Journal of Wine Economics, Volume 10, Number 1, 2015, Pages 75–107
doi:10.1017/jwe.2015.1

Wine Tasters, Ratings, and *En Primeur* Prices*

Philippe Masset [a], Jean-Philippe Weisskopf [b] and Mathieu Cossutta [c]

Abstract

This paper examines the ratings of 12 influential wine critics on the Bordeaux *en primeur* market over the vintages 2003–2012. We hypothesize that wine experts differ significantly in their rating approach and influence on prices. We find that European critics are less transparent and in general more severe in their scoring than their American counterparts. Experts also appear to reach a relatively strong consensus on overall wine quality but have more diverse opinions on wines that achieve a surprising level of quality given the vintage, the ranking, or the appellation from which they originate. Our evidence also suggests that Robert Parker and Jean-Marc Quarin are the most influential critics, as a 10% surprise in their scores leads to a price increase of around 7%. We further find that their impact is higher for appellations and estates that are not covered by the official 1855 classification and for the best vintages. (JEL Classifications: C60, G11, Q11)

Keywords: *en primeur*, expert ratings, fine wine, wine futures.

I. Introduction

Every spring, wine enthusiasts and professionals turn to Bordeaux for the *en primeur* market. From March to April, tastings take place while the *en primeur* campaign itself usually runs from late April to June. Over the past decade, however, the *en primeur* market has evolved. More and more wine tasters and journalists take part in the *en primeur* tastings and create enthusiasm for a vintage. Buyers have become a heterogeneous mix of wine lovers, collectors, and investors, who are

*We thank participants at the Sixth Annual Meeting of the American Association of Wine Economists (AAWE), Karl Storchmann (the editor), and an anonymous referee for valuable comments and suggestions. We are indebted to Alain Bradfer, Anne-Christine Erlekam-Vogel for sharing their data and expertise on wine prices and Bertrand Le Guern and François Mauss for sharing their data and expertise on wine scores with us. All remaining errors are those of the authors.
[a] Ecole Hôtelière de Lausanne, Route de Cojonnex 18, CH-1000 Lausanne 25, Switzerland; e-mail: philippe.masset@ehl.ch (corresponding author).
[b] Ecole Hôtelière de Lausanne, Route de Cojonnex 18, CH-1000 Lausanne 25, Switzerland; e-mail: jean-philippe.weisskopf@ehl.ch.
[c] Ecole Hôtelière de Lausanne, Route de Cojonnex 18, CH-1000 Lausanne 25, Switzerland; e-mail: mathieu.cossutta@ehl.ch.

sometimes more concerned about the brand or the potential for price appreciation than consumption. Deciding which wines to buy has therefore become challenging. This is reinforced as the taste and quality of Bordeaux wines can strongly vary from one estate to another and from one vintage to another. Thus, Bordeaux wines are considered experience goods. One has to experience (i.e., consume) the wine to evaluate its quality.[1] Storchmann (2012) argues that the very nature of wine as an experience good generates an information void for customers that wine experts can help to fill. Rosen (2012) illustrates this influence that wine critics enjoy, writing that Robert Parker "has been called guru, emperor, pope … his name is now a verb (parkerize) used for the creation of a style of wine that pleases his taste." In this article, we therefore study wine critics, their ratings, and the influence they exert on *en primeur* prices.

So far, the existing academic literature has mostly addressed questions on the relation between producers, vintage, climate, or expert opinions and wine prices. Most papers have also focused on the relation between scores and prices and, more specifically, the influence of Robert Parker on the *en primeur* market. This article goes further by taking a broader perspective. Before examining the influence of the different expert scores on prices, we want to specify the way in which experts rate wines. Our dataset includes the scores of 12 leading critics for the vintages 2003–2012. We first study the rating approach of these experts. This analysis encompasses both their rating scale and philosophy. We then identify and quantify variables that may explain scores to gain a better understanding of the predictability of expert scores and their added value. We further investigate the degree of consensus among the different experts. Good experts may disagree when it comes to a particular wine, but on average they should display converging opinions. In a last step, we test whether expert surprises in scores correlate with prices.

We find that European critics are less transparent and use scoring systems that are often more difficult to interpret than do their American counterparts. Over the past decade, several critics have decided to change their rating scale, which has a direct influence on the score distribution. In general, European critics appear to be more severe than their American counterparts. Experts also reach a relatively strong consensus on overall wine quality, but have more diverse opinions on wines that achieve a surprising level of quality given the vintage, the ranking, or the appellation from which they originate. Our evidence also suggests that Robert Parker and Jean-Marc Quarin are the most influential critics, as a 10% surprise in their scores leads to a price increase of around 7%. We further find that their impact is higher for appellations and estates that are not covered by the official 1855 classification and for the best vintages (2005, 2009, 2010). Both appear to be complementary, catering to the French and U.S. market, respectively.

Our main contributions are twofold. First, we examine how 12 of the most respected experts in the world rate wines over a 10-year period. To the best of our

[1] See Gergaud et al. (2015) for a review of the literature on consumer behavior for experience goods.

knowledge, this article is the first to consider a dataset with such a large number of experts over such a long period. This allows us to investigate the precise differences in the rating methodology of each expert. We believe it is an important step toward better understanding the critics' work and influence on the wine market. Second, we analyze the relation between wine scores and prices using an approach that is robust toward endogeneity and unaffected by different distributional characteristics of the scores. Moreover, as we study multiple critics, we are able to contrast the various tasters and their influence on prices. Notably, our paper contributes to the literature by demonstrating not only that Robert Parker and Jean-Marc Quarin scores and *en primeur* prices are strongly correlated but more importantly that this correlation is much higher than it is for other experts.

The remainder of the paper is structured as follows. The next section reviews the existing literature while section 3 presents our dataset. Section 4 examines how experts rate wines. Section 5 displays an analysis of the relation between expert scores and *en primeur* prices, and section 6 concludes.

II. Literature

To date, the literature has focused mainly on the relation between climate or expert opinions and wine prices. Different studies, such as Ashenfelter (2010) or Ashenfelter et al. (1995), show that the final quality of a wine depends strongly on the climatic conditions that prevailed during the growing season. They further indicate that climate appears to predict wine quality better than do wine experts. Haeger and Storchmann (2006) suggest that for U.S. Pinot Noir, expert ratings improve the explanation of wine prices by weather data only marginally. Ashenfelter and Jones (2013) study the relation among climate, expert opinion, and prices. They illustrate that expert opinions on Bordeaux wines not only correctly reflect available climatic data but also display an independent effect and private information for the highest-rated wines.[2]

Jones and Storchmann (2001) specifically look at Parker ratings and show that they have stronger influence on prices for wines from small châteaux cultivating Cabernet-Sauvignon that have been highly rated in the past. In a theoretical paper, Mahenc and Meunier (2006) indicate that asymmetric information could be a source of inefficiency that leads *en primeur* prices to be set too high by the châteaux. In particular, they argue that the usual assumption of perfect competition, while acceptable for bottled wine, is not appropriate for *en primeur* wine. Ali and

[2] Studies have also examined the relation between expert opinions and prices for other products. For example, Reinstein and Snyder (2005) show that movie reviews display no overall effect but that positive reviews increase box office revenues. Berger et al. (2010) indicate that both favorable and negative *New York Times* reviews have a positive relation with book sales but that this is larger for positive reviews. Vaillant and Wolff (2011) conclude that cigar aficionados benefit from the information on cigar quality given in expert comments and ratings.

Nauges (2007) demonstrate that the marginal impact of *Wine Spectator* and Parker scores on *en primeur* prices is small but significant. Ali et al. (2008) examine the impact of Parker by exploiting the fact that he rated the 2002 Bordeaux vintage after its price release and confirm a limited but significant impact of Parker points on prices. Hay (2010) finds evidence that each additional Parker point is worth GBP 10 a bottle while Dubois and Nauges (2010) also find a significant influence of Parker on prices. Roberts and Reagans (2007) find evidence that new world wines sold on the U.S. market display a positive rating-price relationship that increases for producers with critical exposure. San Martín et al. (2008), by contrast, indicate that expert opinions have less influence on prices of Argentinian wines than labeling, location, or variety. Friberg and Grönqvist (2012) look at the effect of favorable reviews on wine demand in Sweden. They document that, following positive reviews, demand rises by 6% and remains significant for the following 20 weeks. Hilger et al. (2011) refine these results and suggest that demand decreases for wines with low scores, while it increases for wines with average or high scores. They argue that expert opinion labels therefore transmit not only shelf visibility but also some information on quality to consumers. Ramirez (2010), in a study on Napa Valley wines, shows that not only expert ratings but also their comments have a positive impact on wine prices.

Another avenue that has gained attention considers the way in which wine prices act as quality signals on the market. To convince consumers to buy their products, producers can send quality signals in different forms (price, advertising, brand name, warranty).[3] This also holds true on the market for fine wine in which consumers may use the label (Gluckman, 1990) or awards (Orth and Krška, 2001) to infer product quality. Lockshin and Rhodus (1993) show that wine prices are also used as strong quality signals. Evidence by Ashenfelter (2010) and Ashenfelter et al. (1995) indicates that, at release, wine prices diverge from their full information prices. However, as additional information is released, these wine prices tend to move toward their full-information equilibrium. Schnabel and Storchmann (2010) find some support for the hypothesis put forward by Bagwell and Riordan (1991) that prices can be used to signal product quality. Their results show that price signals display a negative relation with additional information and a positive link to wine quality. Gergaud and Livat (2007) analyze a sample of 6,000 European wine consumers and find that prices substitute for umbrella branding as a signal. These prices acting as signals of wine quality are finally less used by wine connoisseurs than nonconnoisseurs.

III. Data

Our dataset includes information on *en primeur* scores and prices for a total of 122 Bordeaux estates over the vintages 2003 to 2012. These estates have been selected on the basis of the number of available price and score observations as lesser-known estates are rated by very few experts. These 122 estates represent more than 90%

[3] See Kirmani and Rao (2000) for a detailed literature review on this topic.

of the *en primeur* market.[4] More specifically, our sample includes 67 reds from the Médoc, 10 reds and 6 whites from Pessac-Léognan and Graves, and 39 reds from Pomerol and Saint-Emilion.[5]

A. En Primeur *Prices*

In Bordeaux, the relation between producers and final customers is complicated. Producers typically use the services of *courtiers* to help them set their release prices. They then sell their wines to *négociants* (also known as the *négoce* or the trade), which take a percentage before selling the wines to merchants all over the world. Merchants then add their own margin, before selling it to the general public. In this article, we use ex-VAT (value-added tax) final customer prices that we have obtained from different sources.

Table 1 reports descriptive statistics on prices. As apparent from the first line, our dataset contains release prices for almost all wines included in the sample, except for some rare wines that were not released *en primeur* (e.g., in 2012, Latour stopped selling its wines *en primeur*). The average is always higher than the median because a minority of top estates sell their wines at high prices. One can also observe an asymmetric relationship between price changes and quality from one vintage to another. Excellent vintages, such as 2005 or 2009, lead to high price increases. However, if subsequent vintages are of lower quality, prices decrease slowly. The gap between the most expensive and cheapest wines has also widened since 2003.

B. En Primeur *Scores*

We have hand-collected *en primeur* scores from 12 internationally acknowledged wine critics. Table 2 provides basic information on each critic, including name, number of wines tasted (in our sample), country of origin, and a brief description. Not all critics have rated all wines, and some have skipped entire vintages. For instance, the Grand Jury Européen has not rated the 2003, 2011, and 2012 vintages. Similarly, no information is available for Bettane and Desseauve scores for the 2003 vintage. James Suckling only started rating *en primeur* Bordeaux in 2010, while Farr Vintners only provides ratings for a limited number of wines. Among these 12 critics, four are from France, four from the United Kingdom, and four from the United States.

Magazines such as *La revue des vins de France, Decanter,* or *Wine Spectator* employ different experts to rate the different Bordeaux appellations or vintages.

[4] Though it is difficult to find precise data, some evidence supports this claim. According to Liv-ex, in May 2013, 87% of the trading volume of 2012 Bordeaux was due to just five estates, namely, Lafite Rothschild, Carruades de Lafite, Haut-Brion, Mouton Rothschild, and Trotanoy. All these estates are included in our dataset.

[5] An overview of the different Bordeaux appellations and a complete list of estates used in this study are in Appendixes 1 and 2.

Table 1
Wine Prices (in Euros): Descriptive Statistics

	2003	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*
# Producers	117	122	122	122	120	122	122	122	122	118
Average	52.3	41.4	91.7	89.8	72.7	65.3	158.1	165.9	103.5	88.6
Median	30.7	25.5	50.9	38.8	35.6	32.7	63.8	69.2	49.4	43.2
10th percentile	16.1	15.1	21.4	19.7	18.7	17.3	24.6	28.1	23.7	22.3
90th percentile	128.5	80.5	200.5	170.6	133.0	111.7	285.2	262.4	157.7	178.6

Notably, James Suckling was the Bordeaux expert at *Wine Spectator* before launching his own website in 2010. Similarly, Michel Bettane and Thierry Desseauve were the editors of *La revue des vins de France* before leaving the magazine in 2004. The Grand Jury Européen is also special as it is not a single expert but an association that includes several well-known experts. Each tasting session involves at least 15 experts who rate each wine and is subject to strict conditions. According to François Mauss, the founder of the association, the scores emanating from their tasting sessions should be less subjective than those from individual critics as they include and aggregate the opinions of different experts, thereby reducing idiosyncratic biases.

IV. Analysis of Expert Scores

A. How Do Experts Rate the Wines They Taste?

(1) Rating Systems Used by Experts

Comparisons are difficult to make, as not all experts use the same scale to establish their scores. Table 3 shows the various rating systems that are currently used. It is apparent that two distinct approaches are followed: the American 100-point scale and the European 20-point scale.

All American critics (including Neal Martin [NM]) use a similar scoring approach, with six or seven quality tiers and scores ranging from 50 to 100. James Suckling (JS) states on his website: "I rate wines using the 100-point scale. I have used this point system for close to 25 years. I still believe it is the simplest way to rate a wine, with its origins from grade school in the United States."[6] The definitions of the various tiers have only minor differences . In particular, Stephen Tanzer is more severe, as can be seen from his scale. For instance, a wine deemed excellent will get a score marginally lower than or equal to 89 from Stephen Tanzer (ST), while the same wine gets a better score from Robert Parker (RP) or James Suckling (who do not have a specific category for excellent wines: they are either very good or outstanding). The Grand Jury Européen (GJE) uses the same scale as its American counterparts with some differences in the definition of the various tiers.

[6] http://www.jamessuckling.com/about-us.html.

Table 2
Presentation of the Wine Experts

Expert	*Origin*	*Obs.*	*Brief description*
Grand Jury Européen (GJE)	F	834	Group of influential experts. *En primeur* score for each wine is calculated as the average score of 15 experts. Scores and vintage analysis are published on the GJE website (gje.mabulle.com).
Michel Bettane and Thierry Desseauve (BD)	F	1,015	Former editors of *La revue des vins de France* (until 2004). Their tasting notes are available on their website (www.mybettanedesseauve.fr) and in *Le guide Bettane & Desseauve des vins de France* (a book updated once a year).
Jean-Marc Quarin (JMQ)	F	1,162	Independent critic. Widely considered one of the best experts on Bordeaux wines. He publishes his tasting notes in his *Carnets de dégustation* (available on his website: www.quarin.com).
La revue des vins de France (RVF)	F	1,071	Leading French wine magazine. It launched a Chinese edition in 2011. Every year, a special edition is devoted to *en primeur* Bordeaux wines.
Decanter (DEC)	UK	1,163	Widely considered one of the most influential wine magazines. Michael Broadbent has contributed to the magazine since 1975 and developed *Decanter*'s scoring system.
Farr Vintners (FV)	UK	861	Britain's largest wholesale fine wine merchant. Known for the quality and the objectivity of the information (including tasting notes and scores) they provide to their customers.
Jancis Robinson (JR)	UK	1,188	Often considered the British counterpart to Robert Parker. She has a deep understanding of the wine business and writes a weekly column for the *Financial Times*. Her tasting notes are available on her website (www.jancisrobinson.com).
Neal Martin (NM)	UK	1,102	After having run his own website for about four years, he was invited by Robert Parker in 2006 to become a contributor to *The Wine Advocate*. Since then, he has had great autonomy and is often considered Parker's potential successor.
James Suckling (JS)	United States	328	The main competitor to Robert Parker, at least in the United States (he is less known in Europe). He specializes in the wines of Bordeaux and Tuscany. After having worked for *Wine Spectator*, he launched his own website in 2010 (http://www.jamessuckling.com).
Robert Parker (RP)	United States	1,045	Launched *The Wine Advocate*, a bimonthly publication, in 1978. He is considered the reference for Bordeaux and Californian wines and, to a lesser extent, for wines from the Rhône Valley. His scores are believed to have an influence on both demand and prices. (Website: http://www.erobertparker.com)

Continued

Table 2
(Continued)

Expert	*Origin*	*Obs.*	*Brief description*
Stephen Tanzer (ST)	United States	1,066	Launched the *International Wine Cellar,* a bimonthly publication available in English, French, and Japanese, in 1985. He is known for his independence and his intransigent opinions (his score being typically lower than those of RP and JS).
Wine Spectator (WS)	United States	1,115	American wine magazine, with 15 issues per year. Well known for its Top 100 wines of the year. James Molesworth replaced James Suckling as the Bordeaux specialist.

In Europe, most critics apply a 20-point scale. However, their scores are more difficult to compare than those from their American counterparts. For instance, a wine rated 16 points would be considered excellent by Jean-Marc Quarin (JMQ), very good by the *Revue des vins de France* (RVF) and Bettane and Desseauve (BD), and just good by Decanter (DEC) and Jancis Robinson (JR). In general, British critics (DEC and JR) can be considered less severe than French critics (JMQ, RVF, and BD). Theoretically, scores could be lower than 10, but, in practice, this possibility does not occur. This ambiguity and cultural specificities explain why scores by French critics are often lower than those of American critics. Moreover, some scales may sometimes be inconsistent or have gaps. In particular, some critics, like Jean-Marc Quarin and *Decanter,* while primarily using a 20-point scale, also provide the equivalent scores on a 100-point scale. The relationship between the two scales is, however, incoherent at times.

Another key difference between American and European critics regards the degree of transparency of their scoring method. The former clearly disclose and explain their rating system. In Europe, it is more complex. First, three critics (GJE, Farr Vintners [FV], and BD) do not disclose their scoring system. It is, however, possible to infer some basic information on the basis of the tasting notes that are published with the observed scores.

Second, over the past decade several critics have changed their rating scale, sometimes multiple times . The Grand Jury Européen used a 5-star scale for 2004 and 2007 vintages but a 100-point scale in 2005 and 2006. The *Revue des vins de France* switched from a 5-star system to a 20-point system in 2006 after having used a 5-point scale in 2003–2005. *Decanter* used a 5-star rating scheme in 2003, 2004, 2006, and 2008 but a 20-point scale for all other vintages. Neal Martin is the only critic who used a 25-point scale, in 2003 and in 2005. The remaining six critics have never changed their scoring system (JMQ, FV, JR, RP, ST, and Wine Spectator [WS]).

(2) Effective Distribution of Expert Scores

Table 4 reports key statistics on the scores delivered by the various experts. Several critics have opted either for the 20-point scale (RVF and DEC) or the 100-point scale

Table 3
Rating System Used by Wine Experts

	American 100-point rating system			*European 20-point rating system*				
	ST	*NM, RP*	*WS, JS*	*RVF*	*BD*	*JMQ*	*DEC*	*JR*
Extraordinary	95–100	96–100	95–100	20	20	18–20(96–100)	18.5–20(95–100)	20
Outstanding	90–94	90–95	90–94	17.5–19.5	18–19.5			19
Excellent	85–89			15.5–17	16.5–17.5	16–17.5(90–95)	17–18.25(90–94)	18
Very Good		80–89	85–89		15–15.5	15–15.75(87–89)		17
Good	80–84		80–84	13.5–15	13.5–14.5	14–14.75(85–87)	15–16.75(83–89)	16
Average	75–79	70–79		11.5–13		10–13.75(70–84)	13–14.75(76–82)	15
Below average	70–74	60–69	75–79				11–12.75(70–75+)	13–14
Avoid	<70	50–59	50–74	≤11		<10(<50)	<11(66–69)	12

Notes: FV: Information is not available, but the scale used is apparently similar to that of JR.

GJE: Information is not available, but the scale used is apparently similar to that of RP.

BD: The scale is based on the one provided in their book (*Le grand guide des vins de France*). However, it is not clear whether this score is valid for both in-bottle and *en primeur* wines.

Table 4
Key Statistics on Expert Scores

	ST	*NM**	*NM*	*RP*	*WS*	*JS*	*GJE**	*GJE*	*RVF**	*RVF*	*BD*	*JMQ*	*DEC**	*DEC*	*JR*	*FV*
Scale min.	50	0	50	50	50	50	0	50	0	10	10	10	0	10	10	10
Scale max.	100	25	100	100	100	100	5	100	5	20	20	20	5	20	20	20
Mean (m)	90.27	20.17	91.05	91.94	91.35	92.93	3.58	90.84	4.10	16.83	17.07	16.25	3.83	17.39	16.69	16.24
Median	90.00	20.00	91.00	92.00	90.50	92.50	3.50	91.00	4.00	17.00	17.00	16.25	4.00	17.50	16.50	16.00
STD (s)	2.48	2.09	3.41	3.31	3.40	2.71	0.85	2.71	0.66	1.33	1.21	0.96	0.66	0.97	0.94	1.26
CV	0.03	0.10	0.04	0.04	0.04	0.03	0.24	0.03	0.16	0.08	0.07	0.06	0.17	0.06	0.06	0.08
Skewness	0.52	0.02	−0.40	−0.31	0.11	0.34	0.17	0.06	−1.59	−0.20	−0.30	0.42	−0.04	0.17	−0.11	0.56
Kurtosis	3.07	2.64	3.34	4.17	2.67	3.16	1.76	3.13	7.41	3.18	3.17	3.92	2.74	3.00	3.67	3.10
Min.	85.5	15.0	76.5	73.5	81.5	86.5	2.0	82.0	1.0	11.0	12.0	13.0	2.0	14.5	12.0	13.0
Max.	98.5	25.0	99.5	99.0	98.5	100.0	5.0	98.0	5.0	20.0	20.0	20.0	5.0	20.0	19.5	20.0
<m – 1s	87.8	18.1	87.6	88.6	87.9	90.2	2.7	88.1	3.4	15.5	15.9	15.3	3.2	16.4	15.7	15.0
	15%	22%	15%	15%	17%	13%	21%	18%	9%	20%	13%	16%	29%	11%	14%	8%
<m – 2s	85.3	16.0	84.2	85.3	84.5	87.5	1.9	85.4	2.8	14.2	14.7	14.3	2.5	15.5	14.8	13.7
	0%	1%	3%	2%	2%	2%	0%	3%	3%	2%	3%	2%	1%	1%	2%	1%
>m + 1s	92.8	22.3	94.5	95.2	94.8	95.6	4.4	93.6	4.8	18.2	18.3	17.2	4.5	18.4	17.6	17.5
	15%	14%	16%	14%	15%	11%	30%	15%	12%	16%	16%	14%	13%	19%	12%	20%
>m + 2s	95.2	24.3	97.9	98.6	98.2	98.4	5.3	96.3	5.4	19.5	19.5	18.2	5.1	19.3	18.6	18.8
	4%	2%	1%	2%	1%	5%	0%	2%	0%	3%	2%	3%	0%	3%	2%	5%

Notes: STD denotes standard deviation, and CV denotes coefficient of variation (defined as the ratio of s/m); the line on the right of "m – 1s" shows the average (m) minus one standard deviation (s), and the line below it reports the percentage of observed scores that are lower or equal to this score.

Several critics have changed their rating scheme over the last decade. A star (*) has been added next to the names of those critics in order to distinguish their old scheme from the new one.

(GJE and NM). For those critics, we report two sets of statistics, one for each scale used. Figure 1 shows the entire distribution of scores for each critic.

As can be observed in Table 4, European critics sometimes give perfect scores to *en primeur* wines. At the exception of James Suckling, American experts generally wait until the wines are bottled before eventually awarding such a score. Scores from European experts can also be quite low. They are often lower than 16 (or 80 points on the American scale) and some of them even below 14 (or 70). The higher coefficient of variation further indicates that their scores are more variable than those of American tasters.[7] These observations do not necessarily mean that European experts are more severe than Americans but, rather, indicates that they use a larger proportion of the available score range. These observations reflect a cultural difference between American and European experts.

Among the critics who use a 100-point scale, James Suckling and Robert Parker have the highest average scores, while Stephen Tanzer has the lowest scores. Neal Martin is in the middle of these two extremes. Scores from the Grand Jury Européen are relatively low on average and have a limited degree of variability, which is due to its aggregation of scores from several critics. Figure 1 further confirms that Stephen Tanzer is, on average, more severe than his U.S. colleagues. For critics who rate on a 20-point scale, *Decanter* is the most generous and Farr Vintners the most severe. This last observation is surprising given the fact that Farr Vintners is a wine merchant, but it is consistent with its reputation of impartiality.

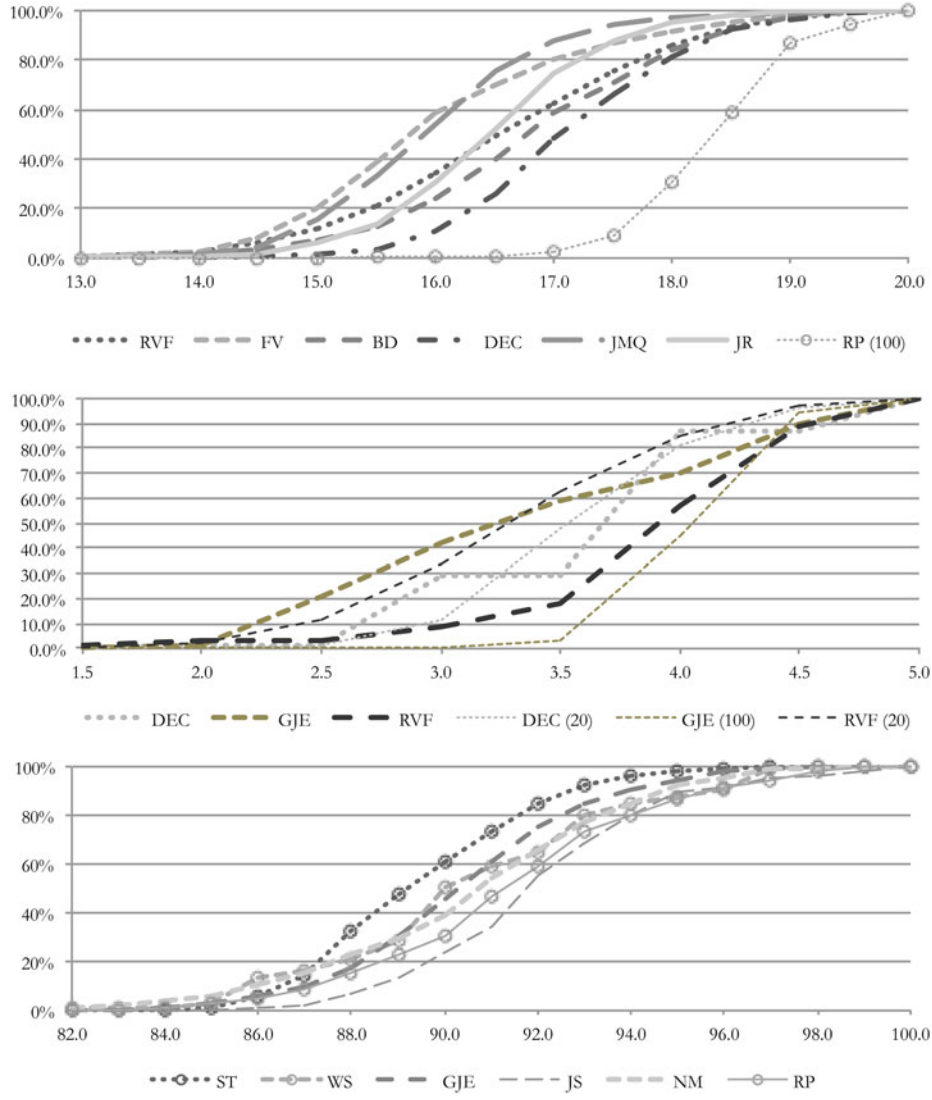
Expert scores have a kurtosis that is often slightly above 3, and the skewness strongly varies from one expert to another. The assumption of normality has to be rejected (at the 5% level) for most experts, except for the Grand Jury Européen and *Decanter*. Scores from these two experts are distributed almost linearly along the portion of the scale they use (see Figure 1). The *Revue des vins de France* exhibits the most negative skewness due to the larger number of rather low scores relative to the number of high scores. Stephen Tanzer presents the opposite situation. Being more severe than some of his counterparts on average, he also awards high scores to wines he considers exceptional. The high kurtosis exhibited by Robert Parker scores is explained by a significant number of very low and very high scores, although most of his scores are centered around the mean.

The rating scheme has a direct impact on the way in which experts rate wines. For instance, the Grand Jury Européen has used a 5-point scale for the 2004 and 2007 vintages and has opted for a 100-point scale for all the others. Figure 1 shows that this change led to a shift in the scores toward the right-hand side of the graph, thereby indicating that scores, when expressed on a similar scale, look more generous today than in 2004 or 2007. Interestingly, if one considers the case of the *Revue des vins de France*, the situation is reversed, with an increase of the number of low scores

[7] Overall, the most variable scores are from critics that use a five-star scale.

Figure 1

Cumulative Distribution of Expert Scores



Notes: The upper panel shows the distribution for experts rating on a 20-point scale. The middle panel shows the distribution for experts that were rating on a 5-star scale and contrasts it with the distribution based on their current scores. The lower panel is for experts rating on a 100-point scale.

and a reduction of high scores. Finally, the distribution of *Decanter* scores did not significantly change after its decision to change its scoring approach. Overall, these observations reinforce the impression that scores expressed on a 20-point scale look less generous than those expressed on a 100-point scale.

B. Determinants of Expert Scores

In this section, we examine the determinants of expert scores. To make scores comparable for the different critics and facilitate the interpretation of the results, we have rescaled all scores to a range between 1 and 100 points. We consider variables such as the appellation (used as a proxy for natural endowments), the ranking[8] of a wine (used as an estimate of reputation and also to complement the information provided by the appellation), and its vintage (used to take into account climatic conditions and their impact on quality). We further assume the relation between all these explanatory variables and expert scores to be linear and of the form

$$Score_{i,t}^{(j)} = \alpha_0^{(j)} + \sum_{v=2003}^{2012} \alpha_v^{(j)} D_{\{vint_t=v\}} + \sum_{a=1}^{9} \alpha_a^{(j)} D_{\{app_i=a\}} + \sum_{r=1}^{8} \alpha_r^{(j)} D_{\{rank_i=r\}} + \varepsilon_{i,t}^{(j)} \quad (1)$$

where $D_{\{vint_t=v\}}$ takes the value 1 if the wine is from vintage v and 0 otherwise; similarly, $D_{\{app_i=a\}}$ and $D_{\{rank_i=r\}}$ take the value 1 if the wine is from the ath appellation and is part of the rth ranking category. We estimate the model using a conventional ordinary least squares (OLS) regression. We also consider a model, in which all scores are pooled together:

$$Score_{i,t} = \alpha_0 + \sum_{v=2003}^{2012} \alpha_v D_{\{vint_t=v\}} + \sum_{a=1}^{9} \alpha_a D_{\{app_i=a\}} + \sum_{r=1}^{8} \alpha_r D_{\{rank_i=r\}} + \varepsilon_{i,t} \quad (2)$$

Model 1 enables us to assess how the different experts appraise wines from various appellations, vintages, and ranking categories. Model 2 provides additional information on the *average* preferences of all experts. These two models can be refined to determine which wines attract the best scores. Thus, we replace the appellation and ranking dummies in models 1 and 2 by a wine producer dummy to obtain:

$$Score_{i,t}^{(j)} = \alpha_0^{(j)} + \sum_{v=2003}^{2012} \alpha_v^{(j)} D_{\{vint_t=v\}} + \sum_{k=1}^{122} \alpha_k^{(j)} D_{\{ID_i=k\}} + e_{i,t}^{(j)} \quad (3)$$

$$Score_{i,t} = \alpha_0 + \sum_{v=2003}^{2012} \alpha_v D_{\{vint_t=v\}} + \sum_{k=1}^{122} \alpha_k D_{\{ID_i=k\}} + e_{i,t} \quad (4)$$

where $D_{\{ID_i=k\}}$ takes the value 1 if the wine under consideration is the kth from the list and 0 otherwise; $D_{\{vint_t=v\}}$ is defined as before. Table 5 reports detailed results for the estimation of models 1 and 2. All coefficients are to be interpreted as the score

[8] The official classification dates back to 1855 and was ordered by Napoléon III for the Paris Universal Exhibition. With the exception of Château Mouton Rothschild, it has not been altered since. It classifies red wines from the Médoc (and Château Haut-Brion) from first (1GCC) to fifth growth (5GCC) and white wines from Sauternes and Barzac into superior first, first and second growth. The Saint-Emilion classification was first proposed in 1954. It is issued by the and is more or less updated every 10 years (2012 was the last time). It includes first growth A (1GCC A), first growth B (1GCC B) and growth (GCC) wines.

Table 5

Determinants of Expert Scores (Results from the Regression Analysis)

	FV	*BD*	*DEC*	*GJE*	*JMQ*	*JR*	*JS*	*NM/WJ*	*RP*	*RVF*	*ST*	*WS*	*All*
Intercept	86.19***	92.62***	94.68***	99.29***	74.93***	89.05***	82.05***	94.44***	93.09***	89.51***	73.46***	96.06***	88.31***
Médoc	−16.12***	−12.43***	−14.94***	−13.54***	−12.03***	−11.85***	−19.03***	−8.92***	−5.64*	−7.18**	−12.66***	−17.13***	−11.43***
Margaux	−10.58***	−10.41***	−13.22***	−12.92***	−6.38**	−6.05***	−8.98***	−8.94***	−6.08***	−7.40***	−10.35***	−9.56***	−8.50***
Saint-Estèphe	−3.76	−2.54	−5.10	−2.36	−3.58	−1.33	−6.81**	−7.55**	−4.03*	−0.40	−2.01	−6.39*	−3.13
Saint-Julien	−4.85*	−0.41	−5.54	−1.79	−2.90	−1.93	−2.21	−0.62	−3.27	0.71	−0.93	−1.09	−2.01
Pessac-Léognan	−0.98	−0.46	−5.37	−0.94	0.40	−5.53***	4.78	−0.43	3.23	7.13*	4.12	−2.22	0.38
Pomerol	12.07***	2.81	2.90	10.21**	7.48**	1.29	14.82***	7.24**	7.60***	5.52	14.78***	9.98***	8.36***
Saint-Emilion	−3.70	−0.01	−0.89	4.23	−2.76	−5.06	5.00	−3.80	7.18*	5.42	4.68	−0.73	0.27
Graves White	19.48***	9.39**	10.86**	15.5***	7.32**	3.00	33.58***	6.63*	14.88***	13.58***	24.98***	15.47***	13.79***
2GCC	−19.34***	−12.39***	−19.82***	−22.38***	−13.6***	−18.42***	−16.35***	−14.10***	−7.33***	−8.74**	−20.74***	−15.87***	−15.47***
3GCC	−28.78***	−18.05***	−24.32***	−31.74***	−17.16***	−19.21***	−17.15***	−17.06***	−11.39***	−11.95***	−25.18***	−20.49***	−19.93***
4GCC	−37.22***	−23.18***	−36.13***	−41.01***	−24.42***	−22.64***	−26.33***	−23.71***	−16.90***	−19.64***	−34.41***	−28.57***	−27.36***
5GCC	−35.05***	−28.03***	−41.31***	−43.73***	−28.21***	−24.36***	−27.73***	−25.67***	−21.91***	−21.63***	−37.92***	−30.52***	−29.68***
1GCC A	−14.19	−4.80	−14.28**	−9.62	−2.58	−9.76*	−8.37	−2.85	−7.17	−4.03	−10.02	−4.07	−6.95
1GCC B	−31.80***	−22.22***	−31.76***	−31.30***	−14.77***	−20.98***	−25.15***	−19.76***	−18.92***	−20.04***	−33.96***	−21.17***	−23.31***
Others	−36.82***	−22.97***	−34.64***	−39.28***	−26.01***	−21.47***	−27.92***	−22.00***	−18.69***	−20.53***	−37.16***	−27.53***	−27.81***
2003	−8.64***	—	−1.08	—	−9.45***	—	—	−32.09***	−10.06***	1.70	−11.04***	−10.92***	−9.58***
2004	−17.26***	−15.65***	0.60	−7.73***	−9.41***	−5.35***	—	−13.56***	−8.15***	2.18	−11.27***	−24.08***	−10.18***
2005	−2.42**	−7.79***	−3.49***	−25.43***	−0.58	—	—	−15.21***	−2.53***	4.57***	−1.32	−4.40***	−4.94***
2006	−9.38***	−12.08***	0.96	−13.21***	−7.44***	−2.90***	—	−13.38***	−6.12***	−13.90***	−8.85***	−21.45***	−9.50***
2007	−14.76***	−9.18***	−18.87***	−16.88***	−10.73***	−8.95***	—	−14.76***	−17.23***	−10.45***	−16.88***	−31.65***	−14.75***
2008	−14.98***	−8.26***	2.27	−9.71***	−8.81***	−2.54***	—	−12.7***	−5.34***	−12.36***	−12.42***	−22.4***	−8.79***
2009	1.44	2.77**	−2.11**	−2.43***	−1.66*	0.18	—	−4.52***	2.46***	−1.09	3.44**	2.32*	0.42
2011	−14.11***	−6.18***	−13.21***	—	−6.43***	−4.03***	−20.02***	−9.73***	−11.34***	−15.66***	−9.79***	−13.23***	−11.07***
2012	−8.44***	−8.71***	−9.44***	—	−7.38***	−2.57***	−20.54***	−9.60***	−9.12***	−11.78***	−10.04***	−13.31***	−9.47***
Observations	861	1015	1163	834	1162	1188	328	1102	1045	1071	1066	1115	11950
Adj. R^2	0.54	0.43	0.42	0.53	0.44	0.29	0.55	0.38	0.48	0.42	0.46	0.53	0.26

Notes: The first row indicates expert names for the first 12 columns and is estimated with model 1. The last column, All, pools all experts and is estimated with model 2.

The reference is a first growth from Pauillac, 2010 vintage.

Others indicate second wines, cru bourgeois, and grand cru classé (in Saint-Emilion).

— denotes that the respective expert has not tasted the specific vintage

***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

difference between a wine from any ranking, any appellation, any vintage, and a 2010 Pauillac first growth (used as the reference).

In Bordeaux, status, tradition, *terroir,* and climate have great influence on both perceived and effective quality. However, unobservable variables (e.g., specific viticultural or oenological decisions made by a winemaker or the influence of a particular microclimate in a given vintage) may lead to positive or negative surprises. Customers expect wine critics to be able to identify such surprises and to be sufficiently precise in their tastings to ensure that their scores are truly reliable. Cao and Stokes (2010) argue that wine experts should be judged on the basis of three characteristics: bias (difference between the scores of an expert and the average score of a panel of experts), variation (degree of randomness in an expert's scores), and discrimination (ability to distinguish wines on the basis of their quality).

In the present context, a low R^2 for model 1 would indicate that scores are erratic (or random), which may cast doubt on their reliability. However, an R^2 that is too high would suggest that scores do not provide information on the unobservable quality of a wine and thereby do not carry added value (low discrimination ability). Scores from the Grand Jury Européen being aggregated from the opinions of several experts should only have a minimal bias and exhibit a low degree of randomness (the noise imbedded in the appreciation of different judges is cancelled out). Moreover, they are obtained under strict blind tasting conditions, these scores cannot be influenced by information such as appellation or ranking and should therefore facilitate the identification of genuinely successful wines in a given vintage.[9] We therefore use the R^2 obtained on the Grand Jury Européen scores as a benchmark. Indeed, the Grand Jury Européen exhibits an R^2 of 0.53, the third highest among all critics.[10] Farr Vintners and James Suckling display similar R^2, at 0.54 and 0.55. The other R^2 are generally close to 0.40–0.50, indicating that about 40% to 50% of expert scores can be explained by generic variables related to quality and reputation. Jancis Robinson and Neal Martin seem to be the most independent critics, with R^2 of 0.29 and 0.38, respectively. This low R^2 suggests that their scores are less reliable than those from other critics.

Model 1 also allows us to assess whether an expert favors wines from a particular appellation or rank. If expert opinions were true measures of a wine's objective quality, they should all lead to similar conclusions concerning the quality of a wine and therefore share the same preferences. In general, wines from the right bank, in particular, those from Pomerol, obtain better scores than those from the left bank. On the left bank, Pauillac, Saint-Estèphe, and Saint-Julien are the

[9] Other experts do not necessarily perform strict blind tastings. There is some opacity on how wines are tasted by most experts and on how many tastings are conducted before attributing scores.

[10] As a robustness test, we estimated the average score of all critics who use a 20-point scale. We then used this average score in the regression. The R^2 is slightly higher than 0.60. This demonstrates that the aggregation of scores almost mechanically leads to an increase in the R^2.

preferred appellations of most critics. Margaux and Médoc are less appreciated. The rare whites from the Graves appellation enjoy high scores from all critics.

Although there is a general consensus regarding the best and the worst appellations, some significant differences among critics can be observed. American critics are more negative about wines from the northern part of the Médoc (in particular, Saint-Estèphe) but more positive about wines from the right bank and Pessac-Léognan. Robert Parker and, to a lesser extent, Stephen Tanzer are the most positive about these wines, while Jancis Robinson exhibits almost opposite preferences. These results are in line with Cardebat and Livat (2013), who conclude that American experts (Robert Parker, Stephen Tanzer, and *Wine Spectator*) favor wines from the right bank and Pessac-Léognan, while Jancis Robinson gives on average better scores to wines from the left bank.

The relation between the 1855 classification and expert scores is statistically and economically significant. That is, a second growth from the Médoc will typically obtain a lower score than a first growth. First growths "A" from Saint-Emilion are generally less appreciated than first growths from the Médoc, but the difference is statistically insignificant. When it comes to first growths "B," the difference widens and becomes statistically significant. Overall, critics appear to demonstrate the same preferences in terms of ranking and appellation, but the amplitude of those preferences varies from one critic to another.

We also examine vintage regression coefficients in order to determine the qualitative hierarchy of the last ten vintages. In general, the preferred vintage is 2009, which is followed closely by 2010. Most experts also rank 2005 quite high, except for *Decanter* and the Grand Jury Européen; 2006, 2008, and 2012 are good to very good but relatively heterogeneous. Appreciation varies quite significantly from one expert to another. In some instances, 2003 can be outstanding but is more generally only average to good while 2004, 2007, and 2011 seem to be weaker.

The estimation of models 3 and 4 allows us to identify the wines that are the most praised by experts.[11] Latour displays the best scores over the sample period. It is followed by Margaux, Haut Brion, and Lafite Rothschild, with Mouton Rothschild and Léoville Las-Cases not far behind. Among the first growths "A" from St-Emilion, Ausone, and Cheval Blanc are preferred and are close to Latour. Angélus and Pavie, which have been recently promoted, attract lower scores. In general, the results are in line with the official classification but nevertheless some striking differences can be observed. For instance, Palmer (third growth) and Pontet-Canet (fifth growth) fare much better than their ranking suggests, while Rauzan-Gassies or Durfort-Vivens (both second growth) have disappointing scores. This illustrates the inherent limitations of the rigid 1855 classification.

[11] In order to save space, results for models 3 and 4 are not reported here but are available from the authors upon request.

Moreover, models 3 and 4 allow a finer analysis of the Pomerol appellation, for which no classification exists. Pétrus and Lafleur are clearly the preferred wines from this appellation, followed by Eglise-Clinet, Vieux Château Certan, and Trotanoy. The higher degree of precision in models 3 and 4 is further reflected in their R^2, which are on average close to 0.75 for model 3 and 0.40 for model 4 versus 0.50 for model 1 and 0.26 for model 2.

C. Consensus Among Wine Experts

In this subsection, we concentrate on the consensus criterion put forward by Ashton (2012, 2013). As stated by Hodgson (2008), "good judges [should] agree with each other." In the case of wine, an objective measure of quality does not exist which makes consensus an essential criterion to distinguish good from average experts.

In order to examine the degree of agreement among the various wine experts, we compute the correlations between their scores and also between their residual scores of models 1 and 2. By definition, a residual is the part of an observation that a model cannot explain and can therefore be interpreted as a surprise. In the present context, a positive residual would thus indicate that the expert has been positively surprised by the quality of a wine. Residuals from model 1 capture the score surprises, while those from model 2 capture both these surprises and the preferences (in terms of appellation and ranking) of each expert compared to the others. The corresponding correlation coefficients are reported in Table 6. We use Spearman's rank correlations, as the distribution of scores is not normal and varies from one expert to another.[12] Panel A indicates correlation coefficients of expert scores. Panel B uses the residuals of equation 1 (lower triangle) and of equation 2 (upper triangle).

All correlation coefficients in Panel A are positive and statistically significant. The average correlation is close to 0.60. This indicates a strong consensus among all experts on overall wine scores. U.S. experts seem to agree more with one another than European critics do. James Suckling has the highest correlation with other critics. Jancis Robinson and the *Revue des vins de France*, by contrast, are much less correlated with their colleagues.

The significantly positive correlation coefficients in the upper triangle of Panel B further demonstrate that experts generally share similar preferences in terms of appellations, ranking categories and vintages. These correlations are, however, lower than in Panel A, which suggests that subtle differences in preferences among the various experts nevertheless exist. This is in line with the results in the previous subsection. James Suckling and Stephen Tanzer are the critics whose preferences are the most correlated with the others. Jancis Robinson exhibits the most idiosyncratic preferences.

[12] As a robustness test, we redo the calculations using Pearson's correlation. Results remain qualitatively similar.

Table 6
Correlations Among Expert Scores

Panel A: Correlation of scores

	FV	*BD*	*DEC*	*GJE*	*JMQ*	*JR*	*JS*	*NM/WJ*	*RP*	*RVF*	*ST*	***WS***
FV		0.67***	0.62***	0.63***	0.67***	0.55***	0.72***	0.61***	0.61***	0.54***	0.67***	0.69***
BD	0.67***		0.58***	0.68***	0.64***	0.53***	0.67***	0.60***	0.60***	0.54***	0.65***	0.65***
DEC	0.62***	0.58***		0.57***	0.59***	0.54***	0.74***	0.50***	0.62***	0.57***	0.59***	0.50***
GJE	0.63***	0.68***	0.57***		0.64***	0.50***	0.77***	0.68***	0.59***	0.54***	0.64***	0.58***
JMQ	0.67***	0.64***	0.59***	0.64***		0.51***	0.68***	0.59***	0.64***	0.54***	0.68***	0.59***
JR	0.55***	0.53***	0.54***	0.50***	0.51***		0.48***	0.49***	0.44***	0.38***	0.52***	0.50***
JS	0.72***	0.67***	0.74***	0.77***	0.68***	0.48***		0.73***	0.76***	0.75***	0.64***	0.80***
NM/WJ	0.61***	0.60***	0.50***	0.68***	0.59***	0.49***	0.73***		0.57***	0.43***	0.60***	0.55***
RP	0.61***	0.6***	0.62***	0.59***	0.64***	0.44***	0.76***	0.57***		0.56***	0.72***	0.67***
RVF	0.54***	0.54***	0.57***	0.54***	0.54***	0.38***	0.75***	0.43***	0.56***		0.57***	0.52***
ST	0.67***	0.65***	0.59***	0.64***	0.68***	0.52***	0.64***	0.60***	0.72***	0.57***		0.68***
WS	0.69***	0.65***	0.50***	0.58***	0.59***	0.50***	0.80***	0.55***	0.67***	0.52***	0.68***	

Panel B: Correlation on residuals

	FV	*BD*	*DEC*	*GJE*	*JMQ*	*JR*	*JS*	*NM/WJ*	*RP*	*RVF*	*ST*	*WS*
FV		0.43***	0.50***	0.49***	0.45***	0.39***	0.51***	0.41***	0.37***	0.43***	0.45***	0.48***
BD	0.34***		0.47***	0.49***	0.43***	0.34***	0.46***	0.36***	0.39***	0.41***	0.48***	0.44***
DEC	0.40***	0.33***		0.37***	0.45***	0.38***	0.56***	0.41***	0.44***	0.41***	0.49***	0.43***
GJE	0.36***	0.44***	0.37***		0.44***	0.37***	0.53***	0.45***	0.40***	0.45***	0.49***	0.50***
JMQ	0.36***	0.40***	0.35***	0.34***		0.36***	0.52***	0.41***	0.42***	0.39***	0.51***	0.37***
JR	0.27***	0.33***	0.33***	0.20***	0.35***		0.34***	0.32***	0.26***	0.26***	0.35***	0.34***
JS	0.44***	0.33***	0.49***	0.59***	0.36***	0.11		0.48***	0.50***	0.56***	0.47***	0.54***
NM/WJ	0.31***	0.38***	0.28***	0.43***	0.38***	0.33***	0.41***		0.34***	0.37***	0.41***	0.39***
RP	0.23***	0.34***	0.38***	0.28***	0.39***	0.25***	0.43***	0.33***		0.41***	0.55***	0.47***
RVF	0.26***	0.31***	0.35***	0.26***	0.32***	0.20***	0.50***	0.23***	0.36***		0.46***	0.39***
ST	0.44***	0.40***	0.41***	0.40***	0.44***	0.24***	0.42***	0.33***	0.46***	0.33***		0.51***
WS	0.44***	0.36***	0.24***	0.30***	0.28***	0.21***	0.55***	0.28***	0.37***	0.26***	0.48***	

Panel A indicates correlation coefficients of expert scores.

Panel B uses the residuals of equation 1 (lower triangle) and of equation 2 (upper triangle).

***, **, and * denote significance respectively at the 1%, 5%, and 10% level.

The lower triangle of Panel B shows that the divergence among critics on unexpected quality is more pronounced, as the correlation coefficients calculated on the basis of score surprises are much lower than those obtained earlier. This indicates that score surprises differ substantially among the various critics. Thus, while experts appear to reach a relatively strong degree of consensus on overall wine quality, they tend to have more diverse opinions on wines that achieve a surprising level of quality given the vintage, the ranking, or the appellation from which they originate. As before, Jancis Robinson appears to be the most independent and least correlated with respect to her colleagues while James Suckling and Stephen Tanzer display the highest correlations.

V. Expert Scores and *En Primeur* Prices

In the previous section, we showed that a few key variables explain a substantial part of the variation in expert scores. We now focus on the relationship between expert scores and *en primeur* prices. Variables such as appellation, ranking, and vintage account for more than 65% of the variation in *en primeur* log prices. If we add individual producer information, this figure increases to 95%.[13] These variables capture a very large proportion of the variation in both expert scores and prices, as they are intrinsically related to climate, natural endowments, vineyards, and winemaking techniques, or, in sum, to *quality*. Moreover, the fact that these variables appear to explain prices better than scores is due to the prevalence of status and tradition in Bordeaux: between two wines of similar quality, the one from the more renowned appellation or with the higher ranking will sell for a higher price.[14]

A. Methodology

Even if a substantial part of quality can be captured by a limited set of observable variables, expert scores may still be informative and potentially exercise influence on prices. That is, part of the quality of a wine is not directly observable and known to the general public. In any vintage, some wines achieve a surprisingly high (or low) quality level that cannot be fully explained by observable variables. Presumably, only the wine producer and some critics who have tasted the wine should be capable of fully appraising its quality. Ratings of trustworthy experts should therefore reflect the overall quality of a wine and contain information related to both observable and unobservable quality attributes.

A common specification used in literature regresses wine prices on a set of explanatory variables, including expert ratings (see Haeger and Storchmann, 2006). In the present context, we use the same explanatory variables as in model 3 to control for observable quality attributes and propose a model that relates wine prices to the

[13] These figures are based on two sets of linear regressions: appellation, ranking, and vintage dummies are used as explanatory variables in the first set, while vintage and wine producer dummies are used in the second.
[14] See Hay (2010) for a discussion of status and its influence on Bordeaux prices.

vintage, the estates, and expert ratings:

$$p_{i,t} = \theta_0^{(j)} + \sum_{v=2003}^{2012} \theta_v^{(j)} D_{\{vint_t=v\}} + \sum_{k=1}^{122} \theta_k^{(j)} D_{\{ID_i=k\}} + \gamma^{(j)} Score_{i,t}^{(j)} + \varepsilon_{i,t}^{(j)} \quad (5)$$

where $p_{i,t} = log\,(P_{i,t})$. For low correlations between the scores and the exogenous variables (the dummies for wine producers and vintages), the coefficients $\theta^{(j)}$ should be similar for all experts. In this article, we focus on a modified version of equation 5. Instead of focusing on prices, we analyze their evolution. We define the percentage change in price (the log-return) between year *t-1* and *t* as $r_{i,t} = p_{i,t} - p_{i,t-1}$. By taking the first difference of equation 5, we obtain the following specification:

$$r_{i,t} = \beta_0^{(j)} + \sum_{v=2004}^{2012} \beta_v^{(j)} D_{\{vint_t=v\}} + \gamma^{(j)} \Delta Score_{i,t}^{(j)} + \varepsilon_{i,t}^{(j)} \quad (6)$$

where $\Delta Score_{i,t}^{(j)} = Score_{i,t}^{(j)} - Score_{i,t-1}^{(j)}$. To ensure that the β coefficients capture the average return between each pair of vintages, we further define $\Delta^s Score_{i,t}^{(j)}$ as $\Delta Score_{i,t}^{(j)} - \frac{1}{N}\sum_{i=1}^{N} \Delta Score_{i,t}^{(j)}$ and substitute it into equation 6 to obtain:

$$r_{i,t} = \beta_0^{(j)} + \sum_{v=2004}^{2012} \beta_v^{(j)} D_{\{vint_t=v\}} + \gamma^{(j)} \Delta^s Score_{i,t}^{(j)} + \varepsilon_{i,t}^{(j)}. \quad (7)$$

Equation 7, while conceptually similar to equation 5, displays advantages. First, $\Delta^s Score_{i,t}^{(j)}$ captures the part of expert ratings that is not related to observable variables (e.g., wine producer or vintage), which is the truly informative and valuable part of expert scores. This can also be considered a surprise in quality, estimated by a particular expert. We therefore refer henceforth to this variable as the surprise in score. If the expert is correct or if his scores have an influence on customer choices, such surprises should correlate positively with prices. Second, this approach reduces multicollinearity[15] issues, as it uses a more parsimonious set of explanatory variables, which are uncorrelated with the surprise in score. As such, the coefficients in model 7 are not only easier to interpret but also estimated in a more robust manner than those in model 5.

A positive relation between $\Delta^s Score_{i,t}^{(j)}$ and $r_{i,t}$ does not necessarily imply that ratings influence prices. It may simply indicate that experts, as wine producers, are able to identify good and bad wines and to score, and respectively price, them accordingly. Prices and scores are both related to quality and may therefore lead to potential endogeneity problems that can be solved using one of the following two methods. First, equation 7 can be estimated using an instrumental variable technique. Instruments used in the existing literature are climate-related variables (see Cardebat et al., forthcoming; Haeger and Storchmann, 2006) or the lag of the

[15] In section 4, we show that up to 80% of the variation in a particular expert score can be explained by a set of dummies for wine producers and vintages. Thus, coefficients estimated on the basis of specification 5 are likely to be unstable because of multicollinearity problems.

potentially endogenous variable (see Malter, 2014). In the present case, it is not possible to find an appropriate instrument as the variable, surprise in score, cannot be anticipated or instrumentalized by definition. A second approach is based on mediation theory (see Baron and Kenny, 1986; Hayes, 2013), which is widely used in human and social sciences. The underlying idea is to introduce a variable that controls for the surprise in quality into equation 7 to obtain:

$$r_{i,t} = \beta_0^{(j)} + \sum_{v=2004}^{2012} \beta_v^{(j)} vint_t + \gamma^{(j)} \Delta Score_{i,t}^{(j)} + \lambda^{(j)} x_{i,t} + \varepsilon_{i,t}^{(j)} \quad (8)$$

$x_{i,t}$ serves as a proxy for the effective surprise in quality. In order to estimate it, we exploit the fact that our dataset contains scores of the most acknowledged experts in the world. As talented as an individual expert can be, he may nevertheless sometimes be overly optimistic (or pessimistic) about a wine, thereby leading to a positive (or negative) surprise in score for this wine. But if a wine receives positive surprises-in-scores from most experts, this suggests that its quality is surprisingly good. We therefore use the average surprises in scores as an estimate of $x_{i,t}$. This average should not only be less noisy but, more importantly, also more informative than individual surprises in scores about the effective quality surprise. We, however, do not invoke the law of large numbers and do not argue that this average is equal to the surprise in quality. We only assume that it can serve as a proxy for it. Models 7 and 8 are further estimated using OLS and robust standard errors (see Arellano, 1987).

B. Main Results

Table 7 reports results on the relationship between surprises in scores and price changes following model 7. The different vintage coefficients give indications on the price evolution over the 2004–2012 vintages. Prices have strongly increased for the 2005 and 2009 vintages. In 2010 they increased slightly and, for all other vintages, they decreased. The last row displays the effect of the various experts on price changes. The surprise in score of eight critics significantly correlates with price changes while four critics (GJE, JR, NM, and RVF) do not appear to have an influence on prices. Robert Parker and Jean-Marc Quarin are the most influential experts, with a 10% increase in their scores leading to about 7% higher prices. These results confirm previous results on the strong influence that Robert Parker has on Bordeaux prices. Jean-Marc Quarin, though less known, has in-depth knowledge of the Bordeaux market, a region in which he lives, and is widely regarded as one of the most authoritative expert on Bordeaux wines in France. For the remaining six critics, prices go up by around 2% to 4.5% for a 10% surprise in score. The last column of Table 7 reports the results with the average surprise in score of all experts used as an explanatory variable in model 7. Results suggest that the average surprise has a stronger impact on price than any individual surprise, as a 10% higher average score leads to a 10% increase in prices. This is consistent with our hypothesis that the average surprise in score can be used as a proxy for the unobservable surprise in quality. When quality is especially high, producers will be

Table 7
Score Surprises and Price Changes

	FV	*BD*	*DEC*	*GJE*	*JMQ*	*JR*	*JS*	*NM/ WJ*	*RP*	*RVF*	*ST*	*WS*	*Average*
Intercept	0.15***	0.12***	0.14***	0.13***	0.14***	0.13***	−0.39***	0.13***	0.14***	0.11***	0.14***	0.14***	0.13***
2004	−0.38***	—	−0.32***	—	−0.31***	−0.30***	—	−0.32***	−0.37***	−0.30***	−0.32***	−0.33***	−0.31***
2005	0.60***	0.51***	0.49***	0.48***	0.49***	0.49***	—	0.50***	0.67***	0.55***	0.51***	0.51***	0.48***
2006	−0.30***	−0.24***	−0.25***	−0.24***	−0.25***	−0.25***	—	−0.24***	−0.26***	−0.24***	−0.26***	−0.26***	−0.25***
2007	−0.30***	−0.25***	−0.26***	−0.25***	−0.26***	−0.26***	—	−0.26***	−0.29***	−0.23***	−0.27***	−0.27***	−0.25***
2008	−0.32***	−0.24***	−0.25***	−0.25***	−0.26***	−0.25***	—	−0.25***	−0.25***	−0.23***	−0.27***	−0.27***	−0.25***
2009	0.66***	0.56***	0.53***	0.53***	0.54***	0.53***	—	0.52***	0.54***	0.56***	0.54***	0.54***	0.53***
2011	−0.53***	−0.49***	−0.50***	—	−0.50***	−0.51***	—	−0.50***	−0.52***	−0.48***	−0.54***	−0.51***	−0.50***
2012	−0.26***	−0.23***	−0.25***	—	−0.25***	−0.25***	0.27***	−0.24***	−0.25***	−0.23***	−0.26***	−0.26***	−0.25***
Score surprise (γ)	0.0045***	0.0024***	0.0019***	−0.0005	0.0067***	−0.0002	0.0041***	−0.0002	0.0071***	0.0004	0.0035***	0.0036***	0.0093***
Observations	674	844	1017	693	1004	1037	193	938	820	915	907	933	1081
Adj. R^2	0.75	0.69	0.68	0.65	0.70	0.68	0.35	0.68	0.72	0.68	0.71	0.70	0.70

Notes: Column "Average" reports the results for the surprises-in-score averaged over all experts.

2004 to 2012 denote the respective vintages with vintage 2010 as reference year.

— indicates that the respective expert has not tasted the specific vintage.

***, **, * shows significance at the 1%, 5%, and 10% level respectively.

inclined to increase their prices to send a signal to the market and presumably also because they believe their wine deserves to be sold at a higher price.

In order to determine whether the positive relation between surprises in scores and price changes is due to a genuine score effect or to a quality effect, we resort to model 8. The purpose of this model is to distinguish the influence of an individual expert (captured by γ in model 8) and of quality (λ). Results are reported in Table 8. Coefficients attached to each vintage are again very similar and in line with the results in Table 7. Consistent with the results in model 7, λ is always positive and statistically significant. Controlling for surprises in quality, we find that only four critics have a positive and significant influence on prices. The two most influential critics are Jean-Marc Quarin and Robert Parker, who lead to a 5% price increase per 10% surprise in score. They are followed by Stephen Tanzer and *Wine Spectator,* which lead to 1.75% higher prices for a 10% surprise in score.

In order to further refine our findings, we re-estimate model 8 and allow γ to be different, depending on the specific region, appellation, vintage, or ranking considered. The results indicate that in general few experts exert genuine influence on prices of wines from the left bank (Saint-Estèphe, Pauillac, Saint-Julien, Margaux, and Haut-Médoc). The situation is quite different, however, on the right bank (Saint-Emilion and Pomerol) and in Pessac-Léognan, as, in these areas, Robert Parker and Jean-Marc Quarin strongly affect prices. *Wine Spectator* also has a significant, albeit more limited, impact on the prices of wines from these appellations.

Wines included in the 1855 classification are generally not very sensitive to surprises in scores. By contrast, unclassified wines and those from Saint-Emilion (where the classification is regularly updated) show a strong relation between surprises in scores and prices. A most pronounced effect is once again found for Robert Parker and Jean-Marc Quarin and, to a lesser extent, for Stephen Tanzer and *Wine Spectator*. These results suggest that in appellations such as Pomerol, Saint-Emilion, and Pessac-Léognan and for unclassified wines in general, expert scores may substitute for the presence of a long-standing and rigid classification system, such as the 1855 classification that exists on the left bank, and may therefore affect prices more heavily.

Finally, the relation between surprises in scores and prices is dependent on the vintage. The effect is most pronounced in the best vintages (2005, 2009, and 2010). It thus appears that producers justify part of their price increases on the basis of the scores and comments of experts. Robert Parker is the expert who has an effect most often, as only in 2008 and 2012 does he fail to display significant influence. The year 2008 constitutes an interesting context and explains the finding obtained. For that vintage, many producers started selling their wines before Parker published his scores, and thus it is not surprising that his scores do not have an impact on prices.

In the same vein, it is possible that the relationship between surprises in scores and price changes is nonlinear. Very positive surprises in scores could have a larger

Table 8

Average Surprises, Score Surprises and Price Changes

	FV	*BD*	*DEC*	*GJE*	*JMQ*	*JR*	*JS*	*NM/WJ*	*RP*	*RVF*	*ST*	*WS*
Intercept	0.15***	0.12***	0.13***	0.13***	0.14***	0.13***	−0.39***	0.13***	0.14***	0.11***	0.14***	0.14***
2004	−0.37***	—	−0.30***	—	−0.31***	−0.30***	—	−0.30***	−0.37***	−0.30***	−0.33***	−0.33***
2005	0.60***	0.51***	0.49***	0.49***	0.49***	0.49***	—	0.50***	0.66***	0.55***	0.50***	0.51***
2006	−0.29***	−0.25***	−0.25***	−0.25***	−0.25***	−0.25***	—	−0.24***	−0.26***	−0.24***	−0.26***	−0.25***
2007	−0.30***	−0.26***	−0.26***	−0.25***	−0.26***	−0.25***	—	−0.25***	−0.29***	−0.24***	−0.27***	−0.27***
2008	−0.31***	−0.24***	−0.25***	−0.25***	−0.26***	−0.25***	—	−0.24***	−0.26***	−0.24***	−0.27***	−0.27***
2009	0.65***	0.55***	0.53***	0.53***	0.54***	0.53***	—	0.53***	0.54***	0.56***	0.53***	0.54***
2011	−0.53***	−0.50***	−0.50***	—	−0.50***	−0.51***	—	−0.49***	−0.52***	−0.49***	−0.54***	−0.51***
2012	−0.26***	−0.23***	−0.25***	—	−0.26***	−0.25***	0.27***	−0.24***	−0.26***	−0.23***	−0.27***	−0.27***
Average surprise (λ)	0.0093***	0.0136***	0.0107***	0.0147***	0.0061***	0.0116***	0.0104*	0.0138***	0.0073***	0.0121***	0.008***	0.0086***
Individual surprise (γ)	0.0022*	0.0000	−0.0001	−0.002***	0.0051***	−0.0024***	0.0017	−0.0027***	0.0049***	−0.0018**	0.0017**	0.0018**
Observations	674	844	1017	693	1004	1037	193	938	820	915	907	933
Adj. R^2	0.76	0.71	0.70	0.68	0.71	0.70	0.37	0.71	0.72	0.70	0.71	0.71

Notes: 2004 to 2012 denote the respective vintages with vintage 2010 as the reference year.

— indicates that the respective expert has not tasted the specific vintage

***, **, * shows significance at the 1%, 5%, and 10% level respectively.

impact on prices than negative surprises. We therefore re-estimate our model to account for the possibility of a nonlinear relation. Adding a quadratic term or a dummy taking the value 1 for positive surprises in scores to our model does not alter our results. Only Jean-Marc Quarin displays some slight nonlinearity, but the economic impact on the score-price relation is nearly insignificant.

Finally, we run a regression in which we include only the four critics (JMQ, RP, ST, WS) that have been found to influence wine prices significantly. This allows a comparison of their relative influence using a sample of wines that they have all rated. Once again, Robert Parker and Jean-Marc Quarin show very strong influence. Their influence remains stable compared to their respective univariate regressions, indicating that their scores are complementary. This suggests that Robert Parker may reflect the taste of American customers and Jean-Marc Quarin that of French consumers. *Wine Spectator*, by contrast, loses influence, and the impact for Stephen Tanzer turns insignificant. This tends to indicate that the two other critics dominate them.

Overall, our results suggest that Jean-Marc Quarin and Robert Parker are the two most influential experts on the Bordeaux market. Both appear to be complementary, catering to the French and U.S. markets, respectively confirming results by Cardebat and Livat (2013).

C. Robustness Tests

To verify the robustness of the results obtained in the previous section, we perform a series of tests.[16]

1. Model misspecification: our results might be affected by the fact that we examine the relation between *differences* in scores and log prices instead of using these variables expressed in levels. To ensure that our results are robust to the particular specification used, we re-estimate them using equation 6 (levels) instead of equation 8 (differences). The results remain consistent, with Robert Parker and Jean-Marc Quarin still having a positive and significant effect on prices. It should be noted, however, that the coefficients associated with the producer and vintage dummies are less stable due to multicollinearity problems.
2. Definition of the surprise in score variable: the definition we use to estimate the surprises in scores may also influence the results. We consequently construct an additional surprise in score measure using the residuals of model 3. To account for the fact that the surprises are estimated in a first regression, we estimate the coefficients of the second regression and their standard errors using a bootstrapping approach. The results, and, in particular, the relation between surprises in scores and returns remain similar to those obtained with our initial definition of the surprise in score variable.

[16] Detailed results are not reported here but are available from the authors upon request.

3. Definition of the surprise in quality variable: our results could also be affected by the proxy used to control for the unobservable surprise in quality. We therefore re-estimate model 8 using two additional proxies. We first use the median surprise in score instead of the average. We then use a common factor extracted from all surprises in scores using a principal component analysis (PCA). Replacing the mean surprise in score by either proxy does not affect the results (RP, JMQ, and, to a lesser extent, WS and ST still have a significant influence on prices). However, the relation between surprises in quality and changes in log-price becomes stronger when a proxy based on a PCA is used. This suggests that this proxy is more closely related to surprise in quality than the proxy based on the average surprise in score.
4. Influence of extreme observations: some experts may award extremely high or low scores. We examine whether such scores have an impact on our results by either discarding or winsorizing extreme values at the 5% and 10% level and re-estimating our different specifications. Overall, these two procedures do not alter our results (RP, JMQ, ST, and WS are the only experts that influence prices in both settings). More precisely, (i) when extreme surprises are first winsorized, the relation between surprises and price changes becomes slightly stronger (for all experts), but (ii) when extreme surprises are discarded, the influence of RP and ST increases, while JMQ and WS see their influence on prices decrease.
5. Sample size and missing values: we analyze the impact our sample and its definition has on results. In a first test, we restrict our sample and consider only the wines that have been rated by the four experts who have a significantly positive influence on prices (WS, JMQ, RP, and ST). The results are essentially unchanged. In a second test, we assess if the fact that four experts (DEC, GJE, NM, and RVF) have changed their rating scale during the sample period alters the results of the analysis. We thus re-estimate the regressions for these four experts using a restricted sample that only includes the data based on the scoring system each expert currently uses. In all cases, the results remain qualitatively similar.

VI. Conclusion

In this article, we provide an empirical analysis of 12 international wine experts on the market for Bordeaux wines over the period 2003–2012. We first analyze their rating approach and their influence on *en primeur* prices. We illustrate that score comparisons between various experts are difficult, as they use different approaches. European critics are less transparent and use scoring systems that are often more difficult to interpret than their American counterparts. Scores generally follow complicated distributions, and the assumption of normality has to be rejected for all critics, except for the Grand Jury Européen and *Decanter*. Over the past decade, several critics changed their rating scale. We show that this change has a direct influence on the score distribution, with the *Revue des vins de France* looking more severe than it used to be, and the Grand Jury Européen appearing to have

become more generous. In general, European critics appear to be more severe than their American counterparts.

All critics demonstrate similar preferences in terms of ranking and appellation, but the amplitude of those preferences varies from one critic to another. In particular, American critics generally prefer Pessac-Léognan and Saint-Emilion to Pauillac, while European critics are more indecisive. Experts also reach a relatively strong consensus on overall wine quality, but have more diverse opinions on wines that achieve a surprising level of quality given the vintage, the ranking, or the appellation from which they originate.

Wine experts play an important role on the wine market by delivering valuable information for an experience good for which true quality cannot be precisely determined. We use a mediation technique as proposed in Baron and Kenny (1986) to discriminate between the quality surprise in wines and pure surprises in scores by critics that measure the true influence a critique has on wine prices. We finally estimate different alternative specifications to test the robustness of our results.

Compared to most studies that focus on Robert Parker, we explicitly contrast and compare his influence with other critics for a sample of 122 Bordeaux estates. Our evidence suggests that Stephen Tanzer and *Wine Spectator* display significantly positive albeit limited influence on prices. By contrast, Robert Parker and Jean-Marc Quarin are the most influential critics, as a 10% surprise in their scores leads to a price increase of around 7%. We further find that Robert Parker and Jean-Marc Quarin are especially influential under certain conditions. Their impact is higher for appellations and estates that are not covered by the official 1855 classification, and for the best vintages (2005, 2009, 2010). Overall, our results suggest that Jean-Marc Quarin and Robert Parker are the two most influential experts on the Bordeaux market and appear to be complementary in catering to the French and U.S. markets, respectively.

References

Ali, H.H., and Nauges, C. (2007). The pricing of experience goods: The example of en primeur wine. *American Journal of Agricultural Economics*, 89(1), 91–103.

Ali, H.H., Lecocq, S., and Visser, M. (2008). The impact of gurus: Parker grades and en primeur wine prices. *Economic Journal*, 118(529), F158–F173.

Arellano, M. (1987). Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics*, 49(4), 431–434.

Ashenfelter, O. (2010). Predicting the quality and prices of Bordeaux wine. *Journal of Wine Economics*, 5(1), 40–52.

Ashenfelter, O., and Jones, G.V. (2013). The demand for expert opinion: Bordeaux wine. *Journal of Wine Economics*, 8(3), 285–293.

Ashenfelter, O., Ashmore, D., and Lalonde, R. (1995). Bordeaux wine vintage quality and the weather. *Chance*, 8(4), 7–14.

Ashton, R.H. (2012). Reliability and consensus of experienced wine judges: Expertise within and between? *Journal of Wine Economics*, 7(1), 70–87.

Ashton, R.H. (2013). Is there consensus among wine quality ratings of prominent critics? An empirical analysis of red Bordeaux, 2004–2010. *Journal of Wine Economics*, 8(2), 225–234.

Bagwell, K., and Riordan, M.H. (1991). High and declining prices signal product quality. *American Economic Review*, 81(1), 224–239.

Baron, R.M., and Kenny, D.A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182.

Berger, J., Sorensen, A.T., and Rasmussen, S.J. (2010). Positive effects of negative publicity: When negative reviews increase sales. *Marketing Science*, 29(5), 815–827.

Cao, J., and Stokes, L. (2010). Evaluation of wine judge performance through three characteristics: Bias, discrimination, and variation. *Journal of Wine Economics*, 5(1), 132–142.

Cardebat, J.-M., and Livat, F. (2013). *Wine experts rating: A matter of taste?* Paper presented at the 4th I-CHLAR Conference, Lausanne, Switzerland.

Cardebat, J.-M., Figuet, J.-M., and Paroissien, E. (2014). Expert opinion and Bordeaux wine prices: An attempt to correct bias in subjective judgements. *Journal of Wine Economics*, 9(3), 282–303.

Dubois, P., and Nauges, C. (2010). Identifying the effect of unobserved quality and expert reviews in the pricing of experience goods: Empirical application on Bordeaux wine. *International Journal of Industrial Organization*, 28, 205–212.

Friberg, R., and Grönqvist, E. (2012). Do expert reviews affect the demand for wine? *American Economic Journal: Applied Economics*, 4(1), 193–211.

Gergaud, O., and Livat, F. (2007). How do consumers use signals to assess quality? American Association of Wine Economists, AAWE Working Paper No. 3.

Gergaud, O., Storchmann, K., and Verardi, V. (2015). Expert opinion and product quality: Evidence from New York City restaurants. *Economic Inquiry*, 53(2), 812–85.

Gluckman, R.L. (1990). A consumer approach to branded wines. *European Journal of Marketing*, 24(4), 27–46.

Haeger, J.W., and Storchmann, K. (2006). Prices of American Pinot Noir wines: Climate, craftsmanship, critics. *Agricultural Economics*, 35(1), 67–78.

Hay, C. (2010). The political economy of price and status formation in the Bordeaux *en primeur* market: The role of wine critics as rating agencies. *Socio-Economic Review*, 8(4), 685–707.

Hayes, A.F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: Guilford Press.

Hilger, J., Rafert, G., and Villas-Boas, S. (2011). Expert opinion and the demand for experience goods: An experimental approach in the retail wine market. *Review of Economics and Statistics*, 93(4), 1289–1296.

Hodgson, R. (2008). An examination of judge reliability at a major U.S. wine competition. J*ournal of Wine Economics*, 3(2), 105–113.

Jones, G.V., and Storchmann, K.H. (2001). Wine market prices and investment under uncertainty: An econometric model for Bordeaux crus classés. *Agricultural Economics*, 26(2), 115–133.

Kirmani, A., and Rao, A.R. (2000). No pain, no gain: A critical review of the literature on signaling unobservable product quality. *Journal of Marketing*, 64(2), 66–79.

Lockshin, L.S., and Rhodus, W.T. (1993). The effect of price and oak flavor on perceived wine quality. *International Journal of Wine Marketing*, 5(2/3), 13–25.

Mahenc, P., and Meunier, V. (2006). Early sales of Bordeaux grands crus. *Journal of Wine Economics*, 1(1), 57–74.

Malter, D. (2014). On the causality and cause of returns to organizational status evidence from the grands crus classés of the Médoc. *Administrative Science Quarterly*, 59(2), 271–300.

Orth, U.R., and Krška, P. (2001). Quality signals in wine marketing: The role of exhibition awards. *International Food and Agribusiness Management Review*, 4(4), 385–397.

Ramirez, C.D. (2010). Do tasting notes add value? Evidence from Napa wines. *Journal of Wine Economics*, 5(1), 143–163.

Reinstein, D.A., and Snyder, C.M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *Journal of Industrial Economics*, 53(1), 27–51.

Roberts, P.W., and Reagans, R. (2007). Critical exposure and price-quality relationships for new world wines in the U.S. market. *Journal of Wine Economics*, 2(1), 84–97.

Rosen, M. (2012). Critics and pricing: Parker's nose is not as big a pointer as it was. *Financial Times*, London, June 19, http://www.ft.com/cms/s/0/f29748d6-9dc6-11e1-9a9e-00144feabdc0.html#axzz3SruzS4Y8.

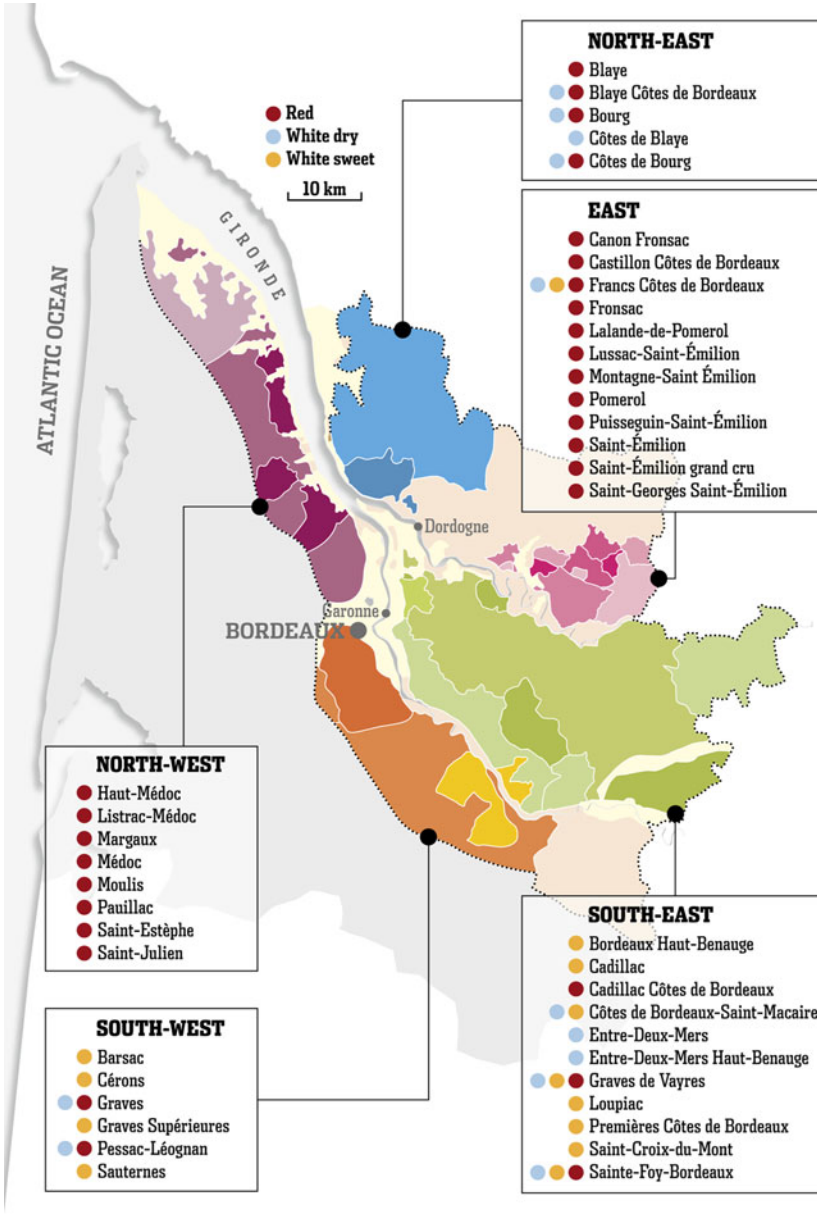
San Martín, G.J., Troncoso, J.L., and Brümmer, B. (2008). Determinants of Argentinean wine prices in the U.S. *Journal of Wine Economics*, 3(1), 72–84.

Schnabel, H., and Storchmann, K. (2010). Prices as quality signals: Evidence from the wine market. *Journal of Agricultural & Food Industrial Organization*, 8(1), article 2.

Storchmann, K. (2012). Wine economics. *Journal of Wine Economics*, 7(1), 1–33.

Vaillant, N.G., and Wolff, F.-C. (2011). Understanding how experts rate cigars: A "Havanometric" analysis. *Applied Economics*, 45(1), 99–109.

Appendix 1: Map of Bordeaux



Source: Map created by Ralf Powierski (www.info-graphic.de).

Appendix 2: List of Estates

Name	*Appellation*	*Rank*
Camensac	Haut Médoc	5GCC
Cantemerle	Haut Médoc	5GCC
La Lagune	Haut Médoc	3GCC
La Tour Carnet	Haut Médoc	4GCC
Alter Ego de Palmer	Margaux	
Brane-Cantenac	Margaux	2GCC
Cantenac Brown	Margaux	3GCC
Dauzac	Margaux	5GCC
d'Issan	Margaux	3GCC
du Tertre	Margaux	5GCC
Durfort Vivens	Margaux	2GCC
Ferrière	Margaux	3GCC
Giscours	Margaux	3GCC
Kirwan	Margaux	3GCC
Lascombes	Margaux	2GCC
Malescot Saint-Exupéry	Margaux	3GCC
Margaux	Margaux	1GCC
Marquis de Terme	Margaux	4GCC
Palmer	Margaux	3GCC
Pavillon Rouge de Margaux	Margaux	
Prieuré-Lichine	Margaux	4GCC
Rauzan Gassies	Margaux	2GCC
Rauzan-Ségla	Margaux	2GCC
Potensac	Médoc	CB Exc.
Chasse-Spleen	Moulis	CB Exc.
Poujeaux	Moulis	CB Exc.
Batailley	Pauillac	5GCC
Carruades de Lafite	Pauillac	
Clerc Milon	Pauillac	5GCC
Croizet-Bages	Pauillac	5GCC
d'Armailhac	Pauillac	5GCC
Duhart-Milon	Pauillac	4GCC
Grand-Puy Ducasse	Pauillac	5GCC
Grand-Puy-Lacoste	Pauillac	5GCC
Haut-Bages Libéral	Pauillac	5GCC
Haut-Batailley	Pauillac	5GCC
Lafite Rothschild	Pauillac	1GCC
Latour	Pauillac	1GCC
Le Petit Mouton de Mouton-Rothschild	Pauillac	
Les Forts de Latour	Pauillac	
Lynch-Bages	Pauillac	5GCC
Lynch-Moussas	Pauillac	5GCC
Mouton-Rothschild	Pauillac	1GCC
Pichon Baron	Pauillac	2GCC
Pichon Comtesse	Pauillac	2GCC
Pontet-Canet	Pauillac	5GCC

Continued

(**Continued**)

Name	*Appellation*	*Rank*
Domaine de Chevalier	Pessac-Léognan	
Domaine de Chevalier blanc	Pessac-Léognan	
Haut-Bailly	Pessac-Léognan	
Haut-Bergey	Pessac-Léognan	
Haut-Brion	Pessac-Léognan	1GCC
Haut-Brion blanc	Pessac-Léognan	
La Chapelle de La Mission Haut-Brion	Pessac-Léognan	
La Mission Haut-Brion	Pessac-Léognan	
La Mission Haut-Brion blanc (formerly known as Laville Haut-Brion)	Pessac-Léognan	
Les Carmes Haut-Brion	Pessac-Léognan	
Malartic Lagravière	Pessac-Léognan	
Malartic Lagravière blanc	Pessac-Léognan	
Pape Clément	Pessac-Léognan	
Pape Clément blanc	Pessac-Léognan	
Smith Haut Lafitte	Pessac-Léognan	
Smith Haut Lafitte blanc	Pessac-Léognan	
Certan de May de Certan	Pomerol	
Clinet	Pomerol	
Clos l'Eglise	Pomerol	
Gazin	Pomerol	
Hosanna	Pomerol	
La Conseillante	Pomerol	
La Fleur de Gay	Pomerol	
La Fleur Petrus	Pomerol	
Lafleur	Pomerol	
Latour à Pomerol	Pomerol	
Le Bon Pasteur	Pomerol	
L'Eglise-Clinet	Pomerol	
L'Evangile	Pomerol	
Petit-Village	Pomerol	
Petrus	Pomerol	
Trotanoy	Pomerol	
Vieux Château Certan	Pomerol	
Angélus	Saint-Emilion	1GCC A
Ausone	Saint-Emilion	1GCC A
Beau-Séjour Bécot	Saint-Emilion	1GCC B
Beauséjour Duffau-Lagarrosse	Saint-Emilion	1GCC B
Canon	Saint-Emilion	1GCC B
Canon La Gaffelière	Saint-Emilion	1GCC B
Cheval Blanc	Saint-Emilion	1GCC A
Clos Fourtet	Saint-Emilion	1GCC B
Figeac	Saint-Emilion	1GCC B
La Chapelle d'Ausone	Saint-Emilion	
La Gaffelière	Saint-Emilion	1GCC B
La Mondotte	Saint-Emilion	1GCC B
Larcis Ducasse	Saint-Emilion	1GCC B

Continued

(Continued)

Name	*Appellation*	*Rank*
Le Petit Cheval	Saint-Emilion	
Magdelaine	Saint-Emilion	1GCC B
Monbousquet	Saint-Emilion	GCC
Pavie	Saint-Emilion	1GCC A
Pavie-Decesse	Saint-Emilion	GCC
Pavie-Macquin	Saint-Emilion	1GCC B
Troplong Mondot	Saint-Emilion	1GCC B
Trottevieille	Saint-Emilion	1GCC B
Valandraud	Saint-Emilion	1GCC B
Calon-Ségur	Saint-Estèphe	3GCC
Cos d'Estournel	Saint-Estèphe	2GCC
Cos Labory	Saint-Estèphe	5GCC
Haut-Marbuzet	Saint-Estèphe	CB Exc.
La Dame de Montrose	Saint-Estèphe	
Lafon-Rochet	Saint-Estèphe	4GCC
Les Ormes de Pez	Saint-Estèphe	CB Exc.
Montrose	Saint-Estèphe	2GCC
Phélan Ségur	Saint-Estèphe	CB Exc.
Beychevelle	Saint-Julien	4GCC
Branaire-Ducru	Saint-Julien	4GCC
Clos du Marquis	Saint-Julien	
Ducru-Beaucaillou	Saint-Julien	2GCC
Gruaud Larose	Saint-Julien	2GCC
La Croix de Beaucaillou	Saint-Julien	
Lagrange	Saint-Julien	3GCC
Langoa Barton	Saint-Julien	3GCC
Léoville Barton	Saint-Julien	2GCC
Léoville Las-Cases	Saint-Julien	2GCC
Léoville Poyferré	Saint-Julien	2GCC
Talbot	Saint-Julien	4GCC

Note: GCC stands for "Grand Cru Classé" and "CB Exc." stands for "Cru Bourgeois Exceptionnel." For instance, "1GCC A" corresponds to "Premier Grand Cru Classé A," i.e. "First Classified Growth A" in English.