INSEEC Research
PARIS • BORDEAUX • LYON • ALPES-SAVOIE
MONACO • LONDON • CHICAGO

Working Paper
n°: 2014-106-01

Expert Opinion and Bordeaux Wine Prices: An Attempt to Correct Bias in Subjective Judgments

Jean-Marie Cardebat [a,b], Jean-Marc Figuet [a], Emmanuel Paroissien [a]

a Larefi, Bordeaux University and Bordeaux Wine Economics
b INSEEC Business School, 27 avenue Claude Vellefaux 75010 Paris, France

March 2014

An ulterior version of this article appeared in Journal of Wine Economics, vol 9, Issue 3, December 2014, ISSN 1931-4361
It can be purchased at: http://journals.cambridge.org/action/displayJournal?jid=JWE

Abstract: The focus of the paper is to assess the role of expert opinion on Bordeaux wine prices. We construct an original dataset including retail prices, scores attributed to wines from France, Spain and United States by 19 experts from 2000 to 2010 and the corresponding meteorological conditions. This dataset aims to measure the bias rooted in subjectivity of the experts. We assess that the scores given by the experts have two components: an objective one driven by the fundamentals of wine quality (soils quality, producers' skills and weather conditions), and a subjective one due to the individual opinions of the experts. We use control function techniques to compare the respective impacts of these two components, finding that prices are mainly driven by the fundamentals. Furthermore, we notice the great impact of the maximum score on price, which we interpret as a "marketing effect": the most favorable score is likely to be the most publicized, and therefore the most influent.
Keywords: expert opinion, price, Bordeaux wine
JEL codes: C21, D89, L15

Expert Opinion and Bordeaux Wine Prices:
An Attempt to Correct Bias in Subjective Judgments

Whenever consumers have access to perfect information, the Bertrand Model indicates that the equilibrium price of goods and services is equal to its marginal cost. In practice, however, this result is seldom seen and price dispersion, due to incomplete information, can be observed. For consumers, finding information about products is a costly business. In the case of experience goods, the methods used to inform the public about the quality of products they might purchase are particularly important, since consumers are only able to determine their true quality once they have purchased and consumed them. Pioneering research in this field by Akerlof (1970) and Nelson (1970) established that information asymmetries pertaining to the quality of a product negatively influence consumer demand.

Brands (Montgomery and Wernerfelt, 1992), advertising (Ackerberg, 2003), quality labelling (Jin and Leslie, 2003), and expert endorsement (Salop, 1976) all constitute transmission channels to provide consumers with information about a product's quality. Although experts are found in a vast variety of domains—art, economics, weather forecasting, sport, gastronomy, cars, and electronic material—it is extremely difficult to assess their influence and the optimality of the opinions they express. Reinstein and Snyder (2005) concluded that movie reviews did not affect a film's box office earnings. Sorensen and Rasmunssen (2004) demonstrated that book reviews, whether favorable or unfavorable, led to boosted sales, thereby confirming the old adage "there is no such thing as bad publicity." Hilger et al. (2011) considered how experts' influence on demand is difficult to quantify. Indeed, empirical studies face a major methodological problem: high quality products obtain high scores since they are, in fact, of high quality. Thus, it becomes difficult to determine to what extent expert endorsements stimulate demand for them. Experts, according to Sinkey (2012), are not Bayesian because they accord too much weight to certain pieces of information and not enough to others, being just as subject to psychological bias as amateurs. For Budescu et al. (2003) consumer confidence in expert endorsement rose in relation to the number of experts involved, the convergence of their opinions, and the asymmetrical way in which product information was distributed. On the other hand, consumer confidence dropped when experts' conclusions diverged, leading to great variance in their ensuing scores. Monic (2007) insisted on the role played by variance: the average score of a product is certainly important, but this is equally true of variance. A high level of variance indicates that the experts do not agree, which complicates consumer choice. Recent papers by Hodgson (2008, 2009) questioned the consistency of expert wine judges in a wine competition setting and demonstrated that wine experts make mistakes. Lecocq and Visser (2006) and Ashton (2011) pointed out, however, how having the opinions of several experts reduced errors in judgment.

Bordeaux wine represents an experience good that requires a great deal of expertise in order to determine each wine's final quality and, hence, its price. This paper seeks to establish whether the experts, including the most renowned of them, Robert Parker, provide pertinent information for consumers, and whether these experts, by way of the scores they ascribe, influence wine prices.
For Ashenfelter (1989), the fallibility of Parker's judgment allows buyers to profit from his errors of judgment when wines are sold at auction. According to Ashenfelter, a wine's age, the average temperature from April to September, the rainfall in August and September and then from October to March, as well as vintage, are the main factors behind price variations. Ashenfelter and Jones (2000) considered how the hypothesis about the effective influence of experts' scores is rarely verified when it comes to predicting Bordeaux wine prices. The

scores provided no useful information on poor years and only correlated, at best, with good years. The experts overlook such key data as climate conditions even though these are extremely important for the ultimate quality of a wine. As detailed information about local weather conditions, in particular, is known privately to each individual château (Di Vittorio and Ginsburgh, 1994), the experts merely transmit publicly available information to the consumer. Ginsburgh et al. (1992), applying the hedonic pricing model to a sample of 102 Médoc wines, showed that expert ratings do not provide a better explanation for price than climate conditions, the 1855 classification, terroir, or production technique. Sixty-six percent of price variations could be explained by weather conditions or differences in vineyard practices. This percentage rose to 85% when the 1855 classification was taken into account. Di Vittorio and Ginsburgh (1994) came to the same conclusion. A hedonic function, calculated on the basis of the auction prices of 58 Médoc *crus classés*, indicated that the 1855 classification plays a greater role in explaining a wine's price than any alternative rating system drawn up by experts.

For Jones and Storchmann (2001), Parker's scores influence prices in a differentiated fashion. A rise of 1 point engendered a rise in price of between 4 and 10%, with an average increase of 7%. This result, obtained from prices for 21 prestigious Bordeaux wines, indicated that the sensitivity of a wine's price relative to Parker's scores is greater for wines made from Cabernet-Sauvignon than for those made from Merlot. Hilger et al. (2011), adopting a more experimental approach, also showed the impact of expert ratings. They analysed wine sales in a supermarket by choosing a random sample of 150 wines from 476 rated wines and displaying each wine's score on supermarket shelves. Sales of the selected wines increased by an average of 25%, and sales of those with the best scores increased more quickly than those with lower scores. This led them to conclude that the advertising surrounding expert endorsement produces a positive effect on global demand as it reduces information asymmetry. Storchmann et al. (2012) argued that expert opinion has a negative effect on the price dispersion of American wines tested by the *Wine Spectator* between 1984 and 2008. The authors showed that expert opinion distorts the relationship between quality and price, especially in the case of poor quality wines. Roma et al. (2013) constructed a hedonic price model to determine the variables influencing the prices in a sample of Sicilian wines. They concluded that price depends on traditional objective variables and sensorial variables as well as on the ratings published in specialized reviews. Using five years of data on expert opinions published in six Swedish periodicals, Friberg and Grönqvist (2012) showed how a positive review induced an increase in demand of 6% the week after publication. This positive effect then declined but was still significant 20 weeks later**.** A neutral expert opinion led to a small increase in demand, whereas a negative one had no effect.

The debate about the impact of expert opinion on price is even more complicated for Bordeaux wines. Bordeaux *crus classés* can be sold *en primeur* in the futures market six months after harvesting, and are only delivered to the purchaser two or three years later. This incurs a great deal of uncertainty concerning the wine's ultimate quality. It is the expert's role to ascertain that ultimate quality, which, consequently, influences the sale prices of *primeur* wines. Hadj Ali and Nauges (2007), using a sample of 108 châteaux for vintages from 1994 to 1998, showed that the price *en primeur* is determined chiefly by reputation. Parker's scorings have a significant but marginal effect—a rise of 1 point triggered a rise in price of 1.01%. Hadj Ali et al. (2008) measured the effect of Parker's scorings on *en primeur* prices by exploiting the fact that, in 2003, Parker's ratings came out after the wine producers had published their prices. In this case, the overall increase in price was estimated at 2.80 € per bottle. [M1]

Simply put, the role of experts in influencing the price of any wine remains uncertain and differs from one study to another. Wine is not homogeneous, but varies according to a set of characteristics. Some of these characteristics, color or grading, for example, are easy to measure inexpensively. Others, such as sensorial or taste characteristics, are difficult to measure before consumption. Expert opinions are purported to summarize quality characteristics of wine. These opinions may convey less information than a complete description of characteristics. In addition, these opinions may be imperfect because they may fail to capture the quality experienced by consumers, for example. This imperfection may have implications for prices and consumer welfare.

The present research aims at further exploring the question of the impact of expert opinion on fixing the retail price of wine. It is based on exhaustive data concerning the scores attributed to different wines by a broad panel of 19 experts for wines from three different countries over a period of 11 years (2000 to 2010). Our main objective is to reduce the systematic econometric bias bound up with regard to expert opinion and to test the impact on prices of a consensus or divergence among experts. This bias has been evidenced by Lecocq and Visser (2006) or Oczkowski (2001).

As a first attempt to correct the measurement error bias, we aggregate a solid body of information from 19 experts to reduce the risk of error from any one expert. Since we use their average score, such a risk was reduced, thereby minimizing individual bias. Most other research uses data from a single expert, so this methodological approach allows us to reduce such errors of judgment (Ashton, 2011). More, examining the opinions of several experts allows us to underline the specific impact of each in regard to prices. Additionally, since the key role played by Robert Parker is often highlighted, we can compare the impact of his opinion with that of other experts.

As shown by Lecocq and Visser (2006), the use of the average score might not prevent the measurement error. Assuming the measurement errors are independent and zero-mean, the average tends to zero when the number of expert tends to infinity. The key issue is: how many experts are required to correctly estimate the impact of scores on price? If the number of observed scores for each wine is not enough, then scores are endogenous and the estimates are downward biased. In order to address this issue we use weather data as identifying instruments for the scores. This method allows us to extract both the objective component of the scores, and the measurement errors, then to estimate their respective impacts on price in an augmented regression.

Further, the gathered data permits to test the influence of the scores' dispersion on the prices. One argument for this influence is that consumers might be wary of the true quality of a wine when its scores are very different among the experts. This uncertainty is negatively perceived by risk averse consumers, decreasing their demand, thus decreasing the equilibrium price. We test that hypothesis by using the standard deviation of scores for each wine as a determinant of the price.

We first examine the methodology adopted and the data used before presenting the econometric results obtained and then concluding briefly.

1. Model

1.1 The naïve model

Rosen's hedonic model (1974) is traditionally used to determine the price of agricultural produce (Costanigro and McCluskey, 2011). A hedonic function is the relation between differentiated prices for a given good and the quantity of constituent characteristics contained in that good (Triplett, 2004)[1]. In the case of wine, prices are determined by factors including appellations, vintage, climatic conditions, expert opinions, reputation, etc. (Combris et al., 1997; Landon and Smith, 1998; Oczkowski, 1994 and 2001; Cardebat and Figuet, 2004 and 2009; Benfratello et al., 2009, etc. For a survey, see Costanigro and McCluskey, 2011). Our design aims at giving structure to the relation between price and its determinants. The focus is here on the relation between quality, experts' grades and prices.

We consider here that wine prices are determined by intrinsic quality, age, and reputation of the producers. Several variables are available to control for these factors: the names of the producers, the vintage, the experts' scores, and the weather conditions of growth. We consider that the impact of reputation is captured by a producer-specific fixed effect, instead of lagged scores as in Oczkowski (2001). As the data are a cross section on year 2011, the reputation of the producers is the same for all vintages: it is the reputation of the producer on year 2011. The issue related to the use of fixed effects model has been addressed by Dubois and Nauges (2010). They explain why those fixed effects cannot be used for the purpose of controlling for quality. Therefore, we interpret these fixed effects in terms of reputation. The age is easily calculated with the vintage. The real issue is the quantitative estimation of quality. Our best indicators of quality are the scores, but they need to be corrected (see Oczkowski, 2001; Lecocq and Visser, 2006; or Dubois and Nauges, 2010). To deal with this issue, we use the following measurement error model:

$score_{ite} = q_{it} + o_{ite}$ (1)

Where $score_{ite}$ is the score of producer i for the vintage t with the expert e, q_{it} is the objective quality of this wine, and o_{ite} is the personal opinion of the expert e on this wine. We assume that the quantity q_{it} is the objective component of the score, and that o_{ite} is its subjective component. Since the experts aim at evaluating the intrinsic quality of wine, their opinions o_{ite} are seen as measurement errors.

We still need to evaluate the qualities q_{it}. A first naïve method is to assume that o_{ite} are independent and identically distributed (iid), with zero mean. Under this hypothesis, we can apply the law of large numbers (LLG):

$$\operatorname*{plim}_{n\to+\infty} \frac{1}{n}\sum_{e=1}^{n} o_{ite} = 0$$

In this design, the average score among the 19 experts for each wine is thus a consistent estimator of the objective score. We estimate the following price model:

$\ln(p_{it}) = \gamma\overline{score}_{it} + \delta t + \mu_i$ (2)

Where p_{it} is the price of the wine of age t of producer i, $\overline{score}_{it}$ is its average score among the 19 experts, and μ_i is the fixed effect of producer i[2]. These fixed effects aim at capturing the effect of reputation on prices given the score and the age. The coefficient δ measures the storage cost, the quality improvements due to the keeping and the scarcity value all at once. Remember that the data are a cross section, which means that the vintage, t of wine i solely determines the age.

We estimate equation (2) with the generalized least squares (GLS). We use the Newey-West variance estimator, since the residuals faced both heteroscedasticity (the variance of the errors differ among producers) and autocorrelation (there is some inertia across vintages). The $\hat{\gamma}$ obtained with the average score is compared to those obtained when we replace the average

---

[1] This method has been used for cars (Court, 1939; Griliches, 1961; Triplett, 1969; Arguea and Hsiao, 1993), real estate (Taylor, 2003), computers (Triplett, 1989), the environment (Freeman, 1993), corn (Espinosa and Goodwin, 1991), cereals (Stanley and Tschirhart, 1991), apples (Carew, 2000), and even for the French vaulting stallion semen market (Vaillant et al., 2010).

[2] Notice that these producers-specific fixed effects forbid the use of the ranks or the appellations as additional control variables, because of perfect multicolinearity issues among the dummy variables. Yet, using fixed effects at the level of the producer should be more efficient.

score with the score of a few major experts. We then use a subsample of wines that have been graded by at least the five major experts (these have graded at least 1,000 wines of our whole sample). One of these experts is Robert Parker (*The Wine Advocate*, WA), who enjoys the reputation of being a wine guru with great influence on prices (Hadj Ali et al., 2008; Jones and Storchmann, 2001). Others include the *Wine Spectator* (WS), Jancis Robinson (JR), Jean-Marc Quarin (JMQ) and Stephen Tanzer (*International Wine Cellar*, IWC). This cut leaves us with 700 prices of wines from 129 Châteaux of Bordeaux, from vintage 2001 to 2011. The use of this subsample allows us to conduct a multi-expert regression that controls for the correlations between the experts' grades. In the expert-specific regressions, the impacts of the different experts are not taken into account simultaneously, although the real impacts are indeed linked to each other. We therefore obtain more accurate estimates of experts' respective influences. This also provides an idea of the error made when only one expert is considered.

1.2 The two-stage model

This naive model faces some major limits. The first one is the application of the LLG with at best 19 experts for each wine. As Lecocq and Visser (2006) pointed out, it seems hardly acceptable that the opinions of the 19 experts correct each other perfectly. Worse, the opinions o_{ite} must be iid in order to validate the LLG. That is problematic since experts' grading behavior depend on both their taste and their grading scale. As we shall see, some experts grade systematically under the average score and some other systematically above the average. Given that information, this first model should be abandoned.

Another key limit of the naïve model is that it assumes that the opinions have no influence on price. The underlying hypothesis is that the price is solely determined by the objective component of the scores, while the subjective component is irrelevant in the price equation. This point is mostly unacceptable, since the only observable score contains both the objective and the subjective components. More, some consumers might have great interest in the differentiated opinions of the experts, acknowledging that each expert represent a certain taste. This point has been highlighted by Lecoq and Visser (2006). When a consumer feels well represented by one expert, he is likely to be very influenced by the subjective opinion of this expert, and might not look at the comments of the other experts.

We have shown the necessity of integrating both the objective and subjective components in the price equation. Of course, this is achieved by using the raw scores, but this specification implies that the two components have the same coefficient. Testing this hypothesis is another goal of the present article. Hence, we need to disentangle q_{it} from o_{ite} in equation (1). To this end, we use the weather data as determinants of quality and identifying variables.

It seems reasonable to consider that quality is solely determined by the soil quality, the skills of the producer (including the viticulture ability, the precision during the harvest, the maturing process, and the variety blend) and the weather conditions of growth. Making the assumption that the two first factors can be captured by producers-specific fixed effects and a trend, we design the following model for objective quality:

$$q_{it} = \beta w_{it} + \rho t + \nu_i \qquad (3)$$

Where w_{it} is the vector of the weather variables, and ν_i is the fixed effect of the producer i on quality. In this design, the fixed effects are indicators of both soil quality and producer skills. The trend aims at taking into account the global improvements in technology. In order to limit the number of coefficients, we assume that the impact of the weather variables on the objective scores is the same for all producers.

We obtain the reduced form of scores by combining (1) and (3):

$$score_{ite} = \beta w_{it} + \rho t + \mu_i + o_{ite} \qquad (4)$$

(4) can be estimated with the GLS, minimizing the variance of the opinions. Again, we use the Newey-West variance estimator, as the opinions o_{ite} showed heteroscedasticity and autocorrelation.

This first stage regression gives us an estimate $\hat{o}_{ite}$ of the opinions of the experts. Let $\hat{o}_{it}$ be the vector that contains the $\hat{o}_{ite}$. Adding this variable to the model (2), and replacing the average score by any expert score allow us to estimate the differentiated impacts of quality and experts' opinions on prices. Formally it writes:

$$\ln(p_{it}) = \gamma score_{ite_1} + \theta \hat{o}_{it} + \lambda t + \mu_i \qquad (5)$$

Where e_1 is the chosen reference expert, and μ_i still aims at estimating the influence of reputation of producer i on price given the scores. Splitting the variable $score_{ite_1}$ into its objective component $\hat{q}_{it}$ and its subjective component $\hat{o}_{ite_1}$, we get the detailed effects of scores on prices:

$$\ln(p_{it}) = \gamma \hat{q}_{it} + (\gamma + \theta_1)\hat{o}_{ite_1} + \theta_2 \hat{o}_{ite_2} + \cdots + \theta_n \hat{o}_{ite_n} + \lambda t + \mu_i$$

Where n is the number of experts. That is why the choice of the reference expert does not matter. The coefficients are estimated by bootstrap in order to correct the sampling error in the ordinary least squares (OLS) variance estimates in the second stage (due to the replacement of (q, o) by $(\hat{q}, \hat{o})$). That is to say, we randomly draw with replacement a same-size subsample from the initial one, we conduct the two-stage estimation with this subsample, and store the second-stage estimates (calculated with the OLS method). We do this procedure 1,000 times, and find the convergent bootstrap estimate as the mean value of each coefficient. The variance of the bootstrap estimates are estimated non-parametrically by the empiric variance of the 1,000 estimates for each coefficient.

At this point, we should precise the link between our approach and instrumental variable techniques. Here, we use the weather data, the age of the wine and producers' dummy variables as instruments for the raw scores. This IV procedure has been used and discussed in Haeger and Storchmann (2006). Our method is slightly different though, since we use the residuals of this first-stage equation as a control variable for the second-stage. This refers to the literature on control function techniques. Some tests are required to check our model:

- an endogeneity test, which is easily obtained by a T-test on the coefficients θ_e,
- an overidentification test, which is given by an F-test for the regression of the second-stage residuals on the instruments.

We provide these tests by bootstrapping the tests statistics.

This procedure has been first conducted with the whole sample by using only the mean score for each wine: this includes all wines that have been graded at least by one expert. Then, we conduct the same analysis separately for each expert. As we did for the naïve model, we finally use the subsample of the 700 wines for further inter-experts analysis and robustness check. We confront these results to the naïve model estimates.

1.3 Consumer defiance versus marketing effect

Our design allows us to provide precise estimates of the individual impacts of each experts, apart from the impact of the objective component of the scores. Yet, it misses one indirect effect of the scores on price. As it is often argued in the literature of marketing and consumer behavior (see, for example, Martin et al., 2007, or Monic, 2007), the standard deviation of grades is likely to affect negatively the trust of the buyers in the scores. Assuming the risk aversion of the consumers, a high dispersion of the scores decreases the equilibrium price because it lowers the demand.

However, another indirect effect of standard deviation on prices might occur. As shown by Hilger et al., (2011), when a retailer exhibits a score for a wine, its sales (or price) increase: the higher the exhibited score, the higher the rise in sales. A high standard deviation in scores for a wine implies than at least one expert liked the wine more than the others and gave it an above-average mark. Retailers know all the scores and can choose to communicate only the best. We call this positive correlation between standard deviation of the scores and wine prices the "marketing effect". The higher the standard deviation, the greater the likelihood for the retailer to exhibit a good score (compared to the average) and the higher the price. Unexpectedly, the lack of consensus among experts allows retailers to improve their marketing and to increase their prices.

Our model allows to test that hypothesis. Since the standard deviation of the scores is the standard deviation of the opinions, we add the latter among the regressors in the second-stage of the multi-expert regression. To this end, we use the empirical standard deviations of estimated opinions for each wine as an estimate of the real deviation of the scores. The estimate of the coefficient and its related significance are obtained by bootstrap. We conduct that analysis on the subsample with only five experts, in order to keep a constant number of grades per observation.

2. DATA

Annual data were obtained for 203 wine producers, mainly located in the Bordeaux area (187 producers from 12 AOC areas, with nine producers from the Napa Valley, USA, and seven from Spain) covering the period from 2000 to 2010. The prices were obtained from the website *winedecider.com*. This website offers prices on a wide range of wines from several countries and AOCs and is representative of the main wine sellers on the Internet, including *Millesima*. The listed price is the average retail price of a bottle packaged in a case of six or 12 bottles in 2011 (constant prices) before VAT and transportation costs. Using the retail price means we can assume that these wines are priced after the experts have published their scores. This point is crucial to the relationship between wine prices and expert opinion. A retailer's pricing behaviour will vary depending upon whether or not he is aware of the expert ratings.

Table 1 provides descriptive statistics about the prices and the scores among the different appellations.

Table 1: Prices and scores across appellations

About here

As in the hedonic approach, we include:

- Objective characteristics: name of the producer and vintage;
- Tasting rating or subjective quality: scores from 19 experts (each wine is graded by at least one expert, and by 4,5 experts on average);
- Weather as a determinant of the objective quality: temperature and rainfall data from several meteorological stations in the heart of the AOC, due to the great heterogeneity of local weather conditions across the vast wine-producing area of Bordeaux (discussed below)

Table 2 examines the evolutions of the descriptive statistics of table 1 on the whole sample, across the different vintages.

Table 2: Prices and scores across vintages

About here

As for weather data, we obtain details of daily weather conditions for the three main areas of the Bordeaux region and for one area in the Napa Valley. We define the three main climate areas of the Bordeaux appellations: Médoc, Saint-Emilion/Pomerol and Graves. Meteorological studies related to wine reveal significant weather variability within the Bordeaux appellation (Bois, 2007; Bois and Van Leeuwen, 2008). Table 3 shows the average temperatures and rainfalls across the four available stations.

Table 3: Weather variables descriptive statistics

About here

In agreement with Bois and Van Leeuwen's (2008) climate observations, this information is crucial to our study. It is essential to correlate meteorological data from each of these three areas and not only information from the main meteorological station based in Mérignac. Even if Lecocq and Visser (2006) show that the Mérignac station provided a reasonably acceptable proxy of the weather for the Bordeaux appellation as a whole between 1993 and 2002, they do note that certain differences appear: "The climate conditions prevailing in the main weather station [Mérignac] are thus clearly not representative of the Bordeaux wine region as a whole" (Lecocq and Visser, 2006, p. 6). Our design aims at using meteorological data as an instrument for the scores. Consequently, we cannot use only one station if we want to maintain some heterogeneity in our fitted scores.
We have managed to gather the monthly temperatures and rainfalls from three stations representative of the three main wine regions of Bordeaux. For the Médoc region, we use weather data from Château Latour (which is very close to Pauillac). In the Graves region, we use weather data from Château Haut-Bergey (in Léognan). For Saint-Emilion/Pomerol we use data from Château Grand Barrail. As for the Napa Valley, the data come from Oakville meteorological station. We do not have weather conditions for the two Spanish appellations.

3. Empirical Results

3.1 Results from the naïve model

Table 4 gathers the γ estimates for the naïve model[3]. We have first estimated the equation (2) using the average of all available scores, then for each of these experts. For these expert-specific regressions, we provide the numbers of observations and the coefficient of determination (R^2). Table 4 also contains the estimates of the multi-expert regression, including the average score. The last column shows the VIFs, indicating potential weak multicolinearity issues for the average score coefficient.

Table 4: Naïve Model γ estimates
About here

The main observation is that the estimate of γ is very dependent on the specification. The results displayed in the third column underline the importance of using several experts in order to model wine prices, since they all have differentiated impacts[4]. According to the naïve model, a one point-increase in the objective score lead to a 4.8% increase in prices. Notice that the coefficient of determination is not maximum for the average score model though it should be the best model. This suggests that the subjective opinions of the experts do contribute to determine wine prices, which would invalidate the naïve model.

The hierarchy between the experts' influence is remarkably the same in the two designs. In particular, Robert Parker (WA) is not the most influential expert: he comes after Stephen Tanzer (IWC). Both designs conclude to a minimal influence for Jancis Robinson (JR), and an average impact for Jean-Marc Quarin (JMQ) and Wine Spectator (WS). This is consistent with the results of Ashton (2013) about the correlations between experts: he found that JR was the most "out of line" expert, and that she embodied the most different taste from Robert Parker (WA). The coefficient of the average score is very low in the multi-experts regression. Its non-significance can be questioned in light of its VIF which is greater than 10, indicating potential weak multicolinearity issues.

3.2 The two-stage model results

a. First stage

Table 5 shows the first stage estimates related to equation (4), for each single-expert regression, for the average score regression, and for the multi-experts regression. We have tried numerous aggregated specifications of the weather data, since the estimates of the raw monthly temperatures and rainfalls coefficients were found quite inconsistent. We have chosen to display the results using the following aggregates, which give the most robust and significant results[5]:

---

[3] The estimates of the trend and the fixed effects are available upon request.

[4] The differences among the $\hat{\gamma}$ in the third column are not entirely due to the differences in the samples (see the number of observations in the second column), since the distribution of the observations across appellations and vintage are actually similar for each expert. Still, we tested that assumption by conducting the six different regressions on the same sample of 700 wines. The results still indicated that the respective impacts of the experts are not equal.

[5] The several specifications we tried for the first stage all led to quasi-identical results in the second stage, comforting our conclusions.

- the total rainfalls during the first part of the growth season: April to July (referred as Rainfalls 1),
- the total rainfalls during the last part of the growth season: August and September (referred as Rainfalls 2),
- the average of the monthly average temperatures during the growth season (from April to September, referred as Temperature).

We divided the growth season into two parts for the rainfalls, because of the major expected impact of the rainfalls in the last months of the growth. At this point of the growth, intense rainfalls cause the grapes to rot, which jeopardize the quality of the crop. To a lesser extent, rainfalls still negatively impact the quality of the vintage during the rest of the growth season. The temperature showed important multicolinearity issues that is why we focused on the average temperature during the growth season. We also show the estimated coefficients related to the trend, which is seen as the impact of the overall progress in technology.

Table 5: First stage estimates
About here

The signs of the coefficients are in keeping with the literature in phenology: a good vintage is caused by a dry summer, with high temperature. As expected, rainfalls during August and September have greater negative impact than rainfalls during the beginning of the growth season - except for the JMQ model. The trend is also very significant, indicating great progress in the technology through the trend in the scores.
As a clue of the accuracy of the first-stage fitted values, we provide in Table 6 some descriptive statistics of the residuals of the multi-expert regression. The second column displays the mean opinion for each expert, the third displays the mean of the absolute values, and the fourth gives the average deviation from the real value in percentage.

Table 6: Descriptive statistics of the opinions from the multi-expert regression
About here
The first columns shows that JR is on average far below the mean score. This is due to the fact that she originally grades out of 20, and that we converted these grades out of 100 for the homogeneity of the results. On the opposite, WA, WS and IWC are generally above the mean score, while JMQ is notably close to the objective score. This regression is rather accurate, since scores are estimated with an average error going from 1.8% for JMQ up to 7.4% for JR. Remember that our goal here is not the precise estimation of the scores, but rather the estimation of the residuals. The accuracy of the first stage regression is not the issue, since we actually expect some heterogeneity in the residuals.

b. *Second stage*

We now comment the estimates of equation (5). Table 7 displays the bootstrap estimates for each expert-specific regression, for the average score regression, and for the multi-expert regression. None of the non-reported Fisher tests for overidentification allow to reject the exogeneity of the instruments at any level (all P-values are greater than 0.85). Hence, the exclusion condition for the validity of our instruments cannot be refuted.

Table 7: Second stage estimates
About here

The results definitely reject the naïve model as a suitable design for assessing the impact of the objective component of scores. Indeed, each of the subjective components included in any of the specifications have a significant impact on wine price. The naïve model aims to get rid of the opinions in order to focus on the objective scores, but these opinions do have a significant impact on prices. The naïve model is then flawed: the $\hat{\gamma}$ displayed in Table 4 are all negatively biased. The two sources of bias are observed: the LLG is not valid because the opinions are not iid and the subjective components have a significant impact on prices. Besides, the systematic significance of the opinions also confirms the endogeneity of the raw scores, and backs our two-stage model.

Furthermore, the opinions have different impacts. This means that in the aim of properly estimating the impact of the objective component, one should use different experts since they all influence wine prices their own way. This is illustrated by the dispersion among the $\hat{\gamma}$ obtained in the average score and expert-specific regressions (upper part of Table 7). Notice that the hierarchy between the experts is still robust: IWC and WA are the most influent experts, while JR is the least influent.

All the $\hat{\gamma}$ obtained with the two-stage procedure are way bigger than those obtained with the naïve model. This can be explained both by the downward measurement error bias (see Chesher, 1991; or Lecocq and Visser, 2006), and by the omitted variable bias as the opinions are positively correlated with wine price, and mainly negatively correlated with the objective scores. It supports the use of our design in order to avoid those two biases. Notice that the $\hat{\gamma}$ are also greater than their related $\hat{\theta}$. This is good news as it indicates that the objective component of the scores is more influent than the subjective one. Another way to say it is that the prices are more precisely determined by the fundamentals than by the subjective opinions of the experts.

Another feature of these estimates is that, in the multi-expert regression, the expert's respective influences are a little lower than those estimated with the naïve model. This is consistent with Dubois and Nauges (2010), who also found an upward bias of the estimated influences when not controlling for the unobserved quality, or the objective score as we name it here. In our design, a one-point increase of the objective score is estimated to have a 15.1% impact on the price, whereas a one-point increase in the experts' opinion has a maximum impact of 4.2% on the price for IWC, 3.7% for WA, and no significant impact for JR. Indeed these are pure mental projections, as we cannot observe the two components of the scores. Finally, the trend is also very significant: in the multi-expert regression we estimate that a wine is each year 1.9% more expensive, due to the storage cost, the maturing and the scarcity value.

3.3 Marketing effect

As an application of our model, we tested the impact of the deviation among the scores on prices. Slightly digressing from our structural model, we added the empirical standard deviation of the opinions to equation (5), in order to assess the significance of its coefficient, and the sign of the latter. The upper part of Table 8 gathers the results of this estimation.

Table 8: Standard deviation and Maximum score
About here

The model indicates a strong positive impact of the standard deviations of the opinions on wine prices. That correlation might result from what we introduced as the "marketing effect" in section 1.3. In that case, we should observe a major impact of the maximum score, which is supposed to be the most publicized, hence the most important in the price equation.

We test that hypothesis by confronting the maximum opinion to all the individual opinions in the two-stage model. The results of this estimation are displayed in the lower part of Table 8. The coefficient of the maximum opinion is estimated as the most important one. This is an argument in favour of the "marketing effect" interpretation. The maximum score is the most often exhibited score, so that it is the only hint of quality for the consumers that haven't searched for the other experts' grades (Hilger et al., 2011). Therefore, the maximum score obtains the maximum influence on the price[6].

The revealed impact of the maximum score sheds a new light on the respective influence of the experts. The idea of the marketing effect about the diffusion of the scores might explain why JR is granted so little influence on price. Since she often grades under the average, the sellers might not be eager to exhibit her scores. On the opposite, the fact that WA and WS are on average above the mean score might have a role in their major influence. That interpretation is consistent with the results of Table 8: the introduction of the maximum score in the regression has lowered the coefficients of the above-the-mean-score experts. Also, JMQ is the closest to the objective score and his coefficient has barely changed.

Conclusion

This research aims to assess the role of expert opinion on Bordeaux wine prices using a methodology which, by including detailed meteorological data, fixed-effects models and the systematic use of numerous expert scores, avoids endogeneity and bias rooted in errors of judgment. As Oczkowski (2001), Lecocq and Visser (2006), or Dubois and Nauges (2010), we assume that the observed scores result from an error measurement model: they can be split into an objective component, shared by all experts, and a subjective component specific for each expert and each wine. The latter is often seen as something that should be corrected because it confuses the signal on quality embodied by the objective component. We provide evidence that in a price equation though, one should not try to get rid of the subjective component because it does significantly impact wine prices. Worse, if not handled specifically, it leads to downward biased estimates of the impact of the scores as a quality indicator for wine prices. This result is consistent with Lecocq and Visser (2006).

The nice feature of our results is the light shed on the role of the standard deviation in the price equation. We find a strong positive correlation between wine price and the standard deviation of the scores. Our interpretation is based on the fact that a great standard deviation means that at least one score is above the others. In line with the marketing literature, this maximum score might be used as a promoting argument by the sellers. Hence, this particular score is likely to be the most publicized. As a result this is certainly the only score that the average consumers have heard of. This is what we call the "marketing effect": the maximum score is the most influent because it is the best-known among consumers. Our interpretation is comforted by the empirical analysis, since the maximum opinion is granted the highest impact on prices.

We have to be cautious with that interpretation though. There is actually another interpretation of our results: the consumers might be risk seeking. In that case, the standard deviation of the scores is also supposed to have a positive influence on demand, thus on prices. It is mainly agreed among economists that the ordinary consumer is rather risk averse, but the market we discuss here is very specific. The prices of our data go from 8$ to 3000$, and the average price is 83$. That is no market for the uninitiated. The consumers involved in that market are either connoisseurs, professionals, or investors, and at least the latter are likely to be risk seeking. We maintain our interpretation though, assuming that the market prices are likely to be more impacted by the marketing effect than by risk seeking behaviors.

[6] This final result holds with all the specifications of the instruments we tried, and with the naïve design, as including the maximum scores in equation (2) for a multi-expert regression leads to the same conclusion.

References

Ackerberg, D. (2003) Advertising, learning, and consumer choice in experience good markets: a structural empirical examination, *International Economic Review,* **44,** 1007-1040.

Akerlof, G. (1970) The market for `lemons': quality uncertainty and the market mechanism, *Quarterly Journal of Economics,* **84,** 488-500.

Arguea,, N. and Hsiao, C. (1993) Econometric issues of estimating hedonic price functions: with an application to the US Market for automobiles", *Journal of Econometrics,* **56(1-2),** 243-67.

Ashenfelter, O. (1989) How auctions work for wine and art, *Journal of Economic Perspectives,* **3(3),** 23-36.

Ashenfelter, O. (2008) Predicting the prices and quality of Bordeaux wines, *Economic Journal*, **118,** 40-53.

Ashenfelter, O., Ashmore, D. and Lalonde, R. (1993) Wine vintage quality and the weather: Bordeaux, *Paper presented at the* 2nd *International Conference of the Vineyard Data Quantification Society*, Verona, February 1994.

Ashenfelter, O. and Jones, G. (2000) The demand for expert opinion: Bordeaux wine, *Cahiers de l'OCVE,* **3,** 1-17.

Ashton, R. H. (2011) Improving experts' wine quality judgments: two heads are better than one, *Journal of Wine Economics,* **6(2),** 160-178.

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(02)**, 225-234.

Benfratello, L., Piacenza M. and Sacchetto S. (2009) Taste or reputation: what drives market prices in the wine industry? Estimation of a hedonic model for Italian premium wines, *Applied Economics,* **41,** 2197-99.

Bois, B. (2007) Cartographie agroclimatique à méso-échelle : méthodologie et application à la variabilité spatiale du climat en Gironde, Ph. D. Thesis, University of Bordeaux 1 and INRA.

Bois, B. and Van Leeuwen, C. (2008) Variabilité climatique dans la zone de production des vins de Bordeaux, *Terroirs viticoles*, ENITA Bordeaux - décembre.

Brochet, F. (2000) La dégustation : étude des représentations des objets chimiques dans le champ de la conscience », *Ph. D. thesis*, Institut Scientifique de la Vigne et du Vin (ISVV) and University Bordeaux II.

Brown, J. and Rosen, H. (1982) On the Estimation of Structural Hedonic Price Models, *Econometrica*, **50(3),** 765-768.

Budescu, D., Rantilla, A., Yu, H.-T., Karelitz, T. (2003) The effects of asymmetry among advisors on the aggregation of their opinions, *Organizational Behavior and Human Decision Processes*, **90(1),** 178-191.

Cardebat, J.-M. and Figuet J-M. (2004) What explains Bordeaux wine prices?, *Applied Economic Letters,* **11(5),** 293-296

Cardebat, J.-M. and Figuet J-M. (2009), Estimation of a hedonic price equation for Alsace, Beaujolais and Provence wines, *Applied Economics Letters,***16(7-9),** 921-927.

Carew, R. (2000) A hedonic analysis of apple prices and product quality characteristics in British Columbia, *Canadian Journal of Agricultural Economics,* **58(1),** 93-108.
Chanel O., Gérard-Varet L-A.and Ginsburgh V. (1996) The relevance of hedonic price indices. The case of paintings, *Journal of Cultural Economics,* **20(1),** 1-24.
Chevet, J.-M., Lecocq, S. and Visser, M. (2011) Climate, grapevine phenology, wine production and prices: Pauillac (1800-2009), *American Economic Review, Papers and Proceedings,* **101,** 142-146.
Chesher, A. (1991), The effect of measurement error, *Biometrika,* 78**(3),** 451-462.
Clemons, E. K., Guodong, G.G. and M. Hitt, L. (2006) When online reviews meet hyperdifferentiation: a study of the craft beer industry, *Journal of Management Information Systems*, **23(2),** 149–171.
Combris, P., Lecocq, S. and Visser, M. (1997) Estimation of a hedonic price equation for Bordeaux wine : does quality matter ?, *Economic Journal*, 107, 390-402.
Costanigro, M. and McCluskey, J. J. (2011) Hedonic price analysis in food markets. In Lusk, J.L., Roosen, J., Shogren, J. F. (eds) *The Oxford Handbook of the economics of food consumption and policy*, Oxford University Press, 152-180
Court, A.T. (1939) Hedonic price indexes with automotive examples, *in The dynamics of automobile demand, New York,* The General Motors Corporation, 99-117.

Di Vittorio, A. and Ginsburgh, V. (1994) Des enchères comme révélateurs du classement des vins, *Journal de la Société Statistique de Paris,* **137,** 19-49.

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(3),** 205-212.

Espinosa, J. and Goodwin, B. (1991) Hedonic prices evaluation for Kansas wheat characteristics, *Western Journal of Agricultural Economics,* **16,** 72-85.

Frankel, J., and Rose, A. (2010) Determinants of Agricultural and Mineral Commodity Prices, *Research Working Paper Series*, **10-038,** John F. Kennedy School of Government, Harvard University.

Freeman, A. (1993) *The measurement of environment and resource values: theory and methods,* Resources for the future, Washignton D.C.

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

Ginsburgh, V., Monzak, M. and Monzak, A. (1992) Red wines of Médoc : what is wine tasting worth?", *Paper presented at the 2nd International Conference of the Vineyard Data Quantification Society*, Verona, February 1994.

Golan A. and Shalit H. (1993) Wine, quality differentials in hedonic grape pricing, *Journal of Agricultural Economics,* **44**, 311-321.

Griliches, Z. (1961) Hedonic price indexes for automobiles: an econometric analysis of quality change, in: price statistics review Committee, National Bureau of Economic Research, *The prices statistics of the Federal Government*: *Review, Appraisal, and Recommendations,* General Series, **73,** New York 173-96, Hadj Ali H., Nauges, C. (2007) The pricing of experience goods: The example of *en primeur* wine, *American Journal of Agricultural Economics*, 89 **(1)**,91-103.

Hadj Ali, H., Lecocq, S.and Visser, M. (2008) The impact of gurus: Parker grades and *en primeur* wine prices, *Economic Journal*, **118,** 158-173.

Haeger, J.W. and Storchmann, K. (2006), Prices of American Pinot Noir Wines: Climate, Craftsmanship, Critics, *Agricultural Economics,* 35(1), 67-78.

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),** 1289-96.

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

Hodgson, R. (2009), An analysis of the concordance among 13 U.S. wine competitions, *Journal of Wine Economics*, 4(1), 1–9

Huglin, P. (1978) Nouveau mode d'évaluation des possibilités héliothermiques d'un milieu viticole, *Symposium International sur l'Écologie de la Vigne,* **l,** 89-98, Constança, Roumanie.

Huglin, P. and Schneider, C. (1998) *Biologie et écologie de la vigne.* Paris, Lavoisier, 370p.

Jaeger, D. and Storchmann K-H. (2012) Wine Retail Price Dispersion in the United States: Searching for Expensive Wines? *American Economic Review, Papers and Proceedings,* **101,** 136-141.

Jin, G., and Leslie, P. (2003)**,** The effects of information on product quality: evidence from restaurant hygiene cards, *Quarterly Journal of Economics,* **118,** 409-451.

Jones, G. 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,

Jolie, M., Barron, G. and Norton, M. I. (2007) Choosing to be uncertain: preferences for high variance experiences. *Working Paper Harvard Business School.*

Landon, S. and Smith, CE. (1998) The use of quality and reputation indicators by consumers: the case of Bordeaux wine", *Journal of Consumer Policy,* **20,** 289-323.

Lecocq, S. and Visser, M. (2006) Spatial Variations in Climate and Bordeaux Wine Prices, *Journal of Wine Economics,* **1(1),** 42-56.

Monic, J.S. (2007) The informational role of consumer disagreement, *Department of Economics Boston University*.

Montgomery, C., and Wernerfelt, B. (1992) Risk reduction and umbrella branding, *Journal of Business,* **65,** 31-50.

Nelson, P. (1970) Information and consumer behavior, *Journal of Political Economy,* **78,** 311-329.

Nerlove, M. (1995) Hedonic price functions and the measurement of preferences: the case of Swedish wine consumers, *European Economic Review,* **39,** 1697-1716.

Oczkowski, E. (1994) A hedonic price function for Australian premium table wine, *Australian Journal of Agricultural Economics,* **38(1),** 93-110.

Oczkowoski, E. (2001) Hedonic wine price function and measurement error, *Economic Record,* **239,** 374-382.

Parker, R. (2008), *Parker's wine buyer*'s *guide*, Simon and Schuster, New York.

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

Roma, P., Di Martino, G. and Perrone G. (2013), What to show on the wine labels: a hedonic analysis of price drivers of Sicilian wines, *Applied Economics,* **45(19),** 2765-2778.
Rosen, S.M. (1974) Hedonic prices and implicit markets: product differentiation in pure competition, *Journal of Political Economy*, **82(1),** 34-55.

Salop, S. (1976) Information and monopolistic competition, *American Economic Review, Papers and Proceedings*, **66(2),** 240-245.

Sinkey, M. (2012) How do experts use Bayes' rule? Lessons from an incentive-free environment, *Working Paper Ohio State University.*

Sorensen, A., and Rasmussen, S. (2004) Is any publicity good publicity? A note on the impact of book reviews, *Working Paper.*

Stanley, L. and Tschirhart, J., (1991), Hedonic prices for a non-durable goods: the case of breakfast cereals, *Review of Economics and Statistics,* **73(3),** 537-541.

Storchmann, K., Mitterling, A. and Lee, A. (2012), The detrimental effect of expert opinion on price quality dispersion: evidence from the wine market, *AAWE Working paper,* **118.**

Taylor, L. (2003) The hedonic model, in Champ, P., Boyle K. and Brown C. (eds), *A primer on non-market evaluation,* Kluwer Academic Publishers.

Tonietto, J. and Carbonneau, A. (2004) A multicriteria climatic classification system for grape-growing regions worldwide, *Agricultural and Forest Meteorology,* **124(1-2),** 81-97.
Triplett, J. (1969) Automobiles and hedonic quality measurement, *Journal of Political Economy,* **77(3),** 408-17.

Triplett, J. (1989) Price and technological change in a capital good: a survey of research on computers, in Jorgenson, D. and Landau, R. (eds.), *Technology and Capital Formation,* Cambridge, MIT Press, 127-213.

Triplett, J. (2004) *Handbook on hedonic indexes and quality adjustments in price indexes: special application to information technology products,* OECD, Paris.

Vaillant, N., Lesot P., Bonnard, Q. and Harrant V. (2010), The use of expert opinion, quality and reputation indicators by consumers: evidence from the French vaulting stallion semen market, *Applied Economics,* **42,** 739-745.

Annex:

Professionals

AP	Andrè Proensa
BH	Burghound (Allen Meadow)
DEC	Decanter
FD	Franck Dubourdieu
IWC	International Wine Cellar (Stephen Tanzar)
GM	Gault Millau
HAC	Guide Hachette
JMQ	Jean-Marc Quarin
JP	José Penin
JR	Jancis Robinson
MB	Michael Broadbent
PL	Ignacio Pérez Lorenz
RVF	Revue des Vins de France
WA	The Wine Advocate (Robert Parker)
WS	Wine Spectator

Amateurs

EPI	Epicuvin
920R	920-Revue
WD	Winedecider

BlueWine

TABLES

Table 1: Prices and scores across appellations

AOC	# châteaux	Min Price €	Max Price €	Avg Price €	SD Price €	Min Note/100	Max Note/100	Avg Note/100	SD Note/100
Medoc	2	8	23	14.68	4.18	65	91	87.62	4.65
Saint Estephe	21	8	292	33.40	36.99	67.5	100	87.25	4.83
Pauillac	24	16	1604	74.71	239.97	65	100	89.05	5.03
Saint Julien	16	10	247	55.89	43.74	72.5	100	89.09	4.36
Listrac	4	8	44	14.23	7.02	70	92	83.39	4.53
Moulis	3	14	34	21.61	4.50	67.5	92	85.29	5.18
Margaux	30	11	829	58.07	100.46	65	100	87.68	4.79
Haut-Medoc	11	9	37	16.53	5.53	65	93	84.30	5.28

Pessac-Leognan	24	10	756	82.67	149.25	65	100	87.95	5.08
Sauternes	17	11	566	46.40	79.80	65	100	88.41	5.11
Pomerol	18	14	3359	176.39	485.58	70	100	26.31	5.47
Saint-Emilion	17	11	1501	124.62	253.65	60	100	88.04	5.51
Ribera del Duero	4	42	155	88.08	33.67	75	99	91.08	4.11
Rioja	3	15	88	37.24	19.46	80	97	90.63	4.24
Napa Valley	9	17	543	142.33	124.07	69	100	91.02	4.86
Total	**203**	**8**	**3359**	**83.12**	**203.57**	**60**	**100**	**87.94**	**5.2**

As shown in table 1, the range of prices, from 8€ to 3359€ per bottle, and the range of notes, from 60 to 100, is comparatively wide.

Table 2: Prices and scores across vintages

Year	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Score (/100)											
Min	79	81	79	80	83	85	81	82	84	83	85
Max	97	100	98	97	97	98	96	98	97	98	99
Avg	89.4	88.6	87.8	88.6	88.6	90.2	88.9	88.4	89.3	90.9	90.8
SD	4.12	4.67	4.36	5.24	4.30	4.29	4.55	4.19	3.36	4.49	4.63
Price (in €)											
Min	10	10	9	10	9	9	9	8	9	10	9
Max	3359	1415	1332	1439	1274	2680	1242	1164	2008	2741	2449
Avg	113.9	71.4	64.5	76.7	63.3	101.6	69.6	64.4	72.1	112.4	105.4
SD	311.5	135.7	132.2	165.2	121.6	253.7	136.3	127.3	178.8	280.7	265.5

Note: min and max are calculated on average scores, rounded to the nearest integer.

Table 3: Weather variables descriptive statistics

	Médoc	Saint-Emilion / Pomerol	Graves	Napa Valley
Temperatures				
April	13.89	12.79	12.39	13.00
May	17.15	16.17	15.81	16.47
June	21.48	20.14	19.86	18.83
July	22.21	21.02	20.89	19.43
August	22.37	20.68	20.48	18.99
September	19.42	17.32	17.29	18.12
Rainfalls				
December to March	311.90	202.00	233.00	665.40
April to June	213.50	231.20	228.20	100.33
August to September	95.38	113.90	101.55	2.13

Table 4: Naïve Model γ estimates

	Average score and expert-specific regressions			Multi-expert regression	
	Number of observations	$\hat{\gamma}$	R^2	$\hat{\gamma}$	VIF
Average Score	2172	0.048***	0.945	0.013**	9.810
IWC	1169	0.099***	0.961	0.042***	5.361
JR	1723	0.016***	0.944	0.004*	2.838
WA	1644	0.063***	0.959	0.039***	5.573
WS	1758	0.047***	0.951	0.022***	4.466

Note to the reader: The columns $\hat{\gamma}$ contain the estimated influence of the average score and of the experts on price, for the two different specifications.
Significance: *** : level 0.01 ; ** : level 0.5 ; * : level 0.1 ; none: non-significant at the 0.1 level

Table 5: First stage estimates

	Average Score	IWC	JR	WA	WS	Multi-expert
Rainfalls 1	-0.004***	-0.004***	-0.003	-0.005***	-0.015***	-0.010***
Rainfalls 2	-0.007**	-0.010***	-0.013***	-0.017***	-0.017***	-0.014***
Temperature	0.354***	0.430***	0.383*	0.537***	0.375***	0.446***
Trend	0.315***	0.379***	0.441***	0.351***	0.453***	0.444***

Significance: *** : level 0.01 ; ** : level 0.5 ; * : level 0.1 ; none: non-significant at the 0.1 level

Table 6: Descriptive statistics of the opinions from the multi-expert regression

Expert	Mean	Mean of the absolute values	Average absolute deviation from the real value
IWC	1.474	1.873	2.1%
JR	-5.765	5.875	7.3%
WA	2.212	2.656	2.9%
WS	2.079	2.399	2.6%

Table 7: Second stage estimates

Average score and expert-specific regressions					
	Average score	IWC	JR	WA	WS
Trend	-0.044***	-0.034***	-0.047***	-0.025***	-0.021***
$\hat{\gamma}$	0.170***	0.183***	0.128***	0.109***	0.088***
$\hat{\theta}$	0.034***	0.077***	0.007***	0.048***	0.032***
Multi-expert regression					
Trend	$\hat{\gamma}$	$\hat{\theta}_{IWC}$	$\hat{\theta}_{JR}$	$\hat{\theta}_{WA}$	$\hat{\theta}_{WS}$
-0.022***	0.137***	0.045***	0.004***	0.041***	0.015***

Note to the reader: In the upper part, the fourth line $\hat{\gamma}$ contain the estimated influence of the objective score using either only the average score or only a specific expert, leading to six different estimates of the same value. The fifth line $\hat{\theta}$ contains the estimated influences of the subjective scores for each regression. In the lower part are shown the same estimates obtained with the multi-expert regression, leading to only one estimate for the influence of the objective score ($\hat{\gamma}$).
Significance: *** : level 0.01 ; ** : level 0.5 ; * : level 0.1 ; none: non-significant at the 0.1 level

Table 8: Standard deviation and Maximum score

Standard deviation regression						
Trend	$\hat{\gamma}$	$\hat{\theta}_{IWC}$	$\hat{\theta}_{JR}$	$\hat{\theta}_{WA}$	$\hat{\theta}_{WS}$	$\hat{\theta}_{\text{Standard deviation}}$
-0,022***	0,138***	0,035***	0,046***	0,030***	-0,003	0,093***
Maximum score regression						
Trend	$\hat{\gamma}$	$\hat{\theta}_{IWC}$	$\hat{\theta}_{JR}$	$\hat{\theta}_{WA}$	$\hat{\theta}_{WS}$	$\hat{\theta}_{\text{Maximum opinion}}$
-0,023**	0,137***	0,037***	0,003	0,023***	-0,006	0,051***

Note to the reader: The interpretation of the estimates is the same as for the lower part of Table 7 except for the last column. $\hat{\theta}_{\text{Standard deviation}}$ and $\hat{\theta}_{\text{Maximum opinion}}$ refer to the respective influence of the empirical standard deviation of the opinions and of the maximum opinion.
Significance: *** : level 0.01 ; ** : level 0.5 ; * : level 0.1 ; none: non-significant at the 0.1 level