Expert Ratings and Bordeaux *En Primeur* Prices[1]

Charles Wyplosz
The Graduate Institute and CEPR

1. Introduction

Few people feel that they can identify good wines, and many people actually do not care much. But those who do feel a bit helpless when they set out to purchase wines. They seek help and find some from sellers in specialized stores and from expert ratings published in books and reviews. As more people have become interested in quality purchase, expert ratings have been growing fast, witness the number of guides and books that are now available. An interesting question is whether these ratings allow customer to find attractive deals or merely to buy wines at prices that reflect their quality. Since these ratings are publicly available, usually at moderate cost, the market for wine information should be efficient and wine prices should reflect quality is the experts are really able to detect quality. Assuming that experts are competent, wine market efficiency would imply that wine prices are indeed related to expert ratings so that the information asymmetry problem is greatly alleviated.

The information asymmetry is probably greatest for *en primeur* wine sales, which dispatch a significant portion of Bordeaux wine production. This is a futures market where customers buy wines not yet bottled, not even assembled and certainly far from drinking maturity. The wines cannot be tasted by plain customers but are presented to experts before the market opens. This means that the only information that potential customers have is what experts say, in addition to possible knowledge of the year's weather conditions[2].1 Since wine producers too have access to expert ratings, they are likely to use this additional information to set prices. Put differently, expert ratings are expected to affect prices. This paper examines whether this is indeed the case and, if so, how sensitive are prices to the ratings.

Previous papers have examined this question and Hadj Hali (2008) provides a survey. Most of them look at auction prices, which means bottled wines, most of them quite old and all of them prestigious growths. A few papers (Hadj Hali, 2008; Dubois and Nauges, 2006) look at *en primeur* prices and explore the role of one expert, arguably the best know one, Robert Parker and seek to determine his influence. The present paper adds to this literature in two ways. First, it exploits a nearly exhaustive data set that includes wines offered en primeur over eleven vintages. Second it relies on large number of publicly available expert ratings, averaged to represent a sort of "consensus view". This dataset shows that the ratings have a very clear and powerful effect on prices, which can be precisely estimated. As a by product, the analysis also identifies the effect of *appellations* and of vintages. It can be used to determine the price of individual wines that is justified by their attributes and to compare them with the actual *en primeur* prices.

2. The Data Set and Estimation Issues

As well explained in Hadj Hali (2008), Bordeaux wines produced in year *t* are offered in the spring of year *t*+1 – usually from late April to early July – for immediate purchase and delivery in the early months of year *t*+3. At that stage, the wines are far from having reached the evolution that they will have when bottled. They are not even assembled and most of them will undergo *elevage* in oak barrels for another two years, which will undoubtedly change their final taste. Even if practices are reasonable stable from one year to another, the evolution between *en primeur* sales

[1] I am most grateful to Alain and Evelyne Bringolf, the creators of Winemega, for having shared with their database.
[2] Ashenfelter (2008) argues that climate and estate explains almost all of the Bordeauy price variance.

and bottle release is relatively unpredictable since the wine growers can vary the proportion of new barrels and the duration of aging. There are many reasons for the *en primeur* system, which goes back to the 19th century. It allows wine amateurs to ensure that they will be able to obtain wines that are often hard to find once they are bottled. It provides producers with cash long before the bottles leave their cellars. Partly for that reason, *en primeur* prices are normally significantly lower than release prices two years later.

Another reason is that *en primeur* purchase is risky, like any futures contract. Customers cannot taste the wines, so they cannot judge for themselves how good the wines are and which ones better fit their tastes. They have to make a bet, informed by past experience, knowledge of previous vintages and expert opinions. Indeed, experts are invited in March of year *t*+1 to taste the upcoming wines. At that stage, it takes considerable expertise to guess what the wines will be like when they released. Since they not yet assembled, there is no guarantee either that the final product will be identical to the sample presented at that stage. Not infrequently experts revise their judgment. At any rate, several books, reviews and, increasingly, websites, scramble to produce their ratings before the opening of the *en primeur* season.

This means that wine producers know their individual ratings before they set their prices. It stands to reason that this knowledge will affect their pricing decision. Technically, the ratings are clearly exogenous to the prices and the cause-and-effect impact can, in principle, be detected through OLS. However Dubois and Nauges (2006) argue ratings and prices are jointly endogenous to intrinsic wine quality, which invalidates OLS[3]. Their purpose is to detect the influence of Robert Parker over and above other quality indicators. The objective in the present paper is to determine how quality influences prices, using expert judgment as an indicator of quality.

Since 1997, the website Winemega (http://www.winemega.com/) collects each year the ratings of experts on nearly all wines offered en primeur and rated by professional experts. The latest collection, which concerns the 2007 vintage, included 23 experts[4].

3 For each wine, we use the average rating. Not all wines are rated by all experts but each wine is rated by at least 3 experts and, on average, by ten of them. This is important since it reduces idiosyncratic ratings. Averaging the ratings over all experts for each wine provides a measure of the collective judgment. When needed, the ratings are converted into the US scale 50-100, see the website for further details.

Table 1 provides a summary for the latest vintage. There were 320 wines rated by 23 experts (the number of expert ratings collected has increased steadily over the years) with the number of wines rates by each expert indicated in the last lines of each subpanel. The range is pretty narrow, from 79 to 99. Experts vary in severity and some use narrower ranges than others. Pairwise correlations –not shown – are all positive and vary from 0.28 to 0.99. The highest correlation concerns two experts who rated about one hundred wines each, obviously focusing on the best-known ones.

[3] Their own estimates indeed suggest that OLS estimates are biased. Since quality is not observable, one must estimate it. Their solution is to use a two stage non-parametric method which relies on "instruments" that determine quality. Their instruments are annual weather data, which are the same for all wines in a given year and can be captured by dummy variables, and the official classification of wines which raises a number of serious issues. As is well-known, there is considerable heterogeneity of quality within nearly all quality ranks. In addition, better ranking often allows producers to charge higher prices than warranted by quality. In some cases, higher prices are used to undertake more expensive production techniques which may, or may not, result in better wines, especially as "better" is a matter of taste that will vary from one expert to the other. It follows that that the official rankings are noisy signals of wine quality.
[4] The list is available from the website. It excludes Robert Parker who explicitely refused to be part of the panel. Most other well-known and less well-known experts are included.

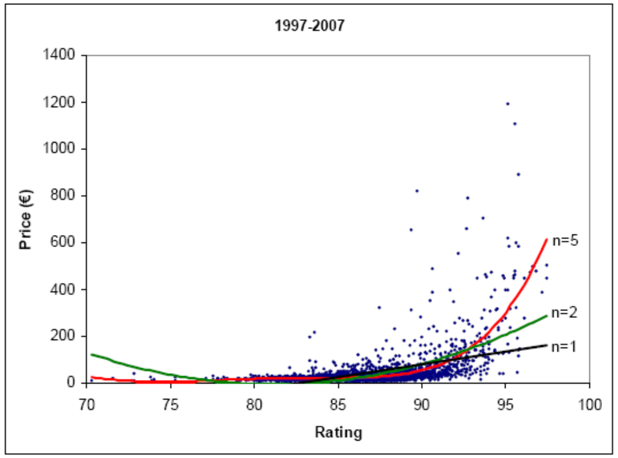
Table 1. Expert ratings for the 2007 vintage

Expert	1	2	3	4	5	6	7	8
Mean	89.3	88.1	89.5	88.9	89.6	89.0	89.2	89.1
Maximum	96.5	95	96.4	96	97.5	97	98.9	96.5
Minimum	81.4	83	83.6	82	83.3	79	82.5	82.7
Std. Dev.	3.1	1.6	3.1	3.3	2.8	2.7	3.0	2.9
Observations	195	67	122	179	98	292	214	222
Expert	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Mean	89.2	90.0	91.7	89.4	91.8	90.4	91.8	89.3
Maximum	96.8	100	100	93.9	98	95	99	95.3
Minimum	81.2	86	85	81.4	85	83	86.3	81.7
Std. Dev.	3.2	2.8	2.6	3.0	3.1	3.0	3.0	3.0
Observations	120	216	284	104	136	107	63	151
Expert	**17**	**18**	**19**	**20**	**21**	**22**	**23**	
Mean	89.5	89.3	88.5	88.5	87.6	89.3	89.4	
Maximum	95.9	95.3	98.5	95	94	96	97	
Minimum	84.1	82	84.6	82.5	76	85	83	
Std. Dev.	3.2	2.4	3.0	2.4	3.0	2.9	3.1	
Observations	216	116	194	132	125	46	99	

Winemega also collects *en primeur* prices from VinorumCodex (http://www.vinorumcodex.com/). They are quoted in euros, using the conversion rate between the French franc and the euro for wines on sale before the advent of the single currency. The prices are not adjusted for inflation. The reason is that the average price of Bordeaux wines has increased by 41% over the sample period while the Consumer Price Index only rose 16.4%. Such an important relative price increase means that adjusting for inflation does not serve much of a purpose.

For the eleven years of the sample, the sample includes complete information on 2899 wines, with a minimum of 196 for the 1999 vintage and a maximum of 320 for the vintages since 2001. Figure 1, which collect the 2899 observations, shows what is to be explained.

Figure 1. Prices and ratings



Note: All wines for vintages from1997 to 2007. The trends are polynomials of degree 1, 2 and 5.

Œnometrie XVI – Namur – 2009
Session 2

The link between prices and ratings is very strong but highly non linear. This is seen with the three trend lines also displayed on the figures. The trends are calculated as polynomials of degrees 1, 2 and 5, respectively. We have to go (at least) to the fifth degree to avoid a nonsensical negative relationship between price and rating on part of the trend. This non-linearity reflects the fact that some wines fetch astronomical prices as ratings become high. It can be noted that the highest prices do not always correspond to the highest ratings. For example, four wines cost more than € 800: Le Pin 2001 at € 820 for a rating of 89.7 and in 2006, Petrus (€ 890 for a rating of 95.8), Le Pin (€ 1200, rating: 95.1) and Ausone (€ 1100, rating: 95.6).

This non-linearity means that elasticity of prices to ratings increases with the rating (or the price) and that the functional form of the regression is an important issue to explore.

3. Estimates for the Whole Sample

In addition to the ratings, quality and therefore prices should be driven by the vintage, which reflects the weather conditions and the *appellation*, which may be an indication of quality and a source of market segmentation[5].4 Both are identified by dummy variables. The wines are classified in ten appellations:

- GR: Graves, including Pessac-Leognan
- MA: Margaux
- ME: Medoc and Haut Medoc
- PA: Pauillac
- PO: Pomerol and satellites
- SA: Sauternes and Barsac
- SEM: Saint Emilion and satellites
- SES: Saint Estèphe
- SJ: Saint Julien
- Others: Bordeaux, Bordeaux Supérieur, Fronsac, Canon-Fronsac, Listrac, Moulis.

The sharply non-linear relationship visible from Figure 1 requires a search for the most appropriate specification. If *pit* is the price, *zit* the average rating, *X* the vintage and *appellation* dummy variables, we examine the following specifications:

Linear: $p_{it} = c + \beta z_{it} + X\gamma + \varepsilon_{it}$ (1)

Exponential: $\ln p_{it} = c + \beta z_{it} + X\gamma + \varepsilon_{it}$ (2)

Double exponential: $\ln(\ln p_{it}) = c + \beta z_{it} + X\gamma + \varepsilon_{it}$ (3)

Log-log: $\ln p_{it} = c + \beta \ln z_{it} + X\gamma + \varepsilon_{it}$ (4)

Exponential-power function: $\ln p_{it} = c + \beta (z_{it})^{\alpha} + X\gamma + \varepsilon_{it}$ (5)

All equations are estimated by OLS. Equation (5) is estimated with a grid search over α, using the value that maximizes the log-likelihood. The choice of the specification will be

based on the standard deviation between the actual and fitted prices.

The results of estimation of Equations (1) to (5) are shown in Table 2. The variables Dxx are the annual dummies, all other variables are as indicated above. Most variables are highly significant. For each equation, the table reports the standard deviation of the difference between actual prices and their corresponding fitted values (not the standard deviation of the dependent variable which differs across equations). The standard deviation clearly indicates that (5), where $\alpha = 14.2$, offers the best fit. The better fit of (5) is readily confirmed by Figure 2, which displays simulations

[5] One could add the ranking (growth). This is left for future study.

of the estimated Equations (1) to (5). This specification implies that predicted price rises very sharply as the rating increases, the curve becoming nearly vertical for ratings close to 100[6].

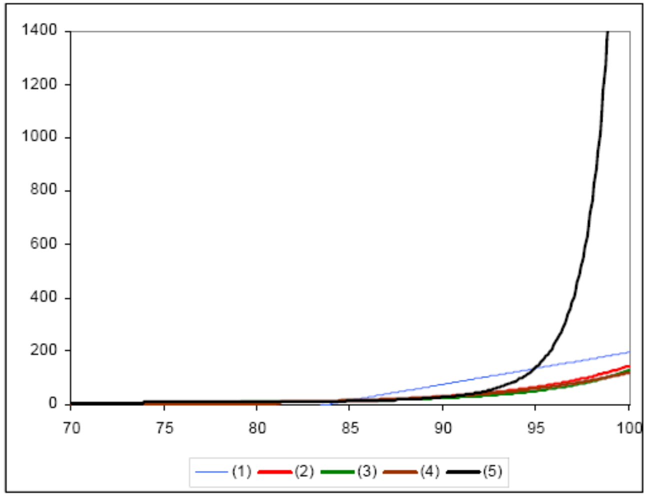
The table also reports the estimated semi-elasticity $(1/\ p)(\partial p\ /\ \partial n)$ computed at the sample median value z = 86.75 . The preferred Equation (5) indicates that a one point increase in the rating is accompanied by a 13.7% price increase for the median wine. The semi-elasticity $\alpha\beta n\alpha -1$ ranges from 4.7% for z = 70 to 78.4% for z = 99. This explains why experts are given red-carpet treatment in the top estates.

Table 2. Estimation results for the whole sample

	(1)			(2)			(3)			(4)			(5)		
	Coefficient	t-Statistic	Prob.	Coefficient	t-Statistic	Prob.	Coefficient	t-Statistic	Prob.	Coefficient	t-Statistic	Prob.	Coefficient	t-Statistic	Prob.
Rating	12.001	15.277	0.000	0.158	31.821	0.000	0.043	35.133	0.000	13.408	29.692	0.000	2.5E-28	38.476	0.000
GR	5.594	1.371	0.170	0.351	7.626	0.000	0.123	8.832	0.000	0.358	7.667	0.000	0.344	8.824	0.000
MAR	11.574	3.223	0.001	0.436	10.610	0.000	0.149	12.019	0.000	0.443	10.611	0.000	0.410	11.863	0.000
ME	3.273	1.071	0.284	-0.042	-1.013	0.311	-0.025	-1.854	0.064	-0.038	-0.919	0.358	-0.075	-2.138	0.033
PA	9.468	2.153	0.031	0.448	10.271	0.000	0.152	11.859	0.000	0.461	10.363	0.000	0.402	10.889	0.000
PO	27.706	5.733	0.000	0.607	12.842	0.000	0.189	13.401	0.000	0.617	12.921	0.000	0.581	13.653	0.000
SA	-18.590	-4.464	0.000	0.153	3.393	0.001	0.083	6.195	0.000	0.170	3.731	0.000	0.117	2.907	0.004
SEM	14.346	3.887	0.000	0.525	13.333	0.000	0.174	14.462	0.000	0.535	13.426	0.000	0.509	14.847	0.000
SES	0.142	0.029	0.977	0.251	5.424	0.000	0.093	6.667	0.000	0.258	5.513	0.000	0.259	6.388	0.000
SJ	-9.507	-2.575	0.010	0.335	7.371	0.000	0.130	9.414	0.000	0.346	7.447	0.000	0.349	8.945	0.000
D98	-14.548	-3.209	0.001	-0.112	-2.084	0.037	-0.019	-1.219	0.223	-0.109	-2.025	0.043	-0.082	-1.511	0.131
D99	12.972	2.253	0.024	0.295	4.631	0.000	0.093	5.259	0.000	0.297	4.643	0.000	0.288	4.667	0.000
D00	17.545	3.350	0.001	0.365	7.019	0.000	0.114	7.684	0.000	0.365	6.981	0.000	0.350	6.850	0.000
D01	25.232	4.760	0.000	0.497	9.737	0.000	0.153	10.597	0.000	0.497	9.691	0.000	0.460	9.238	0.000
D02	3.832	0.896	0.371	0.134	2.655	0.008	0.048	3.247	0.001	0.131	2.599	0.009	0.157	3.099	0.002
D03	-9.489	-2.455	0.014	-0.103	-1.994	0.046	-0.026	-1.688	0.092	-0.102	-1.978	0.048	-0.076	-1.460	0.145
D04	-6.204	-1.611	0.107	-0.043	-0.880	0.379	-0.006	-0.437	0.662	-0.042	-0.865	0.387	-0.031	-0.645	0.519
D05	1.065	0.210	0.834	0.110	2.122	0.034	0.043	2.938	0.003	0.114	2.200	0.028	0.138	2.712	0.007
D06	-17.041	-2.612	0.009	-0.186	-3.154	0.002	-0.035	-2.079	0.038	-0.170	-2.860	0.004	-0.163	-2.911	0.004
D07	-25.430	-4.878	0.000	-0.248	-4.636	0.000	-0.054	-3.556	0.000	-0.238	-4.413	0.000	-0.189	-3.645	0.000
N. obs.	2899			2899			2899			2899			2899		
Adjusted R2	0.275			0.549			0.568			0.535			0.634		
St. Dev.	64.044			63.740			63.420			64.570			57.560		
Semi-elasticity	0.328			0.158			0.118			0.155			0.137		

Notes: Constant not reported. White heteroskedasticity-consistent standard errors and covariances. The vintage base is 1997, the appellation base is "all others". Standard deviation of *pit* − *p*ˆ*it* where *p*ˆ*it* is the fitted value of wine prices. The elasticity is computed at the sample median.

Figure 2. Simulated estimated equations



[6] Note that one cannot "rank" *à priori* the steepness of Equations (1) to (5) without knowing the values of the parameters.

Œnometrie XVI – Namur – 2009
Session 2

The estimation also sheds light on other component of prices. Based on Equation (5), the *appellations* bring a premium over the "other" category that serves as benchmark, which can be substantial as further illustrated in Figure 3. For instance, everything else equal, a Pomerol wine is estimated to be 58% more expensive than the benchmark. The exception is the Medoc/Haut Medoc region, arguably a large mixed bag, which suffers from a 7% average discount, relatively small but significant at the 5% confidence level.

Figure 3. Appellation effect (%)

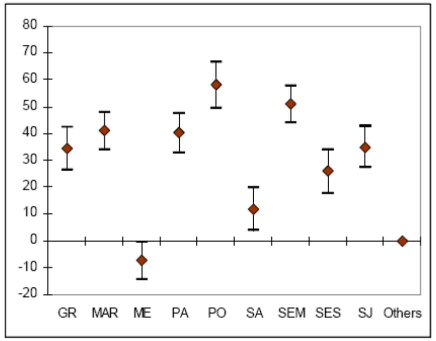
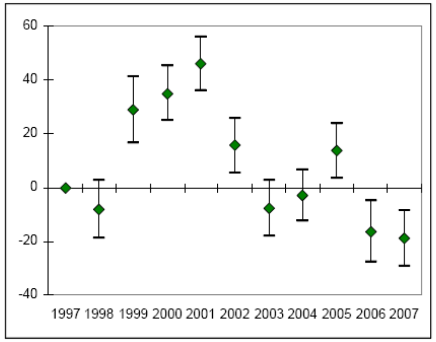


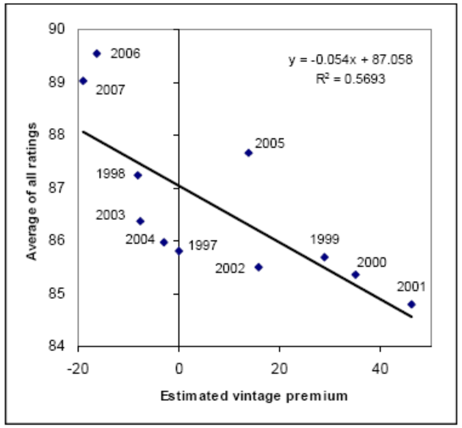
Figure 4. Vintage effect (%)



Note: Using Equation (5), this figures show the point estimates and two-standard error bands.

The effect of vintages is illustrated in Figure 4, which is patterned after Figure 3. This figures shows, for example, given the ratings and the appellation, a wine was on average 46% more expensive than in 1997, taken as the base year. Thus 1999, 2000, 2001 and 2005 are found to be pricy vintages while 2006 and 2007 are found to be underpriced. Unsurprisingly perhaps, 2000 and 2005, the two vintages identified as "vintage of the century" by the professional rumor mill, command a premium but they are relatively moderate given the hype that surrounded the corresponding *en primeur* seasons. Since the estimates of the vintage effect takes ratings as given, an interesting question is whether experts are tougher in good years and inflation-prone in difficult years. This hypothesis can be tested by looking at the correlation between the vintage premium shown in Figure 4 and the average of all wine ratings in every year. Figure 5 shows that this is indeed the case. Average ratings tend to be markedly higher in modest years, like 2006 and 2007, although 2000 stands as an outlier. The discomforting implication is that expert ratings are systematically biased to offset intrinsic quality changes from vintage to vintage. For instance, ratings for the vintage 2006 would have to be lowered by a full 1.5 point in 2006 to be brought down to the trendline. According to the estimated semi-elasticity, this would translate into prices too high by about 20% for the median wine in that year, a not insignificant impact. Note also that the last three years, 2005 to 2007 tend to lie above the trendline, which could signal a lasting grade inflation tendency.

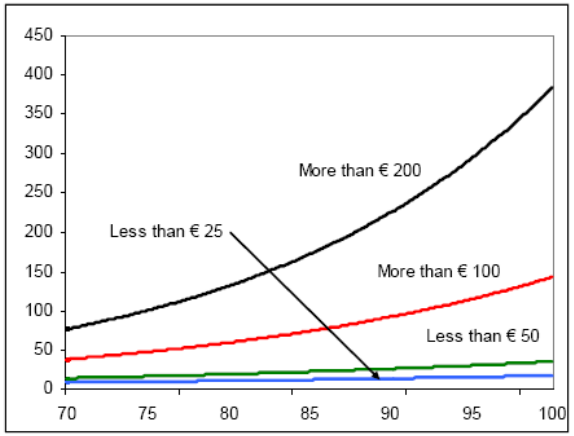
Figure 5. Rating inflation in bad years?



4. Subsample estimates

Figure 1 suggests that prestige wines, which command very high prices, may be in a different category. This would be in line with market segmentation, which is designed to attach special value to products whose brand name can be easily recognized. It would also be decline dramatically relative to those in Table 2. Simulations of the corresponding relationship between prices and ratings are shown in Figure 7. The differences are dramatic. For example, for an average vintage year and no *appellation* effect, a wine that was sold € 200 would fetch as much as € 77 if its ratings would be a disastrous 70 while a wine that sells less than € 25 would only fetch € 9 (still a respectable amount!). Of course, such a wine would never receive such a poor rating, if only because top estates have access to resources that essentially guarantee quality. The point, here, is rather that we face difference categories of wines, which calls for separate estimations.

Figure 7. Estimated slopes



Œnometrie XVI – Namur – 2009
Session 2

For each subsample ($p \le 25$, $p \le 50$, $p > 100$ and $p > 200$), all five equations (1) to (5) were reestimated and the best fits, determined as before with the standard deviation of $pit - p\hat{}it$ where $p\hat{}it$ is the fitted value of wine prices, are reported in Table 3. The simulated relationships are shown in the Appendix, Figure A.1. Over the range of ratings 70-100, there is little difference between the relationships found for $p \le 25$ and $p \le 50$ even though Equation (5) is the best fit for $p \le 50$ while the linear Equation (1) dominates for $p \le 25$. For both $p > 100$ and $p > 200$, the highly non-linear Equation (5) dominates. Interestingly, the parameter α increases sharply as the as we move from subsample $p \le 50$ to subsample $p > 100$ and to susbsample $p > 200$. This explains why the semi-elasticities, measured for the median rating in each subsample, increase as we move to higher price subsamples, except for the susbsample $p > 200$ where it starts rising steeply for ratings slightly above the median (See Figure A.1). This confirms that the higher is the price the more prices react to ratings.

5. Conclusions

The effect of expert ratings on *en primeur* prices is clear, even upon visual inspection, but it is more subtle than meets the eyes. It is extremely nonlinear, limited for cheaper wines and extraordinary powerful for the expensive wines. This is pretty intuitive. Customers who spend € 200 per bottle or much more, for enjoyment or speculation, have every reason to collect as much information as possible since most of them cannot taste the wine before they commit to *en primeur* purchases.

In fact, the market seems highly segmented. The link between prices and ratings differs considerably between cheaper and expensive wines. Not only do elasticities differ but does the influence of vintages and even *appellation*. Pomerol, Saint-Emilion and to a lesser extent Pauillac benefit from a powerful luxury image not just at the top end of the market. This may be another segmentation that should be examined in subsequent work.

Vintages, which capture weather and, maybe economic conditions matter most at the low end of the market. The prices of wines in the € 100 - € 200 range do not seem to much affected by the vintage effect, which surprisingly resurfaces at the top end.

About 85% of *en primeur* wines sell for less than € 50. The price of the median bottle in that market segment stands to rise by 7.3% when the average expert rating increases by one point. Experts do have a highly significant and not negligible effect. This could well create a conflict of interest for them. In principle they work to inform the broad public and their judgments should be beyond suspicion. The fact that their ratings may raise prices by huge amounts at the top end of the market is a source of concern. Also intriguing is the fact that ratings tend to be higher in poor years. Another puzzling conclusion is that expert ratings make it less likely that "cheap gems" can be found in the *en primeur* market since favorable ratings are immediately priced.

This study is a first exploration of a large database. There are many ways in which the analysis can be extended: a more detailed analysis of market segments, the differential impact of individual experts, the possible effect of growth ranking (not yet in the data base), the time response of prices to ratings, and much more. More work is also needed to improve upon the admittedly crude econometric analysis.

Appendix

Table A1.

	(2)			(3)			(4)		
	Coefficient	t-Statistic	Prob.	Coefficient	t-Statistic	Prob.	Coefficient	t-Statistic	Prob.
Rating	0.037	18.085	0.000	0.012	16.317	0.000	2.488	14.097	0.000
Less than € 25	-0.007	-55.123	0.000	-0.002	-52.028	0.000	-0.132	-55.335	0.000
Less than € 50	-0.007	-39.957	0.000	-0.002	-36.762	0.000	-0.142	-41.029	0.000
More than € 100	0.007	24.744	0.000	0.001	20.210	0.000	0.143	26.620	0.000
More than € 200	0.010	19.514	0.000	0.002	17.591	0.000	0.208	19.883	0.000
N. obs.	2899			2899			2899		
Adjusted R2	0.910			0.868			0.910		
St. Dev.	32.939			34.508			33.262		

Note: Constant and other variables not shown. Equations (2), (3) and (4) indicated in text. The second to fourth rows correspond to the interaction of the dummy indicated in the first column with the ratings. White heteroskedasticity-consistent standard errors and covariances. The vintage base is 1997, the appellation base is "all others". Standard deviation of $p_{it} - \hat{p}_{it}$ where $\hat{p}_{it}$ is the fitted value of wine prices.

Figure A.1. Slopes estimate on different samples