



AMERICAN ASSOCIATION OF WINE ECONOMISTS

AAWE WORKING PAPER

No. 91

Economics

THE BUYER'S DILEMMA --
WHOSE RATING SHOULD A WINE
DRINKER PAY ATTENTION TO?

Omer Gokcekus and Dennis Nottebaum

September
2011

www.wine-economics.org

The buyer's dilemma - To whose rating should a wine drinker pay attention?

Omer Gokcekus (Seton Hall University) & Dennis Nottebaum* (University of Münster)*

Abstract:

A wine buyer is faced with a multitude of often diverging expert reviews and ratings of one and the same wine. In this article we attempt to shed light on the question to whose rating a regular buyer should pay attention. We do so by comparing the taste of regular consumers, captured by community tasting notes to the expert ratings of Robert Parker, the Wine Spectator and Stephen Tanzer. We find that for a randomly selected sample of 120 2005 Bordeaux wines Stephen Tanzer's scores are most closely associated with the community ratings; and more interestingly, compared to expert ratings, average price paid for a bottle of wine is more highly correlated with median community score. The latter finding possibly confirms the "acquired taste" of experts, but may also be explained by cognitive dissonance.

Key words:

Expert ratings, Bordeaux, consumer choice, Robert Parker, Wine Spectator, Stephen Tanzer

* Omer Gokcekus: John C. Whitehead School of Diplomacy and International Relations, Seton Hall University, USA. omer.gokcekus@shu.edu

* Dennis Nottebaum: Centre for Interdisciplinary Economics (CIW), University of Münster, Germany. nottebaum@wwu.de

Acknowledgements: We would like to thank Joe Roberts, Karl Storchmann, Edward Tower, Alder Yarrow and participants in the Session #1B: Wine Judging at the 2011 American Wine Association Meeting in Bolzano, Italy for their comments and suggestions.

The buyer's dilemma - To whose rating should a wine drinker pay attention?

1. Introduction

Wine is an experiential good; until one buys and opens a bottle the content and quality remain unknown.¹ Since wine drinkers tend to be risk-averse people, they often count on expert ratings when making their choice.² A handful of experts have established a reputation for their reviews and ratings that are published in various widely circulated magazines; some of the most well-known being Robert Parker of the *Wine Advocate*, various experts at the *Wine Spectator*, and Stephen Tanzer of the *International Wine Cellar*.³ Yet, despite having access to these ratings wine drinkers face the second choice of picking one expert whose ratings to follow. This is because expert opinions quite often differ profoundly, both in terms of their ratings as well as their descriptions of the same wine.⁴

Let's say a wine drinker has decided to buy a certain type of wine from a particular region. Although this narrows down his range of options considerably he is still left with further decision-making to do: Not only are there hundreds of wines to choose from within a wide price range, for each one of these wines there are several – often diverging – ratings and reviews as well.

In an ideal world, a consumer would try different wines and choose the one he prefers. But since this first-best option may not be available, wine drinkers often consider expert ratings, usually conveniently placed next to each wine in the store. But given the profound variations in expert ratings and descriptions, to which rating should a regular risk-averse wine drinker pay attention? This is a question that we can't address directly, but we can infer an evaluation from an analysis of differences between expert and consumer ratings, assuming that consumer ratings depict the preferences of an average buyer. Accordingly, this paper poses the following research question: Which expert rating is more closely correlated with regular drinkers' evaluations and is thus best suited as a guide for a regular buyer?⁵

The remainder of this paper will be structured as follows: The ensuing part looks at several examples of differences in expert ratings and descriptions to illustrate our point that expert

opinions profoundly differ from one another. On this basis, the subsequent part analyzes these differences and compares them to community tasting notes, our proxy for the taste of regular drinkers. The final part looks at possible explanations for our findings, drawing on insights from a rich body of literature in social psychology.

2. Illustrating the buyer's dilemma

In order to exemplify our observation that expert ratings and descriptions differ from one another when looking at the same vintage, we examine three 2005 Bordeaux wines in more detail. The first one is a 2005 Château La Confession, a \$60 St. Émilion Grand Cru. It is rated by both Robert Parker and the Wine Spectator, with Parker's score being 95 and the Wine Spectator's 89. The second wine is a 2005 Château Rauzan-Ségla, a \$110 Margaux which is rated 97 by the Wine Spectator, 94 by Robert Parker, and 92 by Stephen Tanzer. The third wine, a 2005 Château La Vieille Cure is a \$30 Fronsac. Robert Parker rates it with 93 points, while the Wine Spectator awards 90 and Stephen Tanzer 87-89. These ratings not only differ quite strongly from one another, the reviews and evaluations are also rather varied.

For instance, Robert Parker describes the 2005 Château La Vieille Cure as follows: "La Vieille Cure's amazing 2005 is even better than their terrific 2003 and 2000. The 2005 boasts an inky/ruby color as well as a gorgeously sweet perfume of charcoal, black cherries, black currants, and spring flowers as well as an underlying mineral component. Superb concentration, full-bodied power, wonderful symmetry, purity, and texture, and a multidimensional mouthfeel are all found in this fabulous sleeper of the vintage."

Less enthusiastic, Stephen Tanzer - who rated this wine 87-89 - has the following evaluation: "Full ruby. Kirsch and licorice on the nose. Fat, dense and sweet, if a bit youthfully monolithic. At once plump and primary, with lively flavors of black fruits and bitter chocolate. Finishes with nicely buffered tannins and a rich chocolaty nuance."

An average, risk-averse wine drinker now finds himself confronted with the question which of these ratings and descriptions he should follow. We obviously cannot give a definite answer here - we are talking about taste after all, a fundamentally subjective issue. However, assuming that one can

objectify "taste" to some extent - which the use of ratings explicitly does - we can shed some light on this question.

3. Analysis

In order to do this, we seek to answer the question of which expert rating is more closely correlated with regular drinkers' evaluations, by comparing expert and consumer evaluations from the web page www.cellartracker.com. CellarTracker is a wine cellar management software that has by now become the world's largest wine social networking site, in both number of catalogued bottles and number of listed tasting notes. Currently, its database contains more than 1.8 million "community tasting notes" in addition to 300,000 professional reviews. Given the size and reputation of CellarTracker, we believe their data to be a reasonable proxy for consumer opinion, although users engaged in a dedicated wine community are probably more experienced and critical than Joe Average.

For our sample we randomly select 120 wines out of 4,212 (2005 vintage) Bordeaux listed at www.cellartracker.com.⁶ Community tasting notes and ratings are available for all wines in our sample. In addition to community tasting notes (a) average score, (b) median score, (c) the average number of ratings, and (d) average price paid are available as statistical information. As summarized in Table 1, 2,928 bottles of the wines in our sample were purchased by the evaluating community members, which yielded 35 evaluations on average per wine with an average price paid of \$109 and average and median scores of 91.4 and 91.8, respectively.

- Table 1 about here -

In our analysis we first compare the community ratings with the scores of three experts, namely Wine Advocate's Robert Parker (RP), International Wine Cellar's Stephen Tanzer (ST), and the Wine Spectator (WS). In a subsequent step, we examine the relationship between prices and ratings.

3.1. Community scores versus Robert Parker

Out of 120 wines in our sample, Robert Parker evaluates 107. For these 107 wines, Table 2 shows that the average community score is 91.66 and average RP score is 93.24. We run a two sample t-

test (with unequal variances) to check if the 1.58 points difference between community and RP is statistically significant. The t-statistic is 4.58 with a critical t-value of 1.97. We therefore reject the null hypothesis that there is no difference between community and RP average scores.

- Table 2 about here -

Having observed significantly higher RP scores, we also examine whether this uniformly holds at different RP score levels. We sort 107 wines according to their RP scores in an ascending order to calculate the moving average (10) for the difference between each wine's RP and community scores (RP-community). As presented in Figure 1, there is a positive relationship between RP scores and RP-community. In other words, the wines that RP scores as 89 or 90 receive higher scores from the community evaluators, while the higher the RP scores get the lower the community scores are. For instance, the community awards 91 to wines that RP scores at 90, 91.5 to wines with a 93 RP score, 92 to wines with 95 RP score, and 95 to wines with 98 RP score.

- Figure 1 about here -

Consequently, two interesting findings emerge from Figure 1: (a) wines with low RP scores receive higher scores from the community; and (b) the higher the RP score the bigger the difference between RP and community scores.

3.2. Community scores versus Wine Spectator

Out of 120 wines in our sample, the Wine Spectator evaluates 104. For these 104 wines, Table 2 shows that the average community score is 91.73 and average WS score is 93.24. We run a two sample t-test (with unequal variances) to check if the 1.51 point difference between community and RP is statistically significant. The t-statistic is 4.30 and the critical t-value is 1.97. We therefore reject the null hypothesis that there is no difference between the community and WS average scores.

Having observed significantly higher WS scores, we also examine whether this uniformly holds at different WS scores. We follow the same procedure as in Figure 1 but find no systematic relationship.

3.3. Community scores versus Stephen Tanzer

Out of 120 wines in our sample, Stephen Tanzer evaluates only 61. For these 61 wines, Table 2 shows that the average community score is 92.16 and average ST score is 92.08. We run a two sample t-test (with unequal variances) to check if the 0.08 points difference between community and ST is statistically significant or not. The t-statistics is 0.18 and the critical t-value is 1.98; therefore we cannot reject the null hypothesis that there is no difference between the community and ST average scores.

3.4. Further analysis

In line with the t-test results, the calculated correlation coefficients also indicate that community scores are closer aligned with ST scores than RP and WS scores. As Table 3 summarizes the correlation coefficients for community scores versus RP, WS, and ST are 0.77, 0.77, and 0.83, respectively.

- Table 3 about here -

Finally, we check the relationship between average prices and scores. Table 3 shows that average prices are (a) highly correlated with median community and ST scores, 0.76 and 0.73 respectively; and (b) not highly correlated with RP and WS scores, 0.53 and 0.64, respectively.

4. Concluding Remarks

Our analyses indicate that (a) both the average and median community scores are lower than the expert ratings of Robert Parker and the Wine Spectator; (b) the correlation between the expert ratings is higher than the correlation between the community and the expert ratings; (c) among the experts, Stephen Tanzer's scores are most highly correlated with the community ratings; and (d) perhaps more interestingly, compared to expert ratings, average price paid for a bottle of wine is more highly correlated with median community score. What do we make of these findings?

Robert Parker evaluates most of the 120 wines that the community scored in our sample. This could be seen as an indication that wine buyers would pay attention to Robert Parker more than the

other experts simply because of the availability of his ratings, at least in deciding whether or not to buy a particular bottle of wine. Yet, we find the strong positive relationship between RP scores and the difference between RP and the community scores puzzling.

One possible explanation for this puzzle is that perhaps regular drinkers are unable to differentiate nuances in wines as well as Robert Parker. Indeed a relatively small variance among community members' scores compared to a higher variance for RP scores supports this explanation. Yet, is there another story here to explain this emerging picture? One plausible explanation could be that regular wine drinkers (although the fact that these wine drinkers are posting their opinions at a widely visited webpage suggests they probably see themselves as "wine enthusiasts" if not amateur wine experts) resent Robert Parker's influence - or shall we say hegemony over the wine community - and systematically challenge his ratings by either giving higher scores to the wines with low RP ratings and lower scores to the wines with high RP ratings.

This story of "iconoclastic" behavior may also explain the strong correlation between prices and community scores. First of all, the community scores are usually produced after expert ratings have already been available and they are not the result of blind tasting. Moreover, community members already possess the price information when they taste and prepare their tasting notes, and similar to most other commodities, price is perceived as an indicator of quality, so that higher priced wines are expected to be of higher quality. Furthermore a psychological phenomenon coined by Leon Festinger (1957) in the 1950s, cognitive dissonance, could be in effect when people assign a low score to a less expensive wine and a high score to a more expensive wine. Cognitive dissonance is the uncomfortable feeling that people experience when they have two conflicting cognitions (in our case having bought an expensive wine and not liking it that much). To eliminate or reduce the unpleasant feeling they either change their attitude or their behavior. Since they cannot change the fact that they have already purchased and opened an expensive bottle of wine they convince themselves that they actually liked that expensive wine even if it goes against the wine experts' opinions.

References

- Cox, D. (2009), "Predicting consumption, wine involvement and perceived quality of Australian red wine", *Journal of Wine Research*, Vol. 20 No. 3, pp. 209-229.
- Festinger, L. (1957), *A theory of cognitive dissonance*, Row & Peterson, Evanston, IL.
- Gergaud, O. and Livat, F. (2007), "How do consumer use signals do assess quality?", AAWE Working Paper No. 3., New York, April 2007.
- Ginsburgh, V. (2003), "Awards, success and aesthetic quality in the arts", *Journal of Economic Perspectives*, Vol. 17, pp. 99–111.
- Hadj Ali, H., Lecocq, S. and Visser, M. (2010), "The impact of gurus: Parker grades and *en primeur* wine prices", *Journal of Wine Economics*, Vol. 5 No. 1, pp. 22–39.
- Lindley, D. V. (2006), "Analysis of a wine tasting", *Journal of Wine Economics*, Vol. 1 No.1, pp. 33–41.
- Lockshin, L.S., and Rhodus, W.S. (1993), "The effect of price and oak flavor on perceived wine quality", *International Journal of Wine Marketing*, Vol. 5 No.2, pp. 13-25.
- Quandt, R. E. (2006), "Measurement and inference in wine tasting", *Journal of Wine Economics*, Vol. 1 No.1, pp. 7–30.
- Schiefer, J. and Fischer, C. (2008), "The gap between wine expert ratings and consumer preferences: Measures, determinants and marketing implications", *International Journal of Wine Business Research*, Vol. 20 No. 4, pp. 335 – 351.

Table 1: Summary Statistics

	<u>COMMUNITY</u>				<u>EXPERT RATINGS</u>		
	# of Bottles Sold	# of Notes	Average Score	Median Score	Wine Spectator (WS)	Wine Advocate (RP)	International Wine Cellar (ST)
Average	2928	35.3	91.4	91.8	93.2	93.2	92.1
Standard Deviation	1884.1	30.5	2.4	2.4	2.7	2.7	2.5
No of Observation	120	120	120	120	104	107	61
Average Price		\$ 108.86			\$ 119.81	\$ 109.97	\$ 145.60

Table 2: Average Community Scores versus Expert Ratings

<i>Average Community Score versus Robert Parker's Rating</i>			
	Community	RP	RP-Community
Mean	91.66	93.24	1.6
Variance	5.14	7.52	
# of observation	107	107	
H₀: Mean Difference = 0			
Degrees of freedom	205		
t Stat	-4.58		
P(T<=t) two-tail	0.00		
t Critical two-tail	1.97		
<i>Average Community Score versus Wine Spectators' Rating</i>			
	Community	WS	WS-Community
Mean	91.73	93.24	1.5
Variance	5.46	7.49	
# of observation	104	104	
H₀: Mean Difference = 0			
Degrees of freedom	201		
t Stat	-4.30		
P(T<=t) two-tail	0.00		
t Critical two-tail	1.97		
<i>Average Community Score versus Stephen Tanzer's Rating</i>			
	Community	SP	SP-Community
Mean	92.16	92.08	-0.1
Variance	5.73	6.32	
# of observation	61	61	
H₀: Mean Difference = 0			
Degrees of freedom	120		
t Stat	0.18		
P(T<=t) two-tail	0.86		
t Critical two-tail	1.98		

Figure 1: Robert Parker's versus Community's Average Ratings – Moving Average (10)

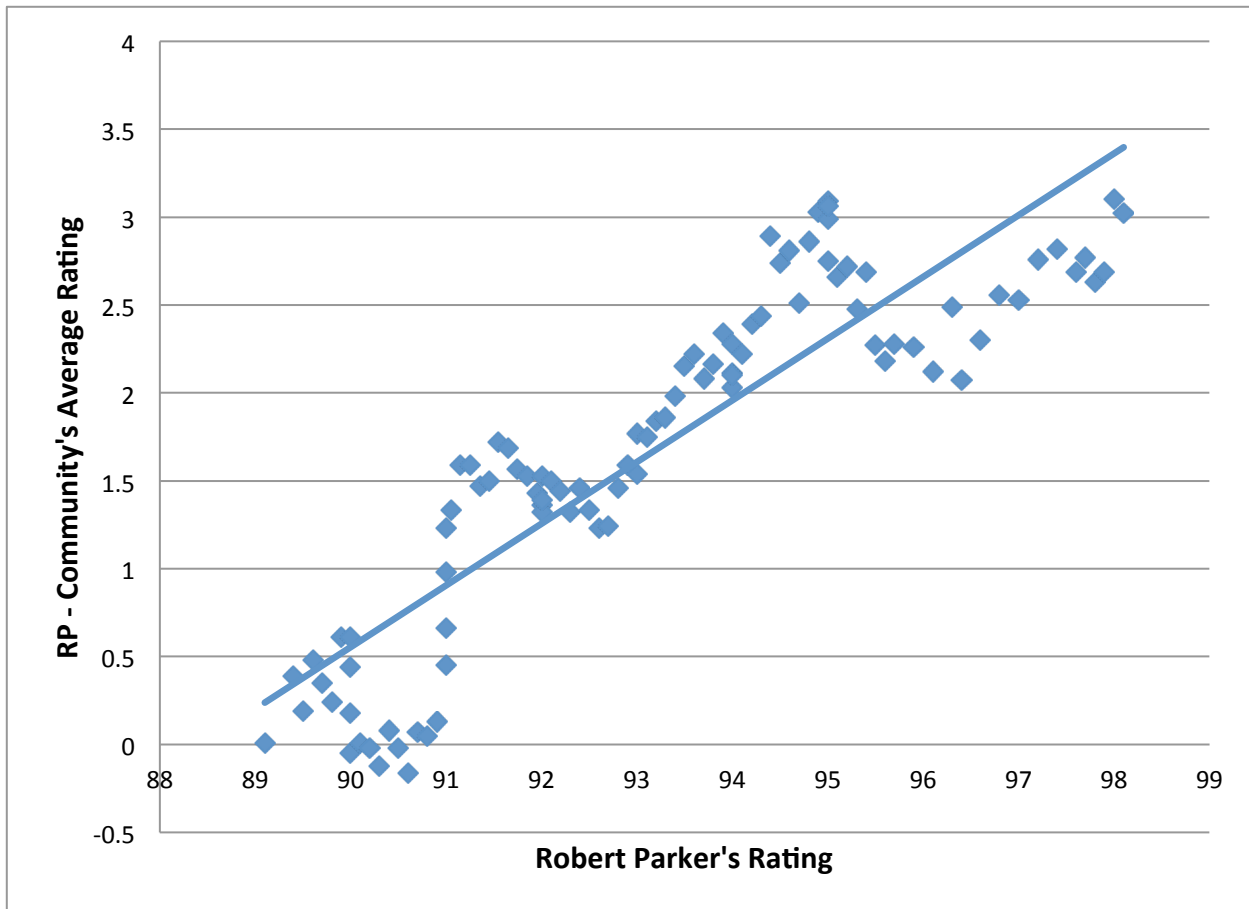


Table 3: Correlation Coefficients

	Average Price	Average Score
Median Score	0.76	
WS Score	0.64	0.77
RP Score	0.53	0.77
ST Score	0.73	0.82

End Notes

¹ Gergaud and Livat (2007) investigate the way consumers use signals such as the price collective brand name or umbrella, goodwill, and their past consumption experiences.

² There is an extensive literature on experts' importance in different markets, e.g., see Ginsburg (2003).

³ Hadj Ali *et. al* (2010) quantify the effects of Robert Parker's opinion on Bordeaux wine prices.

⁴ There are a number of studies on the consistency and reliability of expert opinions, e.g., Lindsay (2006) and Quandt (2006).

⁵ Schiefer and Fischer (2008) compare German consumer ratings obtained in a sensory laboratory with German Agricultural Society's quality competition.

⁶ www.cellartracker.com (accessed January 15, 2011). Out of a small sample of vintages and wine regions for which a sufficient amount of tasting notes was available on Cellar Tracker, 2005 Bordeaux was randomly chosen.