Parker v. Prial: The Death of the Vintage Chart
[Lighten Your Wallet]

Roman L. Weil
Co-Chairman, Oenonomy Society of the US [1]

By February 2000, Frank Prial had declared the vintage chart dead.[2] "Rarely," he also wrote, "does a year go by that doesn't produce good wine …. the winemakers of the world have rendered the vintage chart obsolete."[3]

Here, I report my tests of Prial's implied hypothesis, that winemakers have rendered the vintage chart obsolete in that they make wine of such uniform high quality that the wine drinker either

- cannot distinguish in blind tastings the wines of years rated high from those of years rated low, or if they can
- do not agree with the vintage chart's preferences.

Bill Marsano[4] and several tasters in my controlled experiments suggested that vintage charts will help distinguish good from bad vintages that have had time to age. The oldest wines in these experiments were 17 years beyond the vintage date when tasted. The youngest were only 4 years beyond their vintage.

Conclusion. In a word, Prial appears *correct* for most wines: the 240 wine drinkers on whom I've systematically tested Prial's hypothesis cannot distinguish between wines of good and bad vintages, except for Bordeaux, and even when they can distinguish, their preferences and the chart's do not match better than a random process would imply.

Methods

---

[1] Also, V. Duane Rath Professor of Accounting, Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago IL 60637; 773.702.7261; fax 206.202.2114; roman@uchicago.edu. Thanks to Doug Hanna for help in experiment preparation and data recording. I presented this paper at the 8th Oenometrics Meeting of the Vineyard Data Quantification Society, Napa Valley, May 22, 2001, whose attendees provided my final data set. Some of these attendees asked why is an accountant doing such work. I respond as follows. First, note that accounting is an intellectual discipline even though you probably think of it as bookkeeping and tax reporting. Accounting records in aggregating numbers information about complex transactions. Then, it attempts to enable users of that information to deduce from the numbers the underlying reality and how to use the data in decision making. Similarly, vintage charts record in aggregating numbers useful information about complex sensory experience. This paper attempts to help the user decide how to use the reported data in making decisions.
[2] Prial, Frank J. "Wine Talk: So Who Needs Vintage Charts," *New York Times,* 9-Feb-2000, B1 continuing to B14.
[3] What a coincidence—on my flight from Chicago to San Francisco, to present this research at the Conference, I read Bill Marsano's "Vintage Nonsense," in the May 2001 issue of United Airlines' *Hemispheres* magazine. One of the testers at the 19-May tasting called this article to my attention. Marsano says, "For those who rather enjoy their wine, rather than assess it, here are some reasons to ignore [vintage charts] and be happy…. Winemakers now have the technology and skills to make good and even very good wines in undistinguished years, although [the wines] won't be long lived." He suggests that vintage charts are useful only for old wines, which age gracefully,
[4] See preceding footnote.

First, and most difficult, I located pairs of wines with the following characteristics: the tasters could afford them [$40 or less per bottle on average], the pairs have identical features[5] in all respects except vintage, and Robert Parker rated one the vintages of those two wines *Average* to *Appalling* while he ranked the other *Excellent* to *The Finest*.[6]

To be clear, look at Exhibit 1, Bordeaux—Pomerol row, columns for 1991 and 1994.   Parker rates these two Pomeral vintages 58 [Appalling] and 89 [Excellent]. I found both 1991 and 1994 Pensees de LaFleur, a Pomerol, at Brown Derby Wines of Springfield MO for this test. The vintage chart in Exhibit 1 indicates the wines used in the experiments reported here.

Second, I split the wine into four containers labeled A, B, C, and D, putting each of the wines into exactly two of the four containers. I gave each taster wines from three of the four containers.[7]   I asked each taster to say which of the three wines differed from the other two. Then, which of these did you prefer, the singleton or the doubleton? A typical taster, say one tasting from containers A, B, and D, would respond something like, "wine A differs from B/D, and I prefer B/D."

I tallied how many of the tasters correctly distinguished the wines and which they preferred, although we don't later care which they say they prefer if they didn't correctly distinguish the wines.

After the tally, I matched the vintages to the letter labels and counted how many of the tasters correctly distinguished vintages. I counted which of the vintages the taster preferred if and only if that taster correctly distinguished the vintages. I announced the result; e.g., "Fifteen of you got this distinction right and of those fifteen, eight preferred the higher ranked vintage and seven preferred the lower ranked vintage."

I announced the result of each pairwise tasting before going on to the next pair. Each taster, except those of the experts on 07-May and the conferees at the VDQS meeting, who tasted one pair of wines, tasted three pairs. Exhibit 1 shows the order in which the subjects tasted the wines on each date.

Most of the tasters were either MBA students at the Graduate School of Business of the University of Chicago or its alumni, alumnae, and their companions. They are primarily upper middle-class, experienced and enthusiastic wine drinkers, but not experts. All tasters paid an entry fee for the tasting, which fee covered full costs of the tasting, and in the case of some of the alumni, more. How often do experimenters get their subjects to pay to participate?

---

[5] Common features include all label items (such as shipper, vineyard, producer) except vintage, retail source, and date of purchase.

[6] Parker, Jr., Robert M. "The Wine Advocates's Vintage Guide 1970 -- 1999," dated1/1/2001; see Exhibit 1 in this paper.

[7]In the tastings with 39 or more individuals, I had two bottles of each wine/vintage. I did not mix those two bottles and then split them in half, but gave the two bottles separate labels.   I wonder if you'd prefer the experimenter to mix identical wines and then split into two containers, which controls somewhat for bottle variation but would not present wines the way individual buyers and drinkers face them.

Results

Exhibits 2 and 3 report the results, Exhibit 2 arranged by tasting and Exhibit 3 by wine tested. Each individual who tests a vintage pair has a one-third probability of correctly distinguishing the vintages by chance: three glasses, one is different; one in three chance of guessing that glass. With n tasters and random choices, the expected number of correct distinctions is $n/3$.[8]

Overall Results. Look at the bottom of Exhibit 2, the totals. A total of 241 testers have tasted 593 pairs of wines. One-third of 593 is 198, so we'd expect 198 correct distinctions if the process is random, so that testers cannot distinguish the wines. The observed number is 246, 41 percent. This differs from the expected relative frequency of 33.3 percent by over 4 standard deviations.   Note, however, that Exhibit 3 shows *all* this significant difference derives from the ability of half the tasters' ability to distinguish between the Finest and the Appalling vintages of Bordeaux Pomerol.

Refer to the last pair of columns in Exhibits 2 and 3, which report the preferences of those who correctly distinguished the wines. Of the 246 pairs where the taster could distinguish correctly, the taster preferred the higher-rated vintage in just over half [51% = 125/246] the pairs and the lower-rated vintage in half. So, even the tasters who can distinguish vintages have only an even chance of agreeing with Parker about which is better. For the Pomerol, 55 percent of those correctly distinguishing the 1991 from the 1994 preferred the higher ranked 1994.

The Exhibit does not show this fact, but at the tastings held in 2001, I counted the number of tasters who made all three distinctions correctly. You'd expect 1/27 [= 1/3 x 1/3 x 1/3] of the tasters to get all three right at random. Eight tasters of 114 correctly identified the different wine in all three pairs, exactly twice as many as we'd expect if the process is random.   Given that 8 made three correct distinctions, we expect exactly one of those eight [1 = 8 x ½x ½x ⅓] to agree with all three of the vintage-chart preferences if the process is random. Only one of the eight did. That one, from the 14-May-01 tasting, described himself as a long-time wine buff.

In response to criticism of early drafts of this report, criticism that I included no experts in the tests, I ran a smaller version, only one pairing of wines, the Excellent/Appalling Pomerol pairing, with a group of self-professed experts, who have been drinking great wine together monthly for more than 15 years. There results were exactly as random: 4 of 12 tasters correctly picked the different wine.   Of those 4, three preferred the higher rated vintage.   At the Conference with wine-academics, 2/3 correctly distinguished.   Seven of those conferees identified themselves during the tally as French wine academics. All seven of these correctly distinguished the 191 from the 1994; 4 preferred the higher ranked 1994 and 3 preferred the 1991.

---

[8] In the tastings with experts and Conference attendees, who study wine-related issues for a living, I asked for help in generating a null hypothesis. Assume, I said, that a group of enthusiastic, experienced amateurs will attempt to distinguish between two wines, one rated 89 and one rated 58 [the ratings of the two Pomerols in my tests]. What fraction of them would you expect to get the distinctioin correctly.   Estimates ranged from 1/3 [the number we'd see from these tests if the process were random] to ¾.   I judge the median estimate of these experts to be about 40 percent.

Caveats

Many of us winos think that Bordeaux wines have the largest variance in vintages (with Burgundy having the largest variance within vintages). Even Prial, in his article suggesting the death studied here, says about Bordeaux, "the weather is as risky as a dot-com stock." The data suggest that tasters can distinguish the Excellent from the Appalling vintage in Bordeaux, even if they don't agree with which is the Appalling one. You can see from Exhibit 3 that a larger portion of the tasters correctly distinguished the wines than of any other wine. We can see that while 55 percent of 170 testers could distinguish the 1991 from the 1994, only 55 percent of those preferred the higher-rated 1994. That does not significantly differ from the 50 percent we'd expect if the process were random. That is, if I tell you I have identified 170 people who can reliably distinguish vintages labeled *Excellent* and *Appalling,* you'd expect half of those to prefer the Appalling vintage if the labels were arbitrarily devised. So, even those who can tell the difference can't be sure that they can use the vintage chart to help them find a wine they will like.[9]

I did not keep track of tasters' sex. About half the tasters have been women.

Implications

What to do with these results? Sterling Pratt[10] recommends removing the vintage chart from your wallet, lightening it, while providing the subtitle to this paper. If, however, the marketplace believes in vintage charts, then carrying one will help you find the good buys in wine. Buy wines from the Appalling years.

Some experienced wine drinkers have commented that one requires practice and guidance to identify wine tastes. William Wecker [who devised the taxonomy that distinguishes wine drinkers from wine fondlers] suggests one should use a vintage chart and tasting notes to tell you what to expect when tasting so that you can learn to align your evaluations with your sensations. He says use the vintage chart as a teaching guide.

Maybe so, but don't expect the vintage chart to lead you to be able to select wines you will enjoy from those you won't.

---

[9] Discussant Ashenfelter, who conceived the Oenonomy Society and invited me to join him as co-chair, to my everlasting pleasure, suggests that the weather data imply little difference between the 1991 and 1994 vintages. He goes on to guess that, because yields were small in 1991, Pomerol wine makers, rather than produce a small quantity of wine, produced none, claiming the vintage was bad and 'we don't make wine in bad years,' hoping to burnish their reputations. Ashenfelter suggests that such tricks deceived Parker and tricked his palate into rating 1991 Pomerol as Appalling. Whatever the reason for the Appalling rating of the 1991, the implication remains that the vintage chart has little use, at least for the pairs tested here.

[10] Pratt is head of the wine department of Schaefer's in Skokie IL and supplied several of the wines tasted in the tests reported here.

Exhibit 1

THE WINE ADVOCATE'S VINTAGE GUIDE 1970 - 1999® Date: 1/1/2001

REGION	1999	1998	1997	1996	1995	1994	1993	1992	1991	1990	1989	1988	1987	1986	1985	1983	1982	1981	1980	1979	1978	1975	1970
St-Julien-Pauillac-St-Estephe	87E	86T	87E	94T	93T	87T	86T	79E	75R	98T	90E	87T	82H	94T	90R	86R	98R	85R	78R	85R	87R	89T	97R
Margaux	89E	88T	66E	88T	88E	86T	85T	75E	74R	80E	86E	65E	78R	90T	66R	95R	86R	82R	79C	87R	87R	78E	85R
Graves	89E	89T	87E	88E	89E	88E	84T	75E	74R	90R	89E	69E	84R	89E	90R	69R	88R	84R	78C	88R	88R	88T	87R
Pomerol	89E	95T	88R	95E	92T	91T	87T	82R	58C	95E	86R	89T	86C	87T	88R	90R	98R	88R	79C	88R	84R	94R	90R
St-Emilion	89E	96T	87R	87T	88E	86T	84C	75R	59C	98T	68E	88E	74C	85E	87R	89R	94R	82R	72R	84R	84R	85R	85R
Barsac/Sauternes	88E	87E	89E	87E	85E	78E	70C	70C	70C	98T	90E	98T	70R	94T	85R	88T	75R	85R	85R	75R	73R	90T	84R
Côte de Nuits (Red)	?	83E	50E	92T	90T	84E	87T	78R	88T	92R	87R	79R	85R	65C	87R	75C	75C	50C	84C	77C	88C	50C	82C
Côte de Beaune (Red)	?	84E	93E	92T	88T	84E	87T	82R	72E	80R	88R	86R	79C	72C	87R	78C	80C	74C	78C	77C	88R	50C	82C
White	90E	86R	89R	92T	91E	87R	72C	90R	70C	87A	82R	92R	78R	99R	89R	85C	88C	86C	75C	88C	88C	85C	83C
Mâcon-Côte Rôtie	95T	90T	80E	86T	90T	88E	58C	78E	92E	92T	96E	88R	86E	84T	90R	89T	85R	75C	83R	87R	98E	73C	90R
South-Châteauneuf-du-Pape	90E	96E	82R	82R	92T	85T	85T	78A	65C	95E	94T	88R	80C	73C	88R	87R	70C	88H	77C	88R	97R	80C	88R
Beaujolais	89R	84R	87R	82C	87C	85C	80C	77C	88C	88C	92C	88C	85C	84C	87C	86C	75C	83C	50C	80C	84C	—	—
Alsace	?	90E	87R	87R	89R	90R	87R	85R	75R	93R	93R	88H	83C	82C	88R	93R	82C	86C	80C	84C	80C	82C	80C
Loire Valley (Sweet White)	84E	84E	88E	91E	88R	67R	88R	60R	76R	90R	82R	88R	82R	87R	66R	64C	84C	82C	72C	63C	80C	—	—
Champagne	?	86C	86R	91E	87E	N.V.	88E	N.V.	N.V.	98E	90R	68E	N.V.	83R	85R	84R	90R	84R	N.V.	88R	85C	98R	86C
Piedmont	80T	95E	84E	95T	87C	77C	86E	74C	76E	95E	96E	90T	85E	73R	90R	76C	92R	80R	70C	88R	95T	65C	84R
Tuscany	88T	86C	65E	78R	88T	85C	86T	72C	85T	90E	72C	69T	73R	84R	93R	80R	86R	82C	70C	76C	85C	84R	84C
Germany	?	86E	88E	83T	87R	90R	87R	90R	86E	92E	90E	88R	82R	83R	88R	90R	80R	82R	85R	84R	72C	85R	80C
Vintage Port	N.V.	N.V.	89T	N.V.	N.V.	92T	N.V.	95E	90E	N.V.	N.V.	N.V.	N.V.	N.V.	96E	92E	88T	N.V.	84T	N.V.	93E	82R	90R
Rioja	86E	82C	88R	85E	90E	90E	87E	85E	75E	87E	90E	87E	82E	82E	82R	74R	92R	83C	75R	79R	84R	84R	90R
Penedès	88E	73C	86R	82E	89E	90E	87E	82E	74E	87E	84E	67E	66E	77R	85R	85R	87R	84R	86R	—	—	—	90R
New So. Wales & Victoria	88E	95E	88R	90E	87E	90E	87R	87R	89E	88E	83E	85E	87E	93E	86R	76R	83C	85C	88C	—	—	—	—
Cabernet Sauvignon	80T	85R	44E	90T	94T	95E	93T	93E	94T	94E	84E	76E	80E	99R	90T	76R	88R	86R	87R	88R	82R	65R	92R
Chardonnay	89E	89R	89R	87C	92C	88C	90C	92C	65C	90C	76C	85C	75C	99C	84C	85C	85C	86C	65C	83C	40C	86C	83C
Zinfandel	90E	86C	85E	89E	87R	92R	90R	90R	91R	91R	83C	82C	90C	87C	80C	78C	90C	82C	82C	83C	85C	80C	80C
Pinot Noir	90E	89R	90E	88R	88R	92R	88R	88R	86R	86R	85C	87C	86C	81C	86C	85C	84C	83C	86C	80C	84C	—	89C
Pinot Noir	90E	89T	85C	88T	78C	90E	87E	89E	87C	87R	66C	86C	72C	83C	87C	90C	84C	66C	86C	—	—	—	—
Cabernet Sauvignon	?	90C	88T	88T	66E	90E	87E	85R	85C	87R	92A	88R	90R	78R	86T	92R	78C	78C	—	—	—	—	—

ABOUT VINTAGE CHARTS

This vintage chart should be regarded as a very general overall rating slanted in favor of what the finest producers were capable of producing in a particular viticultural region. Such charts are filled with exceptions to the rule ... astonishingly good wines from skillful or lucky vintners in years rated mediocre, and thin, diluted, characterless wines from incompetent or greedy producers in great years.

NOTE: For subscription information, call 410-329-6477 or fax us at 410-357-4504

KEY (General Vintage Chart)
90-100 = The Finest
80-89 = Above Average to Excellent
70-79 = Average
60-69 = Below Average
Below 60 = Appalling

Explanations of
C = Caution, too old or irregular in quality
E = Early maturing and accessible
T = Still tannic or Youthful
R = Ready to drink
N.V. = Non-Vintage
? = No Impression Yet Formed

Exhibit 2
Prial v. Parker: Data from 7 Tastings, Involving 241 Tasters
      Sorted by Tasting Date/Tasters

Tasting Date	Wines [In Order Tasted]	Year 1	Parker Rating & Price	Year 2	Parker Rating	Absolute Difference in Ratings & Price Ratio	Number of Tasters	Actual Number Getting Distinction Right	Number Getting Distinction Right Who Prefer Higher Ranked Wine	
Feb-00	Northern Rhone	1993	58/$36	1995	90/$36	32 // 1.00	21	10	5	Guigal Hermitage
[MBAs]	Italy Tuscan	1996	78/$11	1997	95/$11	17 // 1.00	21	4	3	LeCorti Chianti Riserva
	California Cabernet	1983	76/[a]	1990	94/[a]	18 // [a]	21	17	12	Ridge Santa Cruz
May-00	Northern Rhone	1993	58/$36	1995	90/$36	32 // 1.00	41	15	7	Guigal Hermitage
[GSB Alums]	Italy Tuscan	1996	78/$11	1997	95/$11	17 // 1.00	41	16	6	LeCorti Chianti Riserva
	California Cabernet	1983	76/[a]	1990	94/[a]	18 // [a]	40	3	2	Ridge Santa Cruz
1-Apr-01	Oregon Pinot Noir	1995	76/$29	1998	89/$34	13 // 1.17	24	6	2	St. Innocent Freedom Hill
[MBAs]	Italy Tuscan	1996	78/$27	1997	95/$33	17 // 1.22	26	6	1	Savignola Paolina Riserva
	Bordeaux Pomerol	1991	58/$44	1994	89/$64	31 // 1.45	26	14	8	Pensees de LaFleur
7-May-01	Bordeaux Pomerol	1991	58/$44	1994	89/$64	31 // 1.45	12	4	3	Pensees de LaFleur
[Experts]										
14-May-01	Oregon Pinot Noir	1995	76/$29	1998	89/$34	13 // 1.17	55	27	15	St. Innocent Freedom Hill
[GSB Alums]	Italy Tuscan	1996	78/$27	1997	95/$33	17 // 1.22	54	23	11	Savignola Paolina Riserva
	Bordeaux Pomerol	1991	58/$44	1994	89/$64	31 // 1.45	51	31	18	Pensees de LaFleur
19-May-01	Oregon Pinot Noir	1995	76/$29	1998	89/$34	13 // 1.17	40	14	2	St. Innocent Freedom Hill
[GSB Alums]	Italy Tuscan	1996	78/$27	1997	95/$33	17 // 1.22	39	12	8	Savignola Paolina Riserva
	Bordeaux Pomerol	1991	58/$44	1994	89/$64	31 // 1.45	39	16	7	Pensees de LaFleur
21-May-01	Bordeaux Pomerol	1991	58/$44	1994	89/$64	31 // 1.45	42	28	15	Pensees de LaFleur
[VDQS]										

Totals …………………………………………………………………………………………………							593	246	125

Expected Frequency if Process is Random …………………………………………………………………………	33.3%	50.0%
Observed Relative Frequency …………………………………………………………………………………	41.5%	50.8%
Standard Deviation of Relative Frequency if Process is Random …………………………………………………	1.94% [b]	3.19% [c]
Z Score: (Observed - Expected) Divided by Standard Deviation …………………………………………………	4.2	0.3

Sources of Wines: Guigal and LeCorti--Schaefer's Wines and Spirits of Skokie IL; Ridge--from Vineyard; St. Innocent and Pensees de LaFleur--Brown
      Derby of Springfield MO; Savignola--Portland Wine Merchants of Portland OR.

Note a. I acquired the wines from Ridge in a basket purchase, so I cannot disentangle the costs.   On April 15, 2000, lots of Mondavi Cabernet Reserve
traded at Sotheby's auction in New York. The ratio of the prices for 1990/1983 was $62.73 to $56.29 or 1.11 to 1.00.

Note b: Square Root (1/3 x 2/3 x 1/593)
Note c: Square Root (1/2 x 1/2 x 1/246)

Prial v. Parker: Data from 7 Tastings, Involving 241 Tasters
Sorted by Wine // Listed in Order of Largest Parker Differences Between the Two Vintages

Tasting Date	Wines [In Order Tasted]	Year 1	Parker Rating & Price	Year 2	Parker Rating	Absolute Difference in Ratings & Price Ratio	Number of Tasters	Getting Distinction Right Number	%	Number Getting Distinction Right Who Prefer Higher Ranked Wine	%	
Feb-00	Northern Rhone	1993	58/$36	1995	90/$36	32 // 1.00	21	10		5		Guigal Hermitage
May-00							41	15		7		
Totals	Northern Rhone	1993	58/$36	1995	90/$36	**32 // 1.00**	62	25	40%	12	48%	Guigal Hermitage
	Z Score: (Observed - Expected) Divided by Standard Deviation								1.2		(0.2)	
1-Apr-01	Bordeaux Pomerol	1991	58/$44	1994	89/$64	31 // 1.45	26	14		8		Pensees de LaFleur
7-May-01							12	4		3		
14-May-01							51	31		18		
19-May-01							39	16		7		
21-May-01	[All 7 French Experts in this Testing Correctly Distinguished the Singleton, but 3 of those 7 Preferred the 1991]						42	28		15		
Totals	Bordeaux Pomerol	1991	58/$44	1994	89/$64	**31 // 1.45**	170	93	55%	51	55%	Pensees de LaFleur
	Z Score: (Observed - Expected) Divided by Standard Deviation								5.9		0.9	
Feb-00	California Cabernet	1983	76/[a]	1990	94/[a]	18 // [a]	21	17		12		Ridge Santa Cruz
May-00							40	3		2		
Totals	California Cabernet	1983	76/[a]	1990	94/[a]	**18 // [a]**	61	20	33%	14	70%	Ridge Santa Cruz
	Z Score: (Observed - Expected) Divided by Standard Deviation								(0.1)		1.8	
Feb-00	Italy Tuscan	1996	78/$11	1997	95/$11	17 // 1.00	21	4		3		LeCorti Chianti Riserva
May-00							41	16		6		
Totals	Italy Tuscan	1996	78/$11	1997	95/$11	**17 // 1.00**	62	20	32%	9	45%	LeCorti Chianti Riserva
	Z Score: (Observed - Expected) Divided by Standard Deviation								(0.2)		(0.4)	
1-Apr-01	Italy Tuscan	1996	78/$27	1997	95/$33	17 // 1.22	26	6		1		Savignola Paolina Riserva
14-May-01							54	23		11		
19-May-01							39	12		8		
Totals	Italy Tuscan	1996	78/$27	1997	95/$33	**17 // 1.22**	119	41	34%	20	49%	Savignola Paolina Riserva
	Z Score: (Observed - Expected) Divided by Standard Deviation								0.3		(0.2)	
1-Apr-01	Oregon Pinot Noir	1995	76/$29	1998	89/$34	13 // 1.17	24	6		2		St. Innocent Freedom Hill
14-May-01							55	27		15		
19-May-01							40	14		2		
Totals	Oregon Pinot Noir	1995	76/$29	1998	89/$34	**13 // 1.17**	119	47	39%	19	40%	St. Innocent Freedom Hill
	Z Score: (Observed - Expected) Divided by Standard Deviation								1.4		(1.3)	
Totals	....						593	246	41%	125	51%	
	Z Score: (Observed - Expected) Divided by Standard Deviation								4.2		0.3	
Totals *Excluding* Bordeaux Pomerol	....						423	153	36%	74	48%	
	Z Score: (Observed - Expected) Divided by Standard Deviation								1.2		(0.4)	