Are We Tasting the Same Thing? The Relevance of “Magnitude Matching” to Normalizing Fine Wine Ratings

Denton Marks
Department of Economics
University of Wisconsin-Whitewater

Abstract

Wine has been produced for millennia, and wine markets have existed for almost as long. The efficiency of those markets has increased over time with advancements as varied as low-cost, high-volume, temperature-controlled shipping to inventions that could perform various winemaking tasks at a cost per unit much lower than traditional methods to the refinement of packaging that enhanced the preservation of wine—even down to the bottle closure (e.g., Taber 2007).

The global wine industry seemed to advance toward increased information efficiency with the appearance and rising popularity of systematic expert wine evaluation within the last 40 years. The evaluation of wine...
characteristics became more systematic as one or a few wine aficionados—and particularly one “wine advocate”—began offering evaluation of a large number of wines from a variety of origins ostensibly without any commercial ties to the producers and with only consumer education and protection as a professional goal (McCoy 2005).

The market for everyday table wine as a daily beverage was as old as the industry, but consumer ignorance likely inhibited the growth of the fine wine market. The demand for fine wine was driven by those within “sensory communities” (e.g., Vannini, Waskul, and Gottschalk 2012) in traditional wine-growing and/or consuming areas who had grown up with wine appreciation and could afford the best examples of local wines and those outside those communities who might have the experience to know the wines or purchased them, perhaps, as Veblen goods. The increasing affluence of the industrialized world and the growth of the middle class meant that more households could afford fine wine, but the consumer’s problem of knowing relatively complex and potentially durable fine wines available only at high prices relative to beverage wine dampened demand. We can only speculate how much.

The recent appearance of numerous, perhaps objective evaluations of fine wines seems motivated by an interest in addressing the consumer’s problem in order to expand the wine market. Some of the motivation, especially from the pioneers in this field such as Parker, seems motivated by an interest in promoting informed wine buying generally, and consumer willingness to purchase such evaluations reveals that such an information market exists. However, the opportunity to influence wine purchases through ratings toward particular regions or even producers and vendors could complicate and cloud the process (e.g., vendors providing their own in-house ratings of otherwise unrated wines).

A considerable literature reports attempts to measure the extent to which expert ratings correlate with fine wine prices—usually auction prices for Bordeaux—with the testable hypothesis that better ratings correlated positively with prices (see, for example, the discussion in Storchmann (2012) and the more recent paper by Cardebat, Figuet, and Paroissien 2014). However, the evidence is mixed. It is naïve to conclude that expert ratings provide a straightforward guide to wine quality that consumers can use to form their willingness to pay with the result that, cet. par., higher rated wines are better and therefore more expensive. A recent paper (Marks 2015) provides a discussion of many of the difficulties of evaluating and comparing ratings and how little “knowledge” of a wine they likely impart.

This paper examines an hedonic measurement technique from food science that could make wine ratings more informative or at least help us understand better why wine ratings are not and perhaps cannot be more useful as they are currently produced. One of the structural difficulties with interpreting and comparing ratings is that, as a form of hedonic index of quality, interpersonal comparisons are questionable—for example, between two putative experts or between one’s chosen expert and oneself. In examples of comparing tastes, one difficulty is that the standards against which different tasters measure the appeal of a given taste differ. One of the clearest examples of this kind of difference is the existence of “supertasters” whose scale of intensity of taste differs markedly from the rest of us. Moreover, we can imagine that there are degrees of supertasting. If this applies to wine ratings, then the logical basis for saying that one or more experts can tell consumers what they will find in the bottle and that guides their willingness to pay would seem to fall apart. To the extent that we find positive ratings-price correlations, we need to look for other explanations such as the desire to “create collectibles” (Ashenfelter and Jones 2013).

This difficulty with simple hedonic rating comparisons has been discussed at length by food scientists such as Bartoshuk (e.g., Kalva et al. 2014). They have attempted to provide a method normalizing hedonic evaluation using the technique of “magnitude matching” (e.g., Marks (L.E.) et al. 1988) which, in effect, establishes a common scale of taste appeal and intensity so that tasters are providing hedonic ratings against a putative common yardstick. From this has emerged the hedonic general Labeled Magnitude Scale (gLMS).

Given the prominence of wine ratings in wine marketing and in consumer efforts to learn what is in the bottle,
applying the logic of this technique to the evaluation of wine ratings should be valuable. The paper provides a brief critique of the difficulties of interpreting wine ratings and focuses upon the structural difficulties of hedonic ratings generally. It then discusses the logic of gLMS and the case for its providing a common scale against which to measure taste quality as well as the critique of its reliability. The application of gLMS in food science has been under conditions considerably more controlled than those that produce wine ratings so I discuss how it might be applied in that case and some difficulties. The concluding section also considers other popular rating schemes (e.g., films, restaurants), how they compare to wine rating, and whether they might benefit from adopting the gLMS technique.