

A Modest Proposal for a Supply-Weighted Wine Scoring System

Jamus Jerome Lim[†]

I. Introduction

The present trend in wine scoring systems – as exemplified by *Wine Spectator* or Robert Parker – tends to focus on the demand side of the equation; that is, scoring is often based entirely on the quality of the wine and the individual preferences of the taster. Such a system has clear advantages: it does not color the score of an individual wine with a taster's biases; it retains a level of objectivity that allows comparison across all available wines; and it allows for reporting that is clear and easily interpreted. However, while one may argue that a focus on the demand side is supremely relevant for trade professionals and merchants, the value of this system is diluted as it moves downstream towards consumers. Indeed, these very ratings might exercise an endogenous effect on the price of a wine; the vast majority of consumers, who make wine purchases from retailers and restaurants, are then faced with the unfortunate prospect of paying an excessive premium for what might have only been an above average bottle of wine.

Hence, in order for a ratings system to remain meaningful and relevant, it is virtually a necessity that such a system considers supply side effects as well. Therefore, it is no longer the quality of the wine alone that should factor into the wine's score – influences such as absolute and relative prices, overall vintage quality, and various markups (including, but possibly not limited to, alcohol taxes and restaurateur and retailer margins) now come into play.

This short note describes a simple proposal for such a system. Although technically more elaborate, this extended system can mark the beginning of a revolution that places the *consumer* back into the wine ratings picture, and lead to wines that are not simply well-rated due to inherent characteristics alone, but rated according to *value for money*.

[†] Advanced Certificate in Wines and Spirits, Wine and Spirit Education Trust, and panel member, WineAmigo.com. I thank, without implicating, Peter Werry and Tommy Lam for useful comments and suggestions. All errors, naturally, remain my own.

The note is structured as follows. After this introduction, several more prominent methodologies that are currently being used are reviewed (Section II). Section III will describe the proposed system in some detail, before a final section closes the article.

II. Existing Methodologies

Although the types of ratings systems are extremely varied, this section will examine a few which are more or less representative of the systems employed by wine enthusiasts and trade professionals in the market today. These are: hedonistic scoring; European scoring (20 point); American scoring (100 point). Clearly, variants of these exist, but in order to conserve space (and preserve sanity), only these above-mentioned systems will be discussed.

Hedonistic scoring

Hedonistic scores are related to taster appeal, and not necessarily quality. Such scoring systems are amenable to statistical methods that treat ordinal level data, and summary statistics can then be used to provide quantitative ratings. The U.S. Army Quartermaster General Method, developed in the 1950s and 60s, is the simplest of these, and essentially ranks wines à la consumer preference surveys:

Insert Figure 1 here

An extension of this simple system would attempt to incorporate the various characteristics of the wine into a hedonistic framework:

Insert Figure 2 here

Clearly, the primary drawback of this system is its quantitative limitations. Although, as mentioned above, the system allows for statistical testing, such tabulations require time and effort, and yield results that are not satisfactorily comparable.

European scoring

The fundamental basis for adopting a 20-point system is historical and practical. Previously, scorecards adopted 100-point scales. However, when evaluators noticed that most tasters scored wines between 65-85, a reversion was made to the 20-point system. There is still some basis for arguing the merits of such a smaller range: ease of tabulation, the limited number of outliers, and a minimization of superfluity. Although the definition is somewhat sweeping, such a system will be termed a European scoring system for the purposes of this article.

The basic score card would rank characteristics such as color, clarity, bouquet, and flavor, with a total of 4 points for visuals, 4 for the nose, and 12 for the palate. For example, German wines classified under this system would merit a *Qualitätswein b.A.* for 11 points, *auslese* for 15, and *trockenbeerenauslese* for 17 and above.¹

Insert Figure 3 here

The European system has clear merits: a clear, unambiguous scoring system that allows for inter-wine comparisons to be made; a systematic approach; and a relatively low total score that makes it easy for evaluators to award points. The main disadvantage of this approach, then, stems from its strength: when two excellent wines are sufficiently different to justify different ordinal rankings but not scores, the scoring system is too limited to capture this difference.

*American scoring*²

American ratings allocate a base of 50 points to every wine. The wine's general color and appearance merit up to 5 points. In general, due to advances in winemaking technology and concomitantly quality, most wines tend to receive at least 4 points, and often 5. The aroma and bouquet merit up to 15 points, depending on the intensity level and dimension of the aroma and bouquet as well as the cleanliness of the wine. The flavor and finish merit up to 20 points. Factors evaluated here are the intensity of flavor, balance, cleanliness, and depth and length on the palate. Finally, the overall quality levels, as well as potential for further evolution and improvement – the aging potential – merits up to 10 points.

As a general guide to interpreting the numerical ratings, 90-100 is equivalent to an “A” and is given only for an outstanding or special effort. Wines in this category are the very best produced of their type. There is a big difference between a 90 and 99, but both are top marks. 80-89 is equivalent to a “B” and such a wine, particularly in the 85-89 range, is very, very good; many of the wines that fall into this range often are great values as well. 70-79 represents a “C”, or an average mark, but obviously 79 is a much more desirable score than 70. Wines that receive scores between 75 and 79 are generally pleasant, straightforward wines that lack complexity, character, or depth. If inexpensive, they may be ideal for uncritical

¹ It should be noted, however, that German wines also increase in residual sugar as one progresses up in the classification, which, clearly, is not necessary for other types of wines.

² The discussion that follows is adapted from Robert Parker's *Wine Advocate rating system* (Parker 2001).

quaffing. Below 70 is an “F”. For wine, it is a sign of an imbalanced, flawed, or terribly dull or diluted product that will be of little interest to the discriminating consumer.

III. The System

The proposed system is based on two foundations: an adaptation of the Wine and Spirit Education Trust’s systematic approach to tasting (WSET 2000), and a proprietary weighting scale, which account for the demand and supply sides, respectively. Although the WSET approach encourages a hedonistic rating (poor, acceptable/average, good, excellent), the present system will append scores to each of the three sensory impressions of a wine: sight (appearance), smell (nose), and taste (palate). It then adjusts this given score by various supply-side factors. The overall score of the wine is then the algorithmic calculation of each wine’s demand and supply influences. If tasting in a panel, the final score will be the simple average of these overall scores. The score sheet for the proposed system is produced below.

Insert Figure 4 here

The score sheet is largely self-explanatory. In devising the system, the demand side modeled after the American system (with slight modifications), for a total of 50. The base of 50 can also be added, as per normal. Care was taken to distinguish between purely characteristic/descriptive factors (such as wine sweetness) and quality factors (such as residual sugar balance); consequently, where the translation of these factors from characteristic to quality was ambiguous, a lower weight was allocated.³

On the supply side, in terms of absolute price,⁴ expensive wines were not penalized for their price (although if they were overpriced this was reflected in the relative price adjustment), but cheaper wines were given higher weights to reflect their possibly superior value for money. The risk of this approach is that excessively cheap wines are awarded higher scores simply on the basis of their price. Hence, the price adjustment for the less than \$30 wines to reflect this possible diminution of quality. Note that the calculation of average price (for comparison to obtain the relative price differential) should be, ideally, the average of all the wines in the

³ The table also sought to provide descriptives together with quality scales, where appropriate, in order to assist evaluators in distinguishing the two.

⁴ All values are given in Singapore dollars (the exchange rate, as of Aug 1, 2002, was SGD\$1.76 to USD\$1. Note, also, that the prices chosen take into account the various taxes and transport costs (on average) that apply to wines shipped to Singapore. Clearly, these need to be adjusted according to country specificity.

particular tasting (hence, they should be of approximately the same vintage and region). In the event that this is not possible, i.e. in a vertical tasting, then the average price can be obtained from published data.

The overall vintage quality adjustment is probably the most controversial. Evidently, one does not wish to penalize a good wine from a poor vintage, or concomitantly, to mistakenly award a poor performer from an excellent vintage. However, all things considered, this single point adjustment allows one to capture the possibility that the wine would be superior to other vintages or regions, using established common knowledge.

In terms of implementation, it is conceivable that a spreadsheet program can be used to tabulate the supply-side weightings, such that each taster only need to fill in the demand-side values; the spreadsheet will then calculate the final (weighted) score. The other advantage of employing a spreadsheet is that it can also automate the task of averaging scores when the tasting is performed in a panel.

IV. Conclusion

This article has sought to incorporate supply side considerations into a proposed wine scoring system. Clearly, the limitations of the system provide avenues for extension – for example, more supply-side factors could be considered, or the weightings revised. Also, there would need to be experimental testing of the system in real world contexts. We leave these endeavors for future research.

References

Parker, Robert (2001), *The Wine Advocate Rating System*. Online: <http://www.erobertparker.com/info/legend.html> (accessed Jul 30, 2002).

Wine & Spirit Education Trust (2000). *WSET Student Tasting Notes*. London: WSET School.

Appendix

Quality	Score
Like – extremely	9
Very much	8
Moderately	7
A little	6
Neither like nor dislike	5
Dislike – a little	4
Moderately	3
Very much	2
Extremely	1

Figure 1

Aspect	Rating				
Appearance	Cloudy		Dull		Clear
Aroma	Unpleasant	Nondescript	Clean	Pleasant	Very pleasant
	Very sweet	Sweet	Medium dry	Dry	Very dry
Body	Very light	Light	Medium	Full	Heavy
	Soft		Firm		Harsh
	Acid		Balanced		Flabby
	Unpleasant	Flavorless	Light	Moderate	Full-flavored
	Rough finish		Mild finish		Smooth finish
	Fades quickly	Gone in 5 secs	Lingers 1 min	Linger 1-60 min	Lingers +1 hr

Figure 2

Characteristic	Maximum points
<i>Color</i>	
Pale, overcolored	0
Light	1
Typical	2
<i>Clarity</i>	
Dull	0
Bright	1
Brilliant	2
<i>Bouquet</i>	
Defective	0
Mute	1
Clean	2
Fragrant	3
Fine and flowery	4
<i>Flavor</i>	
Defective	0
Neutral	1-3
Clean but thin	4-5
Balanced	6-9
Ripe and noble	10-12

Figure 3

Demand Side		Supply Side	
Appearance		Absolute price	
Clarity	<i>Dull-Clear</i> (0-2)	> \$150	+0
Intensity	<i>Pale-Deep</i> (0-2)	\$100-150	+1
Rim vs Core	<i>Watery-Full</i> (0-1)	\$50-100	+2
Color	<i>Lemon-Gold</i> <i>Pink-Orange</i>	\$30-50	+3
	<i>Purple-Ruby-Tawny</i> (0-1)	< \$30	+2
Nose		Relative price	
Condition	<i>Clean-Unclean</i> (0-3)	> \$10 more	-1
Intensity	<i>Weak-Pronounced</i> (0-5)	± \$10	0
Development	<i>Youthful-Mature</i> (0-1)	> \$10 less	+1
Character	<i>Fruity/Floral/Vegetal/Spicy</i> <i>Singular-Complex</i> (0-5)	Overall vintage quality	
Palate		Excellent	+1
Sweetness	<i>Excessively Dry/Oversweet-Balanced</i> (0-2)	Average	+0
Acidity	<i>Flabby/Acetic-Balanced</i> (0-2)	Poor	-1
Tannin	<i>Absent/Astringent-Balanced</i> (0-2)		
Alcohol	<i>Absent/Excessive-Balanced</i> (0-2)		
Body	<i>Light-Full</i> (0-2)		
Character	<i>Fruity/Floral/Vegetal/Spicy</i> <i>Singular-Complex</i> (0-5)		
Length/Finish	<i>Short-Long</i> (0-5)		
Conclusions			
Overall Quality	<i>Poor-Excellent</i> (0-6)		
Maturity	<i>Immature-Ready</i> (0-2)		
Aging potential	<i>Low-High</i> (0-2)		

Figure 4