Do More Expensive Wines Taste Better? Evidence from a Large Sample of Blind Tastings*

Robin Goldstein†, Johan Almenberg‡, Anna Dreber§,

John W. Emerson¶, Alexis Herschkowitsch†, and Jacob Katz†

SSE/EFI Working Paper Series in Economics and Finance No. 700

April 24, 2008

Abstract

Individuals who are unaware of the price do not derive more enjoyment from more expensive wine. In a sample of more than 6,000 blind tastings, we find that the correlation between price and overall rating is small and negative, suggesting that individuals on average enjoy more expensive wines slightly *less*. For individuals with wine training, however, we find indications of a positive relationship between price and enjoyment. Our results are robust to the inclusion of individual fixed effects, and are not driven by outliers: when omitting the top and bottom deciles of the price distribution, our qualitative results are strengthened, and the statistical significance is improved further. Our results indicate that both the prices of wines and wine recommendations by experts may be poor guides for non-expert wine consumers.

Keywords: Wine, price/quality relation, expertise.
JEL codes: L15, L66, M30, Q13.

*We thank Jacopo Anselmi, Zoe Chance, Shane Frederick, Richard Friberg, Barry Goldstein, Erik Grönqvist, Daniel Horwitz, Roy Ip, Magnus Johannesson, Thomas Pfeiffer, and Sue Stubbs for helpful comments and suggestions. Johan Almenberg thanks the Ragnar and Torsten Söderberg Foundations for financial support, and Johan Almenberg and Anna Dreber thank the Jan Wallander and Tom Hedelius Foundation for financial support. The Program for Evolutionary Dynamics is sponsored by J. Epstein.

†Fearless Critic Media, 2011B Bouldin Avenue, Austin, TX 78704. Corresponding author (Goldstein). Email: robin@fearlesscritic.com.

‡Department of Economics, Stockholm School of Economics, and the Research Institute of Industrial Economics, Stockholm. Corresponding author. Email: johan.almenberg@hhs.se.

§Department of Economics, Stockholm School of Economics, and Program for Evolutionary Dynamics, Harvard University.

¶Department of Statistics, Yale University.

1 Introduction

When symbolic content is an important part of consumption, the enjoyment of a good might become decoupled from its innate qualities. The symbolic content of a price tag has been emphasized in marketing research (e.g., Cialdini, 1998). At the same time, when goods with similar characteristics differ in price, a reasonable prior is that the more expensive good will, on average, be of a higher quality. People have been shown to expect a positive correlation between price and quality (e.g., Rao and Monroe, 1989). Consistent with this expectation, a meta-analysis reports positive correlations between price and quality ratings for most, but not all, of 1,200 product markets, but also finds that the range of these correlations is very large (Tellis and Wernerfelt, 1987).

For some goods, there is much heterogeneity in consumer tastes, making it harder to infer quality from revealed preferences. Nonetheless, a reasonable prior is that consumers on average will derive more enjoyment from the good with the higher price. Previous research suggests that wine might be a good where consumer tastes are highly heterogeneous (Amerine and Roessler, 1976; Lecocq and Visser, 2006). While individuals may frequently disagree over which wine they prefer, the above hypothesis suggests a positive correlation between the enjoyment of a wine and its price.

Many factors, such as peer consumption and marketing actions, can influence how a good is experienced. Price may in itself be such a factor. Recent research has shown that individuals appreciate the same wine more when they think that it is more expensive (Brochet, 2001; Plassmann et al., 2008). In other words, the price of a good affects the experienced utility derived from that good. Thus, to test the conjecture mentioned above, we need to examine the enjoyment of wine when individuals are unaware of the price.

In this paper, we use a large sample of more than 6,000 US blind tastings, compiled by food and wine critic Robin Goldstein. Blind tastings offer the opportunity to isolate the experience of the wine itself from psychological confounds related to its price, presentation or published expert ratings.

We investigate the relationship between price and subjective appreciation of wines, when the price is unknown to the tasters. Subjective appreciation is measured by overall ratings assigned to wines by individual participants.

Our main finding is that, on average, individuals who are unaware of the price do not derive more enjoyment from more expensive wine. In fact, they enjoy more expensive wines slightly less.

We use an ordered probit estimator, as well as a linear estimator (OLS). In both cases, we use robust standard errors. The ordered probit estimator is particularly well suited to an ordinal dependent variable, but we find that OLS also performs well, and yields estimates that are easier to interpret. In any case, the two models generate highly consistent results. The dependent variable is the overall rating assigned by an individual to a wine. The key independent variable is the price of the wine, expressed as the natural logarithm of the average retail price per 750 mL in US dollars.[1]

In our baseline model, we regress the overall rating on the price of the wine, using both estimators. If individuals found that more expensive wine tasted better, the coefficient on price would be positive. Our baseline model allows us to reject this hypothesis: the coefficient on price is negative, and statistically significant, regardless of which of the two estimators we use.

Next, we extend the model by taking into account that about 12% of the participants in the blind tastings had some form of wine training, such as a sommelier course. A number of studies have reported positive correlations between price and subjective appreciation of a wine for wine experts (e.g., Oczkowski, 1994; Landon and Smith, 1997; Benjamin and Podolny, 1999; Schamel and Anderson, 2003; Lecocq and Visser, 2006).

In the extended model, we allow for the possibility that individuals with wine training (hereafter: "experts") experience wines differently from non-experts. We include a dummy variable for being an expert, as well as an interaction term for price and the expert dummy. In a linear regression, this allows both the intercept and the slope coefficient to differ between the two groups.

Previous research suggests that non-experts may not be particularly sensitive to some of the refinements that are held in high esteem by wine aficionados. Weil (2001, 2005) uses the following experimental setup: two bottles of wine are poured into four containers. Tasters are then given three of the containers and asked to distinguish which one differs from the other two. A random guess has 1/3 chance of being correct. In Weil (2001), the two wines are identical apart from year, but one wine is from a "good" vintage, and the other from a "bad" vintage.[2] The tasters get it right 41% of the time – only marginally better than randomized guessing.[3] In Weil (2005) the wines are a reserve bottling and a regular bottling, from the same producer and year. The fraction of correct answers is merely 40%.

Moreover, Weil finds that even when tasters can distinguish between the vintages, they are about as likely to prefer the good one as the bad one. And among those that can distinguish the reserve bottling from the regular bottling, only half prefer the reserve. In both cases, the wines differ in price by an order of magnitude.

Our data also indicates that experts, unlike non-experts, on average assign as high – or higher – ratings to more expensive wines. The coefficient on the expert*price interaction term is positive and highly statistically significant. The price coefficient for non-experts is negative, and about the same size as in the baseline model. The net coefficient on price for experts is the sum of these two coefficients. It is positive and marginally statistically significant.

[1] If we didn't do this, we would be expecting a one dollar increase to have the same effect at the $5 price level as at the $50 price level. We get similar qualitative results using the dollar prices, but the statistical significance of the coefficients is not as good (but still significant).

[2] Weil uses pairs for which the famous wine critic Robert Parker has rated one of the bottles "average" to "appalling" and the other bottle "excellent" to "the finest".

[3] All the significant difference is driven by the testers' ability to distinguish between the good and bad vintages from Bordeaux Pomerol.

The linear estimator offers an interpretation of these effects. In terms of a 100 point scale (such as that used by *Wine Spectator*), the extended model predicts that for a wine that costs ten times more than another wine, non-experts will on average assign an overall rating that is about four points lower, whereas experts will assign an overall rating that is about seven points higher.

The magnitude of these effects is moderate, but non-negligible given that wine prices cover a large range. In this sample alone, prices range from $1.65 to $150. In wine markets in general, the range is even larger.

We test the robustness of our results by adding individual fixed effects to our model. This does not affect the qualitative results, and the coefficients themselves change only slightly, regardless of whether we use ordered probit or OLS.

To make sure that our results are not driven by wines at the extreme ends of the price range, we also estimate the extended model using a reduced sample, omitting observations in the top and bottom deciles of the price distribution. We use both ordered probit and OLS, with and without individual fixed effects. We find the same qualitative results with the reduced sample. In fact, the effects are larger and the statistical significance improves even further. In other words, our findings are even more pronounced when looking only at mid-range price levels.

Our paper is organized as follows. In section 2, we describe our data. In section 3, we present our econometric model and report the regression estimates. We also perform a robustness check. We conclude in section 4, where we discuss some implications of our results and suggest directions for future research.

2 Data

The data set contains 6,175 observations from 17 blind tastings organized by Robin Goldstein. The blind tastings took place in the US between April 2007 and February 2008. In total, 506 participants tasted wine flights composed from 523 different wines. The wines were presented in a double-blind manner, so that neither the person serving the wine nor the person tasting the wine knew the identity and/or price of the wine. Each taster assigned an overall rating to every wine tasted, prior to discussing the wines with the rest of the group. The rating was the response to the question "Overall, how do you find the wine?", and the available answers were "Bad", "Okay", "Good", and "Great". In the data, these alternatives are coded from 1 to 4, with 1 corresponding to "Bad" and 4 corresponding to "Great".[4] The price per bottle ranged from $1.65 to $150. The prices are average retail prices and were obtained from www.wine-searcher.com. The wines represent a broad variety of types (e.g. red, white, rosé, sparkling), country origins, and grapes.

[4] Tasters ticked one of four boxes. In about 3% of the sample, tasters ticked in between two boxes, suggesting a rating somewhere in between the two responses. For simplicity, we dropped these observations from the regression. Including them makes no difference to our qualititative results, and the changes to the estimates are negligible.

The participants were unpaid volunteers, ranging from 21 to 88 years of age. Selection bias is a concern with any voluntary subject pool, and we have no reason to think that this is an exception. It is quite likely that the sample contains an over-representation of highly educated individuals, and an overrepresentation of individuals working in the food and wine industries. Nonetheless, the size of the sample and the general diversity of the tasters lead us to hope that inference will not be too restricted. For a more extensive description of the blind tastings, please see Chapters 8 and 9 and Appendix 1 in Goldstein (2008).

3 Regression Analysis

Throughout the regression analysis, we use both an ordered probit estimator and a linear estimator (OLS). In both cases, we consistently use robust standard errors. The dependent variable is the overall rating, measured on a scale from 1 to 4, with 4 being the highest rating. The main independent variable is the price variable, expressed as the natural logarithm of the average retail price per 750 mL in US dollars.

In Model 1, we regress the overall rating assigned to wine i, by individual j, on the price of the wine. About 12% of participants had some wine training, such as a sommelier course. In Model 2, we allow for the possibility that these "experts" rate wines in a different manner. We include a dummy variable for being an expert, as well as an interaction term for price and the expert dummy. In a linear regression, this allows both the intercept and the slope coefficient to differ for experts and non-experts. In terms of the linear specification, we can write these two models as

$$y_i = \beta_0 + \beta_1 \ln(P_i) + \epsilon_i \tag{1}$$

and

$$y_{ij} = \beta_0 + \beta_1 \ln(P_i) + \beta_2 EXPERT_j + \beta_3 \ln(P_i) \times EXPERT_j + \epsilon_i \tag{2}$$

where P_i is the price of wine i, and $EXPERT_j$ is a dummy variable indicating if taster j has wine training. If individuals found that more expensive wine tasted better, the correlation between overall rating and price would be positive. In our sample, this is not the case: for both the ordered probit estimates and the OLS estimates, the coefficient on price is negative. In Model 1, the OLS coefficient is about -0.04, implying that a 100% increase in ln(price) is associated with a 0.04 reduction in the overall rating. The negative effect for more expensive wines is statistically significant.

Unlike the non-experts, experts assign as high, or even higher, ratings to more expensive wines. Model 2 shows that the correlation between price and overall rating is positive – or, at any rate, non-negative - for experts. The price coefficient for non-experts is still negative, of about the same size as before, and with greatly improved statistical significance. The coefficient on the ln(price)*expert interaction term is highly statistically significant (ordered probit p-value: 0.017; OLS p-value: 0.015). For experts, the net coefficient on price is the sum of the two, i.e, about 0.11 for the ordered probit and 0.09 for OLS. This net coefficient is marginally significantly different from zero (ordered probit p-value: 0.099; OLS p-value: 0.095). A full set of estimates is shown in Table 1.

Table 1. Dependent variable: overall rating.

	Ordered probit		OLS	
	(1)	(2)	(3)	(4)
ln(Price)	-0.047	-0.061	-0.038	-0.048
	(0.039)**	(0.013)**	(0.038)**	(0.012)**
ln(Price)*Expert		0.171		0.138
		(0.017)**		(0.015)**
Expert		-0.558		-0.448
		(0.001)***		(0.001)***
Constant			2.297	2.337
			(0.000)***	(0.000)***
Observations	5986	5972	5986	5972
R^2/pseudo-R^2	0.000	0.002	0.001	0.005

Robust p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

In sum, we find a non-negative relationship between price and overall rating for experts. Due to the marginal significance of the price coefficient for experts, it remains an open question whether this coefficient is positive, but our results indicate that this is in fact the case.

How large are these price effects? The coefficients are of a moderate magnitude, but non-negligible, given that wine prices cover a large range. In this sample alone, prices range from \$1.65 to \$150. In wine markets in general, the range is even larger. Suppose we have two wines, A and B, and Wine A costs ten times more then Wine B in dollar terms. In terms of a 100 point scale (such as that used by *Wine Spectator*), the OLS estimation of Model 2 predicts that non-experts will assign an overall rating that is four points lower for wine A, whereas experts will assign an overall rating that is seven points higher.[5]

In addition, the expert dummy is negative, quite sizable, and statistically significant (OLS expert dummy coefficient: -0.448; p-value: 0.001). In other words, the OLS estimation of Model 2 consists of two linear relationships, one with a higher intercept but a negative slope and one with a lower intercept but a positive slope. The point where the two lines cross each other is the price level at which experts and non-experts are expected to assign the same rating. If we take the model literally, this point occurs at the price of \$25.70, i.e., ln(price) = approx. 3.25. At this price, the model predicts that both groups will assign a rating of about 2.2. Below this price, the model predicts that experts will assign lower ratings to a wine than non-experts, and vice versa.

[5] If the dollar price increases by a factor of 10, ln(price) increases by about 2.3. Hence the predicted effect on the overall rating of tenfold increase in the dollar price is 2.3 times the ln(price) coefficient for non-experts and experts, respectively, adjusted to a 100 point scale.

We also test a third model, including individual fixed effects. In terms of the linear specification, Model 3 can be written as

$$y_{ij} = \beta_0 + \delta_j + \beta_1 \ln(P_i) + \beta_2 \ln(P_i) \times EXPERT_j + \epsilon_i \qquad (3)$$

where δ_j is a dummy for each individual taster. Including individual fixed effects has very little effect on the qualitative results and the minor differences only serve to reinforce our earlier conclusions, as both the negative effect for non-experts and the positive effect for experts become slightly stronger. These results are presented in Table 2. For each of the four regressions in Table 2, a Wald test rejects that the fixed effects are jointly equal to zero, by a wide margin (p-value < 0.001), suggesting that this is a suitable addition to the model.

Table 2. Individual fixed effects.

	Ordered probit		OLS	
	(1)	(2)	(3)	(4)
ln(Price)	-0.070	-0.089	-0.050	-0.064
	(0.007)***	(0.001)***	(0.009)***	(0.002)***
ln(Price)*Expert		0.209		0.151
		(0.011)**		(0.013)**
Individual fixed effects	Yes	Yes	Yes	Yes
Constant			2.487	2.183
			(0.000)***	(0.000)***
Observations	5986	5972	5986	5972
R^2/pseudo-R^2	0.080	0.081	0.181	0.182

Robust p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

To make sure that our results are not driven by wines at the extreme ends of the price distribution, we also run our regressions on a reduced sample, omitting the top and bottom deciles of the price distribution. Given the broad range of prices in the sample, this is an appropriate precaution. The remaining wines range in price from \$6 to \$15.

Using the reduced sample, we estimate Model 2 with and without individual fixed effects. Doing so produces consistent and even more pronounced estimates. The coefficient on price is still negative, and in each case larger than when using the full sample. The statistical significance of the coefficients improves further, and the R-squared is higher. These estimates are presented in Table 3.

Table 3. Reduced sample[1], with and without individual fixed effects.

	Ordered probit		OLS	
	(1)	(2)	(3)	(4)
ln(Price)	-0.225	-0.173	-0.182	-0.122
	(0.001)***	(0.019)**	(0.001)***	(0.025)**
ln(Price)*Expert	0.523	0.515	0.421	0.364
	(0.002)***	(0.006)***	(0.002)***	(0.009)***
Expert	-1.301		-1.044	
	(0.000)***		(0.000)***	
Individual fixed effects	No	Yes	No	Yes
Constant			2.622	1.910
			(0.000)***	(0.000)***
Observations	4817	4817	4817	4817
R^2/pseudo-R^2	0.003	0.094	0.007	0.206
Robust p-values in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%				

1) The reduced sample excludes the top and bottom deciles of the price distribution.

In sum, we use the reduced sample to check the robustness of our results with regard to mid-range price levels. Based on the above, we conclude that our results are not only robust but in fact even more pronounced when omitting observations at the extremes of the price distribution.

4 Conclusion

The pleasure we get from consuming wine depends both on its intrinsic qualities such as taste and smell and external attributes such as price and presentation. One may argue that the former influences our subjective appreciation through a bottom-up process, where the sensory apparatus plays a key role, and that the latter works through a top-down process, where beliefs and expectations about quality are important determinants.[6] In this paper we have explored the bottom-up effects, by looking at how participants in blind tastings rate wines. We find that, unless they are experts, individuals who are unaware of the price enjoy more expensive wines slightly less.

There is a large relevant literature related to marketing. Lee et al. (2006) look at how knowledge of a beer's ingredients (normal beer with added vinegar) can affect subjective appreciation. They show that the timing of the information plays a substantial role. One group of tasters is told about the vinegar, tastes the beer, and assigns ratings. A second group is told about the vinegar after tasting

[6] This, in turn, might depend on ulterior motives such as status concerns. Wine as a status signal, and the prospect that expensive wine could function as a positional good, is discussed in Goldstein (2008), chapter 5.

the beer, but before the ratings are assigned. On average, individuals in the first group assign significantly lower ratings, suggesting that informing participants about the vinegar influences the experience in itself. Using fMRI, McClure et al. (2004) find that having the subject's favourite brand's name on a drink makes it taste better than if it is unlabeled. In another fMRI study, Plassmann et al. (2008) test whether marketing actions such as changes in the price can influence the experienced pleasantness of a product such as wine. Testers are given different wines that they are told differ in price. In reality, some of the wines are the same but simply presented with different prices. Prices are found to correlate positively with experienced pleasantness, measured through both subjective reports and fMRI scans.

Marketing provides one channel through which consumers can be influenced to buy certain wines. But it is not the only one: wine critics/experts may also play a role in affecting wine prices and shaping consumer preferences. For example, Hadj Ali et al. (2007) find a positive effect of wine critic Robert Parker's ratings on the price of Bordeaux wine.

There is, however, some research expressing skepticism towards wine ratings and their use for the average wine drinker. According to Quandt (2007), many wine ratings do not actually convey any information, nor is there substantial agreement in ratings by experts. Consistent with this view, Weil (2007) investigates whether wine descriptions of experts actually convey any information to wine consumers. This is tested by having testers match wine descriptions to wines. In a similar setup to Weil (2001, 2005), tasters are asked to distinguish the odd one out of three different glasses of wine. Only about 50% of the participants can distinguish the odd one out, and of those that manage to do it, only about half can correctly match a wine critic's description of the wine with the wine itself − which is no better than a random guess.

Our results indicate another reason for why the average wine drinker may not benefit from expert wine ratings: he or she simply doesn't like the same types of wines as experts. This is consistent with Weil (2001, 2005), who finds that even among the subset of tasters who can distinguish between good and bad vintages, or reserve or regular bottlings, they are as likely to prefer the "better" one as the "worse" one.

These findings raise an interesting question: is the difference between the ratings of experts and non-experts due to an acquired taste? Or is it due to an innate ability, which is correlated with self-selection into wine training?[7] Investigating this further would be a fruitful avenue for future research.

In sum, in a large sample of blind tastings, we find that the correlation between price and overall rating is small and negative. Unless they are experts, individuals on average enjoy more expensive wines slightly less. Our results suggest that both price tags and expert recommendations may be poor guides for non-expert wine consumers who care about the intrinsic qualities of the wine.

[7] For a further discussion, see Chapter 4 of Goldstein (2008).

REFERENCES

Amerine, M A and E B Roessler (1976) *Wines: Their Sensory Evaluation*. San Francisco: W.H. Freeman and Company.

Benjamin, B A and J M Podolny (1999) Status, Quality, and Social Order in the California Wine Industry. *Administrative Science Quarterly* 44:563-589.

Brochet, F (2001) Chemical Object Representation In The Field of Consciousness. Working paper, General Oenology Laboratory, France.

Cialdini, R B (1998) *Influence: The Psychology of Persuasion*. New York: Collins.

Goldstein R (2008) *The Wine Trials*. Austin: Fearless Critic Media.

Hadj Ali, H, S Lecocq and M Visser (2007) The Impact of Gurus: Parker Grades and en Primeur Wine Prices. AAEA Working Paper No.1.

Landon, S and C E Smith (1997) The Use of Quality and Reputation Indicators by Consumers: The Case of Bordeaux Wine. *Journal of Consumer Policy* 20: 289-323.

Lecocq, S and M Visser (2006) What Determines Wine Prices: Objective vs. Sensory Characteristics. *Journal of Wine Economics* 1(1): 42-56.

Lee, L, S Frederick and D Ariely (2006) Try It, You'll Like It- The Influence of Expectation, Consumption, and Revelation on Preferences for Beer. *Psychological Science* 17(12): 1054-1058.

McClure, S M, J Li, D Tomlin, K S Cypert, L M Montague, P R Montague (2005) Neural Correlates of Behavioral Preference for Culturally Familiar Drinks. *Neuron* 44: 379-387

Oczkowski, E (1994) A Hedonic Price Function for Australian Premium Wine. *Australian Journal of Agricultural Economics* 38: 93-110.

Plassmann, H, J O'Doherty, B Shiv and A Rangel (2008) Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness. *Proceedings of the National Academy of Sciences* 105(3): 1050-1054.

Quandt, R E (2007) On Wine Bullshit: Some New Software. *Journal of Wine Economics* 2(2): 129-135.

Rao, A R and K B Monroe (1989) The Effect of Price, Brand Name, and Store Name on Buyers' Percetions of Product Quality: An Integrative Review. *Journal of Marketing Research* 36: 351-37.

Schamel, G and K Anderson (2003) Wine Quality and Varietal, Regional and Winery Reputations: Hedonic Prices for Australia and New Zealand. *The Economic Record* 79(246): 357-369.

Tellis, G J and B Wernerfelt (1987) Competitive Price and Quality Under Asymmetric Information. *Marketing Science* 6(3): 240-253.

Weil, R L (2001) Parker v. Prial: The Death of the Vintage Chart. *Chance* 14(4): 27-31.

Weil, R L (2005) Analysis of Reserve and Regular Bottlings: Why Pay for a Difference Only the Critics Claim to Notice? *Chance* 18(3): 9-15.

Weil, R L (2007) Debunking Critics' Wine Words: Can Economists Distinguish the Smell of Asphalt from the Taste of Cherries? *Journal of Wine Economics* 2(2): 136-144.