*Journal of Wine Economics*, Volume 13, Number 4, 2019, Pages 384–393
doi:10.1017/jwe.2018.36

# Does Blind Tasting Work? Investigating the Impact of Training on Blind Tasting Accuracy and Wine Preference*

**Qian Janice Wang** [a] **and Domen Prešern** [b]

## Abstract

We analyzed data from Oxford University Blind Tasting Society's 2018 training season to assess whether blind tasting training improves accuracy. Over time, guesses for grape variety increased in terms of accuracy as well as within-group agreement. Moreover, for grape variety, location, and vintage, the chances of the most common within-group guess being correct were significantly higher than the underlying frequency distribution. Finally, we observed a shift in preference towards older wines, with those with little initial experience gaining a preference for greater acidity and alcohol, and decreasing their preference for oak. Our results have important implications for growing wine markets with an increasingly educated consumer population. (JEL Classifications: C91, C92, D83, L15, L66).

**Keywords:** blind wine tasting, training, wine expertise.

## I. Introduction

There is extensive literature on differences between wine novices and experts when it comes to perceptual sensitivity, the use of descriptions, and wine preference (D'Alessandro and Pecotish, 2013; Gawel, 1997; Hughson and Boakes, 2002; Parr, Heatherbell, and White, 2002; Royet et al., 2013; Weil, 2007). What is relatively unknown, however, is an assessment of how the wine tasting experience changes with the acquisition of expertise. As the Oxford University Blind Tasting Society prepares for its annual Varsity Blind Tasting match against Cambridge (Segal, 2013), we had

* The authors would like to thank the American Association of Wine Economists (AAWE) for awarding our project with the AAWE Research Scholarship, AAWE Conference participants, and an anonymous reviewer for useful suggestions, and the members of the Oxford University Blind Tasting Society for their support.
[a] Crossmodal Research Laboratory, Department of Experimental Psychology, University of Oxford, Anna Watts Building, Oxford, OX2 6BW, UK; e-mail: qian.wang@psy.ox.ac.uk.
[b] Physical & Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, UK; e-mail: domen.presern@chem.ox.ac.uk.

a unique opportunity to follow the progress of a group of students as they go through an intensive five-week training program.

Is blind tasting just nonsense? Rosenheck (2017) performed an analysis of guesses made at the 2017 Varsity Blind Tasting match and demonstrated that most wines were deduced at a significantly better rate than chance. By collecting participants' tasting sheets and guesses during their training program, we aimed to assess the impact of blind tasting training in terms of group accuracy and individual accuracy (including structural elements such as acid and alcohol). We also kept track of any changes in wine preference throughout the training session, to study how the acquisition of expertise may shape buying intentions. Given the rapid increase in the number of knowledgeable wine tasters in developing markets (e.g., China, see Li and Bardaji, 2016), it is crucial to understand how increasing expertise can influence preference for specific wine styles (e.g., less fruit-driven, more bottle maturation, etc.). Because the wines were delivered blind, we were able to track the preferences only based on the intrinsic attributes of the wine, without the participants being biased by the label, price, or origin information (Almenberg and Dreber, 2011; D'Alessandro and Pecotish, 2013).

### *A. Influence of Wine Expertise and Training on the Wine Tasting Experience*

Training improves the ability of people to discriminate the flavors when tasting wine (Owen and Machamer, 1979), possibly because trained panelists and experts can adapt an analytical strategy that helps them to distinguish different components of wine flavors, compared to panelists without training (Arvisenet, Guichard, and Ballester, 2016). Furthermore, several neuroimaging studies involving wine have been conducted with the goal of pinpointing the influence of expertise on multisensory integration in wine evaluation. Sommeliers activate regions of the brain that are involved in high-level cognitive processes, such as working memory and behavioral strategies when tasting wine, a difference from novices who activate the primary gustatory cortex and areas of emotional processing (Castriota-Scanderbeg et al., 2005).

In terms of language usage, wine experts use a different vocabulary when describing wines, using analytical terms whereas non-experts use holistic terms (Weil, 2007). Moreover, there is evidence that experts are more consistent in terms of vocabulary use due to an alignment of sensory concepts (Hughson and Boakes, 2002; Ishii and O'Mahony, 1991).

Taken together, then, these studies suggest that training in wine tasting results in a different way of thinking about and describing wine. Experts in wine are trained to categorize and search for specific flavors or combination of flavors in a wine, and to develop their own prototypes for the representation of wine that includes sensory as well as inferred (such as referring to the winemaking process) and hedonic information (Brochet and Dubourdieu, 2001).

### *B. Aims and Contributions*

We present five hypotheses, related to whether and how training could improve blind tasting accuracy, and, from a commercial prospective, how wine preferences might change with increased wine-tasting experience.

*Hypothesis 1:*
Training will reduce within group variability in guesses (Ishii and O'Mahony, 1991).

*Hypothesis 2:*
Training will make participants more accurate in their guesses (Hughson and Boakes, 2002; Owen and Machamer, 1979; Parr, Heatherbell, and White, 2002).

*Hypothesis 3:*
Training will make participants more accurate in their structural assessment of wines (Arvisenet, Guichard, and Ballester, 2016).

*Hypothesis 4:*
Training will change the way participants write tasting notes, leading to the usage of more specific descriptors and longer tasting notes (Brochet and Dubourdieu, 2001; Gawel, 1997; Weil, 2007).

*Hypothesis 5:*
Training will change the tasters' wine preference, leading them to prefer less fruit-forward and/or more structured wines (D'Alessandro and Pecotish, 2013; Goldstein et al., 2008).

## II. Methods and Materials

### *A. Participants*

A total of 23 people was recruited by the Oxford University Blind Tasting Society. After the training period, we chose to analyze the tasting sheets of 15 participants (3 women, 12 men, mean age = 25.2 years, standard deviation = 3.97), who participated in at least 9 of the 18 training sessions. Nine participants had a very limited exposure to formal wine tasting (hereafter labeled as "novice"). The remaining six participants (hereafter labelled as "experienced") were returning team members or had a significant history of involvement in wine societies (~10 years).

### *B. Experiment Design*

The longitudinal study consisted of a training scheme that lasted for 5 weeks (January 13–February 17), with a total of 18 training sessions. Tasting sheets from participants were collected and analyzed from all sessions (see Figure 1).

### *C. Procedure*

The training sessions were held at the University of Oxford. In each session, participants tasted 10–12 wines in silence on two flights. The wines were served at room temperature in 30-mL samples, in ISO standard 215 mL wine-tasting glasses. In a

*Figure 1*

**Tasting Sheet Used by Participants at Each Training Session**

OUBTS Blind Tasting Sheet

Taster: ..... Date……..

| Wine No. | Predominant Grape Variety (5) | Country of Origin (3) | Main Viticultural Region (2) | Sub-District (3) | Vintage (2) | Notes and Comments Leading to Identification (5) | Preference |
|---|---|---|---|---|---|---|---|
| 1 | | | | | | | |
| 2 | | | | | | | |
| 3 | | | | | | | |
| 4 | | | | | | | |
| 5 | | | | | | | |
| 6 | | | | | | | |

***Preference: 1 = pour it out; 2 = finish the glass; 3 = have a second/third glass; 4 = buy a bottle; 5 = buy a case

Fig. 1 - Colour online, B/W in print

given session, all participants are presented with the same wines in the same order, although they could choose the order of tasting within each flight. Participants were given 30 minutes per flight; for each wine, they gave a guess for the dominant grape variety, place of origin (country, region, and sub-region), and vintage, as well as provided a tasting note that supported their guess and rated their preference for the wine. Preference was measured via a 5-point scale (1 = pour it out; 2 = finish the glass; 3 = have a second/third glass; 4 = buy a bottle; 5 = buy a case).

## III. Results and Discussion

In total, 196 tasting sheets (2,271 tasting notes) were analyzed. A total of 212 wines were included (104 whites, 109 reds) ranging in vintages from 2005–2017. On average, the wines cost £15.38 (standard deviation = 5.08).

### *A. Within-Group Variability (Hypothesis 1)*

We hypothesised that training should reduce within-group variability of guesses; in other words, we expected the participants to agree more in terms of variety of grape, place of origin, and vintage as time passed. Pearson's correlation between time and average number of people per guess showed a significant increase in the number of people who agreed in their grape variety guess over time ($r = 0.13$, $p < 0.05$).

We were also interested in whether implicit group-think was accurate. We calculated the frequency of the most common guess among participants for each correct wine (see Table 1), and compared it with the probability of getting the variety/country/region/vintage correct simply by guessing the most frequently shown attributes. Chi-squared tests of independence showed that the probability of the most common guess being correct was statistically higher than if they had just guessed the most frequently occurring grape variety, country, region, and vintage for each sample.

### *B. Accuracy in Guessing Grape Variety, Place of Origin, and Vintage (Hypothesis 2)*

The Pearson correlation coefficients were calculated for the time-dependency of the average number of correct answers for grape variety, country, region, and vintage. While the accuracy for the grape variety did increase with time ($r = 0.21$, see Figure 2), we did not observe any improvements in accuracy with regard to country and region guesses ($r = –0.01$ and $–0.02$, respectively). In fact, the accuracy of the vintage decreased with time ($r = –0.16$).

### *C. Accuracy in Structural Assessment of Wines (Hypothesis 3)*

With regard to structural elements in the wine, we assessed the accuracy of the participants in the evaluation of acid and alcohol levels. The Pearson correlation

*Table 1*
**Most Common Guesses Among Participants**

| | *Grape Variety* | *Country* | *Region* | *Vintage* |
|---|---|---|---|---|
| Number of occurrences of most common guess being correct | 94 | 105 | 71 | 78 |
| Percent correct | 44 | 50 | 33 | 37 |
| Percent correct by simply guessing the most frequently occurring (grape/country/region/vintage) | 16 | 39 | 11 | 27 |
| $\chi^2$ | 40.3** | 5.06* | 31.5** | 4.33* |

Chi squared tests was conducted to compare the accuracy of the most common guess against the baseline condition of occurrence frequency.
** indicates significance at 0.01 level.

coefficients were calculated between the percentages of estimation error (both signed and unsigned, to consider the magnitude of the error) and time. In terms of acidity, we observed a significant decrease in the magnitude of error over time ($r_{18} = –0.53$, $p = 0.02$), due to an increase in the estimation error over time ($r_{18} = 0.77$, $p < 0.0005$). In other words, participants gradually went from underestimating to overestimating acidity levels, but with the net effect of increasing accuracy over time. In terms of alcohol level assessment, we did not observe any significant changes in error magnitude ($r_{18} = –0.10$, $p = 0.68$) or estimation error ($r_{18} = .03$, $p = 0.92$). This is possibly because the estimation categories (<12%, 12–13.5%, >13.5%) are broad and easy to learn, which means that the estimation error is already fairly small. In fact, one-sample t-tests showed that average acidity and alcohol errors were not significantly different from 0 (for acidity: $t(17) = –1.45$, $p = 0.17$; for alcohol: $t(17) = 0.94$, $p = 0.36$).

## *D. Tasting Note (Hypothesis 4)*

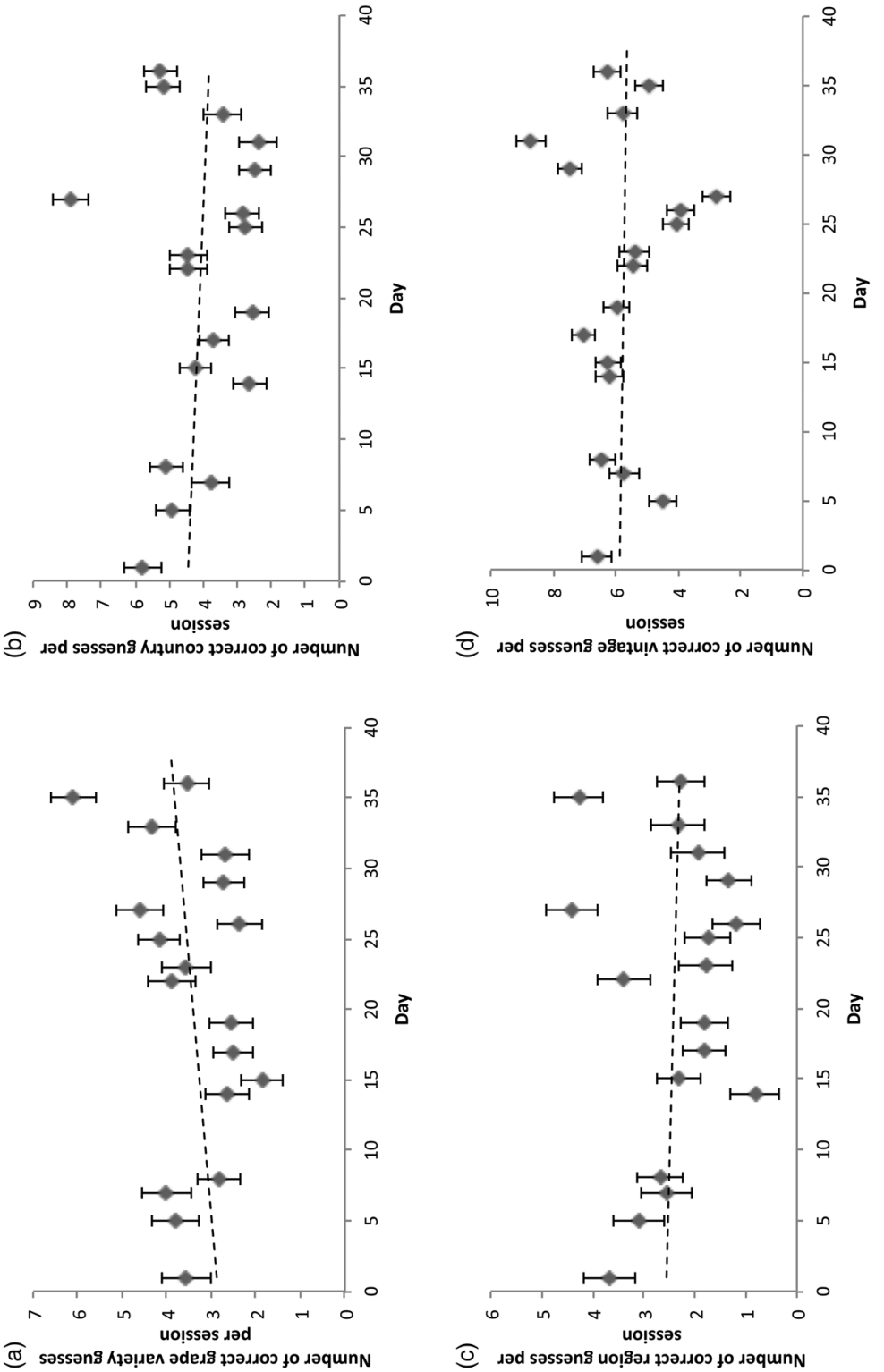
For each participant, we calculated the number of words per tasting note. An ANOVA with prior experience (novices or tasters with experience) and wine preference (1–5) as independent factors and number of words as the dependent measure revealed significant effects of prior experience ($F(1,2109) = 284.14$, $p < 0.0005$, $\eta^2 = 0.12$) and wine preference ($F(4,2109) = 15.15$, $p < 0.0005$, $\eta^2 = 0.03$) on the number of words. In terms of prior experience, novices write shorter tasting notes compared to experienced participants ($M_{novice} = 22.37$, *standard error* = 0.23, $M_{experienced} = 27.59$, *standard error* = 0.21, $p < 0.0005$). In terms of wine preference, tasting notes for the more preferred (those at levels 4 and 5) wines were significant longer (by 23%) than tasting notes for the less preferred wines (those at levels 1–3) wines ($p < 0.005$ for all comparisons).

## *E. Preference (Hypothesis 5)*

To understand which factors drove the participants' preference for wines, Pearson's correlation coefficients were calculated between wine attributes and preference.

*Figure 2*

**Guess Accuracy of Grape Variety (a), Country (b), Region (c), and Vintage (d)**




*Note:* Error bars indicate the standard errors of the mean.

Fig. 2 - B/W online, B/W in print

*Table 2*
**Fixed Effects Regression on Wine Preference**

| *All Participants* | | | | | | |
|---|---|---|---|---|---|---|
| | *Overall (Sessions 1–18)* | | *First Half (Sessions 1–9)* | | *Second Half (Sessions 10–18)* | |
| | *β* | *t-Stat.* | *β* | *t-Stat.* | *β* | *t-Stat.* |
| Vintage | −0.28 | −4.53** | — | — | −0.39 | −4.65** |
| Acidity | 0.30 | 5.13** | 0.30 | 3.58** | 0.35 | 3.99** |
| Sweetness | 0.29 | 4.63** | 0.36 | 4.13** | 0.25 | 2.82** |
| Red (Red = 1, White = 0) | 0.17 | 2.71** | 0.29 | 3.2 ** | — | — |
| Adjusted $R^2$ | 0.27 | | 0.26 | | 0.28 | |
| Observations | 207 | | 103 | | 101 | |
| *Novices Only* | | | | | | |
| | *Overall (Sessions 1–18)* | | *First Half (Sessions 1–9)* | | *Second Half (Sessions 10–18)* | |
| | *β* | *t-Stat.* | *β* | *t-Stat.* | *β* | *t-Stat.* |
| Sweetness | 0.35 | 5.29** | 0.44 | 4.89** | 0.45 | 4.40** |
| Red (Red = 1, White = 0) | 0.25 | 3.69** | 0.26 | 2.75** | 0.15 | 1.47 |
| Vintage | −0.16 | −2.39* | — | — | −0.22 | −2.36* |
| Oak | — | — | 0.18 | 1.99* | — | — |
| Acidity | — | — | — | — | 0.30 | 3.09** |
| Alcohol | — | — | — | — | 0.22 | 2.14* |
| Adjusted $R^2$ | 0.19 | | 0.22 | | 0.25 | |
| Number of Observations | 208 | | 103 | | 99 | |

Results for all 15 participants (top) and for the group of 9 novices (bottom). * indicates significance at $p \leq 0.05$ level. ** indicates significance at $p \leq 0.01$ level. Variables with no value are excluded from the model.

Interestingly, the price of the wines had a significant positive correlation with preference, even though the wines were tasted blind ($r = 0.36$). This was also true for the group of novices ($r = 0.23$), which implies that even novices can some how pick up qualities in the wine that is both preferred and more expensive. Furthermore, novices in general preferred wines with oak ($r = 0.15$) and from the new world ($r = 0.16$), whereas the group overall preferred acidity ($r = 0.26$) (driven by the more expert participants).

In addition, we performed linear regressions to understand how preference can be predicted based on wine factors that can be perceived during tasting. We excluded variables such as the price and the geographical region (new world vs. old world), since these cannot be perceived sensorially from the wine. To assess the effect of training, we conducted regressions separately in the first half of the sessions and in the second half of the sessions (Table 2).

Notably, from the first half of the sessions (1–9) to the second half (10–18), we noticed an overall shift in the importance of vintage (with the older wines being

more preferred) and a decrease in the preference for red wine over white wine. For the group of novices, it is especially interesting to observe that the preference for red wines and oaked wines gave way to the preference for older wines, more acidic wines, and wines higher in alcohol (Table 2).

## IV. Conclusions

The results of the present study demonstrated that blind tasting training can improve accuracy in terms of guesses as well as structural elements. In the span of 18 sessions, we observed an increase in accuracy and within-group agreement when it comes to the variety of the grape, as well as an improvement in the estimation of the acidity. However, we did not observe any improvements in accuracy or within-group agreement when it came to guesses of country or region.

As with country and region guess accuracy, vocabulary size and tasting note length did not change during the course of the training session, which implies that longer-term learning may be needed. In addition, an interesting relationship emerged between length of note and wine preference, with more preferred wines having longer tasting notes. This may help shed some light on the mental process behind writing tasting notes, where more preferred wines may enhance processing fluency (or that wines that are easier to describe are, in turn, more preferred).

The analysis of the preference for wine revealed that, overall, preference was positively correlated with wine age, acidity, sweetness, and color (red wine was preferred to white). More remarkably, when taking into account the price of wine, we demonstrated a significant positive correlation between the price and the preference of the wine, even when the wines were tasted blind. This gives limited evidence that the more expensive wines, broadly speaking, may have sensorial properties which make them appealing to tasters. Over time, we observed a shift in preference towards older wines and a decrease in the importance of wine color as a predictor of preference. Those with little initial blind tasting training also experienced a change in preference toward wines with greater acidity and alcohol, and decreased their preference for oak flavors. These observations have important implications for growing wine markets with an increasingly educated consumer population, where the initial preference for heavily oaked, young red wines may shift in time toward wines with more maturity and structure.

## References

Almenberg, J., and Dreber, A. (2011). When does the price affect the taste? Results from a wine experiment. *Journal of Wine Economics*, 6(1), 111–121.

Arvisenet, G., Guichard, E., and Ballester, J. (2016). Taste-aroma interaction in model wines: Effect of training and expertise. *Food Quality and Preference*, 52(September), 211–221.

Brochet, F., and Dubourdieu, D. (2001). Wine descriptive language supports cognitive specificity of chemical senses. *Brain and Language*, 77(2), 187–196.

Castriota-Scanderbeg, A., Hagberg, G. E., Cerasa, A., Committeri, G., Galati, G., Patria, F., et al. (2005). The appreciation of wine by sommeliers: A functional magnetic resonance study of sensory integration. *NeuroImage*, 25(2), 570–578.

D'Alessandro, S., and Pecotish, A. (2013). Evaluation of wine by expert and novice consumers in the presence of variations in quality, brand, and country of origin cues. *Food Quality and Preference*, 28(1), 287–303.

Gawel, R. (1997). The use of language by trained and untrained experienced wine tasters. *Journal of Sensory Studies*, 12(4), 267–284.

Goldstein, R., Almenberg, J., Dreber, A., Emerson, J. W., Herschkowitsch, A., and Katz, J. (2008). Do more expensive wines taste better? Evidence from a large sample of blind tastings. *Journal of Wine Economics*, 3(1), 1–9.

Hughson, A. L., and Boakes, R. A. (2002). The knowing nose: The role of knowledge in wine expertise. *Food Quality and Preference*, 13(7–8), 463–472.

Ishii, R., and O'Mahony, M. (1991). Use of multiple standards to define sensory characteristics for descriptive analysis: Aspects of concept formation. *Journal of Food Science*, 56(3), 838–842.

Li, Y., and Bardaji, I. (2016). A new wine superpower? An analysis of the Chinese wine industry. American Association of Wine Economists, Working Paper No. 198, June 2016. Available from https://www.wine-economics.org/aawe-working-paper-no-198-business/.

Owen, D. H., and Machamer, P. K. (1979). Bias-free improvement in wine discrimination. *Perception*, 8(2), 199–209.

Parr, W. V., Heatherbell, D., and White, K. G. (2002). Demystifying wine expertise: Olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical Senses*, 27(8), 747–755.

Rosenheck, D. (2017). Think wine connoisseurship is nonsense? Blind-tasting data suggest otherwise. *The Economist*, August 17th. Accessed from https://www.economist.com/blogs/graphicdetail/2017/05/daily-chart-11.

Royet, J.-P., Plailly, J., Saive, A.-L., Veyrac, A., and Delon-Martin, C. (2013). The impact of expertise in olfaction. *Frontiers in Psychology*, 4, Article 928. Available at https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00928/full.

Segal, J., ed. (2013). *Red, Whites and Varsity Blues: 60 Years of the Oxford & Cambridge Blind Wine-Tasting Competition*. London, UK: Pavilion Books.

Weil, R. L. (2007). Debunking critics' wine words: Can amateurs distinguish the smell of asphalt from the taste of cherries? *Journal of Wine Economics*, 2(2), 136–144.