AMERICAN ASSOCIATION OF WINE ECONOMISTS

AAWE WORKING PAPER
No. 1

Editor
Victor Ginsburgh

THE IMPACT OF GURUS: PARKER GRADES AND *EN PRIMEUR* WINE PRICES

Héla Hadj Ali
Sébastien Lecocq
Michael Visser

April 2007

www.wine-economics.org

The impact of gurus:

Parker grades and *en primeur* wine prices *

Héla Hadj Ali†, Sébastien Lecocq‡, Michael Visser§

September 2005

Abstract

The purpose of this paper is to measure the impact of Robert Parker's oenological grades on Bordeaux wine prices. We study their impact on the so-called *en primeur* wine prices, i.e., the prices determined by the château owners when the wines are still extremely young. The Parker grades are usually published in the spring of each year, before the wine prices are established. However, the wine grades attributed in 2003 have been published much later, in the autumn, *after* the determination of the prices. This unusual reversal is exploited to estimate a Parker effect. We find that, on average, the effect is equal to 2.80 euros per bottle of wine. We also estimate grade-specific effects, and use these estimates to predict what the prices would have been had Parker attended the spring tasting in 2003.

Keywords: Expert opinion, natural experiment, treatment effect, Bordeaux wine price.

JEL codes: C21, D89, L15.

*We are grateful to Alan Duncan, Fabrice Etilé, Victor Ginsburgh, Sylvie Lambert, and participants at the VDQS conference (Dijon, 2004), the RES conference (Nottingham, 2005), the JMA meeting (Hammamet, 2005), and the EEA congress (Amsterdam, 2005), for their helpful comments and suggestions.

†INRA, Chemin de Borde Rouge, Auzeville, BP 27, 31326 Castanet Tolosan, France. Email: hadjali@toulouse.inra.fr.

‡INRA, 65 boulevard de Brandebourg, 94205 Ivry, France. Email: lecocq@ivry.inra.fr.

§INRA, 48 boulevard Jourdan, 75014 Paris, France. Email: michael.visser@ens.fr.

1 Introduction

The judgment of experts and gurus matters in many types of markets. For instance, in art markets for books, music, and movies, opinion leaders have an important effect on consumer decisions via their rankings in all sorts of guides and competitions (see Ginsburgh (2003)). Similarly, in auctions of paintings, art experts influence sale prices via the publication of their pre-sale estimates in auction catalogues (see Bauwens and Ginsburgh (2000)). Other markets where expert opinion leaders are very active and influential are sports betting markets (see Avery and Chevalier (1999)), financial stock markets (see Shleifer (1986)), the market for restaurants (see Chossat and Gergaud (2003)), etc...

Given the importance of experts and gurus, it seems of interest to be able to precisely measure their impact on market outcomes. In this paper we consider a particularly well known expert, the American wine critic Robert Parker, and study the effect of his opinion on Bordeaux wine prices. More specifically, we analyze the effect of Parker's opinion on the so-called *en primeur* wine prices. These are prices set by Bordeaux châteaux owners just 6 or 7 months after the grape harvest, i.e., when the wines are still very young and not yet bottled.[1]

Robert Parker is widely regarded as the most influential wine expert in the world.[2] It is often claimed by those working in the wine industry that his wine reports and oenological grades, published in the bimonthly journal *The Wine Advocate* and in his wine books, exercise much power over wine prices: good grades may cause prices to rocket sky-high, while, inversely, very bad grades

[1] Bordeaux *en primeur* wines should not be confounded with Beaujolais *primeur* wines. The latter are specially produced and bottled for immediate consumption (starting in November, just two or three months after the harvest), while the former first go through the usual vinification process (a period lasting between 1 one and two years), in the cellars of the châteaux, before being bottled and put on the market.

[2] One can find many quotes about Robert Parker. Here are some found on the web site http://www.erobertparker.com: "Parker is the most influential wine writer in the world today." (The Los Angeles Times); "Parker is the world's most experienced and trustworthy taster." (Jancis Robinson, another well known wine taster and critic); "The man with the paragon palate... For countless wine lovers Robert Parker's tastes are infallible." (Time Magazine).

may leave wines practically unsaleable unless their prices are adjusted downwards. His power over prices is thought to be particularly important for *en primeur* wines.[3] A possible explanation for this is that consumers do not have easy access to the *en primeur* market (recall that *en primeur* wines are not yet bottled). As a consequence they cannot judge the products themselves, and thus require the judgment of wine connoisseurs.

To assess the effect of Parker's opinion on *en primeur* prices, we exploit the following natural experiment. Each spring (since 1994) Robert Parker comes to Bordeaux to taste and evaluate a number of *en primeur* wines of the latest vintage. He publishes his findings in *The Wine Advocate*, usually in the April issue. The *en primeur* prices are fixed by the château owners in the weeks and months thereafter (the *en primeur* market usually opens at the end of April, and lasts until June), giving them the possibility to incorporate the information contained in the Parker grades. In 2003 things went different because the wine expert did not come in the spring to taste the 2002 vintage. According to Parker himself, the trip was cancelled at the request of his family.[4] The French press however argued that his absence could be explained by the judicial affairs in which one his collaborators was involved at the time (see for example *Libération*, July 1, 2003). Whatever the reasons, it meant that château owners had to determine their prices *without* knowing the grades that Parker would have attributed. So, as Jancis Robinson put it, "the Bordelais are having to re-learn the art of selling a whole vintage, 2002, without his help".[5]

We use this unusual reversal to estimate a Parker effect on *en primeur* prices. Adopting the counterfactual framework introduced by Rubin (1974), this effect is defined as the average treatment effect on the treated. It measures the mean effect of the attribution of Parker grades for those wines that were actually graded. This commonly-used evaluation parameter is identified under a "parallel trend" assumption. Roughly speaking, this assumption states that, had Parker not graded any wine in two subsequent years (which is of course a hypothetical situation), the price evolution over the two years would have been the same for all wines. Under the parallel trend assumption,

[3] See the wine survey "The Globe in a Glass" published in *The Economist*, 18 December 1999, page 102.

[4] Early in 2003, his wife and daughter preferred the trip to be postponed in view of the threat of a war in Iraq.

[5] See http://www.jancisrobinson.com/jr7046.htm.

the treatment effect can be estimated by a difference-in-differences type estimator. Using data on *en primeur* prices established in 2002 (2001 vintage) and 2003 (2002 vintage), we find an estimate of the Parker effect equal to 2.80 euros per bottle of wine. Estimates by appellation show that the treatment effect is the largest for wines from Pomerol, which is quite interesting since these are precisely the wines that Robert Parker appreciates the most. Using data on the grades assigned in 2002, we also estimate grade-specific treatment effects, and show that they steeply increase with the level of grading. Finally we exploit the fact that the grades eventually attributed by Parker to the 2002 vintage (he came to Bordeaux in September 2003) are also observed in our data. Under the hypothesis that these marks have the same interpretation as the ones assigned in the spring (we will explain that this is a plausible hypothesis), we use the autumn grades to predict what the prices would have been had Parker attended the spring tasting in 2003.

The next section summarizes the empirical findings of some other recent papers on the Robert Parker factor. Section 3 describes our data set, Section 4 gives the definition of the Parker effect, explains how it can be estimated, and presents our empirical results. Section 5 concludes.

2 Review of literature

Several papers have looked at the influence of Parker marks on wine prices. They are almost all based on the estimation of so-called hedonic price functions, where prices are regressed on a set of characteristics in order to determine which characteristics have a significant effect. Some of these papers compare the impact on prices of official classifications (such as the 1855 classification in the case of Bordeaux wines) with the impact of Parker's own classification, but most analyze the marginal effect of Parker grades (i.e., the β-coefficient in the regression of prices on grades).

In what is probably the earliest article on the subject, Ginsburgh et al. (1994) apply the hedonic price method to a sample of 102 Médoc wines. Their data set allows them to disentangle the price effects of weather, reputation (as measured by the 1855 classification), natural endowments (soil, exposure of the vineyards or grapes), and all sorts of production factors. They show that technology

and weather conditions explain two thirds of the price variation, and once the reputation variable is added the proportion of explained variance increases to almost 85%. They also show that more recent classifications (such as Parker's classification) do not lead to a better fit of the hedonic price equation than the 1855 classification.

Di Vittorio and Ginsburgh (1996) regress auction prices, gathered during public sales held at Christie's London, of 58 top-classified growths from the Haut-Médoc region on the year of vintage and the name of the château. The estimated hedonic function allows them to calculate, for each château and year of vintage, a price index, and these price indexes are compared with several classifications (those of Parker, the wine critics Tastet and Lawton, and the wine magazine *Wine Spectator*). They find that the 1855 classification does better in explaning their price indexes than any other classification.

Jones and Storchmann (2001) use wine auction prices for 21 prestigious Bordeaux wines to asses the respective effects of the year of vintage, the grape composition (acid and sugar levels), and quality (as measured by Parker points). The authors show that the influence of Parker grades on prices is large and varies between châteaux: the price increase induced by a one-point increase in the Parker rating is 7% on average, ranging from less than 4% to more than 10%. The price sensitivity to Parker grades is found stronger for Cabernet Sauvignon-dominated wines than for Merlot-dominated wines, for wines that have achieved high ratings in the past, and for small properties.

Février et al. (2003) study the role of the buyer's option in multi-unit ascending (or English) auctions using data collected during an important sale of fine wines held in 2000 at Drouot, the largest auction house in Paris. The estimation of an auction model allows them to determine which characteristics influence the willingness to pay for a wine (the production region, the vintage, whether the label is damaged or not, the level of wine in the bottle, the Parker-rating, etc.). Parker grades are taken from a Parker guide published in 1995, wherein each wine is ranked between one star (lowest possible rank) and five stars (highest rank).[6] Because the grading system from this

[6]This contrasts with the other mentioned studies, where wines are graded using a 50-100 point scale.

particular Parker guide reflects the overall and long-term quality of each wine and does not depend on the vintage, the authors also add a grade, taken from a French wine website, reflecting the quality of the vintage. Their results show that the value bidders attribute to a wine significantly increases with Parker's classification.

Hadj Ali and Nauges (2004) exploit a panel data set covering 132 Bordeaux châteaux over sixteen vintages (1982-1983, 1985-1998) to estimate the influence of current quality (as measured by Parker grades and by an overall vintage score given by *Wine Spectator*) and reputation (as measured by the ranking of the wine and by the average of the scores Parker gave to previous vintages) on the pricing of *en primeur* wines. They find that prices mainly depend on the reputation of the châteaux through their ranking. They also find that the marginal impact of Parker grades on *en primeur* prices is significant but small: a one-point increase in the grade leads to a price increase of 1.01%.

Dubois and Nauges (2005) use a subsample of the Hadj Ali and Nauges (2004) data set (108 châteaux over the 1994-1998 vintages). They argue that true quality, which is known by the producer but unobserved by the consumer and the econometrician, will not only influence the pricing of *en primeur* wines but also the Parker-rating, implying an endogeneity issue which may produce biased estimates. Controlling for unobserved quality allows them to disentangle the effect of quality from the effect of Parker grades on prices. They find evidence that Parker grades significantly affect the price of *en primeur* wines: a one-point increase in the grade increases the price by 1.38%. They also show that the impact of Parker's judgment is clearly over-estimated when unobserved quality is not accounted for since in this case it equals 3.95%, which is much larger than the effect found by Hadj Ali and Nauges (2004) using the whole data set.

There are many differences between the previous papers and the present paper, but we will only mention the two most important ones. The first is that most of the previous papers study the marginal impact of Parker grades on prices. We instead focus on the more extreme evaluation parameter "price when wine is graded minus price in the absence of grading". The second difference is that in the previous studies the time lag between the moment the grades are attributed and the moment the wine prices are established is relatively important. For instance in Jones and

Storchmann (2001), the wine prices are the prices fetched at auctions in 1996/1997, whereas the quality index is the Parker-rating of these wines published in 1994/1995. Conversely in Hadj Ali and Nauges (2004), prices are *en primeur* wine prices but grades are those Parker gave once the wines were bottled, a long time after the end of *en primeur* sales. The grades used in this paper are the ones assigned by Parker just a few weeks before the start of *en primeur* sales (2001 vintage) or just a few months after these sales (2002 vintage).

3 Data

Our data set combines two sources of information: the 2002 and 2003 *en primeur* prices for approximately 250 châteaux obtained from a Bordeaux wine broker, and Parker grades extracted from *The Wine Advocate*. Section 3.1 gives a short description of the Bordeaux *en primeur* market and explains how wine prices are defined, Section 3.2 outlines how Robert Parker evaluates and grades *en primeur* wines, and Section 3.3 contains a descriptive analysis of the matched data.

3.1 *En primeur* wine prices

The Bordeaux *en primeur* wine market is a forward market where wines are sold as futures. Traditionally the market opens at the end of April and lasts for just two or three months. The wines that are traded are those from the latest vintage. Thus the wines traded in the spring of 2002 for instance were from the 2001 harvest. Trading therefore takes place when the wines are still very young, unblended, and unbottled. The payments on this market are made when the sale contracts are signed, and delivery occurs once the wines are bottled, one to two years after the grape harvest. The sales are organized by brokers who act as intermediaries between the château owners and the *négociants* (who in turn sell the wine on to importers, retailers and private consumers). The *en primeur* prices are determined by the château owners themselves, and given these prices, the brokers negotiate with potential buyers about quantities, delivery dates, etc... The Bordeaux *en primeur* market exists since the 18th century and has become increasingly important in recent

years. Nowadays it generates a turnover of between 300 and 600 million euros per year, and some châteaux sell up to 80% of their production *en primeur* (see Hadj Ali (2002)).

The *en primeur* prices in our data correspond to first tranche prices (also called opening prices)[7] and they come from the archives of one of the largest brokers in Bordeaux. The full sample contains the prices of all the 375 Bordeaux châteaux traded by this broker during the period 1994-2003. This paper only uses the prices established in 2002 and 2003. Furthermore, we restrict ourselves to the 233 châteaux that were traded in both 2002 and 2003.[8] For each of the 233 châteaux in our sample we thus observe the 2002 (2001 vintage) and 2003 (2002 vintage) *en primeur* prices. Prices are in euros per bottle of 75cl. We also observe for each château its appellation group and its ranking.

Since the broker who gave us the price data is a large and important firm in Bordeaux, there is a good reason to believe that our sample is representative of the population of châteaux active in the *en primeur* market. Evidence for this comes from the fact that the sample covers a large spectrum of châteaux, ranging from the best-known wines such as Ausone, Lafite Rothschild, Latour, Margaux, and Mouton-Rothschild, to less prestigious ones such as Beaulieu, Chantegrive, Fonréaud, Pierdon, etc...

3.2 Parker grades

Each spring, usually late March or early April, Robert Parker comes to the Bordeaux region to taste *en primeur* wines from the latest vintage. Since the wines are not yet bottled, the infantile clarets are extracted from the barrels. The tastings are generally done in peer-group, single-blind conditions. Peer-group tasting means that wines belonging to the same region or appellation are evaluated relatively to each other. Single-blind tasting means that the products are judged without the producer's name being known, which has the advantage that neither the price nor the reputation

[7] Château owners sell their wine on the *en primeur* market in one, two or three tranches. The tranches are sold successively, which allows the producers to eventually adapt prices to market demand. The first tranche prices are by far the most strategic, awaited and scrutinized prices in the *en primeur* market.

[8] A given château is not necessarily traded by the broker in each year. If it is not traded in some year then its *en primeur* price is obviously not observed in that year.

of the château can affect the grading. The Parker grading system uses a 50-100 point quality scale. Unlike the more mature wines, to which Parker assigns single scores (for example 82),[9] he generally gives grade intervals to *en primeur* wines (for example $[88, 90]$). This probably reflects the fact that for him there is an extra element of uncertainty in judging the quality of such young wines. The quantitative rating of each wine is complemented by a verbal tasting note containing an overall description of the wine. The grade intervals and tasting notes of all the wines that are tasted in the spring are published in *The Wine Advocate*, usually in the April issue. These pieces of information can also be found on Parker's web site http://www.erobertparker.com. Using this site, we could determine which of the 233 châteaux in our sample were evaluated by Parker in the spring of 2002, and for those indeed evaluated we recorded the 2001 vintage grade intervals.

As explained in the introduction, things were different in 2003. Robert Parker did not come to Bordeaux in the springtime to taste the latest vintage, but only during the end of the summer (between August 30th and September 8th), so the 2002 vintage was tasted about 11 months after the harvest, unlike the earlier vintages which were evaluated roughly 6 months after the harvest. According to Robert Parker (see issue 149 of the *The Wine Advocate*), there was not much oenological difference in tasting the 2002s in August/September, as opposed to tasting the other vintages in March/April. This suggests that, *ceteris paribus*, we can attach the same meaning to the grade intervals attributed to the 2002 vintage as the ones assigned to the earlier vintages. The grade intervals for the 2002 vintage were published in the October issue of *The Wine Advocate* (issue 149), and can also be found on Parker's web site. We again consulted the web site, this time to check which of the 233 châteaux were evaluated in the late summer of 2003, and copied the relevant 2002 vintage grade intervals.

3.3 Descriptive statistics

Table 1 shows how the 233 châteaux in our data are ranked and from which appellation they originate. Note that about half of the Bordeaux appellations are represented in our sample (there

[9] See the wine reports in *The Wine Advocate* or his wine books.

are 25 different appellations in the sample, while there are roughly 50 of them in the Bordeaux wine region; see Combris et al. (1997)). Note also that there is much diversity in the ranking of the châteaux: about 28% of the châteaux are ranked according to the famous 1855 classification (1er Cru Classé, 2ème Cru Classé,..., 5ème Cru Classé), 14% are classified as Cru Bourgeois, 40% are not classified at all, etc...

Table 1 about here

Table 2 gives some simple descriptive statistics for the *en primeur* prices and the Parker appreciations. To make the 2002 and 2003 prices comparable, we have divided the latter by 1.019, where 1.9% is the evolution of the consumer price index over that period (source INSEE: http://www.insee.fr). Also, all grade intervals have been replaced by their midpoints (for example, the interval $[94, 98]$ is replaced by the midpoint 96).[10]

Table 2. *En primeur* prices and Parker grades

	# of obs.	Mean	Std. dev.	Min	Max
			2001 vintage		
Grade	158	88.52	2.42	82.5	97
Price	233	19.01	17.86	4	125
			2002 vintage		
Grade	121	89.40	2.19	86	96.5
Price	233	15.65	13.23	3.93	98.14

Notes: The grades correspond to the midpoints of the grade intervals.

The prices are per bottle of 75cl, in 2002-euros.

The price statistics are based on all 233 châteaux in our sample. The grade statistics for the 2001 vintage are based on the 158 wines (in our sample) that were evaluated by Robert Parker in the spring of 2002, while those for the 2002 vintage are based on the 121 wines tasted in the late

[10]There is not much variation in the observed intervals. In 91% of the cases the interval is either 1 or 2 points, in 8% it equals 3 points, and in 1% it is either 4 (1 case) or 11 points (1 case).

summer of 2003. As the table shows, the wines from the 2002 vintage received grades that are on average only slightly higher than those received by the wines from the 2001 vintage. Apparently, both vintages were deemed of similar quality. The two vintages differ however markedly in their prices. On average, the 2001 vintage is priced more than 3 euros higher than the 2002 vintage. The variance of the price for the 2001 vintage is also much higher than the one for the 2002 vintage. Recalling that the 2001 vintage prices were released after publication of Parkers' grades and that the 2002 vintages prices were released without this knowledge, this last finding suggests that more information leads to more price variation. We also find that the correlation between grades and prices is stronger for the 2001 vintage than for the 2002 vintage: Spearman's correlation coefficient equals 0.74 for the former and 0.64 for the latter.[11]

4 Estimating the Parker effect

In this section we give the precise definition of the Parker effect, explain how it can be estimated, and present the empirical estimates. The Parker effect will be defined by adopting the counterfactual framework pioneered by Rubin (1974). Let $D_i(t)$ be the treatment status indicator for observation i, for $i = 1, ..., 233$, in year t, for $t = 2002, 2003$. Thus $D_i(t) = 1$ if wine i is graded in the spring of year t, and $D_i(t) = 0$ otherwise. Note that since Robert Parker did not grade the 2002 vintage in the spring of 2003, we have $D_i(2003) = 0$ for all i. Let $P_{i0}(t)$ represent the *en primeur* price for wine i in case it is not graded in the spring of year t (wine i in the control group at t), and let $P_{i1}(t)$ be the price in case the wine is graded (wine i in the treatment group at t). Finally, let $P_i(t)$ be the realized price, i.e., $P_i(t) = D_i(t)P_{i1}(t) + (1 - D_i(t))P_{i0}(t)$. Note that $P_i(2003) = P_{i0}(2003)$ for all i. We assume that in each time period the treatment status of observation i affects only the outcome of observation i in that period (and not the outcomes of other observations). In the treatment literature this assumption is known as the Stable Unit Treatment Value Assumption (see Rubin (1980)). SUTVA is practically always made in empirical evaluation

[11] Pearson's correlation coefficient equals 0.64 for the 2001 vintage and 0.62 for the 2002 vintage.

studies (one exception that we are aware of is Heckman et al. (1998)). We shall do likewise but acknowledge that the assumption may be quite strong in our particular context. Indeed, SUTVA rules out any interference or relationship between observations. SUTVA thus excludes, for instance, that the grading of a particular château influences the wine prices set by other châteaux owners (through general-equilibrium or imitation effects).

A natural definition of the Parker effect for wine i would be the difference $P_{i1}(2002) - P_{i0}(2002)$. Since the 2 potential outcomes, $P_{i0}(2002)$ and $P_{i1}(2002)$, cannot both be observed (in 2002, wine i is either in the control group or in the treatment group), it is impossible to calculate wine-specific Parker effects. This is a fundamental problem in the evaluation of treatments in general. Since it is not possible to determine individual treatment effects, the literature has focussed on estimating average treatment effects, such as the average treatment effect on the treated (see Heckman (1990), Wooldridge (2002, chapter 18), Abadie (2005)). In our context this parameter is

$$ATET \equiv E[P_{i1}(2002) - P_{i0}(2002)|D_i(2002) = 1] \tag{1}$$

which is the expected Parker effect for those wines that were graded in 2002.

For what follows, it is useful to give the mean prices and standard errors, for both vintages, according to whether the wines are in the control or treatment group in 2002. This is done in Table 3.

Table 3. Mean prices (standard errors) in control and treatment groups

	Control group ($D(2002) = 0$)	Treatment group ($D(2002) = 1$)
$t = 2002$	12.27 (7.32)	22.22 (20.35)
$t = 2003$	10.80 (5.94)	17.95 (15.01)

The effect (1) is not identified from the data in the sense that it is not possible to construct a sample counterpart for the counterfactual expectation $E[P_{i0}(2002)|D_i(2002) = 1]$. A possible identifying restriction is: $E[P_{i0}(2002)|D_i(2002) = 1] = E[P_{i0}(2002)|D_i(2002) = 0]$. Under

this assumption the counterfactual expectation is identified since $E[P_{i0}(2002)|D_i(2002) = 0] = E[P_i(2002)|D_i(2002) = 0]$. The last expectation can be estimated consistently from the data (as Table 3 shows, the estimate is 12.27). The restriction is, however, very strong in our context as Parker is more likely to grade the prestigious high-quality wines than the low-quality wines. In the absence of treatment, the expected wine price is therefore probably higher in the treatment group than in the control group. Indirect evidence for this selectivity bias comes from Table 3, where the estimate of $E[P_{i0}(2003)|D_i(2002) = 1] = E[P_i(2003)|D_i(2002) = 1]$ is shown to be much higher than the estimate of $E[P_{i0}(2003)|D_i(2002) = 0] = E[P_i(2003)|D_i(2002) = 0]$ (17.95 vs 10.80).

Fortunately the $ATET$ is also identified under the following, much weaker, "parallel trend" restriction

$$E[P_{i0}(2002) - P_{i0}(2003)|D_i(2002) = 0] = E[P_{i0}(2002) - P_{i0}(2003)|D_i(2002) = 1] \tag{2}$$

Restriction (2) states that, in the absence of treatment, the expected price-evolution would have been the same for the wines in the control group and the wines in the treatment group. It is essentially the same as the identification restriction considered by Heckman et al. (1997), and Abadie (2005).[12] Under restriction (2) and using that $P_i(2003) = P_{i0}(2003)$, the $ATET$ can be rewritten as (see Heckman et al. (1997), and Abadie (2005))

$$\begin{aligned} ATET &= E[P_{i1}(2002) - P_{i0}(2003)|D_i(2002) = 1] \\ &\quad -E[P_{i0}(2002) - P_{i0}(2003)|D_i(2002) = 0] \\ &= E[P_i(2002) - P_i(2003)|D_i(2002) = 1] \\ &\quad -E[P_i(2002) - P_i(2003)|D_i(2002) = 0]. \end{aligned} \tag{3}$$

An estimate of the $ATET$ can be obtained by replacing the expectations in (3) by the corresponding

[12] Although they also consider a framework with two time periods, say $t = 0$ and $t = 1$, the treatment in their model is only administered in period 1 and never in the initial period 0. Thus in their setting $D_i(0) = 0$ for all i, $D_i(1) = 1$ for the treated and $D_i(1) = 0$ for the untreated.

sample averages given in Table 3.[13] This leads to the estimate $\widehat{ATET} = (22.22 - 17.95) - (12.27 - 10.80) = 2.80$ euros. The standard error of the estimate equals 0.52, so the effect we find is strongly significant. Estimating the $ATET$ by the rank of the wine shows that the Parker effect increases with ranking. For instance, we find a Parker effect of 3.73 (1.08) for the top classified wines[14] versus only 2.18 (0.59) for the remaining less prestigious wines. Table 4 reports the estimates of the $ATET$ by appellation. We only give the results for the cases with a sufficient number of observations. The estimates are always positive except for Pessac-Léognan, but for this appellation the effect is not significant. Note that the largest ATET is obtained for Pomerol, which is actually one of the appellations that Robert Parker likes the most.

Table 4. Estimates of ATET by appellation

Appellation	# of obs.	Est. (std. error)
Haut-Médoc	16	0.42 (0.29)
Margaux	24	3.53 (1.72)*
Pauillac	19	6.19 (2.45)*
Pessac-Léognan	26	−0.04 (0.95)
Pomerol	15	8.97 (2.89)*
St-Emilion Grand Cru	51	3.85 (1.27)*
St-Estephe	9	1.32 (0.75)
St-Julien	13	1.36 (0.75)
Sauternes	13	2.78 (0.87)*

Note: *significant at the 5% level.

[13]This estimator is sometimes called the "difference-in-differences estimator". This can be seen by considering the regression model $P_i(t) = \delta(t) + \alpha D_i(t) + \eta_i + v_i(t)$, where $\delta(t)$ is a time-specific component, η_i a wine-specific component, and $v_i(t)$ a transitory shock with mean zero at each period. Differencing, and using that $D_i(2003) = 0$ for all i, we get $P_i(2002) - P_i(2003) = \delta + \alpha D_i(2002) + (v_i(2002) - v_i(2003))$, where $\delta = \delta(2002) - \delta(2003)$. Under the assumption that $v_i(2002) - v_i(2003)$ is mean independent of $D_i(2002)$, the OLS estimator of α converges to the $ATET$.

[14]The group of wines with the following ranking: 1er Cru Classé, 2ème Cru Classé, 3ème Cru Classé, 4ème Cru Classé, 5ème Cru Classé, 1er Grand Cru Classé A, and 1er Grand Cru Classé B.

One may wonder whether the above results are sensitive to the scaling of the outcome variable. It may be that the signs of the estimates of the ATET change when prices are no longer defined in levels. To check this we performed the same kind of calculations as above but now by defining all outcome variables as the natural logarithm of prices. The conclusions obtained with the ATET model in logs are basically the same as those obtained above: the estimate of the ATET is now 0.057 (0.013); the estimate of the ATET for the top classified becomes 0.088 (0.020) and for the less prestigious ones 0.043 (0.016); four out of the five appellations that are significant (and with positive signs) in Table 4 remain so when prices are in logs (St-Emilion Grand Cru is no longer significant however, all appellations with insignificant estimates in Table 4 remain so in the new version of the model), and Pomerol still tops the list with an estimate equal to 0.182 (0.041). Although this is clearly not an exhaustive robustness analysis of the ATET model in levels, it is reassuring to see that our empirical conclusions are not fundamentally modified when the outcome variable is specified in logs. From now on we exclusively focus on the analysis with prices defined in levels.

The ATET is an aggregate and overall measure of the Parker effect as it is the average of grade-specific Parker effects. It is also of interest to study these grade-specific Parker effects. To define these parameters, let $P_{ig}(t)$ now denote the price in the spring of year t had wine i been attributed the grade g, and $P_{i0}(t)$ (as before) the price when the wine is not graded. Let $D_i(t) = g$ if wine i has received the grade g, and $D_i(t) = 0$ (as before) in the absence of treatment. As before $P_i(t)$ denotes the realized price in the spring of year t, and of course we still have $P_i(2003) = P_{i0}(2003)$ for all i. The grade-specific Parker effect is defined as the average treatment effect on the treated at level g

$$ATET(g) \equiv E[P_{ig}(2002) - P_{i0}(2002)|D_i(2002) = g]. \tag{4}$$

Under an identifying assumption analogous to (2),[15] the $ATET(g)$ can be rewritten as

[15]The restriction that identifies $ATET(g)$ is $E[P_{i0}(2002) - P_{i0}(2003)|D_i(2002) = 0] = E[P_{i0}(2002) - P_{i0}(2003)|D_i(2002) = g]$.

$$ATET(g) = E[P_i(2002) - P_i(2003)|D_i(2002) = g] - E[P_i(2002) - P_i(2003)|D_i(2002) = 0]. \quad (5)$$

The second expectation can still be estimated by taking the difference of appropriate sample means, and the first expectation can be estimated using kernel estimation methods. We have

$$\widehat{ATET}(g) = \frac{\sum_{i=1}^{N_1(2002)} (P_i(2002) - P_i(2003)) K\left(\frac{D_i(2002)-g}{h}\right)}{\sum_{i=1}^{N_1(2002)} K\left(\frac{D_i(2002)-g}{h}\right)} - (12.27 - 10.80) \quad (6)$$

where $K(u) = 0.75\ (1-u^2)\ \mathbf{1}_{\{|u|\leq 1\}}$ is the Epanechnikov kernel function, $\mathbf{1}_{\{.\}}$ the indicator function equal to one if the statement between brackets is true and zero otherwise, h the bandwidth parameter, and $N_1(2002)$ the number of observations in the subsample of wines graded in the spring of 2002 (i.e., $N_1(2002) = 158$). The grades $D_i(2002)$, $i = 1, ..., N_1(2002)$, are defined as the midpoints of the grade intervals published in *The Wine Advocate*.[16] The value of the bandwidth parameter is chosen according to Silverman's rule of thumb (see Silverman (1986)): $h = 0.9\ min\left(\sigma, (q_{75} - q_{25})/1.349\right) N_1(2002)^{-\frac{1}{5}}$, where σ is the standard deviation of the grades, and q_{25} and q_{75} are the 25th and 75th percentiles.

When based on all $N_1(2002) = 158$ observations, the estimate of $ATET(g)$ sharply declines for g beyond 94. This is due to the fact that there are very few observations in this range of the sample. The problem could not be resolved by using a variable bandwidth. We therefore dropped 3 extreme observations from the sample, and our estimate is thus based on $N_1(2002) = 155$, and using this sample we find $h = 0.734$. Figure 1 shows the graph of $\widehat{ATET}(g)$ together with the 95% point-wise confidence band.[17] The estimates of $ATET(g)$ tend to increase with the grade, starting around zero for $g = 84.5$ and the maximum is around 14 euros for $g = 93.5$. Note that the null hypothesis "$ATET(g) = 0$" is accepted for grades below 86, but is rejected for grades beyond that value. The fact that the more highly graded wines benefit relatively more from being evaluated by

[16] We have also calculated estimates of $ATET(g)$ by drawing each grade randomly in its corresponding grade interval, but the resulting graphs were very similar to the ones reported below.

[17] The 95% point-wise confidence band is constructed using the bootstrap method described in Härdle (1990, pp. 103-106). Details can be obtained from the authors.

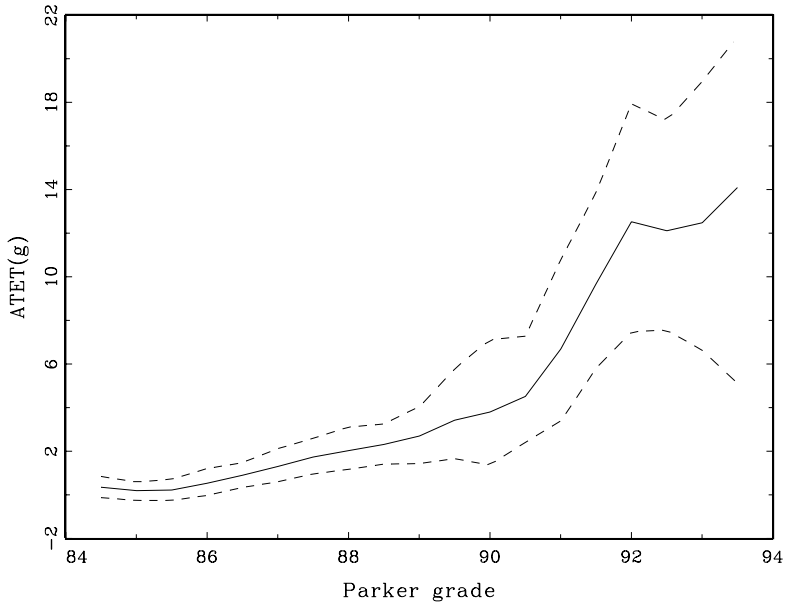

Figure 1: *The estimation of $ATET(g)$ and 95% point-wise confidence band (full sample; $h = 0.734$)*

Parker seems quite intuitive. The fact that $ATET(g)$ does not turn negative for low values of g may come, however, as a surprise. One might expect that very poorly graded châteaux would have been better off had they not received any Parker grade at all. A possible explanation may be that the wines evaluated by Parker in 2002 are simply of too high quality. Indeed, the lowest grade assigned in 2002 was 82.5 (see Table 2), and this is still in the range of wines judged by Parker as "barely above average to very good" (wines with scores between 80 and 89 fall in this category, see *The Wine Advocate*). In other words, had Parker also evaluated châteaux with quality levels well below this range, we might have found $ATET(g)$ to be negative for scores in the 50s ("unacceptable" wines), the 60s ("below average" wines), or may be even the 70s ("average" wines).

Figure 2 displays the estimate of $ATET(g)$ for the top classified châteaux in our sample, and Figure 3 gives the graph for the less prestigious ones. The number of observations $N_1(2002)$ used to calculate the kernel estimate appearing in (6) is 59 for Figure 2 and 96 for Figure 3, and the optimal bandwidths are respectively 0.772 and 0.813. In both figures the estimates start at or slightly above 0, and increase with the grade g. For the top wines the null "$ATET(g) = 0$" is

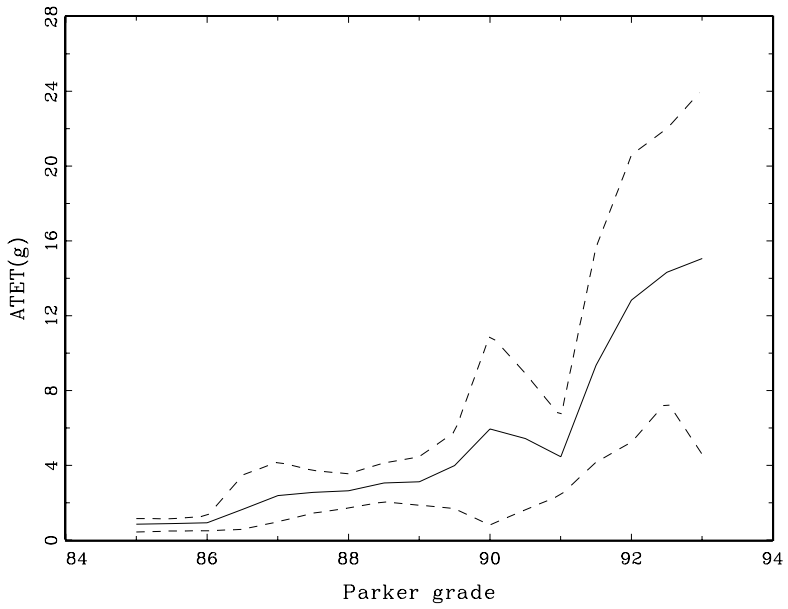

Figure 2: *The estimation of $ATET(g)$ and 95% point-wise confidence band (top classified wines; $h = 0.772$)*

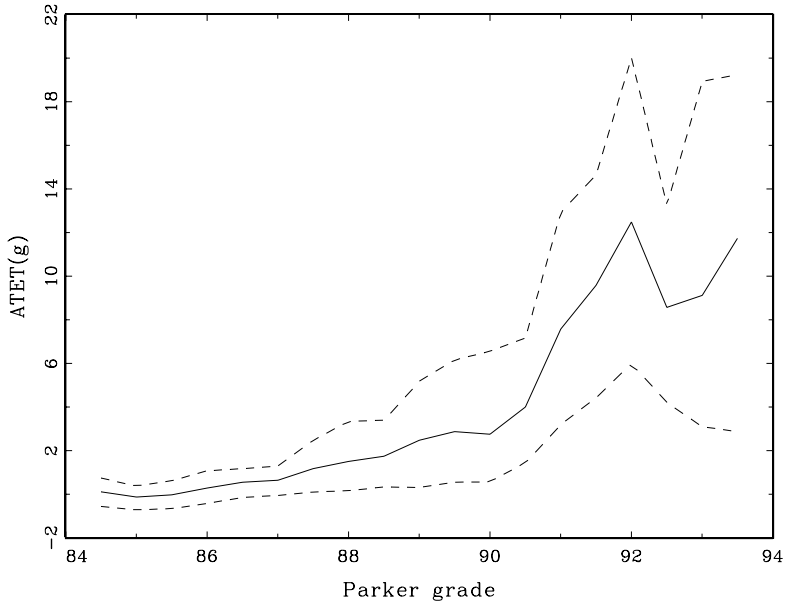

Figure 3: *The estimation of $ATET(g)$ and 95% point-wise confidence band (less prestigious wines; $h = 0.813$)*

always rejected, while for the less prestigious wines it is accepted until g around 89. Note that the curve for the top wines is always above the curve for the less prestigious ones, except for grades between 90.5 and 92.

Our estimates $\widehat{ATET}(g)$ are useful to determine what the wine prices would have been had Parker graded the 2002 vintage in the spring of 2003. For this exercise we use the grades assigned in August/September 2003 (see Table 2), and assume that these grades coincide with the ones Parker would have given in the spring. We argued in Section 3 that this is a reasonable assumption. Let g_i, denote the grade attributed in the late summer of 2003 to wine i. The predicted *en primeur* prices in 2003, $\widehat{P}_{i1}(2003)$, can be written as

$$\widehat{P}_{i1}(2003) = P_{i0}(2003) + \widehat{ATET}(g_i).$$

Figure 4 shows the actual prices $P_{i0}(2003)$ and the predicted prices $\widehat{P}_{i1}(2003)$ for the 117 wines that were graded in August/September 2003 and with grades g_i in the interval $[84.5, 93.5]$.[18] The horizontal axis lists these wines ranked according to their actual prices in 2003. Since $\widehat{ATET}(g)$ is either zero or strictly positive, the predicted prices are always equal or above the actual prices in 2003. The difference between the two series of prices is sometimes substantial (up to around 10 euros) and is more accentuated for the most expensive wines. As the graph shows, there are however some sharp peaks for the less expensive wines as well, suggesting that at least some of the high 2003 grades were assigned to wines that fetched low prices in the *en primeur* market.

Figure 5 shows again the actual and predicted prices, but here the horizontal axis lists the wines by increasing order of the grade they obtained in 2003. Since we generally observe more than one wine per grade level, mean actual and mean predicted prices are calculated for each observed value of the grade. Figure 5 shows a positive relationship between Parker grades and prices. The slope of the grade/price curve is steeper for predicted prices than for actual prices, and the gap between the 2 curves increases from 1 euro for $g = 86$ to 14 euros for $g = 93.5$.

[18] As Table 2 indicates, 121 wines were evaluated in the late summer of 2003, but 4 of them received grades g_i outside the interval, and for such observations $\widehat{ATET}(g_i)$ does not exist.

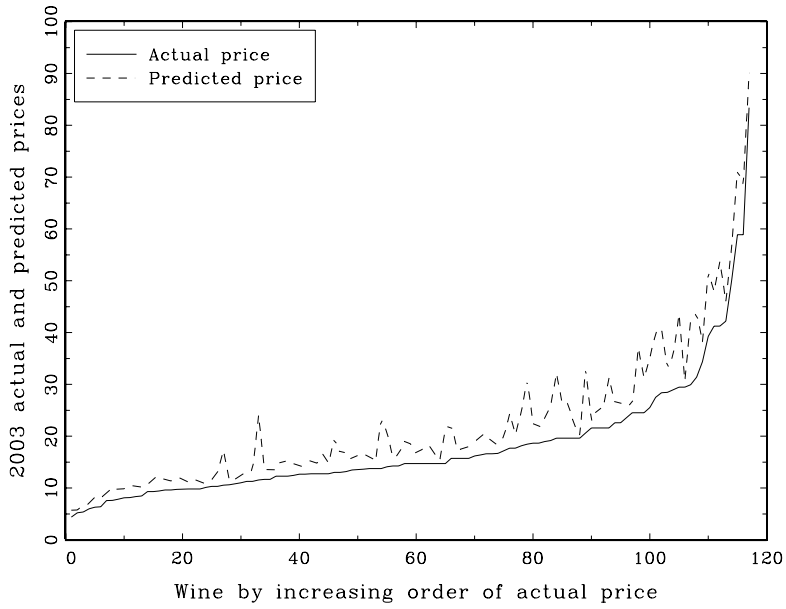

Figure 4: *Actual and predicted prices in 2003*

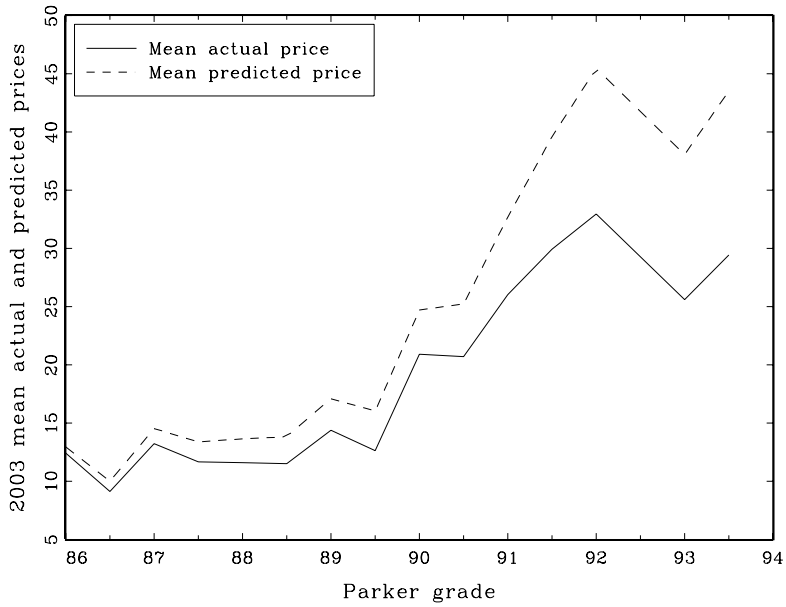

Figure 5: *Mean actual and mean predicted prices in 2003*

5 Conclusion

This paper measures the effect of Robert Parker's oenological grades on *en primeur* wine prices. We exploit the fact that in 2003 prices were determined before publication of these grades, whereas in all previous years the revelation of grades preceded price determination. This unusual reversal allows us to estimate a treatment effect under relatively weak assumptions. Unlike what is stated by Robert Parker ("I doubt that if I had reviewed the wines six months earlier there would have been any change in their pricing, which, by comparison to recent Bordeaux vintages (1995-2000), is extremely low." *The Wine Advocate*, October 2003), we find an overall effect equal to almost 3 euros per bottle. Our estimates of the grade-specific effects show that the impact of the famous wine expert is very important for highly graded wines. The impact vanishes for lowly graded wines, but without actually turning negative.

References

[1] Abadie, A. (2005), "Semiparametric difference-in-difference estimators," *Review of Economic Studies*, 72, 1-19.

[2] Avery, C. and J. Chevalier (1999), "Identifying investor sentiment from price paths: the case of football betting," *Journal of Business*, 72, 493-521.

[3] Bauwens, L. and V. Ginsburgh (2000), "Art experts and auctions. Are pre-sale estimates unbiased and fully informative?," *Louvain Economic Review*, 66, 131-144.

[4] Chossat, V. and O. Gergaud (2003), "Expert opinion and gastronomy: the recipe for success," *Journal of Cultural Economics*, 27, 127-141.

[5] Combris, P., S. Lecocq and M. Visser (1997), "Estimation of a hedonic price equation for Bordeaux wine: does quality matter?," *Economic Journal,* 107, 390-402.

[6] Di Vittorio, A. and V. Ginsburgh (1996), "Des enchères comme révélateurs du classement des vins," *Journal de la Société Statistique de Paris*, 137, 19-49.

[7] Dubois, P. and C. Nauges (2005), "Identifying the effect of unobserved quality and experts' reviews in the pricing of experience goods: empirical application on Bordeaux wine," Working Paper, University of Toulouse.

[8] Février, P., W. Roos and M. Visser (2003), "The buyer's option in multi-unit ascending auctions: the case of wine auctions at Drouot," Working Paper #2003-03, CREST, forthcoming in *Journal of Economics & Management Strategy.*

[9] Ginsburgh, V. (2003), "Awards, success and aesthetic quality in the arts," *Journal of Economic Perspectives*, 17, 99-111.

[10] Ginsburgh, V., M. Monzak and A. Monzak (1994) "Red wines of Medoc: what is wine tasting worth," Verona: Vineyard Data Quantification Society.

[11] Hadj Ali, H. (2002), "La commercialisation des vins *en primeur*," PhD Thesis, University of Toulouse.

[12] Hadj Ali, H. and C. Nauges (2004), "Reputation and quality effects in the pricing of experience goods: the case of *en primeur* wine," Working Paper, INRA-LEERNA.

[13] Härdle, W. (1990), *Applied Nonparametric Regression*. Cambridge, Massachussets: Cambridge University Press.

[14] Heckman, J.J. (1990), "Varieties of selection bias," *American Economic Review, Papers and Proceedings*, 80, 313-318.

[15] Heckman, J.J., H. Ichimura and P.E. Todd (1997), "Matching as an econometric evaluation estimator: evidence from evaluating a job training programme," *Review of Economic Studies*, 64, 605-654.

[16] Heckman, J.J., L. Lochner and C. Taber (1998), "General-equilibrium treatment effects: a study of tuition policy," *American Economic Review, Papers and Proceedings*, 88, 381-386.

[17] Jones, G.V. and K.-H. Storchmann (2001), "Wine market prices and investment under uncertainty: an econometric model for Bordeaux crus classés," *Agricultural Economics*, 26, 115-133.

[18] Rubin, D.B. (1974), "Estimating causal effects of treatments in randomized and nonrandomized studies," *Journal of Educational Psychology*, 66, 688-701.

[19] Rubin, D.B. (1980), "Discussion of paper by D. Basu," *Journal of the American Statistical Association*, 75, 591-593.

[20] Shleifer, A. (1986), "Do demand curves for stocks slope down?," *Journal of Finance*, 41, 579-590.

[21] Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

[22] Wooldridge, J.M. (2002), *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachussets: The MIT Press.

Table 1. Descriptive statistics

Ranking		Appellation	
1er Cru Classé	5.15	Haut-Médoc	6.87
2ème Cru Classé	8.58	Lalande-de-Pomerol	2.58
3ème Cru Classé	4.29	Listrac	1.29
4ème Cru Classé	3.86	Margaux	10.30
5ème Cru Classé	6.44	Médoc	0.43
Cru Bourgeois	14.16	Montagne-St-Emilion	1.72
1er Grand Cru Classé A	0.43	Moulis	1.72
1er Grand Cru Classé B	2.15	Pauillac	8.15
Grand Cru Classé	9.44	Pessac-Léognan	11.16
Cru Classé	6.01	Pomerol	6.44
Cru Non Classé	39.48	1ères Côtes de Blaye	0.86
Total	100.00	1ères Côtes de Bordeaux	0.43
Appellation		Puisseguin-St-Emilion	0.86
Barsac	1.29	St-Emilion Grand Cru	21.89
Bordeaux Blanc	2.15	St-Estephe	3.86
Bordeaux Supérieur	0.43	St-Georges St-Emilion	0.43
Côtes de Castillon	2.58	St-Julien	5.58
Fronsac	1.29	Sauternes	5.58
Graves Blanc	1.29	Total	100.00
Graves Rouge	0.86		

Note: The descriptive statistics are based on the sample of 233 châteaux and give for each variable the percentage of observations that is equal to one.