the structural demand model, we face a possible endogeneity problem of the subjective evaluations. When people form their product quality evaluations, they may consider not only observed products' characteristics but also unobserved product characteristics. These unobserved product characteristics may also directly enter the consumer's utility function. For example, advertising of the wine affects consumer's utility and simultaneously subjective evaluations. People may think that well-known wines have better quality and, thus, give them higher subjective evaluations. At the same time, buying well-known wine may give consumers higher utility because of social prestige. This connection raises the endogeneity problem, because subjective evaluations may be correlated with the error term. In this case, estimated coefficients would be biased. To solve the endogeneity problem, we use exogenous product characteristics of other wines from the randomly formed choice set as instruments. Individuals' behavioral bias allows us to use them as instruments because other products' characteristics influence subjective evaluations of quality. When people evaluate the quality of the product, they not only consider the product itself, but consider other available alternatives and their characteristics. When people evaluate products, they compare them with each other. A product may get lower consumers' quality evaluations in a choice set consisting of products with more favorable characteristics than in a choice set consisting of products with less favorable characteristics. If other alternatives are worse, consumers feel that they are buying something of higher quality and get extra satisfaction. For this model specification, we assume that the characteristics of other products in the choice set affect wine's utility only through its subjective evaluation of quality but not directly.

In real-life situations, consumers make predictions about the quality of the wine by considering its price, because the price and the quality are correlated. In our experiments, all prices are random and do not reflect the wine's quality. Individuals were asked to give their evaluations of wine quality from their choice set before they observed any prices. As a result, in our experimental design, individuals' subjective evaluations of the wines' quality do not depend on prices, and we do not include them into the first stage of two-stage least squares (2SLS). We separate the influencers of consumers' subjective evaluations of the quality: product characteristics and prices. We study how consumers form their evaluations by observing product characteristics of all wines from the consumer's choice set and do not measure the influence of the price. Our experimental design allows us to distinguish these two effects. This procedure is required to solve the price endogeneity problem to estimate unbiased coefficients of the consumer's indirect utility function.

Then, we consider the case where all coefficients and taste shock in the consumer's utility function are purpose-specific. That is, consumers react differently to the product characteristics depending on the purpose of consumption. We study how respondents' choices of the purpose of consumption affect the probability of wine purchase. The wines considered to be suitable for gifts and special occasions have a higher probability of being purchased than wines suitable for other situations. Overall, individuals are more likely to buy wine if they can identify a suitable purpose of use for it. People react differently to price and quality's evaluation when they choose wine for different purposes. They also tend to choose different types of wines for different situations. Individuals care about the general impression of wine when its use is related to other people, for example, when choosing a wine for a party, special occasion or as a gift. Individuals tend to consider wines subjectively evaluated as high quality to be suitable for personal drinking.

To introduce a new product to the market or to change a characteristic of an existing product to increase the profit, firms need to predict how the consumers will react to it. The proposed empirical model of the demand function that captures the behavioral aspects of consumers' choices can help to answer this question. Ignoring the influence of other available alternatives and their characteristics may lead to the misrepresentation of the forecasts and have implications on market demand.

In Chapter 2, we present a design of the experiment and test whether other products' characteristics directly affect utility by using individuals' subjective evaluations. In Chapter 3, we introduce the model of the consumer's product choice and explain the estimation procedures used. We also describe how the obtained results are related to the validity of rival products' characteristics as the instruments for the market prices. In Chapter 4, we estimate consumers' demand model with subjective quality evaluations and show how consumers' reaction to different wines' characteristics depends on the purpose of consumption.

Literature review

In this thesis, we connect empirical industrial organization with behavioral economics in modeling consumers' demand. Thus, our research is related to the literature in these two fields.

Several theories in behavioral economics attempt to explain the phenomenon explored in this thesis – the influence of the other alternatives on the consumer's choice. One of the theories is the model of context-dependent preferences proposed by Tversky and Simonson (1993) that considers the structure of choice set as a context. This approach is close to the reference-dependent preferences model proposed by Kahneman and Tversky (1979), if we look at the characteristics of other available alternatives as a reference point. Sen (1993) criticizes the applicability of independence of the irrelevant alternatives assumption for peoples' decisions. Sen (1997) argues that in the case of limited knowledge, the choice set itself provides information that the individual uses in her decision making. According to Sen, other alternatives affect the evaluation of the product by providing additional information about the quality of the product. Sen calls this channel the epistemic value of the menu (choice set). For example, if the consumer observes that the new product is surrounded by other high-quality products she may conclude that this product is also high-quality because the store manager combines the products from the same range together.

The role of choice set in individuals' decision-making process is reflected in various economic models with the support of the experimental evidence. The choice set dependent preferences are described in such concepts as attraction and compromise effects. Huber, Payne, and Puto (1982) introduce and experimentally prove the existence of attraction effect in consumers' decision-making. They show that the addition of an alternative dominated by at least one alternative in the choice set but not dominated by at least one other increases the share of the item that dominates it. Simonson and Tversky (1992) show that an individual's choice is influenced by the available alternatives under consideration. Their experimental results support the existence of the compromise effect in individuals' decision-making. They show that a brand with an intermediate price and quality tends to take more share from the low-quality, low-priced brand than from the high-quality, high-priced brand. Mazar, Koszegi, and Ariely (2014) experimentally prove that the distribution of the products' prices in the choice set affects consumers' product evaluations. Their results support the idea that preferences depend on the environment that decision-makers face.

The experimental design described in this thesis is different from those in other papers. It allows us to eliminate individuals' heterogeneity and attain more reliable results on the influence of other products' characteristics on an individual's utility of the product. Moreover, the experiment is closer to the real-life environment as individuals were provided with the pictures of real products with all available information.

Several papers in marketing attempt to estimate reference-dependent preferences model using market data. For example, Hardie, Johnson, and Fader (1993) propose an econometric model with loss aversion and reference dependence and estimated it by using scanner data. However, in their model, the reference point is not other available products from the consumer's choice set but the consumer's previous choice. Compared to our research, they do not consider the endogeneity problem. To measure the influence of product's quality, they use Consumer Reports quality ratings that are similar for all individuals. However, people may have different quality evaluations of the same product. In this case, the estimated coefficient does not capture the heterogeneity of the consumers' quality evaluations, while we use individuals' subjective evaluations of the quality.

This thesis also focuses on the relationship between the purpose of the consumption and consumers' choice of wine. Many papers in marketing emphasize the importance of the context effect in consumer's decision-making, for example, Chakravarti et al. (1983) write a literature review on evidence of the existence of the context effects in experimental and social psychology, behavioral decision theory and consumer research. Shocker, Ben-Akiva, Boccara, and Nedungadi (1991) state that the consumers construct their consideration sets based on the purpose of the consumption and they consist of goal-satisfying alternatives. While many alternatives are available to the individual, it is likely that only a few of these are appropriate for a relevant purpose. Day, Shocker, and Srivastava (1979), among other analytical methods in marketing, mention "substitution in use" approach that examines how alternatives relate to specific product usage and how it affects the substitution patterns between them. Ratneshwar and Shocker (1991) empirically study the role of usage context on consideration set. Srivastava, Alpert, and Shocker (1984) describe an empirical application of this approach in the financial and banking services market. Ching and Hayashi (2010) estimate the effects of payment card rewards on consumer choice of payment methods for different retail types. Imai, Ohno, and Nakata (2017) study the relationship between the planned use of wines and the purchase decision in the experiment where they disentangle the informative, signaling and the utility enhancing effect of the brands on the individuals' choices of red wines.

To estimate wines' demand function, we use a multinomial logit discrete choice model. This model was proposed by McFadden (1976). Later to consider consumers' heterogeneity, random coefficients models were employed in demand estimation. For example, Berry, Levinsohn, and Pakes (1995) estimate random coefficient logit model using market-level data. Their method also allows for solving the price endogeneity problem using instrumental variables. We overcome the price endogeneity problem by introducing random prices in our experiment. Later BLP approach was further developed by combining macro and microdata. Petrin (2002) shows the importance of consumer-level data in the estimation of the substitution patterns. Berry, Levinsohn, and Pakes (2004) use microdata to improve differentiated products demand estimation. To capture consumers' preferences heterogeneity in our research, we introduce model specifications and include interaction terms of product characteristics with consumers' demographic characteristics in the demand function estimation.

In BLP model, quality is a part of unobserved product characteristics, but our data set allows us to consider it as a separate term in consumer's utility function and to estimate consumers' reaction to the change in wine's quality evaluation. Even if we include subjective evaluations into consumer's utility function as a proxy for wine quality, we still have some wines characteristics that enter consumer's utility function but are unobservable by an econometrician. For example, advertising of the wine affects consumer's utility and simultaneously subjective evaluations. For example, Ackerberg (2001) and Ocass and Frost (2002) found the impact of the prestige of the brand on consumer behavior. This raises the endogeneity problem, because subjective evaluations may be correlated with wines' unobserved characteristics. To solve the endogeneity problem, we use exogenous product characteristics of other wines from the randomly formed choice set as instruments.

Other approaches are used in empirical research to estimate the parameters of consumers' demand for wine. Many papers on consumers' preferences of wines use hedonic price function estimation. For example, Combris, Lecocq, and Visser (1997) use data on Bordeaux wines to estimate hedonic price equations. They show that the market price depends on the objective characteristics of the wine and does not depend on sensory characteristics. Nerlove (1995) applies a different method to the hedonic price function estimation using Swedish data. Instead of using standard regression of prices on wine characteristics he regresses sold quantities of the wines on prices and product characteristics. Golan and Shalit (1993), by using data on Israeli grapes, study how grape quality affects the price of the wine. Oczkowski (1994) and Schamel and Anderson (2003) estimate hedonic price functions for premium wines from Australia and New Zealand. They show that price premia are different for different wine regions and correlated with wines' ratings. Hedonic price approach allows indirect revealing consumers' willingness to pay for different product characteristics and different quality levels. This method is helpful in determining the factors that affect prices, but it does not consider the strategic interaction of different brands, omitted variable bias, and multicollinearity of the product characteristics.

Another approach to estimate consumers' demand is almost ideal demand system (AIDS). Seale Jr, Marchant, and Basso (2003) estimate demand elasticities of domestic and imported red wines in the United States (US) using the first-difference version of AIDS. Cembalo, Caracciolo, and Pomarici (2014) study consumption of nonpremium wines in Italy. They use Italian households data to estimate the censored demand system (QUAIDS). They show that there is heterogeneity in consumers' reaction to different wine categories. To estimate AIDS, an econometrician must have data on consumers' expenditures among the brands of a given product category, which is not always available.

The dataset we use in our analysis is obtained by stated preference elicitation method. Ben-Akiva, McFadden, and Train (2015) formulated conditions under which this method can contribute to understanding consumer behavior and forecasting market demand. Our survey design is close to the discrete choice experiment which allows researchers to evaluate the relative importance of each product's characteristic for the consumers. This method is widely used in marketing and policy analysis. For example, Ewing and Sarigöllü (2000) study preferences for clean-fuel vehicles versus conventional vehicles. Kjær (2005) reviews the applications of discrete choice experiments in health care. Lockshin, Jarvis, dHauteville, and Perrouty (2006), Mueller, Osidacz, Francis, and Lockshin (2010) and Mtimet and Albisu (2006) apply discrete choice experiments to measure consumer sensitivity in wine choice. The advantage of the discrete choice experiment design is the exogeneity of the product characteristics. In our survey, we partly use discrete choice experiment design for wines' prices to solve the endogeneity problem.

In our experiment, individuals only observe external, not internal, characteristics of the products and evaluate their qualities. Olson (1978) argues that consumers use available information to form their beliefs about the products that affect their choice. Zeithaml (1988) lists the papers that show how extrinsic attributes (e.g., price and brand name) can serve as indicators of a product's quality. Considerable empirical research shows that consumers use price to predict a product's quality when it is the only available information, but when there is another available information the role of price is less clear. Consumers use price as a product's quality signal more when the brand is unfamiliar compared to when they know the brand. Olson (1978) shows that if the information on a product's internal and external characteristics is available, consumers are influenced more by these characteristics than by its price.

In our experiment, consumers use a product's external characteristics as an indicator of its quality as they do not have information on its internal characteristics. This situation occurs in real life when the consumer has little or no experience and does not have enough time or access to get information on internal characteristics. Zeithaml (1985) reports that because of time constraints, working individuals use less product information than other demographic groups and react more to the package and brand name. Huber and McCann (1982) describe the impact of inferential beliefs on product evaluations.

2 Experiment on consumers' choice of wine with quality's subjective evaluations

In this chapter, we describe the dataset that is used for our analysis and the discrete choice experiment from which this dataset was obtained. We provide descriptive statistics for the characteristics of the wines that were chosen for the experiment, demographic characteristics of the individuals participated in the experiment and characteristics related to their wine consumption behavior. Using individuals' subjective evaluations of the wines, we test whether the characteristics of other available wines in the choice set directly affect an individual's utility. We study how the reaction to the characteristics of other products in a choice set varies across different demographic groups. We control for individuals' heterogeneity and do the robustness check of obtained results.

2.1 Experimental design

The survey was designed by Kazuko Nakata (Setsunan University), Susumu Imai (Hokkaido University), and Yuka Ohno (Hokkaido University). The Web survey was conducted by the GMO Research company in Japan. Any person could register to participate in the surveys of the company. Registered participants were paid for answering questions. The participants obtained one point (approximately one yen) from the company by answering each question. GMO Research sent the questionnaires randomly to registered participants (the total number of registered participants was 180,000). Selected participants could complete the survey as they like, and once the number of completed surveys reached the required number, GMO Research stopped accepting the surveys. The selected participants were not incentivized in the survey.

For this, 1,100 individuals above the age of 20 who drink red wine completed the survey (survey questions are in Appendix A.1). Firstly, they were asked about their demographic characteristics such as age, gender, household income, education, marital status and the number of adults (over 20 years old) in the household. Then the individuals answered the questions about their red wine drinking experience: the age when they started drinking, the frequency of drinking, usual place to drink wine, reasons of purchasing wine, and wine characteristics that they consider when making a choice. The individuals were asked about their preferences for the red wines: favorite country of origin, grape variety, etc.

Each individual was provided with a choice set consisting of five wines. These five wines were randomly chosen for each individual from the list of 55 red wines prepared for the experiment (the wine list is in Appendix A.2). Each individual was provided with the pictures of the bottles and labels of all wines from her choice set. An example of the wine's picture and information available to individuals is shown in Figure 1. After observing the information, individuals were asked about their opinion of the wines' qualities, the design of the bottles and general impressions of the wines. Additionally, after expressing their opinion about the wines' qualities, the individuals were asked about the reasons that influenced their evaluations.

A feature of this experimental design is that most individuals were choosing among wines that they have not tried before. As a result, they were in a situation where they had to determine the quality of the product only through its external characteristics and any other additional information available such as characteristics of other wines in their choice sets. To control for the cases where the wines were familiar to the individual, the individuals were asked whether they tried the wine before or knew its market price.

At the next stage, the individuals were provided with the prices of

Figure 1: Example of the wines' pictures and descriptions provided to individuals

Wine Name It's A Game!
Vintage 2010
Country Italy
Region Tuscany
Region (Details)
Taste Full Body
Producer Bibi Graetz
Grape variety Sangiovese

Wine Name Petite Sirene
Vintage 2010
Country France
Region Bordeaux
Region (Details)
Taste Medium Body
Producer Chateau Giscours
Grape variety Merlot, Cabernet Sauvignon

the wines in their choice sets. These prices were randomly generated for each individual and individuals were aware of the randomness. Each individual had to choose one wine from the choice set or an outside option (i.e., none of the wines in the choice set). After this choice was made, each individual was asked about the wine's characteristics that influenced their choices. Each individual was asked to make a choice for the same choice set three times. Each time, the individual faced new random prices of the wines.

The experimental design created high variations in the products' characteristics and the characteristics of other products that subjects faced in a dataset. The variability occurs because individuals face different choice sets. The vector of other products' characteristics may be different for the same product in two different choice sets.

The stated preference elicitation method that was used allowed controlling for the choice sets that people face. In market level data, consumers' choice sets are unobservable. Usually, it is assumed that all consumers face the same choice set consisting of all existing brands. The choice set plays a key role in consumers' purchase decision-making. In logit demand function, a consumer's decision to purchase a particular product depends on the availability of other alternatives. The probability of a consumer purchasing a particular product is lower in the larger choice set. The probability of choosing the product depends on the utility of this product and the utilities of all other products from the consumer's choice set. If assumed choice sets do not coincide with the true consumers' choice sets, the estimated demand function coefficients are far from the true values. Our data set does not have this problem because we observe respondents' choice sets.

In the experiment, individuals were provided with small choice sets consisting of 5 wines. This experimental design not only provides variability in products' characteristics that individuals face but potentially reduces the effect of consideration set formation on the individuals' choices. That is, if the choice set is large, the individual only decide based on a small subset of it, called the "consideration set" in the literature. For example, Fader and McAlister (1990), Roberts and Lattin (1991) and Ben-Akiva and Boccara (1995) propose different models of consumer's consideration set formation. Restricting the choice set of individuals makes it more likely that individuals will include all five wines into their consideration set.

The experimental design is close to the discrete choice experiment as the prices in the experiment vary. There are some weaknesses of the traditional discrete choice experiment method: abstract nature of the proposed products, and respondents have no incentive to make choices in an experiment in the same way they would in the market. In this survey, respondents were provided with the choice sets of existing wines and only prices were randomly chosen which aligns the design closer to the real market environment.

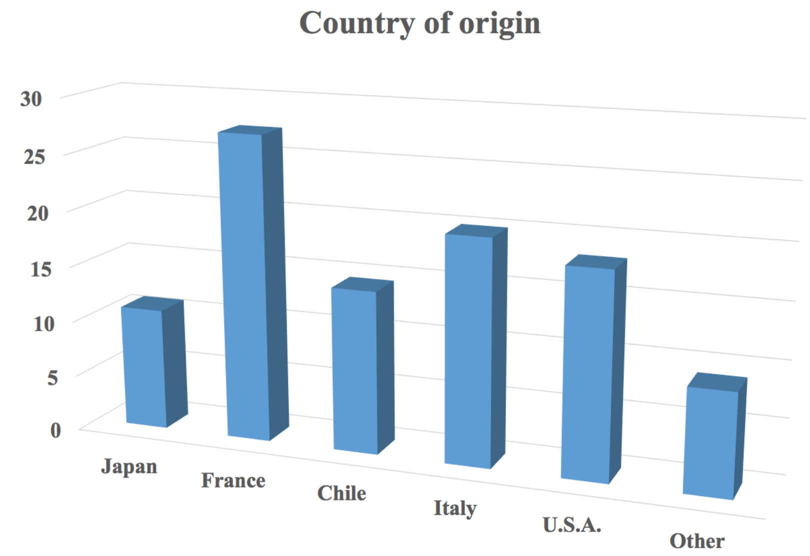
2.2 Data

We used data on individuals' choices, wines' prices and characteristics obtained from the experiment described above. Wines' characteristics observed by individuals from the wines' labels in the experiment include country of wine's origin, region, year, grape variety, and body of the wine.

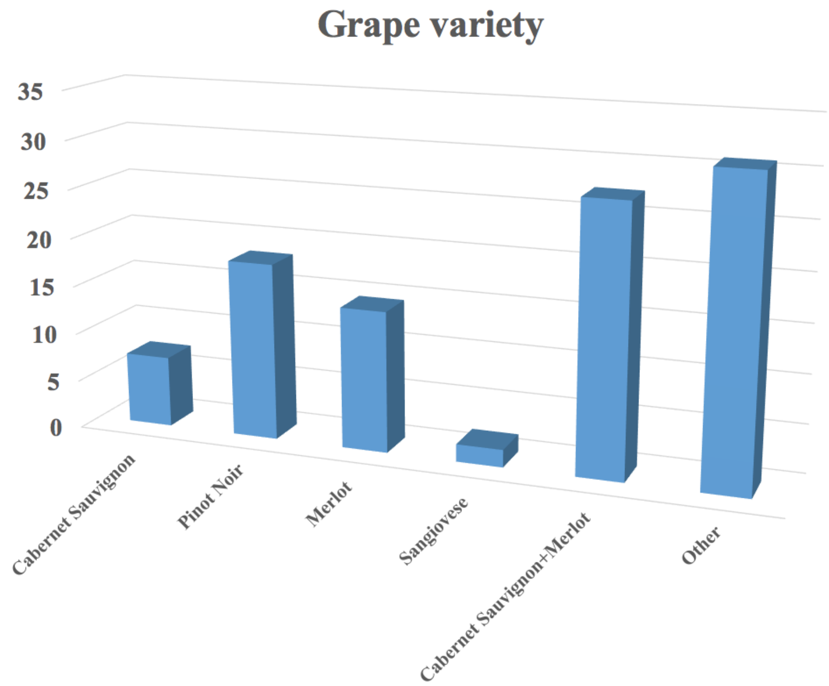
Figures 2–3 present descriptive statistics for the 55 red wines chosen for the experiment. Ten percent of the wines were produced in Japan, and the remainder in foreign countries such as France (27.3%), Chile (14.5%), Italy (20%), and the US (18.2%).

Most wines in the list have Cabernet Sauvignon, Merlot, and Pinot Noir grape varieties. Several wines in the list are blended wines. The wine

Figure 2: Descriptive statistics of the wines: country of origin, grape variety (%)

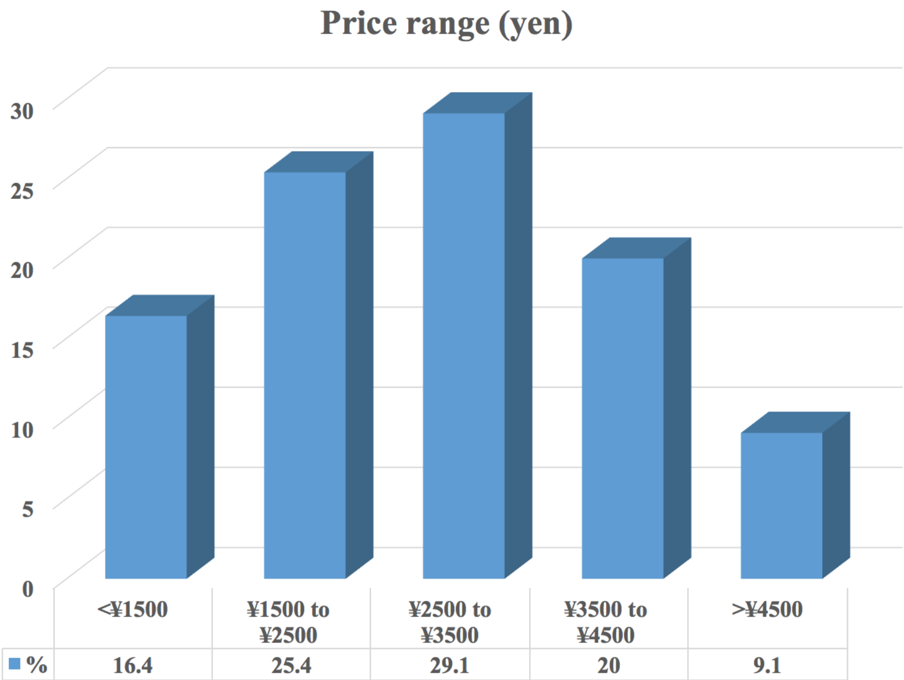


	Japan	France	Chile	Italy	U.S.A.	Other
%	10.9	27.3	14.5	20	18.2	9.1



	Cabernet Sauvignon	Pinot Noir	Merlot	Sangiovese	Cabernet Sauvignon+ Merlot	Other
%	7.3	18.2	14.5	1.8	27.3	30.9

Figure 3: Descriptive statistics of the wines: price range (yen)



brands were chosen from the different price ranges: 29.1% of the wines had a market price in the price range of 2,500–3,500 yen, 25.4% were in the range of 1,500–3,500 yen, and 20% were in the range of 3,500–4,500 yen (wine prices are in Appendix A.3).

The dataset includes the choices of 1,100 individuals made three times for the same choice sets but with different random prices, providing 3,300 observations in total. The average age of participating individuals was 52 years old. The average starting age of drinking red wine was 26 years old. The descriptive statistics for individuals' demographic characteristics presented in Figures 4–7.

Individuals from different income groups are represented in the dataset. Most of the individuals (75.9%) are or have been married. There is asymmetry in the number of female versus male participants (33.8% and 66.2% respectively).

Figure 4: Demographic characteristics: household income (yen)

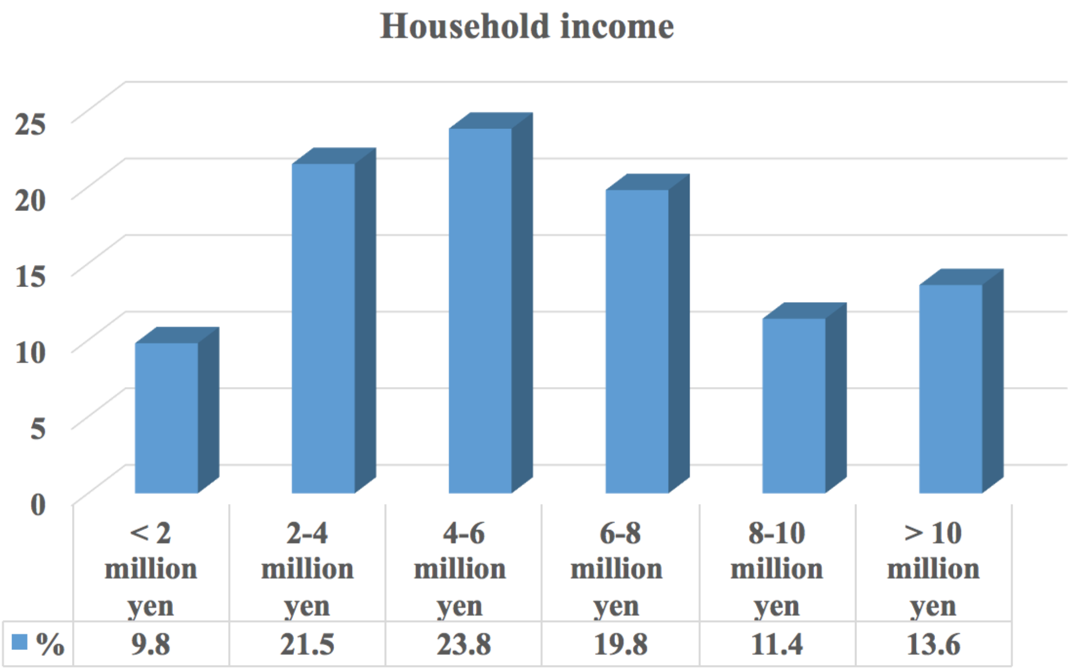
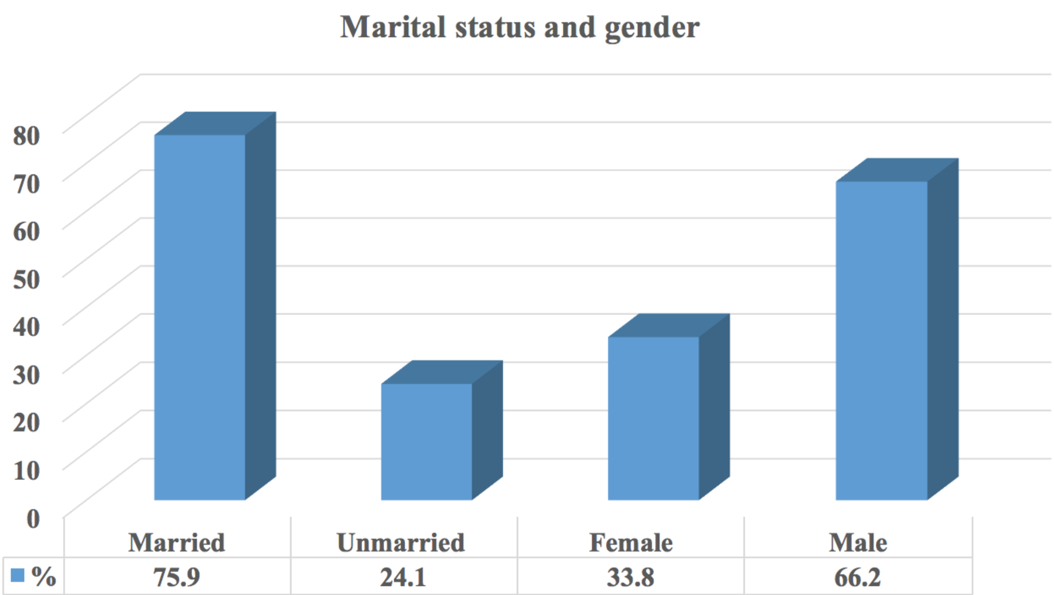


Figure 5: Demographic characteristics: marital status and gender (%)



There is variability in individuals' highest level of education. The largest group includes people with a bachelor degree (47.6%). The second largest group represents people with secondary education: 24.5% of individuals had finished high school, and 20.2% had obtained a college degree. Most of the participants lived in a household consisting of 3–4 adults (58.8%).

Figure 6: Demographic characteristics: education (%)

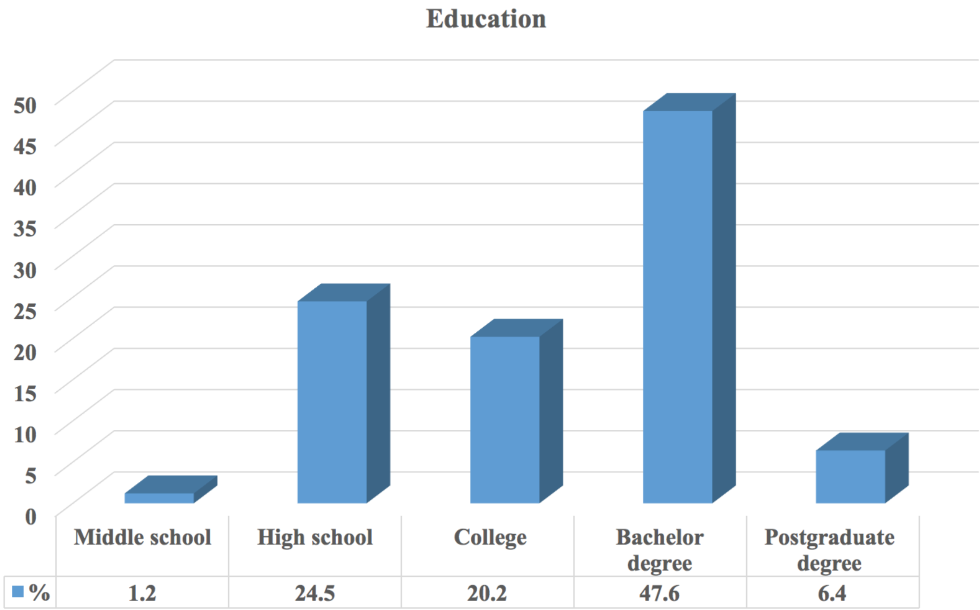
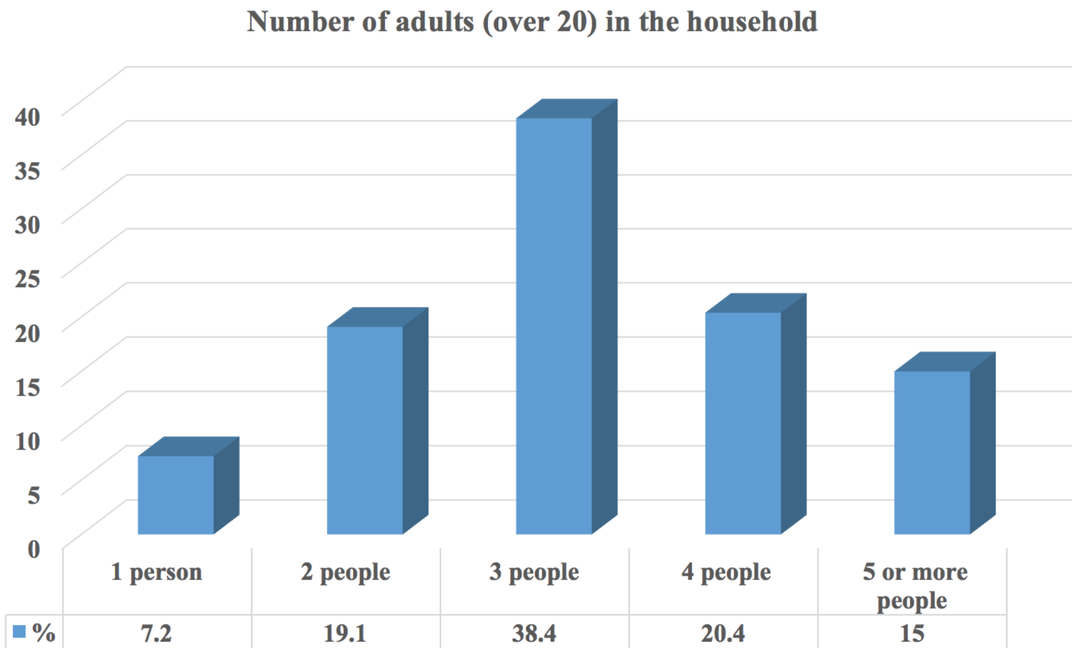


Figure 7: Demographic characteristics: number of adults (persons over 20 years old) in the household (%)



Descriptive statistics for the characteristics related to the wine consumption behavior of the individuals are presented in Figures 8–10.

Figure 8: Frequency of drinking red wine (%)

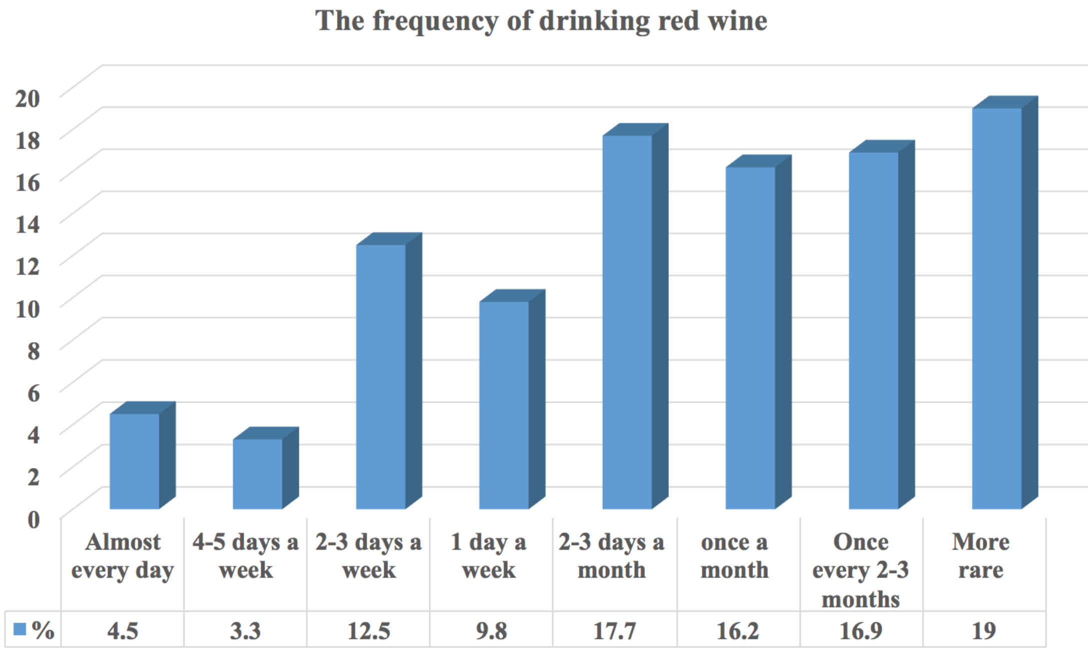
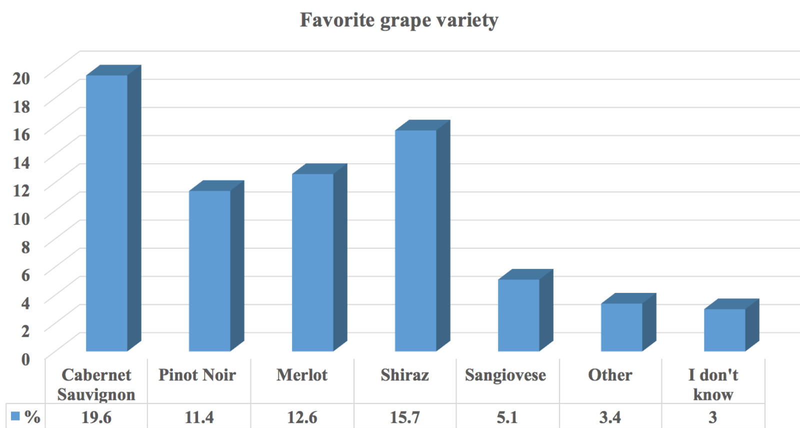


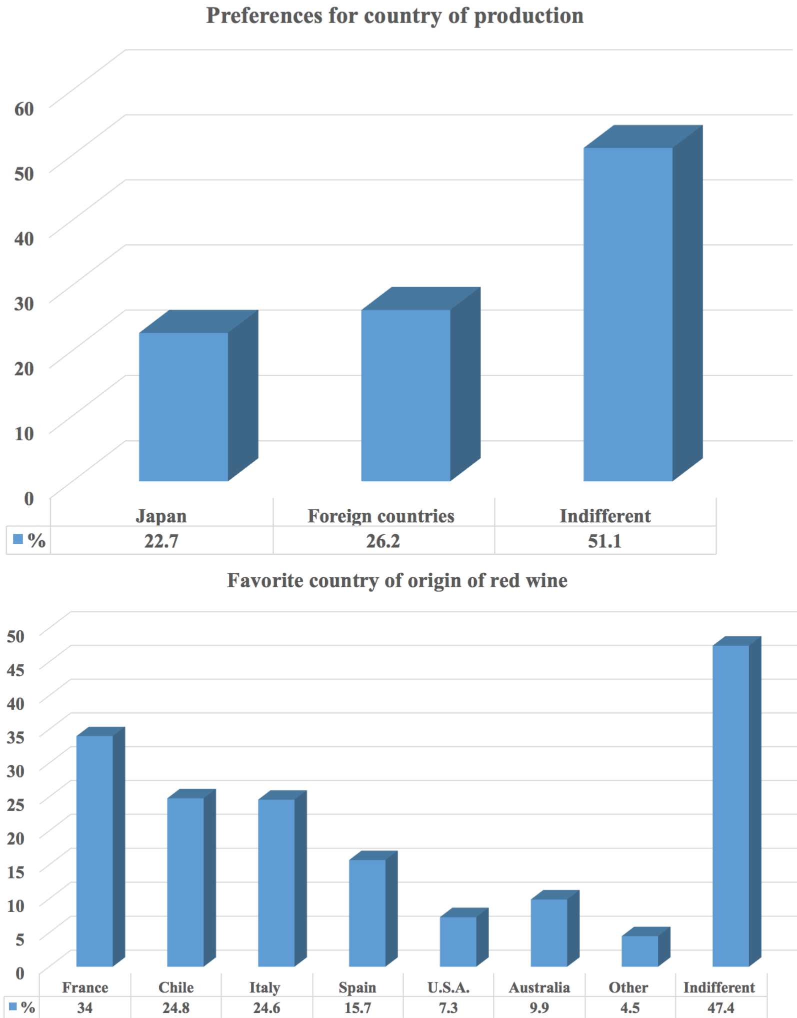
Figure 9: Favorite grape variety (%)



All individuals who participated in the survey drink red wine, but they have different frequencies of drinking: 64% of them drink red wine at least every month, and 30% drink red wine at least once a week. Many individuals choose Cabernet Sauvignon and Shiraz as their favorite grape varieties.

Most (51%) are indifferent to red wines produced in Japan and those

Figure 10: Preferences for country of production (%)

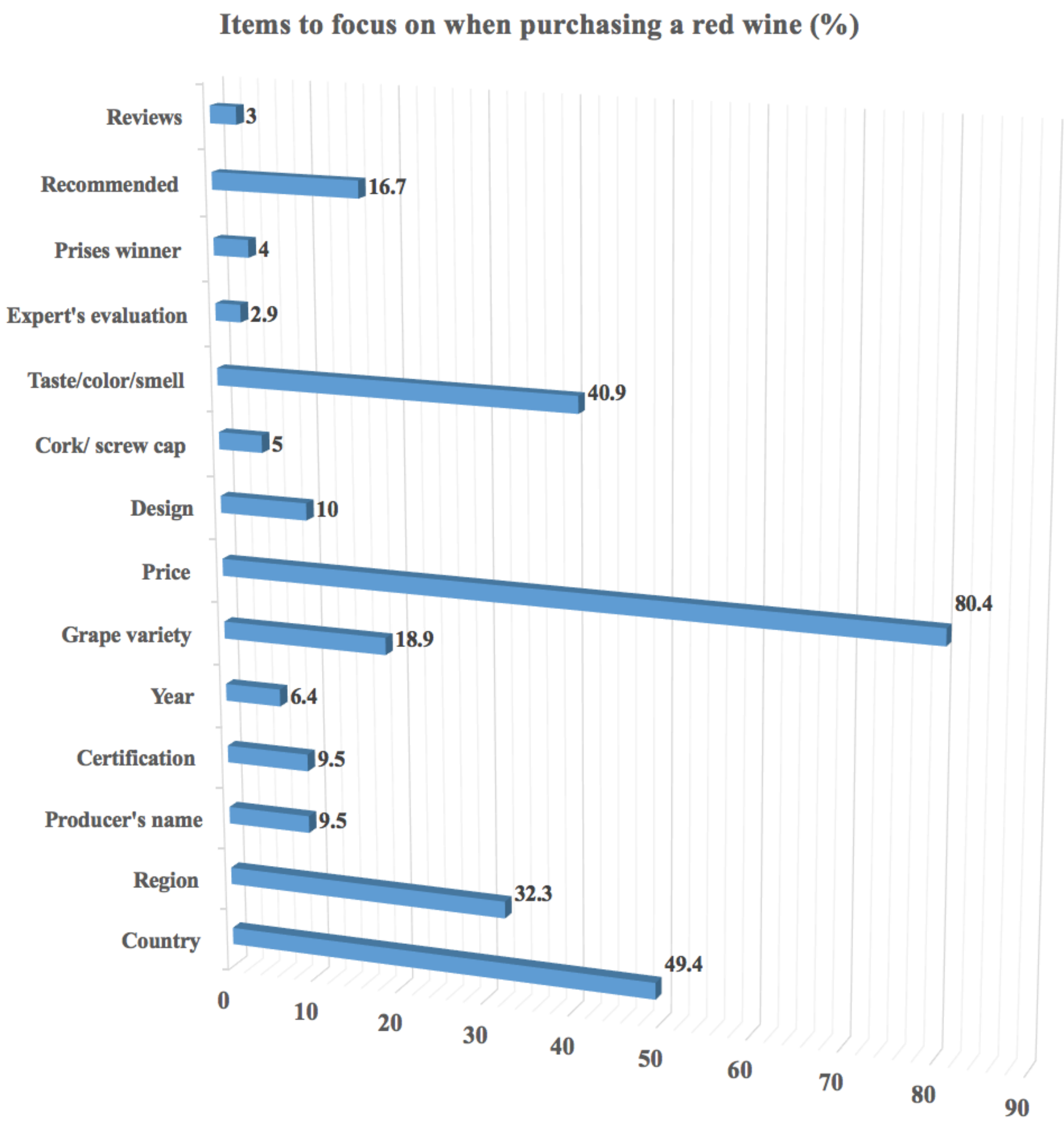


produced in a foreign country, 22.7% prefer wines produced in Japan to foreign red wines, and 26.2% prefer foreign red wines to Japanese red wines. Among those individuals with a preferred country of origin for red wine, France was the most popular, followed by Chile and Italy.

Individuals described the criteria they use when evaluating wine and making purchase choices. Individuals mostly focus on the following characteristics of wine when making a purchase decision: price (80%), country (49%), taste (41%), and region (32%) (see Figure 11).

Individuals consider different aspects when they evaluate the quality

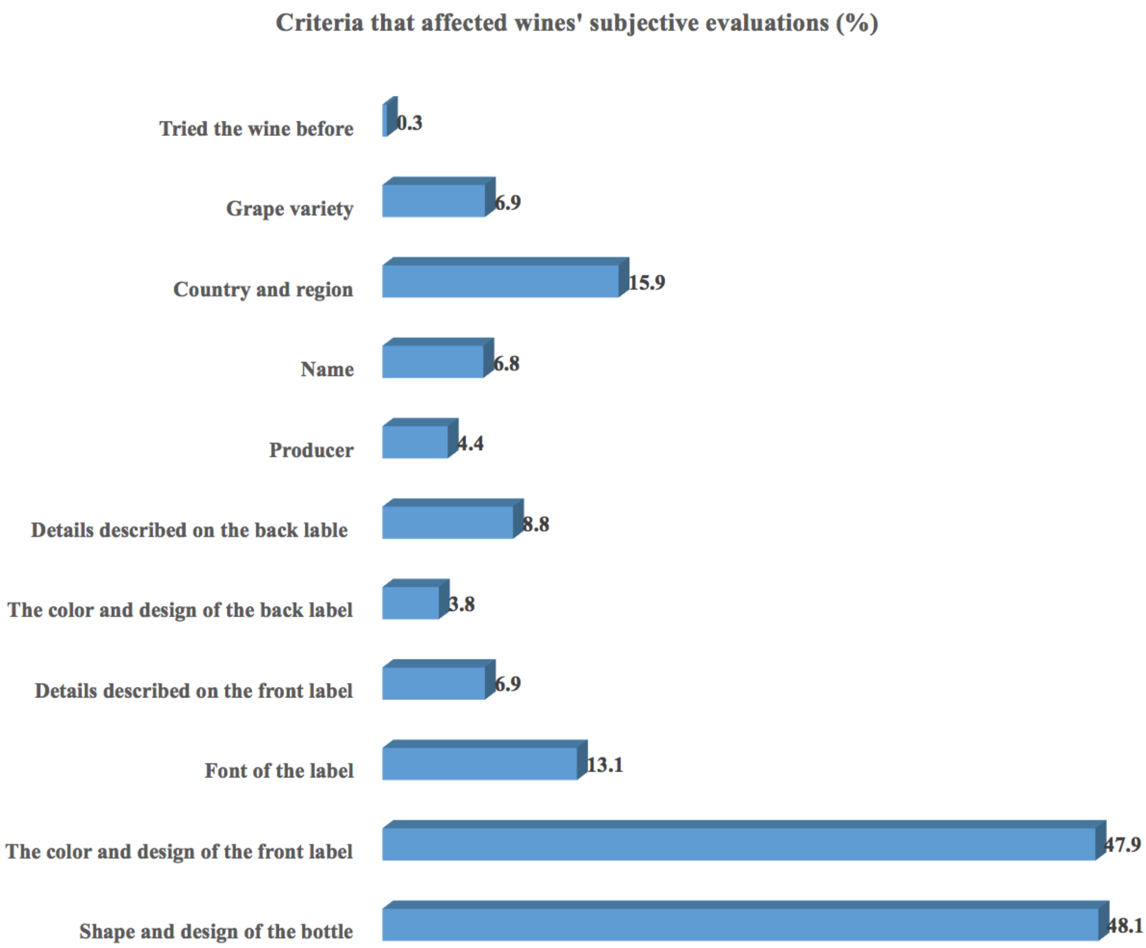
Figure 11: Items to focus on when purchasing a red wine (%)



of a wine that they have not tried before (see Figure 12). Individuals rely mostly on the shape and design of the bottle and its label to evaluate a wine's quality. The country and the region of the wine were also mentioned by the participants as the factors affecting their evaluations of a wine's quality.

The characteristics of the wine that affected the wine choices in the experiment according to the individuals' reports are shown in Figure 13. Most individuals (66.6%) were influenced by the prices of the wines, while 39%

Figure 12: Criteria affecting subjective evaluations of wines' quality (%)



mentioned country as a factor that affected their choices. Only 15% of the participants chose design as a factor that influenced their decision, even though the shape and design was the most commonly chosen factor for the individuals' evaluations of a wine's quality. A potential explanation for the above result is that individuals who assess the quality of the wines based on the label designs do not actually consider wine quality when choosing which wine to buy, they rather look at the price. Also, individuals who care about wine quality when they make purchase choices look at characteristics other than label design for information on quality.

Data related to wine consumption is shown in Figures 14–16.

Figure 13: Criteria affecting wine choices (%)

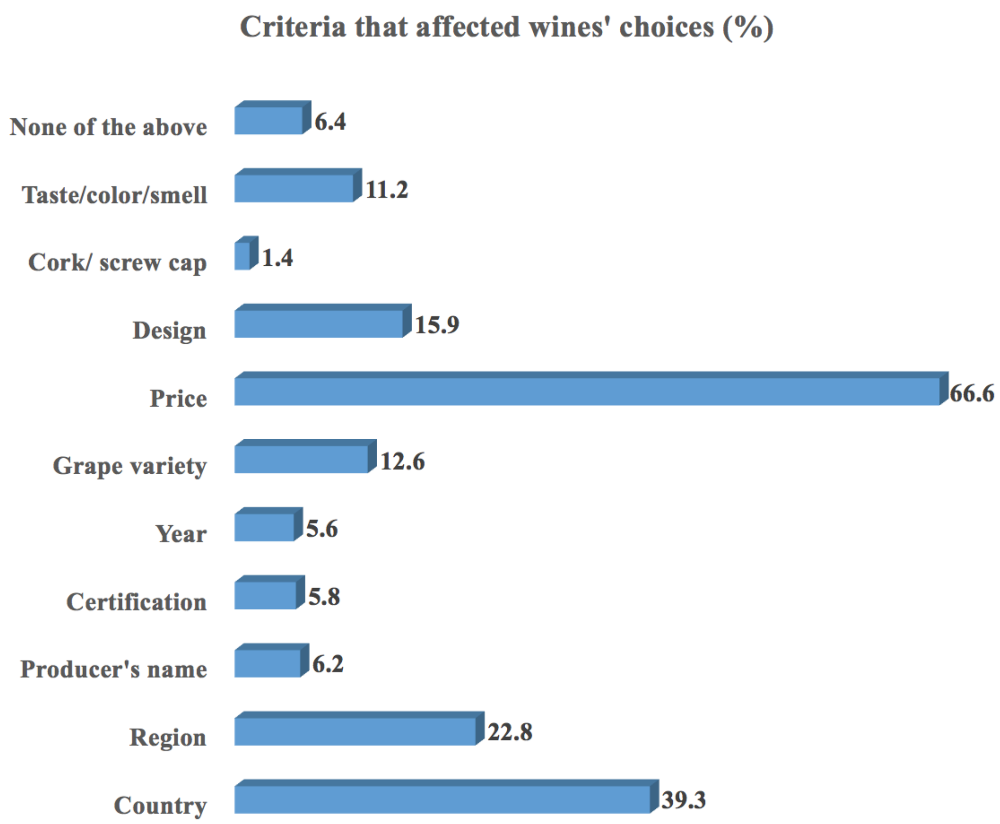


Figure 14: Number of red wines usually in household (%)

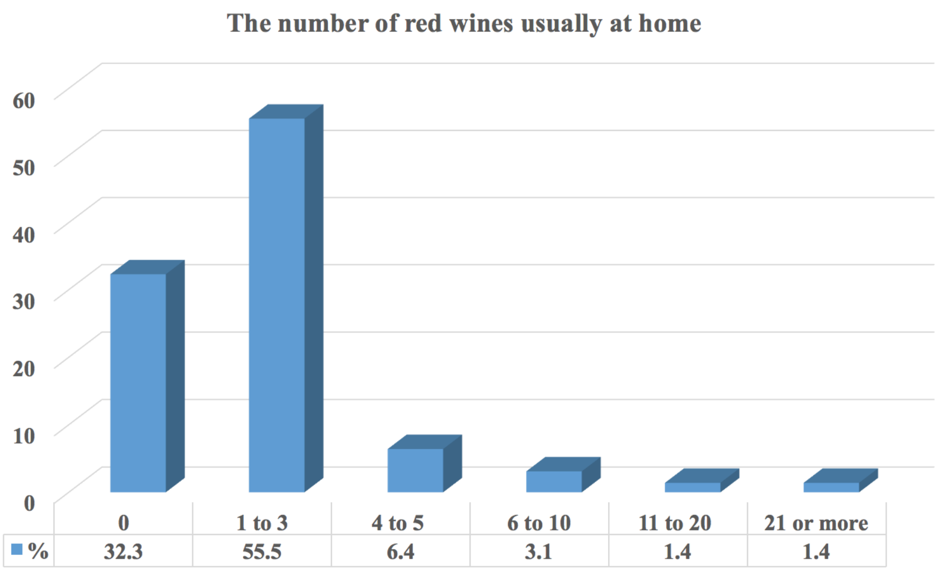
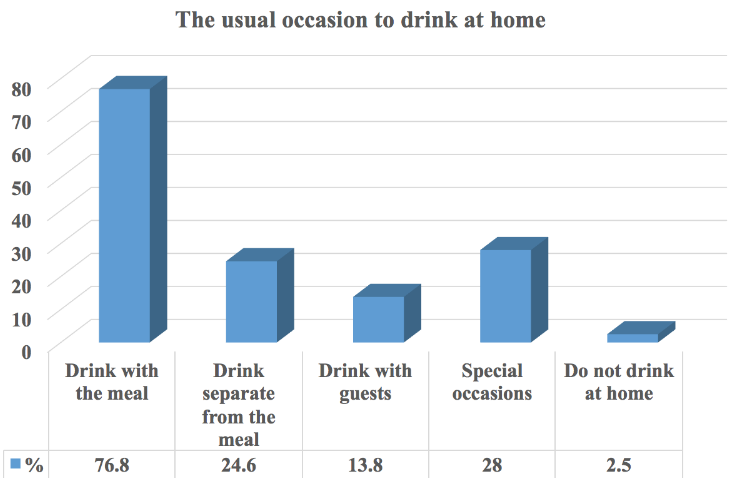
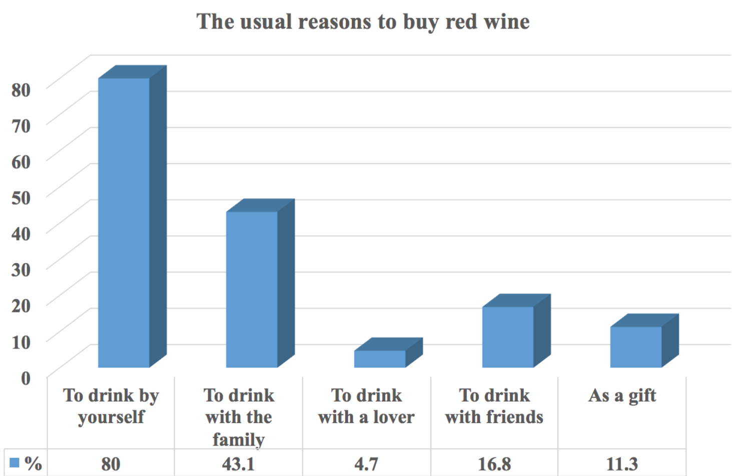


Figure 15: Usual place to drink red wine (%)

Figure 16: Usual occasion to drink at home (%)





Regarding red wine consumption patterns of the individuals, 55.5% usually keep 1–3 bottles of red wine at home and 32.3% do not keep any wine bottles at home. The majority of individuals usually drink at home (92.7%) and in a restaurant (44.7%). The most common occasion to drink at home is drinking with a meal (76.8%), followed by specials occasions such as celebration (28%). Among the usual reasons to buy red wine, individuals mentioned to drink by themselves (80%), to drink with family and friends (59.9%) and to use as a gift (11.3%).

Most individuals mentioned the supermarkets (64.1%) and liquor stores (35.9%) as usual places to buy red wine. The majority of individuals usually purchased red wine at the medium price level.

Figure 17: Usual places to buy red wine (%)

Figure 18: Most common price range for red wine purchase (yen)

2.3 The influence of other products' characteristics on individual's subjective evaluation

In this section, we test whether the characteristics of other products directly enter the consumer's utility function by using individuals' subjective evaluations. There are many ways alternative brands could affect the utility of buying a product, and we provide some (the most important) examples of these effects. First, since individuals only have imperfect information on the actual quality of wines, they may use the comparison brands as signals. For example, consumers may consider any wine in a wine shop that has high-quality brands to be high-quality wine. However, in this case, we believe it is reasonable that individuals to some extent understand the random nature of their choice sets in our experiment and, thus, do not consider the other brands to contain information on the product they are assessing. It is more likely that the comparison effect affects their assessments. For example, a wine that is in a group with high-quality wines may be considered to be very low quality, whereas when other brands are also low quality, consumers may not notice its low quality as much. To separately identify different behavioral mechanisms through which other available alternatives and their characteristics affect consumer's utility is beyond the scope of this paper.

The purpose of this research is to show that consumers' utility of the wine is affected by the characteristics of its alternatives. We hypothesize that individuals' evaluations of the products in the experiment are more likely to be affected by other products' characteristics through the comparison effect than through the signaling effect, because the individuals were aware that the wines in their choice sets were randomly chosen and did not contain any information about the quality of each other. Overall, alternatives and their characteristics may affect consumer's perception of the product through different psychological effects. The purpose of this research is not to distinguish them, but to show that at least one of them affects consumer's evaluation of the product and its utility level.

We use individuals' subjective evaluations of the wines to test whether other products' characteristics directly affect consumer's utility. If characteristics of other wines in a choice set directly affect individual's subjective evaluations, then we can conclude that they also directly affect individual's utility of the wine (as reflected in individuals' subjective evaluations). To test the influence of other products' characteristics on individuals' subjective evaluations we run the following OLS regression:

$$S_{ij} = \kappa + X_j\pi + X'_{ij}\psi + \omega_{ij}, \tag{1}$$

where S_{ij} is individual i's subjective evaluation of product j, which is a discrete variable that varies from one (lowest evaluation) to five (highest evaluation), X_j is a vector of product j's characteristics, X'_{ij} is a vector of other products' characteristics in a choice set of individual i excluding product j, and ω_{ij} is an error term.

$X_j = (X_j^1, X_j^2, ..., X_j^K)$, where X_j^k is a dummy variable that equals one if product j has characteristic k and zero otherwise. For example, if the wine has the Pinot Noir grape variety, the "Pinot Noir" dummy variable equals to one for this wine. $X'_{ij} = (X'^1_{ij}, X'^2_{ij}, ..., X'^K_{ij})$, where $X'^k_{ij} = \frac{1}{J_i - 1} \sum_{l \neq j \in A_i} X_l^k$, and A_i is individual i's choice set. This variable for Pinot Noir grape variety of other products represents the mean of the values for the"Pinot Noir" dummy variable for all other wines in a choice set.

We do not include wines' prices in Eq.(1) because, in our dataset, individuals evaluate wines before they observe any prices and use only information on products' characteristics. The experimental design eliminates the connection between price and quality because the prices in an experiment are random and do not contain any information about the wines' quality.

The results of Eq.(1) estimation are presented in Table 1, Column 1. The results show that individuals' subjective evaluations of the wines' qualities not only depend on wine's own characteristics but on some characteristics of other wines in the choice set. These results support our hypothesis that other products' characteristics directly enter the consumer's utility function.

2.4 Different demographic groups' subjective evaluations

We check whether there are differences in reaction to other products' characteristics across different demographic groups. The level and types of behavioral effects may vary across individuals. For this purpose, we used information on

Table 1: OLS regression with wines' subjective evaluations and rank regression

VARIABLES	OLS	Rank regression
Own characteristics:		
Pinot Noir	0.0962***	0.0535
	(0.0362)	(0.0346)
Sangiovese	-0.447***	0.00510
	(0.0976)	(0.0932)
Cabernet Sauvignon	0.141***	-0.0182
	(0.0498)	(0.0475)
Cabernet+Merlot	0.0735**	0.0562*
	(0.0328)	(0.0313)
Japan	0.0744*	0.0532
	(0.0421)	(0.0402)
United States	-0.0989**	-0.0228
	(0.0387)	(0.0369)
Italy	-0.000983	0.0324
	(0.0377)	(0.0360)
France	0.161***	0.0724**
	(0.0344)	(0.0328)
Other products' characteristics:		
Pinot Noir	-0.0663	-0.0908
	(0.0742)	(0.0708)
Sangiovese	0.469**	0.679***
	(0.194)	(0.186)
Cabernet Sauvignon	0.282***	0.0662
	(0.101)	(0.0964)
Cabernet+Merlot	-0.0256	-0.000249
	(0.0660)	(0.0630)
Japan	0.0190	0.148*
	(0.0850)	(0.0811)
United States	0.0269	-0.0287
	(0.0772)	(0.0737)
Italy	-0.0232	0.0744
	(0.0775)	(0.0739)
France	0.00617	-0.0297
	(0.0697)	(0.0665)
Constant	3.374***	1.471***
	(0.0578)	(0.0552)
Observations	4,754	4,754
R-squared	0.029	0.011

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

individuals' demographic characteristics from our dataset. We created four subsamples based on gender and level of education: women with a high school and a junior college degree (secondary education), women with a bachelor and higher degree (higher education), men with a high school and a junior college degree, and men with a bachelor and higher degree.

We estimated the above OLS regression (Eq.(1)) with wines' subjective evaluations for these four demographic groups. The results are presented in Table 2. On average, women have higher subjective evaluations of wines than men. For the wines' own characteristics, there are differences in reaction across these four groups, but all individuals tend to have higher subjective evaluations of wines produced in France. Most of the wines' own characteristics do not affect their subjective evaluations by women with higher education and men with secondary education. Women with secondary education tend to have higher subjective evaluations of the wines produced in Japan and Italy, and men with higher education have lower evaluations of the wines produced in the US. Regarding grape varieties, women with secondary education and men with higher education have lower subjective evaluations of the wines with the Sangiovese grape variety and higher subjective evaluations of the wines with the Cabernet Sauvignon grape variety. Men with higher education have higher evaluations of the wines with the Pinot Noir grape variety and the blended wines with the Cabernet Sauvignon and Merlot grape varieties.

All four demographic groups react to other products' characteristics when they form the subjective evaluations of the wines but in a different way. Women with secondary education give higher evaluations of the wine if more other wines in the choice set were produced in the US. Women and men with secondary education give lower evaluations if other wines have the Pinot Noir grape variety. Women with higher education give lower evaluations if other wines in the choice set have the Sangiovese grape variety. The more other wines in the choice set have the Pinot Noir, Sangiovese or Cabernet Sauvignon

Table 2: OLS regression with wines' subjective evaluations for different demographic groups

Gender Education	Female Secondary	Female Higher	Male Secondary	Male Higher
Own characteristics:				
Pinot Noir	0.0648	-0.102	-0.0195	0.228***
	(0.0737)	(0.0991)	(0.0794)	(0.0560)
Sangiovese	-0.778***	-0.418	-0.228	-0.407***
	(0.188)	(0.297)	(0.217)	(0.149)
Cabernet Sauvignon	0.183*	-0.000924	0.106	0.187**
	(0.107)	(0.143)	(0.104)	(0.0756)
Cabernet+Merlot	0.0550	-0.0829	0.00116	0.166***
	(0.0685)	(0.0980)	(0.0677)	(0.0502)
Japan	0.176**	-0.00675	0.117	0.0254
	(0.0876)	(0.116)	(0.0936)	(0.0636)
United States	0.108	-0.0827	-0.0192	-0.248***
	(0.0798)	(0.111)	(0.0797)	(0.0605)
Italy	0.200***	-0.111	-0.00146	-0.0746
	(0.0764)	(0.105)	(0.0797)	(0.0591)
France	0.259***	0.176*	0.162**	0.110**
	(0.0701)	(0.0973)	(0.0745)	(0.0526)
Other products' characteristics:				
Pinot Noir	-0.398***	-0.161	-0.400**	0.291**
	(0.146)	(0.200)	(0.167)	(0.118)
Sangiovese	-0.652*	1.275**	0.214	0.844***
	(0.395)	(0.602)	(0.422)	(0.297)
Cabernet Sauvignon	-0.0445	0.319	0.163	0.503***
	(0.215)	(0.287)	(0.220)	(0.152)
Cabernet+Merlot	-0.224*	-0.182	-0.0737	0.124
	(0.135)	(0.199)	(0.141)	(0.102)
Japan	0.0319	-0.181	0.0294	0.114
	(0.180)	(0.242)	(0.183)	(0.131)
United States	0.311**	-0.195	0.249	-0.128
	(0.157)	(0.220)	(0.165)	(0.121)
Italy	0.0930	-0.202	-0.0185	-0.0115
	(0.161)	(0.199)	(0.164)	(0.124)
France	-0.0981	0.180	-0.0284	0.0594
	(0.135)	(0.209)	(0.155)	(0.108)
Constant	3.452***	3.636***	3.379***	3.220***
	(0.114)	(0.151)	(0.127)	(0.0931)
Observations	1,077	589	1,083	2,005
R-squared	0.055	0.045	0.025	0.057

*** p<0.01, ** p<0.05, * p<0.1

grape varieties, the higher the evaluation men with higher education give to the wine.

2.5 Heterogeneity of subjective evaluations

As we have seen, the products' characteristics affect individuals' subjective evaluations differently depending on gender and educational backgrounds. This indicates that individuals' heterogeneity matters in the formulation of subjective evaluations. The heterogeneity may exist not only across these four demographic groups but also on the individual level.

The results presented in Table 1, Column 1 could potentially be caused by the individuals' heterogeneity. The way of perceiving different characteristics of the products may vary across individuals. To capture the influence of other products' characteristics, we need to consider the heterogeneity of consumer preferences in our estimation. To test the robustness of the above results, we need to check whether the influence of the characteristics of available alternatives from the choice set on wines' subjective evaluations still occurs even when we introduce individuals' heterogeneity.

One way to eliminate individuals' heterogeneity in subjective evaluations is to use a fixed effect model. The model assumes that there is an unobserved individual-specific component that affects their subjective evaluations of the wines. We can express this component as individual-specific constant κ_i in Eq.(1). This constant can be removed from the data by demeaning the products' characteristics in the choice set. However, we can not apply this approach to our regression. In Eq.(1) we include the characteristics of other products in the choice set as their mean to eliminate the order effect restrictions. Given this representation of other products' characteristics, estimation of the fixed effect model with both own and other products' characteristics leads to a multicollinearity problem.

For this reason, we used another approach to eliminate the individual-specific fixed effect. We transferred individuals' subjective evaluations of the wines into wines' ranks. For example, individual i has the choice set of the wines $A_i = \{a, b, c, d, e\}$ and her subjective evaluations of these wines are $S_{ia} = 5, S_{ib} = 2, S_{ic} = 2, S_{id} = 4, S_{ie} = 1$. We assign the following ranks to these wines according to their subjective evaluations: $R_{ia} = 4, R_{ib} = 2, R_{ic} = 2, R_{id} = 3, R_{ie} = 1$. The higher rank corresponds to the higher subjective evaluation of the wine in the choice set.

We run the same regression as in Eq.(1), but substitute subjective evaluations with the derived wines' ranks:

$$R_{ij} = \kappa + X_j\pi + X'_{ij}\psi + \omega_{ij}. \qquad (2)$$

The results of Eq.(2) estimation are presented in Table 1, Column 2. Note that many of the wine's own and other wines' characteristics that were significant in the subjective evaluations regression become insignificant in the rank regression. For example, both the wine under consideration and other wines being Cabernet Sauvignon positively and significantly increased the subjective evaluation of the wine under consideration in the regression, but both are insignificant in the rank regression. This indicates the strong fixed effects, which are positively correlated with own or other wines being Cabernet Sauvignon. One interpretation is that individuals are happy to see Cabernet Sauvignons in their choice set, thus they give higher subjective evaluations to all of the wines, regardless of whether they are Cabernet Sauvignons. That is why the relative rankings do not change significantly. Some other wines' characteristics affect wine's utility even if we eliminate individuals' heterogeneity.

2.6 Summary

The dataset that we use in our analysis was obtained from the discrete choice experiment with random prices and choice sets of the wines. The dataset contains information on wines' characteristics, demographic characteristics of the individuals and characteristics related to their wine drinking experience, individuals' choices and subjective evaluations of the wines.

We use individuals' subjective evaluations of the wines as an additional source of information on their utilities. We regress individuals' subjective evaluations of the wines on the wines' own characteristics and characteristics of other wines in a choice set to test whether characteristics of other available alternatives affect individuals' utility. The results reveal that some of the other wines' characteristics directly affect individuals' subjective evaluations of the wines. The results of the estimation for different demographic groups show that all groups react to the characteristics of other wines in a choice set, but there are differences in the reaction across groups.

To show that the obtained results are robust we eliminate individuals' heterogeneity by transferring individuals' subjective evaluations of the wines to the ranks and running a rank regression. The results show that some of the other wines' characteristics directly affect individuals' utilities even if we eliminate the individual specific fixed effect.

3 Consumers' demand model and the price endogeneity problem

In this chapter, we introduce and estimate the multinomial logit model in which the characteristics of other products in a choice set directly enter consumers' utility function. To take into account the heterogeneity in individuals' reaction to the products' characteristics we also estimate a random coefficient multinomial logit model. We test whether the wine drinking experience affects the way how individuals react to other products' characteristics.

In our estimation, we use the pooled data from the three trials as in the experiment individuals had to make a choice three times for the same choice set. There is a possibility that individuals are affected by their previous choices. For the robustness check, we estimate the model using the observations only from the first trial to eliminate the possible influence of the previous choices.

We show how the influence of other products' characteristics on consumers' utility is related to their instrumental validity in solving price endogeneity problem in demand estimation. Using simulations, we show the size of the bias that occurs in price elasticities estimates if the rival products' characteristics are used as instruments for the market prices while the exclusion condition is not satisfied.

3.1 Logit model

We used data on wine choices to test whether the characteristics of other products affect preferences of the products. To do so, we estimated the logit choice model with the following specification. There are N consumers, J products, and T trials. Consumer i's indirect utility function of buying product j in trial t is:

$$u_{ijt} = \beta_0 + \beta_1 p_{ijt} + X_j\alpha + X'_{ij}\gamma + \xi_j + \eta_{ijt}, \tag{3}$$

where p_{ijt} is a price of product j for consumer i in trial t, X_j is a vector of product j's characteristics, X'_{ij} is a vector of characteristics of the products from the choice set of consumer i excluding product j, ξ_j is the mean (across consumers) of product j's unobserved characteristics, η_{ijt} is the idiosyncratic utility shock.

The coefficients in the consumer's utility function reflect her attitude toward the different products' characteristics. Each individual had to make a choice three times, each time she faced the same choice set, but the prices of the wines were different. For this reason, there is a subscript t for the prices, but not for the products' characteristics in the consumer's utility function.

The utility of outside option (not to buy any product) is normalized to zero for all consumers. We consider the logit demand function. Each consumer faces a choice set consisting of several products and an outside option. In our model, we consider a general case, where the consumer has her own choice set. Then, we assume that the consumer chooses the option that gives her the highest utility. If the utility shock for each choice η_{ijt} is i.i.d. extreme value distributed, then the probability of choosing product j by consumer i who faces the choice set A_i in trial t has the following closed form expression:

$$Pr(y_{it} = j) = Pr(u_{ijt} \geq u_{ilt}, \forall l \in A_i) =$$

$$Pr(\eta_{ijt} - \eta_{ilt} \geq (p_{ijt} - p_{ilt})\beta_1 + (X_j - X_l)\alpha + (X'_{ij} - X'_{il})\gamma + \xi_j - \xi_l, \forall l \in A_i).$$

$$Pr(y_{it} = j) = Pr(u_{ijt} \geq u_{ilt}, \forall l \in A_i) = \frac{\exp(\delta_{ijt})}{\sum_{l \in A_i} \exp(\delta_{ilt})}, \tag{4}$$

where

$$\delta_{ijt} = \beta_0 + \beta_1 p_{ijt} + X_j \alpha + X_{ij}^{'} \gamma + \xi_j. \tag{5}$$

3.2 Maximum likelihood estimation

The main identification assumption is that for any product j unobserved product's characteristics ξ are mean independent of the observed product characteristics X. In our dataset unobserved product characteristics ξ are independent of other products' characteristics because individuals' were provided with the random choice sets. Formally:

$$E[\xi_j | X, X^{'}] = 0.$$

To estimate coefficients of the demand function, we use maximum likelihood estimator. We do not face the price endogeneity problem, and there is no need to use instruments because individuals were provided with the random prices and they were aware of it. The likelihood function is:

$$L(\theta) = \prod_{t=1}^{T} \prod_{i=1}^{N} Pr(y_{it}),$$

$$Pr(y_{it} = j) = \frac{\exp(\delta_{ijt}(\theta))}{\sum_{l \in A_i} \exp(\delta_{ilt}(\theta))},$$

where $\theta = (\beta_0, \beta_1, \alpha, \gamma)$ is a vector of parameters, and A_i is consumer i's choice set consisting of J_i products. $\alpha = (\alpha_1, \alpha_2, ..., \alpha_K)$, $\gamma = (\gamma_1, \gamma_2, ..., \gamma_K)$, $X_j = (X_j^1, X_j^2, ..., X_j^K)$, where X_j^k is a dummy variable that equals one if product j has characteristic k and zero otherwise. $X_{ij}^{'} = (X_{ij}^{'1}, X_{ij}^{'2}, ..., X_{ij}^{'K})$, where $X_{ij}^{'k} = \frac{1}{J_i - 1} \sum_{l \neq j \in A_i} X_l^k$.

The log-likelihood function is:

$$\ln L(\theta) = \sum_{t=1}^{T} \sum_{i=1}^{N} \left(\sum_{j \in A_i} I(y_{it} = j) \delta_{ijt}(\theta) - \log(\sum_{l \in A_i} \exp(\delta_{ilt}(\theta))) \right), \qquad (6)$$

where $I(\cdot)$ is an indicator function. We maximize the log-likelihood function and get parameters' estimations:

$$\hat{\theta} = \arg\max_{\theta} [\ln L(\theta)].$$

To estimate standard errors, we use the bootstrap method. First, given data $y, X, X^{'}, p$ with 3,300 observations, we draw a bootstrap sample of the same size with replacement and denote the new sample $y^*, X^*, X^{'*}, p^*$. Second, we obtain $\hat{\theta}^*$ using the bootstrap sample. We repeat these two steps $B = 1,000$ times, generating B independent bootstrapped samples. We have B replications of our estimates $\hat{\theta}_1^*, ..., \hat{\theta}_B^*$. Then, the bootstrap variance of the estimator is

$$s^2_{\hat{\theta}_{Boot}} = \frac{1}{B-1} \sum_{b=1}^{B} (\hat{\theta}_b^* - \bar{\hat{\theta}}^*)^2,$$

where

$$\bar{\hat{\theta}}^* = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}_b^*.$$

3.3 Random coefficients logit model

One of the limitations of the logit demand model is that this model does not consider the heterogeneity of consumers' tastes. One way to model this heterogeneity is to assume that the coefficients in the demand function are different for different consumers, but they follow the distribution with specific parameters that can be estimated. Random coefficients logit model allows this.

Random coefficient model allows capturing individuals' heterogeneity in their reaction to different product characteristics, but there are costs in doing it. As we try to reflect the whole distribution of the utility functions' coefficients, we now have to estimate not only their means but standard errors. As a result, the number of parameters that need to be estimated is doubled. Therefore, we may lose in efficiency in our estimation. Simultaneously, it allows us to see how dispersed the reaction to different product characteristics across the individuals is.

Now we assume that consumers' reaction to the products' characteristics may be different. In this case, consumer i's utility function of buying product j is:

$$u_{ijt} = \beta_{0i} + \beta_{1i} p_{ijt} + X_j \alpha_i + X_{ij}' \gamma_i + \xi_j + \eta_{ijt}. \tag{7}$$

In this specification the coefficients are not fixed, they vary across the consumers. We keep the same assumption on η's iid extreme value distribution. The conditional choice probability of product j by consumer i depends on her reaction to the product's characteristics through the coefficients in her utility function:

$$P(y_{it} = j|\theta_i) = \frac{\exp(V_{ijt}\theta_i)}{\sum_{l \in A_i} \exp(V_{ilt}\theta_i)},$$

where coefficients θ_i are not observable by the econometrician. We assume that coefficients follow the distribution with density $f(\theta|\lambda)$, where λ are the parameters of this distribution. Then choice probability is the integral over θ:

$$P(y_{it} = j) = \int \frac{\exp(V_{ijt}\theta_i)}{\sum_{l \in A_i} \exp(V_{ilt}\theta_i)} f(\theta_i|\lambda) d\theta_i.$$

3.4 Maximum simulated likelihood estimation

Random coefficients logit model can be estimated by simulation methods. We assume that the coefficients of the utility function are normally distributed with density $f(\theta|\lambda)$, where λ are parameters of the distributions of the coefficients. Then, the probability of choosing product j by consumer i is:

$$Pr(y_i = j) = \int P_{ij}(\theta_i) f(\theta_i|\lambda) d\theta_i,$$

where

$$P_{ij}(\theta_i) = \frac{\exp(V_{ijt}\theta_i)}{\sum_{l \in A_i} \exp(V_{ilt}\theta_i)}.$$

We approximated the above probability using simulation. We draw a θ_i from $f(\theta_i|\lambda)$ and denote it θ_i^r where r refers to the number of the draw. Then, we calculate $P_{ij}(\theta_i^r)$ with this draw. We take R draws from the distribution and find the average simulated probability:

$$\hat{Pr}(y_i = j) = \frac{1}{R}\sum_{r=1}^{R} P_{ij}(\theta_i^r).$$

$\hat{Pr}(y_i = j)$ is an unbiased estimator of $Pr(y_i = j)$, and its variance decreases as we increase the number of the draws. This estimator possesses other useful properties. It is smooth and always positive, which is an important requirement to use log-likelihood function, and $\sum_{j\in A_i} \hat{Pr}(y_i = j) = 1$.

Using the same steps, we obtain simulated probabilities for all consumers. Then, we plug in simulated probabilities into the likelihood function and obtain the simulated likelihood function:

$$\hat{L}(\lambda) = \prod_{t=1}^{T}\prod_{i=1}^{N} \hat{Pr}(y_i).$$

By taking logarithm, we obtain the simulated log-likelihood function:

$$\ln \hat{L}(\lambda) = \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j\in A_i} I(y_i = j) \ln \hat{Pr}(y_i = j).$$

The maximum simulated likelihood estimator is the value of λ that maximizes the above function:

$$\hat{\lambda} = \arg\max_{\lambda}[\ln L(\lambda)].$$

3.5 Results

The results of the estimation for the logit demand model are presented in Table 3. The estimated coefficient for the variable that represents the share of other wines with the Cabernet Sauvignon grape variety in the choice set is positive, which means that the higher the share of Cabernet Sauvignon wines in the consumer's choice set, the higher the utility that individual gets from this particular wine. If the wine itself has the Cabernet Sauvignon grape variety, it provides higher utility to the consumer compared to wines with other grape varieties. The same results we observe for the blended wines with the Cabernet Sauvignon and Merlot grape varieties. The coefficients for both variables for own and other wines are positive and significant. The wine with the Sangiovese grape variety, on average, gives lower utility to the individual than wines with other grape varieties. There is no effect of the share of other wines with the Sangiovese grape variety in the choice set on an individual's utility as the estimated coefficient is not significant. If we consider the estimated coefficients for the dummy variables that represent different grape varieties, the Cabernet Sauvignon wines, on average, give the highest utility to the individuals, followed by the Pinot Noir wines and blended wines with the Cabernet Sauvignon and Merlot grape varieties.

The producing country also affects the consumer's utility level. Wine produced in Japan gives higher utility to the consumer compared to those produced in other countries. Simultaneously, the more other wines in the individual's choice set are produced in Japan, the higher utility level the individual gets from the particular wine. The wine would give higher utility if other wines from the individual's choice set were also produced in Japan. The wines produced in the US provide lower utility to the individuals than the wines produced in other countries. The share of other wines produced in the

Table 3: Multinomial logit demand function estimation

VARIABLES	Estimated Coefficients
Own characteristics:	
Pinot Noir	0.394***
	(0.0906)
Sangiovese	-1.218***
	(0.3268)
Cabernet Sauvignon	0.498***
	(0.1298)
Cabernet+Merlot	0.207***
	(0.0827)
Japan	0.68***
	(0.1032)
United States	-0.318***
	(0.0964)
Italy	0.336***
	(0.0969)
France	0.249***
	(0.0815)
Log(Price)	-1.465***
	(0.0420)
Other wines' characteristics:	
Pinot Noir	0.115
	(0.0689)
Sangiovese	0.0182
	(0.1923)
Cabernet Sauvignon	0.246***
	(0.1034)
Cabernet+Merlot	0.141***
	(0.0641)
Japan	0.352***
	(0.0792)
United States	-0.118
	(0.0781)
Italy	0.252***
	(0.0753)
France	-0.0031
	(0.0669)
Constant	10.595***
	(0.3971)
Observations	3300
LLF	-5169.7

*** p<0.01, ** p<0.05, * p<0.1

US in the choice set does not affect an individual's utility level. The estimated coefficients for the Italy and France dummy variables are positive and significant. The wines produced in Italy or France on average provide higher utility to individuals than wines produced in other countries. The higher the share of the other wines produced in Italy in an individual's choice set, the higher the individual's utility level.

The results show that, on average, the individual's utility level depends not only on the wine's own characteristics but on characteristics of other wines in the individual's choice set. These results are consistent with the results obtained for individuals' subjective evaluations of the wines.

Table 4 shows the results of the estimation for the random coefficient model. The values of the estimated coefficients are very close to those obtained in multinomial logit model estimation. However, the standard errors in random coefficient model estimation are high as a result of efficiency loss due to an increase in the number of the parameters.

3.6 Robustness checks

3.6.1 Experience and individuals' reaction

There is a possibility that the behavioral effects responsible for the influence of other available alternatives and their characteristics on choices can play an important role for the inexperienced and negligible role for experienced consumers. For example, List (2004) shows that the influence of the endowment effect on consumers' decision-making process gets smaller as their market experience increases.

To test whether wine drinking experience affects individuals' reaction to other products' characteristics, we divided all individuals into two groups. The first group comprised people who drink red wine at least one day a week, and the second group comprised people who drink red wine less often. We

Table 4: Demand function estimation with random coefficients

VARIABLES	Mean Estimate	Standard Error	Std. Deviation Estimate	Standard Error
Own characteristics:				
Pinot Noir	0.2047	1.3558	1.2356	1.0619
Sangiovese	-1.081	1.4124	1.2238	1.077
Cabernet Sauvignon	0.4942	1.8798	1.2782	1.0825
Cabernet+Merlot	0.4194	1.0916	1.3089	1.0312
Japan	0.7949	1.2242	1.1260	1.0403
United States	-0.1841	1.5052	1.1725	1.0900
Italy	0.3509	1.3761	1.2963	0.9853
France	0.4550	1.5223	1.0644	0.9762
Log(Price)	-1.4662	0.9894	1.2210	1.0706
Other wines' characteristics:				
Pinot Noir	0.4747	1.0528	1.3049	1.0142
Sangiovese	0.2269	1.2945	1.1941	1.1133
Cabernet Sauvignon	0.7047	1.3199	1.2175	0.9656
Cabernet+Merlot	0.4916	1.5602	1.3022	1.004
Japan	1.4769	1.1593	1.2394	1.2634
United States	-0.5772	1.1885	1.1934	1.1286
Italy	1.2169	1.2391	1.2106	0.9847
France	0.0427	1.5118	1.0015	1.2447
Constant	10.5810	1.0053	1.3895	0.9747
Observations	3300	3300	3300	3300

use the frequency of wine drinking as a proxy for wine drinking experience, assuming that people who drink wine more often know more about red wines and their characteristics. We estimated the same multinomial logit model for these two groups of individuals. The results of this estimation are presented in Table 5.

Table 5: Demand function estimation for experienced and inexperienced individuals

VARIABLES	Inexperienced Individuals	Experienced Individuals
Own characteristics:		
Pinot Noir	0.4058***	0.4206***
	(0.1019)	(0.1808)
Sangiovese	-1.3584***	-1.1272***
	(0.4881)	(0.4787)
Cabernet Sauvignon	0.5039***	0.5294***
	(0.1572)	(0.2571)
Cabernet+Merlot	0.2299***	0.1859
	(0.0983)	(0.1684)
Japan	0.6466***	0.8272***
	(0.1229)	(0.2204)
United States	-0.1523***	-0.6962***
	(0.1129)	(0.1807)
Italy	0.5129***	-0.0512
	(0.1134)	(0.2016)
France	0.3642***	-0.0798
	(0.0981)	(0.1677)
Log(Price)	-1.507***	-1.3914***
	(0.0519)	(0.0801)
Other wines' characteristics:		
Pinot Noir	-0.0099	1.6503***
	(0.3132)	(0.6069)
Sangiovese	1.0437	-1.8559
	(1.0204)	(1.3671)
Cabernet Sauvignon	0.7689	1.6254***
	(0.4918)	(0.8389)
Cabernet+Merlot	0.2281	1.3132***
	(0.2971)	(0.5483)
Japan	1.0264***	2.5311***
	(0.3793)	(0.7368)
United States	-0.1586	-1.1726***
	(0.3599)	(0.5629)
Italy	1.0887***	0.7609
	(0.3539)	(0.6713)
France	0.2592	-0.9263
	(0.3118)	(0.5516)
Constant	10.7814***	10.4402***
	(0.4713)	(0.7697)
Observations	2304	996

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

According to these results, experienced and inexperienced consumers have a similar reaction to wine's own characteristics, except for dummy variables that represent blended wines with the Cabernet Sauvignon and Merlot grape varieties and wines produced in Italy and France. The estimated coefficients for these dummy variables are not significant for experienced individuals.

The results show that both groups react to other products' characteristics, but do so in a different way. The grape varieties of other wines affect utility functions of experienced individuals, but do not affect the utility functions of inexperienced individuals. The inexperienced individuals react to other wines that were produced in Japan and Italy, and experienced individuals react to the other wines produced in Japan and the US. Inexperienced individuals are more sensitive to the wine's price than experienced individuals.

When the consumer makes her choice, we cannot distinguish the exact reference point. Most likely, there are several effects that affect her decision. We cannot capture all of them, but our experimental design allows us to test one of them – whether the other alternatives affect consumers' evaluation of the product.

3.6.2 Trials

For the estimations presented above, we used the observations for all three trials. In each trial, individuals were provided with the same choice set and only the wines' prices were changed. In this case, there is a possibility that the choice in the later trial may be affected by the individual's choice in the previous trial. The individuals may be persistent in their choice, and their first chosen alternative may affect the perception of other options in the later trials.

To eliminate the possible influence of the individuals' experience in previous trials, we estimate the same multinomial logit model, but restrict the observations only to the first trials. The results for this estimation are presented in Table 6.

The estimated coefficients for wines' own characteristics are close to those presented in Table 3. However, there are some differences. The estimated coefficient for the Sangiovese grape variety dummy variable becomes insignificant. On average, individuals are less sensitive to the wine's price in the first trial. Regarding other wines' characteristics, the coefficient for the Cabernet Sauvignon variable becomes insignificant if we consider only the first trial. Simultaneously, coefficients for Japan and Italy country of origin variables are still significant, and their values are greater for the first trial subsample.

3.7 Instrument validity of rival products' characteristics

3.7.1 Price endogeneity problem

In this section, we describe how the obtained results are related to the validity of the commonly used approach when characteristics of rival products are used as instruments to overcome the endogeneity problem of market prices. In the demand function, we have aggregate demand or consumers' choices as the dependent variable and market price as the explanatory variable to estimate consumers' reaction to different price levels. Typically, when we use a simple OLS for estimation, the resulting price coefficient estimates are often insignificant, or even significantly positive, implying an upward-sloping demand curve. The reason for this is the potential positive correlation between market prices and unobserved products' characteristics.

Firm j solves the following maximization problem to choose the price

Table 6: Demand function estimation for the first trial

VARIABLES	Estimated Coefficients
Own characteristics:	
Pinot Noir	0.3590***
	(0.1253)
Sangiovese	-0.4020
	(0.4733)
Cabernet Sauvignon	0.3899***
	(0.2018)
Cabernet+Merlot	0.1263
	(0.1276)
Japan	0.5595***
	(0.1374)
United States	-0.3152***
	(0.1385)
Italy	0.2816***
	(0.1417)
France	0.3320***
	(0.1841)
Log(Price)	-1.1130***
	(0.0730)
Other wines' characteristics:	
Pinot Noir	0.5293
	(0.3116)
Sangiovese	1.8384
	(1.0449)
Cabernet Sauvignon	0.3084
	(0.5994)
Cabernet+Merlot	0.2213
	(0.3514)
Japan	0.8989***
	(0.3249)
United States	-0.1333
	(0.2787)
Italy	0.7010***
	(0.3333)
France	0.1123
	(0.6162)
Constant	8.1399***
	(0.5859)
Observations	3300
LLF	-5169.7

*** p<0.01, ** p<0.05, * p<0.1

of its product:

$$\max_{p_j} \Pi_j = (p_j - c_j(\xi_j, X_j))s_j(p, \xi)Q,$$

$$F.O.C.\ s_j(p,\xi) + p_j \frac{\partial s_j(p,\xi)}{\partial p_j} - c_j(\xi_j, X_j) \frac{\partial s_j(p,\xi)}{\partial p_j} = 0,$$

where X_j is a vector of product j's observed characteristics, ξ_j is a vector of product j's unobserved characteristics, $c_j(\xi_j, X_j)$ is a marginal cost of product j, s_j is a probability of choosing good j by a potential consumer, and Q is market size. By solving the system of equations for all firms in the market, we get a vector of prices $p(\xi, X)$ as functions of unobserved product characteristics ξ. As a result, market products' prices are endogenous.

When the firm sets the price for its product, it considers the product's characteristics including those unobserved by the econometrician ξ. For example, ξ can be considered as the quality of the product, then ξ and p are correlated because high-quality products tend to have high costs and high prices respectively. Another way is to consider ξ as a variable that denotes average consumers' value for all unobserved characteristics of the product. In the case of wine choice, ξ may include reputation, the design of the bottle and label of the wine. Individuals may perceive these unobserved characteristics differently, thus, we consider ξ_j as a product-specific mean to highlight the heterogeneity of the consumers' reaction to the different products' attributes. When the firm chooses the price of its product, it considers consumers' reaction to all product characteristics including those unobserved by the econometrician. Firms, usually, conduct marketing research to learn how consumers react to different characteristics of the product and choose the price level respectively. This correlation between prices and unobserved product characteristics create a price endogeneity problem that leads to biased coefficient estimates for the price in the consumers' demand function.

There are several approaches to solve the price endogeneity problem in the estimation of discrete choice models. The first approach is to include product-specific dummy variables to control for unobserved product characteristics. This method is applicable for micro-datasets with a small number of brands as it requires additional parameters to be estimated. If the number of brands is too large, there might not be enough observations to estimate product-specific constants.

The second approach involves the use of instrumental variables for market prices. The most popular instruments are costs' shifters (input prices) or characteristics of rival products. In this thesis, we consider the validity of rival products' characteristics. 2SLS and generalized method of moments (GMM) instrumental variables estimators require linear relationships between outcome and dependent variables, which introduces some limitations for the non-linear discrete choice models. To get a linear relationship Berry (1994) propose an inversion technique to find the implied mean levels of utility for each product. Berry, Levinsohn, and Pakes (1995) use this method for market-level data to estimate demand and supply functions in the US automobile industry. They use the method of moments estimator with characteristics of rival products as instruments. Later this approach was extended by combining market-level data with consumer-level data by Petrin (2002), Berry, Levinsohn, and Pakes (2004), and Goolsbee and Petrin (2004). Control function approach allows using instrumental variables to overcome the endogeneity problem for non-linear models. Petrin and Train (2010) apply the control function approach for solving the price endogeneity problem of households' choices among television options.

Many papers try to find the best instruments for the BLP model. Reynaert and Verboven (2014) using Monte Carlo simulations show that optimal instruments proposed in Chamberlain (1987) reduce small sample bias and increase the estimator's efficiency and stability. However, these optimal instruments depend in a specific way on the product's own characteristics and the characteristics of the other products, which still requires inclusion and exclusion conditions to be satisfied. The validity of rival products' characteristics as price instruments was questioned in Armstrong (2016). He shows that the dependence of prices on other products' characteristics through markups disappears in large markets and estimators based on these are consistent with a large number of small markets and inconsistent in a large market setting. His critique is related to the inclusion condition, while we are arguing that the exclusion condition is not satisfied for these instruments in some cases.

3.7.2 Instrumental validity

The rival product characteristics are valid instruments for the market prices only if relevance and exclusion conditions are simultaneously satisfied. The relevance condition is satisfied if market prices are correlated with rival products' characteristics. This condition can be tested by regressing the market prices on rival products' characteristics. Usually, the relevance condition is satisfied – when the firm sets the price of the product, it takes into account the characteristics of rival products. If rival products are good substitutes, then there is a high level of competition on the market. In this situation, the firm has a low markup. If the level of the competition on the market is low, the firm has a high markup. The level of substitutability between the products is determined by their characteristics. Therefore, a product's market price is correlated with the characteristics of other available products in the market.

The exclusion condition is satisfied if rival products' characteristics are not correlated with the error term in the consumer's indirect utility function. In other words, rival products' characteristics must affect the utility of the product only through the market prices, but not directly.

The exclusion condition may not be satisfied for two reasons. First, when the firm chooses its product's characteristics, including those unobserved by an econometrician, it may consider the characteristics of rival products. In this case, unobserved product's characteristics are correlated with the rival products' characteristics. Second, other products' characteristics may directly enter the consumer's utility function through the behavioral component. This behavioral component captures the direct influence of the other products' characteristics on consumer's utility. In previous sections, we tested and showed the violation of the exclusion condition as a result of the existence of the behavioral component in consumers' choices.

Next, we test whether the exclusion condition is violated as a result of the firm's endogenous choice of product's characteristics that considers the characteristics of rival brands. The exclusion condition is satisfied, and we get unbiased estimated coefficients if other products' characteristics $X^{'}$ are uncorrelated with the error term that includes unobserved product's characteristics ξ.

In a multinomial logit model, we estimate the coefficients of the following indirect utility function:

$$\delta_{ijt} = \beta_0 + \beta_1 p_{ijt} + X_j \alpha + X_{ij}^{'} \gamma + \xi_j. \tag{8}$$

The estimated coefficient for the characteristic k of other products in the choice set is:

$$\hat{\gamma}_k = \gamma_k + \frac{cov(X'^k, \xi)}{var(X'^k)},$$

where γ_k is the true coefficient.

If $cov(X^{'k}, \xi) \neq 0$ the estimated coefficient $\hat{\gamma_k}$ is biased. To test whether unobserved product characteristics are correlated with other products' characteristics we use the following procedure. We estimate the same multinomial logit demand model as we did in Section 3.1, but substitute the characteristics of other products in the individual's choice set with the characteristics of the randomly chosen products outside of the individual's choice set. Instead of Eq.(8), we obtain the following expression for the mean level of utility:

$$\delta_{ijt} = \beta_0 + \beta_1 p_{ijt} + X_j \alpha + X^{''}_{ij} \gamma^{'} + \xi_j, \tag{9}$$

where $X^{''}_{ij}$ is a vector of characteristics of the products outside of the consumer i's choice set. Characteristics of other products outside of individuals' choice sets cannot affect their utility levels and their choices respectively because individuals do not consider them. For this reason $\gamma^{'}_k = 0$.

The estimated coefficient for the characteristic k of other products outside of the choice set in Eq.(9) is:

$$\hat{\gamma}^{'}_k = \gamma^{'}_k + \frac{cov(X^{''k}, \xi)}{var(X^{''k})} = \frac{cov(X^{''k}, \xi)}{var(X^{''k})}.$$

If estimated coefficient $\hat{\gamma}^{'}_k = 0$, then we can conclude that $cov(X^{''k}, \xi) = 0$ and there is no endogeneity problem of other products' characteristics.

The results of Eq.(9) estimations are presented in Table 7. The estimated coefficients for all characteristics of the randomly chosen wines outside of the individuals' choice sets are not significant. The estimated coefficients for the shares of wines with the Cabernet Sauvignon grape variety and blended

Table 7: Demand function estimation with characteristics of the wines outside of the choice set

VARIABLES	Estimated Coefficient
Own characteristics:	
Pinot Noir	0.295***
	(0.0702)
Sangiovese	-1.237***
	(0.2766)
Cabernet Sauvignon	0.300***
	(0.0975)
Cabernet+Merlot	0.0956
	(0.0698)
Japan	0.394***
	(0.0802)
United States	-0.221***
	(0.0761)
Italy	0.134
	(0.0776)
France	0.2549***
	(0.0664)
Log(Price)	-1.467***
	(0.0427)
Other wines' characteristics:	
Pinot Noir	-0.013
	(0.0573)
Sangiovese	-0.154
	(0.1396)
Cabernet Sauvignon	0.0771
	(0.0793)
Cabernet+Merlot	-0.0229
	(0.0536)
Japan	0.0923
	(0.0686)
United States	0.0237
	(0.0602)
Italy	-0.0082
	(0.0576)
France	0.055
	(0.0510)
Constant	11.161***
	(0.3686)
Observations	3300
LLF	-5183.3

*** p<0.01, ** p<0.05, * p<0.1

wines with the Cabernet Sauvignon and Merlot grape varieties are significant when we consider other wines in the individuals' choice sets and they become insignificant if we consider other wines outside of the individuals' choice sets. We observe the same results for the coefficients of the variables that represent the shares of other wines produced in Japan and Italy – coefficients are significant for the wines in individuals' choice sets and insignificant for the wines outside of the individuals' choice sets. According to these results, the characteristics of other products are uncorrelated with the unobserved product characteristics. For this reason, we obtain unbiased estimated coefficients of the characteristics of other products in Eq.(8). Moreover, we can conclude that the exclusion condition is not violated for other products' characteristics as a result of endogenous firm's choice of its product's characteristics based on the characteristics of other products.

Thus, the results show that the exclusion condition is not satisfied because of the existence of the behavioral component but not because of the firm's endogenous choice of its product's characteristics.

3.7.3 Bias in estimates of price coefficient

Below, we show the consequences of using characteristics of rival products as instruments for the market prices in discrete choice models if the exclusion condition is not satisfied.

First, we consider the case of the 2SLS estimators for market level data. Suppose there are J products, I consumers and M markets. The utility of consumer i from product j in market m is:

$$u_{ijm} = \beta_0 + \beta_1 p_{jm} + X_{jm}\alpha + \xi_{jm} + v_{ijm}, \tag{10}$$

where p_{jm} is price of product j in market m, X_{jm} is a row vector of observed characteristics of product j, ξ_{jm} is unobserved part of product j's characteristics and v_{ijm} is an idiosyncratic taste shock.

In the standard model under the assumption that consumers have multinomial logit demand function the share of product j in market m is:

$$s_{jm}(\theta) = \frac{exp(\delta_{jm})}{\sum_{l=0}^{J_m} exp(\delta_{lm})},$$

where $\delta_{jm} = \beta_0 + \beta_1 p_{jm} + X_{jm}\alpha + \xi_{jm}$ is the mean level of utility for product j, $l = 0$ is an outside option (not to buy any product) with $\delta_{0m} = 0, \forall m$, $\theta = (\beta_0, \beta_1, \alpha)$ is a vector of the parameters.

By using the inversion method proposed by Berry (1994) we get a vector of mean utilities for all products in market m:

$$\delta_m(\theta) = s^{-1}(s_m, \theta),$$

where s_m is a vector of observed market shares of the products. Once we obtained mean utilities from the market shares we can estimate the following linear regression:

$$\delta_{jm} = \beta_0 + \beta_1 p_{jm} + X_{jm}\alpha + \xi_{jm}. \tag{11}$$

Suppose we estimate the coefficients of this regression by 2SLS estimator and the characteristics of other products are used as instruments for the market prices.

In the first stage market prices are regressed on the characteristics of other products:

$$p_{jm} = \tau_0 + X_{jm}^{'}\mu + \eta_{jm}, \tag{12}$$

where X'_{jm} is a row vector of the other products' characteristics for product j in market m.

The predicted market prices $\hat{p}_{jm} = \hat{\tau}_0 + X'_{jm}\hat{\mu}$ are obtained from the first stage. Then, the market prices are substituted with the predicted values $\hat{p}$ in Eq.(11):

$$\delta_{jm} = \beta_0 + \beta_1 \hat{p}_{jm} + X_{jm}\alpha + \xi_{jm}. \tag{13}$$

As previously shown, the characteristics of other products may directly enter the consumer's utility function. In this case, they become part of the error term in Eq.(10):

$$v_{ijm} = X'_{jm}\gamma + \epsilon_{ijm}.$$

In this case, the true mean utility of product j in market m is:

$$\delta_{jm} = \beta_0 + \beta_1 p_{jm} + X_{jm}\alpha + \xi_{jm} + X'_{jm}\gamma. \tag{14}$$

As a result, when we estimate Eq.(13) we get a biased price coefficient estimate:

$$\hat{\beta}_1 = \beta_1 + \frac{cov(\hat{p}, X'\gamma)}{var(\hat{p})} = \beta_1 + \frac{cov(X'\hat{\mu}, X'\gamma)}{var(X'\hat{\mu})}.$$

The estimated coefficient is biased because $cov(X'\hat{\mu}, X'\gamma) \neq 0$.

Suppose we now have microdata on individuals' choices of the products instead of market level data. The mean utility of product j for consumer i is:

$$\delta_{ijm} = \beta_0 + \beta_1 p_{jm} + X_{jm}\alpha + \xi_{jm}. \tag{15}$$

To estimate coefficients of the demand function and overcome the price endogeneity problem, we use a control function approach with characteristics of other wines as instruments. We need to use the control function approach because we cannot apply inversion to get a linear expression as we do not observe products' shares.

In the control function approach, we estimate the same regression as we did in the first stage of 2SLS (Eq.(12)) for the market level data, but now we are interested in residuals instead of predicted prices. We insert these obtained residuals $\hat{\eta}_j$ into Eq.(15) to control for unobserved product characteristics;

$$\delta_{ijm} = \beta_0 + \beta_1 p_{jm} + X_{jm}\alpha + \phi\hat{\eta}_{jm}.$$

If the characteristics of other products directly enter consumers' utility functions, then:

$$\delta_{ijm} = \beta_0 + \beta_1 p_{jm} + X_{jm}\alpha + \phi\hat{\eta}_{jm} + X^{'}_{jm}\gamma.$$

Even if we control for other unobserved factors, characteristics of rival firms in the error term are correlated with the market prices which leads to the biased estimated coefficient of the price.

The previous results showed that there is a behavioral component in consumers' choices that captures the direct influence of other products' characteristics on the utility level. We show that some other products' characteristics cannot be used as instruments for the market prices as this will lead to the biased estimated coefficient of the price. This bias appears because the exclusion condition is not satisfied for these instruments.

3.7.4 Simulation

We used simulation to observe the size of the bias of the price coefficient if we use other products' characteristics that enter the utility function directly as instruments for the market prices.

We assumed there are 1,100 different markets. Each market has five wine brands competing with each other. These five wine brands are randomly chosen for each market from the list of wine brands used in our discrete choice experiment. We generate random choice sets of the wines for the markets to have variations in rival products' characteristics that allows using them as instrumental variables for the prices. To generate mean indirect utilities of the consumers, we used the estimated coefficients presented in Table 3. We generate unobserved products' characteristics $\xi \sim U(0,1)$ that enter consumers' utilities functions. The mean indirect utility of product j in market m is:

$$\delta_{jm} = \beta_0 + \beta_1 p_{jm} + X_{jm}\alpha + X^{'}_{jm}\gamma + \xi_{jm}.$$

where p_{jm} is equilibrium price of product j in market m.

At the next stage, we derive the products' shares for each market. Under the assumption of a multinomial logit demand function of the consumers, the share of product j in market m is:

$$s_{jm} = \frac{\exp(\delta_{jm})}{\sum_{l=0}^{J_m} \exp(\delta_{lm})}.$$

We derive the wines' prices for each market as equilibria of Bertrand oligopoly competition. The firm-producer of product j solves the following profit maximization problem for market m:

$$\max_{p_{jm}} \Pi_{jm} = (p_{jm} - c_{jm}(\xi_{jm}))s_{jm}(p, \xi, X, X^{'})Q,$$

$F.O.C.:$

$$s_{jm}(p, \xi, X, X^{'}) + p_{jm}\frac{\partial s_{jm}(p, \xi, X, X^{'})}{\partial p_{jm}} - c_{jm}(\xi_{jm})\frac{\partial s_{jm}(p, \xi, X, X^{'})}{\partial p_{jm}} = 0. \quad (16)$$

The prices of the products in each market is a solution of the system of equations, similar to Eq.(16) for five firm-producers of the products in the market. The generated prices of the wines are correlated with unobserved products' characteristics and characteristics of rival products. In this case, we obtain endogenous market prices. Simultaneously, if we use rival products' characteristics as instruments for the products' prices, the relevance condition would be satisfied.

The ratio of the market share of product j to the market share of the outside option in market m is:

$$\frac{s_{jm}}{s_{0m}} = \exp(\delta_{jm}).$$

When we take logarithm of both parts, we get:

$$\ln(s_{jm}) - \ln(s_{0m}) = \delta_{jm}.$$

First, we estimate the coefficients of the indirect utility function without any correction for the price endogeneity problem by using generated data. For this, we use GMM to estimate the following equation of interest:

$$\ln(s_{jm}) - \ln(s_{0m}) = \beta_0 + \beta_1 p_{jm} + X_{jm}\alpha, \quad (17)$$

where s_{jm} is a market share of product j in market m, s_{0m} is a market share of the outside option in market m. Second, we estimate Eq.(17) by instrumental variables (IV) GMM estimator, where we use other products' characteristics as price instruments.

The results are presented in Table 8. The first column represents the true values of the indirect utility function coefficients that we use for the data generating process. The second column shows the coefficients from GMM estimation without any correction for the price endogeneity. In this case, the estimated price coefficient is upward biased because of the positive correlation between prices and unobserved products' characteristics. The last column represents the results from the IV GMM estimator when we use other products' characteristics as price instruments. The estimated price coefficient is heavily downward biased. This bias occurs because the exclusion condition is not satisfied for other products' characteristics.

Next, we want to observe the size of the influence of the biases in the estimated price coefficients on price elasticities. Unbiased estimation of price elasticity is important for correct predictions of the changes in the market, policy interventions and counterfactual analysis.

Using β_1 from demand function estimation the own-price elasticity of product j in market m is:

$$E^d_{jm} = \frac{\partial s_{jm}}{\partial p_{jm}} \frac{p_{jm}}{s_{jm}}.$$

We need to consider that we estimated the coefficient of the logarithm of the price in our multinomial demand function, therefore, for our simulations

Table 8: Demand function estimations for simulated data

Variables	True Coefficients	Baseline Estimation	IV Estimation
Own characteristics:			
Pinot Noir	0.394	0.2478***	1.1698***
		(0.0285)	(0.0809)
Sangiovese	-1.218	-0.6620***	-2.4718***
		(0.0990)	(0.1585)
Cabernet Sauvignon	0.498	0.4244***	0.7380***
		(0.0408)	(0.1017)
Cabernet+Merlot	0.207	0.2359***	0.4130***
		(0.0259)	(0.0563)
Japan	0.68	0.6059***	1.5916***
		(0.0346)	(0.0946)
United States	-0.318	-0.1131***	0.6784***
		(0.0322)	(0.0827)
Italy	0.336	0.3704***	1.2236***
		(0.0308)	(0.0865)
France	0.249	0.1271***	1.1475***
		(0.0263)	(0.0834)
Log(Price)	-1.465	-0.4708***	-5.3140***
		(0.0351)	(0.3123)
Constant	2.595	3.4503***	8.3079***
		(0.0408)	(0.3164)
Observations		5500	5500

*** p<0.01, ** p<0.05, * p<0.1

we also consider logarithms. Therefore:

$$\frac{\partial s_{jm}}{\partial p_{jm}} = \frac{\frac{\beta_1}{p_{jm}} \exp(\delta_{jm})(\sum_{l=0}^{J_m} exp(\delta_{lm})) - \frac{\beta_1}{p_{jm}} \exp(2\delta_{jm})}{(\sum_{l=0}^{J_m} exp(\delta_{lm}))^2} =$$

$$\frac{\frac{\beta_1}{p_{jm}} \exp(\delta_{jm})(\sum_{l\neq j} \exp(\delta_{lm}))}{(\sum_{l=0}^{J_m} \exp(\delta_{lm}))^2} =$$

$$\frac{\beta_1}{p_{jm}} s_{jm}(1 - s_{jm}).$$

When we insert above expression in price elasticity we obtain:

$$E^d_{jm} = \frac{\partial s_{jm}}{\partial p_{jm}} \frac{p_{jm}}{s_{jm}} = \beta_1(1 - s_{jm}).$$

Table 9 shows the calculated own-price elasticities of demand for the wines in one of the markets for both estimations with and without correction for the price endogeneity. The bias of the estimated price coefficient incurred by using other products' characteristics as instruments when the exclusion condition is not satisfied leads to the inaccurate predictions of the price elasticities.

Table 9: Own-price elasticities of demand

	True Elasticity	Baseline Estimation	IV Estimation
Wine 1	-1.1562	-0.3716	-4.1939
Wine 2	-1.0739	-0.3451	-3.8954
Wine 3	-1.2304	-0.3954	-4.4629
Wine 4	-1.1638	-0.3740	-4.2216
Wine 5	-1.2759	-0.4100	-4.6281

.

3.8 Summary

We estimated the specification of the multinomial logit model with characteristics of other available alternatives in the consumer's utility function. Some of the estimated coefficients of the other products' characteristics are significant, and the results are consistent with the results obtained for the regression with individuals' subjective evaluations.

To account for individuals' heterogeneity, we estimated a random coefficient multinomial logit model. We also checked whether wine drinking experience affects individuals' reaction to other products' characteristics. Both experienced and inexperienced individuals react to other products' characteristics but in a different way. Both experienced and inexperienced individuals react to the country of other wines' origin, and experienced individuals also react to the grape varieties of other wines in a choice set. Inexperienced individuals are more sensitive to the wine's price than experienced individuals.

As the obtained results show that some of the other wines' characteristics directly enter individuals' utility function, the exclusion condition is not satisfied for them, and these characteristics cannot be used as instruments to overcome price endogeneity problem in demand estimation. Otherwise, it leads to biased price coefficients estimates. To show the size of this bias we used simulations. We generated dataset and estimated the parameters of consumers' demand function, first, without correction for the price endogeneity and, then, with correction for the price endogeneity where we used other products' characteristics as instruments. The results of this exercise show that there is a large bias in price coefficient estimates if we use other products' characteristics as instruments while the exclusion condition is not satisfied.

4 Demand model with subjective evaluations of quality and purpose of consumption

We study the role of the individuals' perceptions of the wine's quality in their choice decisions. We propose the multinomial logit model with products' quality subjective evaluations in consumer's utility function. In the experiment, the individuals were provided with the wines that they did not try before and did not know the true quality of the wines. In case of limited information, individuals use wine's own and other wines' characteristics to form their perceptions of the wine's quality. This allows us to use characteristics of other wines as instruments in a control function approach to overcome subjective evaluations endogeneity problem. To compare the effects of product's quality subjective evaluations on individuals' choices we also estimate a linear probability model.

We introduce demand model specification where we add interaction terms of products' characteristics such as price and quality with the individuals' demographic characteristics. We study how the chosen purpose of consumption affects the probability of wine purchase. We estimate the multinomial logit model where coefficients of the consumer's utility function are specific for each purpose of wine consumption.

4.1 Demand model with subjective evaluations of quality

In this chapter, we introduce the model of the consumer's demand function that depends on prices and subjective evaluations of the quality.

There are N consumers, J products, and T trials. Consumer i's indirect utility function of buying product j in trial t is:

$$u_{ijt} = \beta_0 + \beta_1 p_{ijt} + \beta_2 S_{ij} + \eta_{ijt}, \tag{18}$$

where p_{ijt} is a price of product j for consumer i in trial t, S_{ij} is consumer i's subjective evaluation of product j, and η_{ijt} is an error term.

In many models, a product's quality is considered as a part of unobserved product characteristics. Even introducing a product-specific fixed effect does not allow the obtaining of an estimation of the product's quality. Moreover, it does not help to estimate consumers' reaction to the quality level. We take advantage of the available information from the dataset to study this reaction by including individuals' subjective evaluations of quality into consumer's utility function and estimating related coefficients.

The error term in Eq.(18) includes unobserved product's characteristics and a taste shock:

$$\eta_{ijt} = \xi_j + \epsilon_{ijt},$$

where ξ_j is a vector of unobserved characteristics of product j, ϵ_{ijt} is an idiosyncratic taste shock.

Consumer i's subjective evaluation of the product j's quality depend on observed and unobserved characteristics of product j:

$$S_{ij} = \varphi_0 + X_j \varphi_1 + \varphi_2 \xi_j + \nu_{ij}, \tag{19}$$

where X_j is a vector of observed characteristics of product j, and ν_{ij} is an idiosyncratic shock.

We hypothesize that in real-life situations, consumers' subjective evaluations of product quality are formed based on observable product characteristics, unobservable product characteristics, and price. Prices of the products are the signals of their quality levels. In our dataset, individuals evaluated wines' qualities before they observed any prices and used only information on product characteristics. Our survey design eliminated this connection between prices and qualities because the prices in an experiment are random and do not possess any information about the product's quality.

We assume that utility of outside option (not to buy any product) equals zero for all consumers. We consider the multinomial logit demand function. Each consumer faces a choice set consisting of five different products and outside option. Then the probability of choosing product j by consumer i who faces the choice set A_i in trial t is:

$$P(y_{it} = j) = \frac{\exp(\delta_{ijt})}{\sum_{l \in A_i} \exp(\delta_{ilt})}, \tag{20}$$

where

$$\delta_{ijt} = \beta_0 + \beta_1 p_{ijt} + \beta_2 S_{ij} + \xi_j. \tag{21}$$

4.2 Endogeneity problem

In the estimation of the above model, we may face the endogeneity problem of subjective evaluations of quality. Consumers may give high evaluations to well-known wines. If unobserved factors in consumer's utility function include brand recognition, then subjective evaluations are correlated with the error term, and the estimated coefficient of subjective evaluations of quality is biased.

To solve the subjective evaluations' endogeneity problem, we use characteristics of other wines from the consumer's choice set as instruments. We assume that the characteristics of other products in the choice set affect wine's utility only through its subjective evaluation of quality, but not directly. This assumption is consistent with the results presented in Chapter 3, where it was shown that other products' characteristics affect individuals' subjective evaluations of wines' quality.

The individual chooses among the five randomly assigned products and outside option. The characteristics of the other products are randomly allocated and different for different individuals. Thus, the characteristics of the rival products X' can be used as instruments for the consumers' subjective evaluations S. The behavioral bias of the individuals allows us to use other products' characteristics as instruments – when people evaluate something, they compare it with other available alternatives. The individuals' quality evaluations of the wines depend on the characteristics of other wines in the consumer's choice set. For example, the consumer's wine evaluation may be high if other wines in the choice set have unfavorable product characteristics. Simultaneously, characteristics of other wines from the consumer's choice set are uncorrelated with the wine's unobserved characteristics because of the randomness of choice set formation. By providing random and different choice sets for the individuals, we create variability in characteristics of rival products.

For many products such as wines, observed characteristics only convey a relatively small part of the information on product characteristics. Recovering unobserved characteristics ξ_j given a large number of products is difficult. Observed product characteristics X_j and unobserved product characteristics ξ_j may be correlated, because a firm may choose X_j and ξ_j simultaneously. If they are negatively correlated, then the coefficients on observed characteristics may be biased downwards. For this reason, we do not include X_j directly into the consumer's utility function.

4.3 Two-stage least squares

There are two approaches to solve the endogeneity problem of consumers' subjective evaluations. The first is the 2SLS procedure. We use characteristics of other products from the consumer's choice set as instruments for subjective evaluations.

In the first step we estimate the following equation:

$$S_{ij} = \tau + X_{ij}^{'}\gamma + v_{ij}, \tag{22}$$

where τ is a constant and v_{ij} is a shock. We assume $v \sim N(0, \sigma_v^2)$. From the first step, we obtain predicted consumers' subjective evaluations $\hat{S}$.

We insert $\hat{S}$ in Eq.(21) and get:

$$\hat{\delta}_{ijt} = \beta_0 + \beta_1 p_{ijt} + \beta_2 \hat{S}_{ij}.$$

The new log-likelihood function with predicted subjective evaluations is:

$$\ln L(\beta) = \sum_{t=1}^{T}\sum_{i=1}^{N}\left(\sum_{j\in A_i} I(y_{it}=j)\hat{V}_{ijt}\beta - \log(\sum_{l\in A_i}\exp(\hat{V}_{ilt}\beta))\right),$$

where $\hat{V}_{ijt} = (\mathbf{1}, p_{ijt}, \hat{S}_{ij})$ and $\beta = (\beta_0, \beta_1, \beta_2)^T$.

4.4 Control function

Imbens and Wooldridge (2007) argue that the control function approach provides more accurate estimates for non-linear models with endogenous variables than 2SLS. Petrin and Train (2010) apply control function approach to study households' choices among television options, where products' unobserved characteristics are expected to be correlated with the price.

In the control function approach, endogeneity is considered as an omitted variable problem. In our case, the omitted variable is correlated with consumers' subjective evaluations. The probability of choosing the product j by consumer i is

$$Pr(y_{it} = j | p_{ijt}, S_{ij}, \eta_{ijt}).$$

As in 2SLS, we assume the following reduced form for consumers' subjective evaluations:

$$S_{ij} = \tau + X_{ij}^{'}\gamma + v_{ij}. \tag{23}$$

We assume a multivariate normal joint distribution for (η, v). Additionally, we assume the conditional distribution of unobservables $D(\eta|p, S) = D(\eta|v)$. $Pr(y|p, S, v)$ can be obtained from $Pr(y|p, S, \eta)$ by integrating it with respect to $D(\eta|v)$. In this case, we can estimate demand function coefficients by knowing $Pr(y|p, S, v)$.

We estimate Eq.(23) and obtain residuals $\hat{v}$. Then, we insert $\hat{v}$ into the consumer's indirect utility function in the Baseline model and get:

$$\hat{\delta}_{ijt} = \beta_0 + \beta_1 p_{ijt} + \beta_2 S_{ij} + \beta_3 \hat{v}_{ij}.$$

We control for unobservables by using $\hat{v}$ and maximize the following log-likelihood function:

$$\ln L(\beta) = \sum_{t=1}^{T}\sum_{i=1}^{N}\left(\sum_{j \in A_i} I(y_{it} = j)\hat{V}_{ijt}\beta - \log(\sum_{l \in A_i}\exp(\hat{V}_{ilt}\beta))\right),$$

where $\hat{V}_{ijt} = (\mathbf{1}, p_{ijt}, S_{ij}, \hat{v}_{ijt})$ and $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)^T$.

We use a bootstrap method to derive standard errors and critical values for the estimated coefficients.

4.5 Quasi-likelihood function

Since individuals face different choice sets, they may have different outside options. All the above estimation procedures work only under the assumption that all outside options have zero utility. However, even if this assumption does not hold, we still can estimate the price coefficient by only looking at the individuals who made a purchase. We look at the conditional likelihood increment of individuals who chose to buy. That is,

$$Pr(y_{it} = j | j > 0) = \frac{e^{\delta_{ijt}}}{\sum_{l \in \tilde{A}_i} e^{\delta_{ilt}}},$$

where $\tilde{A}_i$ is consumer i's choice set A_i without outside option.

However, given the logit model specification, we get consistent estimations of coefficients, because

$$\frac{Pr(y_{it} = j)}{Pr(y_{it} = l)} = \exp(\beta_1(p_{ijt} - p_{ilt}) + \beta_2(S_{ij} - S_{il})).$$

The log quasi-likelihood function can be written as:

$$\ln L(\beta) = \sum_{t=1}^{T} \sum_{i=1}^{N} \left(\sum_{j \in \tilde{A}_i} I(y_{it} = j) V_{ijt} \beta_j - \log(\sum_{l \in \tilde{A}_i} \exp(V_{ilt} \beta_l)) \right),$$

$$\hat{\beta} = \arg\max_{\beta} [\ln L(\beta)].$$

Independence of irrelevant alternatives is a property assumed by multinomial logit. Our model does not possess this restrictive property. In our model, consumer's utility of the product depends on its price and subjective evaluation of the quality. Subjective evaluations themselves are determined by the product's characteristics and characteristics of the other products from the consumer's choice set. This is how consumer's choice depends on the choice set formation. The quasi-likelihood function allows us to solve the problem of different utilities of outside option for different choice sets. Simultaneously, the inclusion of subjective evaluations of quality in consumer's utility function links the probability of choosing a particular product with the availability of other alternatives.

4.6 Instruments

To solve the subjective evaluations of quality endogeneity problem, we use characteristics of other products from the individual's choice set. We do this because when the individual evaluates the quality of one product, she compares it with other available alternatives. As a result, other products' characteristics affect the evaluation. However, we can use characteristics of other products from the individual's choice set as an instrument for the quality subjective evaluations only if they are not correlated with unobserved product characteristics, $cov(\xi, X'^k) = 0, \forall k$. This possibility exists when the firm decides about the characteristics of the product and considers the characteristics of the rival products. In our dataset, this correlation is unlikely because the wines in the survey's menu are from different categories.

To test whether unobserved product characteristics are correlated with other product characteristics, we use the following procedure. For each individual, we randomly choose five wines outside of her choice set. We then considered a new set $A_i^{'}$ for the individual i consisting of the five randomly chosen wines outside of the $\tilde{A}_i$. We denote $X_{ij}^{''}$ characteristics of other products in the $A_i^{'}$ for product j.

We run the following modified version of the first stage:

$$S_{ij} = \tau^{'} + X_{ij}^{''}\gamma^{'} + v_{ij}^{'}. \tag{24}$$

Then the estimated coefficient for the product characteristic k is:

$$\hat{\gamma_k}^{'} = \gamma_k^{'} + \frac{cov(X''^k, \xi)}{var(X''^k)}. \tag{25}$$

Individual's subjective evaluations of the products' quality should not be influenced by the characteristics of the products outside of her choice set that she does not observe. For this reason $\gamma_k^{'} = 0$ and, as a result, $\hat{\gamma_k}^{'} = \frac{cov(X''^k, \xi)}{var(X''^k)}$. If $\hat{\gamma_k}^{'} = 0$, then we can conclude that $cov(X''^k, \xi) = 0$.

The results of the Eq.(23) and Eq.(24) estimations are presented in Table 10. The estimated coefficients for characteristics of the wines outside of the individuals' choice sets are not significant, and we do not reject a null hypothesis that they equal zero. These results prove that unobserved wines' characteristics are not correlated with the characteristics of other wines from the choice set, and we can use them as an instrument for individuals' subjective evaluations of the quality.

4.7 Results

Results of individuals' demand function estimations are presented in Table 11. In all demand function estimations, we control for the cases when the

Table 10: Subjective evaluations of quality estimation with characteristics of the other wines inside and outside of the individual's choice set

VARIABLES	Inside of the choice set	Outside of the choice set
Cabernet Sauvignon	0.0685***	-0.00143
	(0.0230)	(0.0230)
Pinot Noir	-0.0273*	-0.0157
	(0.0155)	(0.0154)
Sangiovese	0.124***	-0.0283
	(0.0470)	(0.0457)
Sicily	-0.0734**	0.0340
	(0.0309)	(0.0321)
Bordeaux	-0.0369**	-0.0294
	(0.0188)	(0.0196)
Constant	3.484***	3.486***
	(0.0226)	(0.0235)
Observations	4,754	4,754

*** p<0.01, ** p<0.05, * p<0.1

individual had tried the wine before or knew the real price of the wine. We expected subjective evaluations coefficient to be lower for 2SLS and control function (CF) estimation than for Baseline estimation. This is because of the expected positive correlation between wines' unobserved characteristics and individuals' subjective evaluations. However, $\hat{\beta}_2$ is higher in CF and 2SLS estimations than in Baseline estimation. This is explained by measurement errors of individuals' subjective evaluations – the coefficient of respondents' subjective evaluations is positive, and Baseline estimation leads to a negative bias. The measurement error is a part of the error term in the Eq.(18) and creates an endogeneity bias. As a result, $\hat{\beta}_2$ is biased toward zero in Baseline estimation. The measurement errors of subjective evaluations occur because individuals were provided with only six choices of their wines' evaluations, which is small variability. The obtained individuals' wine evaluations are rough measures of their true precise evaluations.

In all cases, subjective evaluations coefficients are positive. This means respondents tend to get higher utility from buying the wine with a higher subjective evaluation of the quality. The individuals will buy the wine with a higher probability if they think that the wine has better quality. Price coefficient in all estimations is negative. This means individuals negatively react to an increase in price. An increase in a wine's price decreases the wine's probability of being purchased by the individual.

Table 11: Demand function estimation

	Baseline	*2SLS*	*CF*	*Baseline*	*2SLS*	*CF*
		Fulllikelihood			*Quasi − likelihood*	
Constant	8.43***	0.47	6.46***	713.34***	693.84***	711.43***
	(0.4402)	(1.5191)	(2.4368)	(21.2400)	(23.2382)	(22.4633)
Log(Price)	−1.60***	−1.42***	−1.60***	−1.56***	−1.39***	−1.54***
	(0.0547)	(0.0478)	(0.0565)	(0.0589)	(0.0496)	(0.0574)
Subjective evaluation	1.25***	3.18**	1.83**	1.20***	3.59***	1.75**
	(0.0523)	(0.4431)	(0.6923)	(0.0537)	(0.8454)	(0.8541)
Observations	2382	2382	2382	2126	2126	2126

In the demand estimation above, we had to eliminate part of the observations for the individuals who replied "I do not know" for some of the wines from their choice set in the subjective evaluation of quality question. However, we can estimate how this uncertainty in individuals' evaluations affects their choices introducing a related dummy variable into the consumer's utility function.

Now consumer i's utility function from choosing product j in period t is:

$$u_{ijt} = \beta_0 + \beta_1 p_{ijt} + \beta_2 D^0_{ij} + \eta_{ijt},$$

where D^0_{ij} is a dummy variable that equals one if individual i answered "I do not know" for the quality's evaluation of wine j and zero otherwise.

Then, we estimate the individuals' demand function as in Baseline specification. The obtained results are presented in Table 12. The estimated coefficient of the dummy variable for unknown subjective evaluations is negative and significant. The wines with unknown subjective evaluations provide lower utility level to the individuals than wines that have subjective evaluations.

Table 12: Demand function estimation with a dummy variable for unknown subjective evaluations

	Baseline	*Baseline*
	Fulllikelihood	*Quasi − likelihood*
Constant	11.57***	716.92***
	(0.3241)	(128.43)
Log(Price)	−1.46***	−1.44***
	(0.0417)	(0.0446)
Dummy for unknown subjective evaluation	−1.15***	−0.76**
	(0.0963)	(0.107)
Observations	3300	3300

*** p<0.01, ** p<0.05, * p<0.1

To compare the effect of different subjective evaluations on the individuals' choices with the effect of unknown subjective evaluations, we estimate the following linear probability model:

$$I_{ijt} = \alpha_0 + \alpha_1 p_{ijt} + \alpha_2 D^1_{ij} + \alpha_3 D^2_{ij} + \alpha_4 D^3_{ij} + \alpha_5 D^4_{ij} + \alpha_6 D^5_{ij} + \eta_{ijt},$$

where I_{ijt} is a dummy variable which equals one if consumer i decides to buy wine j in trial t. The dummy variables represent different subjective evaluations of quality:

- D^1_{ij} is a dummy variable which is equal to one if consumer i evaluates wine j's quality as "tasteless",
- D^2_{ij} -is a dummy variable which is equal to one if consumer i evaluates wine j's quality as "a bit tasteless",
- D^3_{ij} is a dummy variable which is equal to one if consumer i evaluates wine j's quality as "neutral",
- D^4_{ij} is a dummy variable which is equal to one if consumer i evaluates wine j's quality as "delicious",
- D^5_{ij} is a dummy variable which is equal to one if consumer i evaluates wine j's quality as "very delicious",
- D^0_{ij} is a dummy variable which is equal to one if consumer i answers "I do not know" (baseline category) for wine j's quality evaluation.

We also estimate a multinomial logit model that includes all these dummy variables except for baseline category in consumer's utility function. The results of this estimation along with the results of the linear probability model estimation are presented in Table 13.

The coefficient for the "neutral" subjective evaluations is not significant, which means that the effect of "neutral" subjective evaluations on individuals' utility is not different from the effect of unknown subjective evaluations. The estimated coefficients for D^1 and D^2 are negative – wines with "tasteless" and "a bit tasteless" subjective evaluations provide lower utility level to individuals than wines with unknown subjective evaluations. Wines

Table 13: Estimation with dummy variable for all subjective evaluations

VARIABLES	Linear Probability Model	Multinomial Logit Model
Log(Price)	-0.31***	-1.62***
	(0.0105)	(0.0497)
D^1	-0.08*	-1.37***
	(0.0465)	(0.3687)
D^2	-0.08***	-0.44***
	(0.0286)	(0.1808)
D^3	0.02	0.22
	(0.0180)	(0.1235)
D^4	0.23***	1.68***
	(0.0179)	(0.1254)
D^5	0.43***	2.62***
	(0.0284)	(0.1573)
Constant	2.56***	618.11***
	(0.0844)	(3.8305)
Observations	16,500	3300

*** p<0.01, ** p<0.05, * p<0.1

with "delicious" and "very delicious" subjective evaluations provide higher utility level and have a higher chance of being chosen. The dummy variable for unknown subjective evaluations has a negative sign in our first estimation because the effect of all subjective evaluations was aggregated and the effect of the positive subjective evaluations (D^4 and D^5) is stronger than the effect of the negative subjective evaluations (D^1 and D^2). We can conclude that when individuals answer "I do not know" for the evaluations of the wines' qualities, these wines have the same probability of being purchased as wines with "neutral" subjective evaluations.

4.8 Heterogeneous preferences

To introduce heterogeneity of the consumers' preferences, we include interaction terms of prices and subjective evaluations of quality with consumers' demographic characteristics in the utility function. Consumer i's utility function with the interaction terms is:

$$u_{ijt} = \beta_0 + \beta_1 p_{ijt} + \beta_2 S_{ij} + \beta_3 p_{ijt} In_i + \beta_4 S_{ij} In_i + \beta_7 p_{ijt} N_i^a + \beta_8 S_{ij} N_i^a + \epsilon_{ijt}, \quad (26)$$

where In_i is consumer i's income group (higher number corresponds to the higher income level), and N_i^a is the number of adults in consumer i's household.

To estimate Eq.(26), we use a control function approach with a quasi-likelihood function. The results of the estimation are presented in Table 14. Individuals with higher income are less sensitive to the price level but more sensitive to the subjective evaluation of the wine's quality. The more adults an individual has in their household, the more sensitive she is to the price level of the wine. A possible explanation is that these individuals need to manage the family budget more frugally. The coefficient for the interaction term between the number of adults in an individual's household and her subjective evaluation of the wine's quality is not significant.

4.9 Purposes of the wine consumption

In this section, we first show how individuals' choices of the purpose of consumption affect the probability of wine purchase. The wines chosen by the individuals for some situations have a higher probability of being purchased than wines chosen for other situations. People may buy wines for some situations more often than for others.

To examine how the probability of purchase depends on the purpose of consumption we estimate the following equation:

$$I_{ijt} = \alpha_0 + \alpha_1 drink_{ij} + \alpha_2 meal_{ij} + \alpha_3 party_{ij} + \alpha_4 gift_{ij} + \alpha_5 special_{ij} + \alpha_6 dk_{ij} + \epsilon_{ijt}, \quad (27)$$

Table 14: Demand function with demographic characteristics

VARIABLES	*ESTIMATED COEFFICIENTS*
Constant	708.19^{***}
	(4.4939)
Log(Price)	-1.87^{***}
	(0.1421)
Subjective evaluation	1.67^{**}
	(0.8490)
Income $\times$ *Log(Price)*	0.11^{***}
	(0.0369)
Income $\times$ *Subjective evaluation*	0.02^{***}
	(0.0037)
Number of adults $\times$ *Log(Price)*	-0.03^{*}
	(0.0162)
Number of adults $\times$ *Subjective evaluation*	-0.01
	(0.0114)
Observations	2313

*** p<0.01, ** p<0.05, * p<0.1

where I_{ijt} is a dummy variable which equals one if consumer i decides to buy wine j at trial t. The dummy variables represent different purposes of consumption:

- $drink_{ij}$ is a dummy variable which is equal to one if consumer i chooses wine j as a suitable wine for drinking by herself,
- $meal_{ij}$ is a dummy variable which is equal to one if consumer i chooses wine j as a suitable wine for drinking with the meal,
- $party_{ij}$ is a dummy variable which is equal to one if consumer i chooses wine j as a suitable wine for the party,
- $gift_{ij}$ is a dummy variable which is equal to one if consumer i chooses wine j as a suitable wine for the gift,

- $special_{ij}$ is a dummy variable which is equal to one if consumer i chooses wine j as a suitable wine for the special occasions,
- dk is a dummy variable for "I do not know" answer
- nss is a dummy variable for "no suitable situation" (baseline category).

The results of the estimation are shown in Table 15. The wines considered to be suitable for gifts and special occasions have a higher probability to be purchased than wines suitable for other situations. Overall, individuals are more likely to buy wine if they can identify a suitable purpose of use for the wine.

Table 15: Purpose of consumption and purchase decision

VARIABLES	ESTIMATED COEFFICIENTS
$drink$	0.10^{***}
	(0.01)
$meal$	0.13^{***}
	(0.01)
$party$	0.12^{***}
	(0.01)
$gift$	0.20^{***}
	(0.02)
$special$	0.20^{***}
	(0.02)
dk	0.03^{***}
	(0.01)
$const$	0.07^{***}
	(0.01)
Observations	16500

*** p<0.01, ** p<0.05, * p<0.1

Next, we study how individuals' subjective evaluations of wine characteristics affect their choices of the suitable purpose of consumption. For each purpose, we estimate the following linear probability model:

$$I(k)_{ij} = \beta_0 + \beta_1 Im_{ij} + \beta_2 S_{ij} + \epsilon_{ij}, \tag{28}$$

where $I(k)_{ij}$ is a dummy variable which is equal to one if consumer i chooses wine j for the purpose k, S_{ij} is consumer i's subjective evaluation of the wine j's quality, and Im_{ij} is consumer i's general impression of wine j.

The results of the Eq.(28) estimations for different purposes of consumption are presented in Table 16. When people choose wine for personal drinking, they care about the quality and not about the general impression of the wine. When individuals choose wine for a meal, they also do not care about the general impression of the wine, but they care less about the quality than in the situation of personal drinking. The opposite is true for the "gift" and "special occasion" situations – respondents care more about the general impression of the wine and less about the quality. As a result, the general impression of the wine's bottle is essential when respondents choose wine for "social" drinking (when other people are involved). According to these results, it is harder for individuals to choose a suitable purpose for the wine when it has a low subjective evaluation of quality.

4.10 Purpose-specific utility function

Consumer i's indirect utility function of buying product j for purpose k at trial t is:

$$u^k_{ijt} = \beta^k_0 + \beta^k_1 p_{ijt} + \beta^k_2 S_{ij} + \beta^k_3 Im_{ij} + \epsilon^k_{ijt}. \tag{29}$$

We assume that all coefficients and taste shock in the consumer's utility function are purpose specific.

Table 16: Linear probability model for the choice of the purpose

VARIABLES	General impression	Subjective evaluation of quality
drink	-0.116^{***}	0.112^{***}
	(0.01)	(0.01)
meal	0.006	0.028^{***}
	(0.008)	(0.008)
party	0.028^{***}	0.034^{***}
	(0.009)	(0.009)
gift	0.045^{***}	0.013^{**}
	(0.007)	(0.006)
special	0.054^{***}	0.014^{**}
	(0.006)	(0.006)
dk	-0.010	-0.148^{***}
	(0.008)	(0.010)
Observations	4357	4357

*** p<0.01, ** p<0.05, * p<0.1

Suppose that consumer i chooses to buy wine j and she considers this wine to be suitable for purpose k. We then assume that when the individual makes her purchase decision, she chooses the wine for purpose k. In other words, she compares the utility of all wines in her choice set for purpose k. We grouped data according to the purposes of consumption and estimated consumer's utility function coefficients for all purposes.

The probability of choosing product j by consumer i who faces the choice set $\hat{A}_i$ for purpose k is:

$$Pr(y_{it}^k = j) = \frac{e^{\delta_{ijt}^k}}{\sum_{l \in \tilde{A}_i} e^{\delta_{ilt}^k}}. \tag{30}$$

We use a quasi-likelihood function for the estimation because we cannot distinguish choices of outside options for different situations. We also use

Table 17: Results of purpose-specific utility function

	drink	*meal*	*party*	*gift*	*special*	*dk*
Constant	718.28***	715.18***	705.87***	699.16	700.45***	717.81***
	(18.7700)	(26.6981)	(28.3536)	(4239)	(79.6456)	(28.6717)
Log(Price)	−1.64***	−1.48***	−1.40***	−1.76***	−1.52***	−2.02***
	(0.0086)	(0.0345)	(0.0217)	(0.0373)	(0.0579)	(0.0526)
Subjective evaluation	0.16	0.45	2.92	4.62	2.77	0.54
	(0.6729)	(1.8884)	(1.8526)	(2.8608)	(5.5903)	(1.6873)
General impression	−0.11***	0.05	0.21***	0.80***	1.61***	0.48***
	(0.0070)	(0.0189)	(0.0149)	(0.0142)	(0.0567)	(0.0481)
LLF	−1044	−315	−399	−220	−171	−173
Observations	860	255	323	220	198	141

a control function approach. The results are shown in Table 17. Individuals are most sensitive to the price when they choose wine for personal drinking or as a gift. Subjective evaluations' coefficients are not significant for all situations in this model's specification. Individuals do not care about the general impression of the wine when they choose it for personal drinking. Individuals are sensitive to the general impression of the wine when they choose the wine for a party, special occasion or as a gift. These results are consistent with the estimations from the linear probability model for the choice of the consumption's purpose.

4.11 Summary

We estimated demand model where a product's quality subjective evaluation is a part of a consumer's utility function. The results show that consumers positively react to the wine's quality subjective evaluations, but there is heterogeneity in individuals' reaction to wines' characteristics. Individuals in the high-income group are more sensitive to the quality and less sensitive to the price of the wine than people in the low-income group. The results of the estimation of the multinomial logit model with the purpose of consumption specific coefficients of the consumers' utility function show that the individuals are most sensitive to the price when they choose the wine for the personal drinking and as a gift. General impression plays an important role in choices when the purpose of consumption is related to other people (party, special occasion, as a gift), but a negligible role for the personal drinking. Individuals are more likely to buy the wine if they can identify a suitable purpose of consumption for it.

5 Conclusion

In this thesis, we tested the existence of the behavioral component in the consumer's decision-making process that captures the direct influence of other available products and their characteristics on the consumer's utility. For our analysis, we used a dataset on individuals' choices and their subjective evaluations of the quality of red wines obtained from an experiment. To provide evidence of the presence of the behavioral component in the consumer's decision-making process, we estimated whether the characteristics of other products in the choice set affect the individual's subjective evaluation of the wine. We argued that the result obtained using subjective evaluation of the product as a proxy for its utility is more robust to model misspecification than the discrete choice model estimation where only choices are observable.

We proposed the consumer's choice model with characteristics of other products from the choice set in the consumer's utility function. We introduced this modification in our model to reflect a behavioral aspect of consumers' decision-making process. This behavioral aspect can be described as choice set dependent preferences – when the consumer chooses the product, her satisfaction level from the particular product depends on other available alternatives in her choice set and their characteristics.

The results support the proposed modified choice model and provide evidence that other products' characteristics directly affect the consumer's utility level. We found that the behavioral component exists for both experienced and inexperienced individuals. Both groups react to other products' characteristics, but there are some differences in this reaction. Experienced individuals react to both other wines' country of origin and grape variety, while inexperienced consumers react only to the country of origin.

The obtained results are important for solving the price endogeneity problem in demand function estimation. We have shown that the exclusion condition may not be satisfied for some of the other wines' characteristics as they directly enter the consumer's utility function. In this case, other wines' characteristics cannot be used as instruments for the market prices to overcome the endogeneity problem. Moreover, the results show that the exclusion condition is not satisfied because of the existence of the behavioral component, but not because of the firm's endogenous choice of its product's characteristics.

One potential concern of this research is whether the results on the violation of the exclusion condition for other products' characteristics can be generalized to other product categories. The experimental design described in this thesis can be used to test the exclusion condition for other products and more generally, can be adopted for the instrumental variables estimation of the choice models. If there is a need to use instruments to solve the endogeneity problem, the exclusion condition of the instrumental variable can be tested by running a similar experiment with the randomness of the endogenous variable and choice sets of the individuals.

This thesis also studied how the consumers perceive the quality of a wine that they have not tried before from its external characteristics. We estimated the choice model while allowing for the subjective evaluations of wine's quality in the consumer's utility function. This helped us to measure consumers' reaction to the level of the wine's quality. In some cases, individuals could not decide on their evaluations of the wine's quality. We studied how these undetermined quality evaluations influenced individuals' choices. We showed that undetermined subjective evaluations have the same effect on consumers' utility as neutral subjective evaluations.

To solve the subjective evaluations endogeneity problem, we used product characteristics of other wines from the random consumer's choice set as instruments. We assumed that the characteristics of other products in the choice set affect wine's utility only through its subjective evaluation of quality, but not directly. We found that higher subjective evaluation of the wine's quality increases consumer's utility and the probability of purchase. The estimated price coefficient is negative, which is consistent with the results of the previous estimations.

To consider the heterogeneity of consumers' preferences, we used the advantage of the microdata and added interaction terms of the individuals' demographic characteristics with the price and subjective evaluation of the wine's quality to the consumer's utility function. We demonstrated how an individual's sensitivity to the price and quality varies with income and number of adults in their household. We estimated the reaction to the product's own and other products' characteristics in subjective evaluations of quality for different demographic groups.

We examined the influence of the purpose of consumption on individuals' choices of wine. We found that individuals react differently to price and quality when they choose wine for different purposes. Individuals care more about a wine's general impression when they think the wine is appropriate for use related to other people. And they consider wines with high subjective evaluations of quality to be suitable for personal drinking.

To introduce a new product to the market or to change a characteristic of an existing product to increase the profit, firms need to predict how the consumers will react to it. Proposed specifications of the demand model that capture the behavioral aspects of consumers' choices can help to answer this question. Ignoring the influence of other available alternatives and the purpose of consumption may lead to the misrepresentation of the forecasts and market demand implications.

References

ACKERBERG, D. A. (2001): "Empirically distinguishing informative and prestige effects of advertising," *RAND Journal of Economics*, pp. 316–333.

ARMSTRONG, T. B. (2016): "Large market asymptotics for differentiated product demand estimators with economic models of supply," *Econometrica*, 84(5), pp. 1961–1980.

BEN-AKIVA, M., AND B. BOCCARA (1995): "Discrete choice models with latent choice sets," *International Journal of Research in Marketing*, 12(1), pp. 9–24.

BEN-AKIVA, M., D. MCFADDEN, AND K. TRAIN (2015): "Foundations of stated preference elicitation consumer behavior and choice-based conjoint analysis," Working paper.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, pp. 841–890.

——— (2004): "Differentiated products demand systems from a combination of micro and macro data: The new car market," *Journal of Political Economy*, 112(1), pp. 68–105.

BERRY, S. T. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, pp. 242–262.

CEMBALO, L., F. CARACCIOLO, AND E. POMARICI (2014): "Drinking cheaply: the demand for basic wine in Italy," *Australian Journal of Agricultural and Resource Economics*, 58(3), pp. 374–391.

Chakravarti, D., et al. (1983): "A framework for exploring context effects on consumer judgment and choice," *ACR North American Advances*.

Chamberlain, G. (1987): "Asymptotic efficiency in estimation with conditional moment restrictions," *Journal of Econometrics*, 34(3), pp. 305–334.

Ching, A. T., and F. Hayashi (2010): "Payment card rewards programs and consumer payment choice," *Journal of Banking & Finance*, 34(8), 1773–1787.

Combris, P., S. Lecocq, and M. Visser (1997): "Estimation of a hedonic price equation for Bordeaux wine: does quality matter?," *The Economic Journal*, 107(441), pp. 390–402.

Day, G. S., A. D. Shocker, and R. K. Srivastava (1979): "Customer-oriented approaches to identifying product-markets," *The Journal of Marketing*, pp. 8–19.

Ewing, G., and E. Sarigöllü (2000): "Assessing consumer preferences for clean-fuel vehicles: A discrete choice experiment," *Journal of Public Policy & Marketing*, 19(1), pp. 106–118.

Fader, P. S., and L. McAlister (1990): "An elimination by aspects model of consumer response to promotion calibrated on UPC scanner data," *Journal of Marketing Research*, pp. 322–332.

Golan, A., and H. Shalit (1993): "Wine quality differentials in hedonic grape pricing," *Journal of Agricultural Economics*, 44(2), pp. 311–321.

Goolsbee, A., and A. Petrin (2004): "The consumer gains from direct broadcast satellites and the competition with cable TV," *Econometrica*, 72(2), pp. 351–381.

HARDIE, B. G., E. J. JOHNSON, AND P. S. FADER (1993): "Modeling loss aversion and reference dependence effects on brand choice," *Marketing Science*, 12(4), pp. 378–394.

HUBER, J., AND J. MCCANN (1982): "The impact of inferential beliefs on product evaluations," *Journal of Marketing Research*, pp. 324–333.

HUBER, J., J. W. PAYNE, AND C. PUTO (1982): "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis," *Journal of Consumer Research*, 9(1), pp. 90–98.

IMAI, S., Y. OHNO, AND K. NAKATA (2017): "Measuring the different roles of brands with wine-tasting experiment," Discussion paper.

IMBENS, G., AND J. WOOLDRIDGE (2007): *Whats new in econometrics lecture 5*. NBER Summer Institute.

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect theory: An analysis of decision under risk," *Econometrica*, 47(2), pp. 263–291.

KJÆR, T. (2005): *A review of the discrete choice experiment-with emphasis on its application in health care*. Syddansk Universitet Denmark.

LIST, J. A. (2004): "Neoclassical theory versus prospect theory: Evidence from the marketplace," *Econometrica*, 72(2), pp. 615–625.

LOCKSHIN, L., W. JARVIS, F. DHAUTEVILLE, AND J.-P. PERROUTY (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(3), pp. 166–178.

Mazar, N., B. Koszegi, and D. Ariely (2014): "True context-dependent preferences? The causes of market-dependent valuations," *Journal of Behavioral Decision Making*, 27(3), pp. 200–208.

McFadden, D. (1976): "The revealed preferences of a government bureaucracy: Empirical evidence," *The Bell Journal of Economics*, pp. 55–72.

Mtimet, N., and L. M. Albisu (2006): "Spanish wine consumer behavior: A choice experiment approach," *Agribusiness*, 22(3), pp. 343–362.

Mueller, S., P. Osidacz, I. L. Francis, and L. Lockshin (2010): "Combining discrete choice and informed sensory testing in a two-stage process: Can it predict wine market share?," *Food Quality and Preference*, 21(7), pp. 741–754.

Nerlove, M. (1995): "Hedonic price functions and the measurement of preferences: The case of Swedish wine consumers," *European Economic Review*, 39(9), pp. 1697–1716.

Oczkowski, E. (1994): "A hedonic price function for Australian premium table wine," *Australian Journal of Agricultural and Resource Economics*, 38(1), pp. 93–110.

Olson, J. C. (1978): "Inferential belief formation in the cue utilization process," *ACR North American Advances*.

Ocass, A., and H. Frost (2002): "Status brands: examining the effects of non-product-related brand associations on status and conspicuous consumption," *Journal of Product & Brand Management*, 11(2), pp. 67–88.

Petrin, A. (2002): "Quantifying the benefits of new products: The case of the minivan," *Journal of Political Economy*, 110(4), pp. 705–729.

Petrin, A., and K. Train (2010): "A control function approach to endogeneity in consumer choice models," *Journal of Marketing Research*, 47(1), pp. 3–13.

Ratneshwar, S., and A. D. Shocker (1991): "Substitution in use and the role of usage context in product category structures," *Journal of Marketing Research*, pp. 281–295.

Reynaert, M., and F. Verboven (2014): "Improving the performance of random coefficients demand models: the role of optimal instruments," *Journal of Econometrics*, 179(1), pp. 83–98.

Roberts, J. H., and J. M. Lattin (1991): "Development and testing of a model of consideration set composition.," *Journal of Marketing Research*.

Schamel, G., and K. Anderson (2003): "Wine quality and varietal, regional and winery reputations: hedonic prices for Australia and New Zealand," *Economic Record*, 79(246), pp. 357–369.

Seale Jr, J. L., M. A. Marchant, and A. Basso (2003): "Imports versus domestic production: A demand system analysis of the US red wine market," *Review of Agricultural Economics*, 25(1), pp. 187–202.

Sen, A. (1993): "Internal consistency of choice," *Econometrica: Journal of the Econometric Society*, pp. 495–521.

——— (1997): "Maximization and the act of choice," *Econometrica: Journal of the Econometric Society*, pp. 745–779.

Shocker, A. D., M. Ben-Akiva, B. Boccara, and P. Nedungadi (1991): "Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions," *Marketing Letters*, 2(3), pp. 181–197.

Simonson, I., and A. Tversky (1992): "Choice in context: Tradeoff contrast and extremeness aversion," *Journal of Marketing Research*, 29(3), pp. 281.

Srivastava, R. K., M. I. Alpert, and A. D. Shocker (1984): "A customer-oriented approach for determining market structures," *The Journal of Marketing*, pp. 32–45.

Zeithaml, V. A. (1985): "The new demographics and market fragmentation," *The Journal of Marketing*, pp. 64–75.

——— (1988): "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence," *The Journal of Marketing*, pp. 2–22.

A Appendix

A.1 Survey questionnaire, related variables

SC1. Gender
1) Male
2) Female

SC2. Age

SC3. Single or married
1) Single
2) Married (including divorce and widowed)

SC4. Number of adults (over 20 years) in the household

SC5. Do you drink red wine?
1) Yes
2) No

SC6. Whether the answer using any tool to this questionnaire?
1) Laptop, PC
2) Tablet
3) Smartphone or Mobile phone
4) Other

SC7. Household income
1) Less than 2 million yen
2) From 2 million yen to 4 million yen
3) From 4 million to 6 million yen
4) From 6 million yen to 8 million yen
5) From 8 million yen to 10 million yen
6) 10 million yen or more

SC8. Education (in the case of students: what degree are you currently enrolled in?)
1) Middle School
2) High School
3) Professional and Junior College degree
4) Bachelor degree
5) Master's degree and above

Q1S1 Age began drinking red wine

Q2S1. The number of the bottles of red wine that you usually have at home

1) 0
2) 1–3
3) 4–5
4) 6–10
5) 11–20
6) 21 or more

Q3S1. The frequency of drinking red wine
1) Almost every day
2) 4–5 days a week
3) 2–3 days a week
4) 1 day a week
5) 2–3 days a month
6) Once a month
7) Once every 2-3 months
8) Less often

Q4M1. Usual place to drink red wine (home)
1) Yes
0) No

Q4M2. Usual place to drink red wine (restaurant)
1) Yes
0) No

Q4M3. Usual place to drink red wine (wine bar)
1) Yes
0) No

Q4M4. Usual place to drink red wine (Izakaya)
1) Yes
0) No

Q4M5. Usual place to drink red wine (friends and acquaintances home)
1) Yes
0) No

Q4M6. Usual place to drink red wine (other)
1) Yes
0) No

Q5M1. If you drink at home, for what occasion do you usually drink at home? (drink with the meal)
1) Yes
0) No

Q5M2. If you drink at home, for what occasions do you usually drink at home? (drink separately from the meal)
1)Yes
0)No

Q5M3. If you drink at home, for what occasions do you usually drink at home? (drink with visitors)
1) Yes
0) No

Q5M4. If you drink at home, for what occasions do you usually drink at home? (drink on special occasions such as the celebration)
1) Yes
0) No

Q5M5. If you drink at home, for what occasions do you usually drink at home? (do not drink at home)
1) Yes
0) No

Q6M1. Usual reasons to buy red wine (to drink by yourself)
1) Yes
0) No

Q6M2. Usual reasons to buy red wine (for the family)
1) Yes
0) No

Q6M3. Usual reasons to buy red wine (to drink with a lover)
1) Yes
0) No

Q6M4. Usual reasons to buy red wine (to drink with friends and acquaintances)
1) Yes
0) No

Q6M5. Usual reasons to buy red wine (as a gift)
1) Yes
0) No

Q7M1. Where do you usually buy red wine? (liquor stores)
1) Yes
0) No

Q7M2. Where do you usually buy red wine? (wine shops)

1) Yes
0) No

Q7M3. Where do you usually buy red wine? (department store)
1) Yes
0) No

Q7M4. Where do you usually buy red wine? (supermarket)
1) Yes
0) No

Q7M5. Where do you usually buy red wine? (convenience store)
1) Yes
0) No

Q7M6. Where do you usually buy red wine? (discount store)
1) Yes
0) No

Q7M7. Where do you usually buy red wine? (online sales)
1) Yes
0) No

Q7M8. Where do you usually buy red wine? (winery direct sales)
1) Yes
0) No

Q7M9. Where do you usually buy red wine? (other)
1) Yes
0) No

Q8M1. Items you focus on when purchasing a red wine (country of origin)
1) Yes
0) No

Q8M2. Items you focus on when purchasing a red wine (region)
1) Yes
0) No

Q8M3. Items you focus on when purchasing a red wine (producer name, winery name)
1) Yes
0) No

Q8M4. Items you focus on when purchasing a red wine (wine classification)

1) Yes
0) No

Q8M5. Items you focus on when purchasing a red wine (vintage)
1) Yes
0) No

Q8M6. Items you focus on when purchasing a red wine (grape variety)
1) Yes
0) No

Q8M7. Items you focus on when purchasing a red wine (price)
1) Yes
0) No

Q8M8. Items you focus on when purchasing a red wine (design, bottle label)
1) Yes
0) No

Q8M9. Items you focus on when purchasing a red wine (cork /screw)
1) Yes
0) No

Q8M10. Items you focus on when purchasing a red wine (taste, color, smell)
1) Yes
0) No

Q8M11. Items you focus on when purchasing a red wine (experts' evaluation)
1) Yes
0) No

Q8M12. Items you focus on when purchasing a red wine (awards)
1) Yes
0) No

Q8M13. Items you focus on when purchasing a red wine (recommended)
1) Yes
0) No

Q8M14. Items you focus on when purchasing a red wine (consumers reviews)
1) Yes

0) No

Q9S1. In what price range do you buy red wine more often?
1) Less than 500 yen
2) 500 –1,000 yen
3) 1,000–1,500 yen
4) 1,500–2,000 yen
5) 2,000–2,500 yen
6) 2,500–3,000 yen
7) 3,000–5,000 yen
8) 5,000–7,000 yen
9) 7,000–10,000 yen
10) 10,000–20,000 yen
11) 20,000–30,000 yen
12) More than 30,000 yen

Q10S1. Do you like Japanese or foreign red wine?
1) Japanese
2) Foreign
3) Indifferent

Q11M1. Favorite Japanese origin for red wine (Hokkaido)
1) Yes
0) No

Q11M2. Favorite Japanese origin for red wine (Nagano Prefecture)
1) Yes
0) No

Q11M3. Favorite Japanese origin for red wine (Yamagata Prefecture)
1) Yes
0) No

Q11M4. Favorite Japanese origin for red wine (Yamanashi Prefecture)
1) Yes
0) No

Q11M5. Favorite Japanese origin for red wine (Iwate Prefecture)
1) Yes
0) No

Q11M6. Favorite Japanese origin for red wine (other)
1) Yes
0) No

Q11M7. Favorite Japanese origin for red wine (Indifferent)

1) Yes
0) No

Q11M8. Favorite Japanese origin for red wine (I do not like Japanese red wine)
1) Yes
0) No

Q12M1. Favorite country of origin for red wine (other than Japan) [France]
1) Yes
0) No

Q12M2. Favorite country of origin for red wine (other than Japan) [Chile]
1) Yes
0) No

Q12M3. Favorite country of origin for red wine (other than Japan) [Italy]
1) Yes
0) No

Q12M4. Favorite country of origin for red wine (other than Japan) [Spain]
1) Yes
0) No

Q12M5. Favorite country of origin for red wine (other than Japan) [the United States]
1) Yes
0) No

Q12M6. Favorite country of origin for red wine (other than Japan) [Australia]
1) Yes
0) No

Q12M7. Favorite country of origin for red wine (other than Japan) [other]
1) Yes
0) No

Q12M8. Favorite country of origin for red wine (other than Japan) [Indifferent]
1) Yes

0) No

Q12M9. Favorite country of origin for red wine (other than Japan) [I do not like foreign red wine]
1) Yes
0) No

Q13FA1. Favorite foreign origin for red wine 1

Q13FA2. Favorite foreign origin for red wine 2

Q13FA3. Favorite foreign origin for red wine 3

Q14M1. Favorite grape varieties of red wine (Cabernet Sauvignon)
1) Yes
0) No

Q14M2. Favorite grape varieties of red wine (Pinot Noir)
1) Yes
0) No

Q14M3. Favorite grape varieties of red wine (Merlot)
1) Yes
0) No

Q14M4. Favorite grape varieties of red wine (Syrah)
1) Yes
0) No

Q14M5. Favorite grape varieties of red wine (Sangiovese)
1) Yes
0) No

Q14M6. Favorite grape varieties of red wine (other)
1) Yes
0) No

Q14M7. Favorite grape varieties of red wine (I do not have a particular favorite grape variety)
1) Yes
0) No

Q14M8. Favorite grape varieties of red wine (I do not know)
1) Yes
0) No

Q15S1. The degree of preferred bitterness of the taste of red wine (tannins)

1) Weak
2) Slightly weak
3) Moderate
4) Slightly strong
5) Strong
6) I do not care

Q16S1. The degree of preferred acidity of red wine
1) Weak
2) Slightly weak
3) Moderate
4) Slightly strong
5) Strong
6) I do not care

Q17S1. Your preferred alcohol by volume of red wine
1) Low
2) Moderate
3) High
4) I do not care

Q18S1. Preferred type of red wine
1) Light body
2) Slightly light body
3) Medium body
4) Slightly full body
5) Full body
6) I do not care

Q19AM1. Information that can be read from the description or the label of the wine [wine A] (country of origin)
1) Yes
0) No

Q19AM2. Information that can be read from the description or the picture of the wine [wine A] (region)
1) Yes
0) No

Q19AM3. Information that can be read from the description or the picture of the wine [wine A] (producer name, winery name)
1)Yes
0) No

Q19AM4. Information that can be read from the description or the picture of the wine [wine A] (wine name)

1) Yes
0) No

Q19AM5. Information that can be read from the description or the picture of the wine [wine A] (wine classification)
1) Yes
0) No

Q19AM6. Information that can be read from the description or the picture of the wine [wine A] (vintage)
1) Yes
0) No

Q19AM7. Information that can be read from the description or the picture of the wine [wine A] (grape variety)
1) Yes
0) No

Q19AM8. Information that can be read from the description or the picture of the wine [wine A] (taste, color, smell)
1) Yes
0) No

Q20AS1. Your experience with this wine [wine A]
1) I have not tried this wine before
2) I have not tried it, but I know about this wine
3) I do not know about this wine, but I know the producer
4) I do not know about wine or its producers

Q21AS1.What do you think about the design of this wine bottle? [Wine A]
1) I like it very much
2) I like it
3) I neither like nor dislike it
4) I dislike it
5) I dislike it very much
6) I am not interested in design

Q22AFA1. The reason for the answer [wine A]

Q23AS1. Impression who looked at this wine bottle [wine A]
1) Very luxurious
2) There is a sense of quality
3) Normal
4) Cheap
5) Very crappy

6) I do not know

Q24AM2. The reason for the answer [wine A] (color and design of the label)
1) Yes
0) No

Q24AM3. The reason for the answer [wine A] (label font)
1) Yes
0) No

Q24AM4. The reason for the answer [wine A] (details are listed on the label)
1) Yes
0) No

Q24AM5. The reason for the answer [wine A] (color and design of the back label)
1) Yes
0) No

Q24AM6. The reason for the answer [wine A] (details are listed on the back label)
1) Yes
0) No

Q24AM7. The reason for the answer [wine A] (producer)
1) Yes
0) No

Q24AM8. The reason for the answer [wine A] (the name of the wine)
1) Yes
0) No

Q24AM9. The reason for the answer [wine A] (country of origin and region)
1) Yes
0) No

Q24AM10. The reason for the answer [wine A] (grape variety)
1) Yes
0) No

Q24AM11. The reason for the answer [wine A] (have seen the price of this wine)
1) Yes

0) No

Q24AM12. The reason for the answer [wine A] (the color of the foil)
1) Yes
0) No

Q24AM13. The reason for the answer [wine A] (cork/screw)
1) Yes
0) No

Q25AS1. From the impression of a wine bottle, what do you think about this wine? [Wine A]
1) Very delicious
2) Delicious
3) Neutral
4) Slightly tasteless
5) Tasteless
6) I do not know

Q26AM1. The reason for the answer [wine A] (bottle shape and atmosphere)
1) Yes
0) No

Q26AM2. The reason for the answer [wine A] (color and design of the label)
1) Yes
0) No

Q26AM3. The reason for the answer [wine A] (label font)
1) Yes
0) No

Q26AM4. The reason for the answer [wine A] (details are listed on the label)
1) Yes
0) No

Q26AM5. The reason for the answer [wine A] (color and design of the back label)
1) Yes
0) No

Q26AM6. The reason for the answer [wine A] (details are listed on the back label)
1) Yes

0) No

Q26AM7. The reason for the answer [wine A] (producer)
1) Yes
0) No

Q26AM8. The reason for the answer [wine A] (the name of the wine)
1) Yes
0) No

Q26AM9. The reason for the answer [wine A] (country of origin and region)
1) Yes
0) No

Q26AM10. The reason for the answer [wine A] (grape varieties)
1) Yes
0) No

Q26AM11. Reason [wine A] of the answer (tried this wine before)
1) Yes
0) No

Q27AS1. Describe your impression of this wine [wine A]
1) Traditional, classic atmosphere
2) Novel, innovative
3) Neither

Q27AS2. Describe your impression of this wine [wine A]
1) Bold
2) Light
3) Neither

Q28AS1. Situations in which this wine is likely to be suitable [wine A]
1) Good for drinking when eating
2) Good for drinking with meals
3) Good for drinking at parties
4) Good for gifts
5) Good for drinking at special times
6) I do not know
7) There is no situation suitable

Q29S1. If 5 wines were selling at the provided prices, which wine would you buy?

Q29M1. Items that you considered in your choice of the wine (producer)
1) Yes
0) No

Q29M2. Items that you considered in your choice of the wine (region)
1) Yes
0) No

Q29M3. Items that you considered in your choice of the wine (producer name / winery name)
1) Yes
0) No

Q29M4. Items that you considered in your choice of the wine (wine classification)
1) Yes
0) No

Q29M5. Items that you considered in your choice of the wine (vintage)
1) Yes
0) No

Q29M6. Items that you considered in your choice of the wine (grape variety)
1) Yes
0) No

Q29M7. Items that you considered in your choice of the wine (price)
1) Yes
0) No

Q29M8. Items that you considered in your choice of the wine (design of bottles, labels)
1) Yes
0) No

Q29M9. Items that you considered in your choice of the wine (cork/screw)
1) Yes
0) No

Q29M10. Items that you considered in your choice of the wine (taste, color, scent)
1) Yes
0) No

Q29M11. Items that you considered in your choice of the wine (none of them were taken into consideration)

1) Yes

0) No

A.2 Wine list

Wine name	Year	Country	Region	Wine Body	Grape variety
1. Okushiri	2012	Japan	Hokkaido	Medium	Merlot
2. Shiojiri	2010	Japan	Nagano	Full	Merlot
3. Kamino no Meru	2012	Japan	Yamagata	Medium	Merlot
4. Chateau Mercian	2012	Japan	Nagano	Full	Merlot
5. Chateau Mercian	2012	Japan	Nagano	Medium	Merlot
6. Tomi no Oka Red	2011	Japan	Yamanashi	Medium	Merlot, Cabernet Sauvignon
7. Nielson	2012	USA	California	Full	Pinot Noir
8. It's a game!	2010	Italy	Tuscany	Full	Sangiovese
9. Baron de Rothschild	2011	France	Bordeaux	Medium	Merlot, Cabernet Sauvignon
10. Rupert and Rothschild	2011	South Africa	West Cape	Full	Merlot, Cabernet Sauvignon
11. Cheretto barolo	2008	Italy	Piedmont	Full	Nebbiolo
12. Bonterra	2012	USA	California	Full	Cabernet Sauvignon
13. Domaine Weinbach	2009	France	Alsace	Medium	Pinot Noir
14. Chateau Saint Michel	2011	USA	Washington State	Full	Merlot, Cabernet Sauvignon
15. Los Vascos	2013	Chile	Colchagua Valley	Full	Cabernet Sauvignon
16. Stags' Leap	2011	USA	California	Full	Merlot, Cabernet Sauvignon
17. Amiral de Beychevelle	2010	France	Bordeaux	Full	Merlot, Cabernet Sauvignon
18. Chateau Grand Jean	2012	France	Bordeaux	Medium	Merlot, Cabernet Sauvignon
19. Chateau Giscours	2010	France	Bordeaux	Full	Merlot, Cabernet Sauvignon
20. Coteau de Vernon Condrieu	2012	France	Cotes du Rhone	Medium	Schiller
21. Macon-Lugny	2011	France	Burgundy	Medium	Pinot Noir
22. Bertrand Ambroise	2008	France	Burgundy	Medium	Pinot Noir
23. Reserve Belt Mu	2013	France	Languedoc	Medium	Cabernet Sauvignon
24. Delas Freres	2011	France	Cotes du Rhone	Full	Grenache, Schiller
25. Duemani Cifra	2012	Italy	Tuscany	Full	Cabernet Franc
26. Lucente	2011	Italy	Tuscany	Medium	Merlot, Sangiovese
27. Barone Ricasoli Chianti	2013	Italy	Tuscany	Medium	Sangiovese
28. Tenuta San Guido Le Difese	2012	Italy	Tuscany	Medium	Sangiovese, Cabernet Sauvignon

Wine name	Year	Country	Region	Wine Body	Grape variety
29. Castano Solanera	2011	Spain	Yecla	Full	Monastrell, Cabernet Sauvignon, Grenache
30. Marques de Riscal	2009	Spain	Rioja	Medium	Templaranillo
31. Juan Gil Silver Label	2012	Spain	Jumilla	Full	Monastrell
32. Tinto Pesquera	2011	Spain	Ribera Del Duero	Full	Tempraninino
33. Simi Alexander Valley	2012	USA	California	Medium-Full	Cabernet Sauvignon, Merlot+
34. Saint Clement	2011	USA	California	Full	Cabernet Sauvignon, Merlot, Petit Verdot+
35. Silver Stone	2010	USA	California	Medium-Full	Merlot
36. Newton Claret	2012	USA	California	Full	Merlot, Cabernet Sauvignon, Petit Verdot+
37. Cono Sur	2013	Chile	Central Valley	Medium	Pinot Noir
38. Echeverria	2012	Chile	Casablanca Valley	Medium	Pino Noir
39. Cono Sur	2012	Chile	Casablanca Valley	Medium-Full	Pinot Noir
40. Santa Rita	2012	Chile	Central Valley	Medium	Merlot
41. Cono Sur	2011	Chile	Colchagua Valley	Full	Merlot, Schiller, Cabernet Sauvignon
42. Planeta Plumbago	2012	Italy	Sicilia	Medium	Nero d'Avola
43. Migration	2012	USA	California	Medium	Pinot Noir
44. Beringer	2012	USA	California	Light-Medium	Cabernet Sauvignon
45. Paul Jaboulet Aine	2012	France	Cotes du Rhone	Medium	Grenache, Syrah+
46. Morey-Saint-Denis	2009	France	Burgundy	Full	Pinot Noir
47. Chateau Meaume	2010	France	Bordeaux	Full	Merlot, Cabernet Franc, Cabernet Sauvignon
48. Petite Sirene	2010	France	Bordeaux	Medium	Merlot, Cabernet Sauvignon
49. Poggio Ai Ginepri	2010	Italy	Tuscany	Medium	Cabernet Sauvignon, Merlot, Schiller+
50. ValdiVieso	-	Chile	Central Valley	Full	Cabernet Sauvignon, Carmenere, Malbec+
51. Baglio Di Pianetto	2011	Italy	Piedmont	Full	Nero d'Arvora
52. Domaine Henri Boillot	2012	France	Burgundy	Full	Pinot Noir
53. Lamura	2012	Italy	Sicilia	Medium	Nero d'Avola
54. Castel Sallegg	2011	Italy	Alto Adige	Medium	Merlot
55. Santa Carolina	2013	Chile	Maipo Valley	Medium	Merlot

A.3 Market prices of the wines

Wine name	Market Price (Yen)	Wine name	Market Price (Yen)
1. Okushiri	2667	29. Castano Solanera	1500
2. Shiojiri	3590	30. Marques de Riscal	2500
3. Kamino no Meru	3060	31. Juan Gil Silver Label	2850
4. Chateau Mercian	3380	32. Tinto Pesquera	4200
5. Chateau Mercian	1640	33. Simi Alexander Valley	2580
6. Tomi no Oka Red	3540	34. Saint Clement	5500
7. Nielson	2833	35. Silver Stone	1500
8. It's a game!	4444	36. Newton Claret	2900
9. Baron de Rothschild	2276	37. Cono Sur	880
10. Rupert and Rothschild	4050	38. Echeverria	1500
11. Cheretto barolo	4518	39. Cono Sur	2512
12. Bonterra	2485	40. Santa Rita	1200
13. Domaine Weinbach	3020	41. Cono Sur	2512
14. Chateau Saint Michel	3444	42. Planeta Plumbago	3200
15. Los Vascos	1481	43. Migration	4800
16. Stags' Leap	4100	44. Beringer	1100
17. Amiral de Beychevelle	4055	45. Paul Jaboulet Aine	2350
18. Chateau Grand Jean	1407	46. Morey-Saint-Denis	6000
19. Chateau Giscours	2314	47. Chateau Meaume	2380
20. Coteau de Vernon Condrieu	2540	48. Petite Sirene	1481
21. Macon-Lugny	1500	49. Poggio Ai Ginepri	2500
22. Bertrand Ambroise	3000	50. ValdiVieso	4180
23. Reserve Belt Mu	1200	51. Baglio Di Pianetto	1809
24. Delas Freres	1500	52. Domaine Henri Boillot	4000
25. Duemani Cifra	3500	53. Lamura	1290
26. Lucente	5500	54. Castel Sallegg	1990
27. Barone Ricasoli Chianti	1600	55. Santa Carolina	1018
28. Tenuta San Guido Le Difese	4150		