Which experiments in Brescia?

**An Identification Problem:**
(Experimental) Economists at a Wine Tasting Experiment

Matteo M Galizzi  
Univ. of Brescia and CHE, Univ. of York

David J Reiley  
Univ. of Arizona and Yahoo! Research

XV Enometrics  
Collioure  
May 30th 2008

How prices are set in Italian quality wines?

Priced between 7 and 30 € approx (>30 €, top wines segment)

The better the quality perceived by professional wine-tasters, the higher the prices

Producers send samples of bottles to qualified panels of experts (AIS, Seminario Veronelli, Slow Food, Gambero Rosso, etc) for horizontal blind tasting and refereeing (same wine, year etc)

They look at blind “referee” reports and use average evaluation as guidance to pricing strategies within the 7-30 € range

Prices as signals reflecting quality perceived by experts:

Wine experts can perceive differences in wines’ quality which justify differences in prices (at least partially)

But, is it really working for the vast majority of consumers?

Experimental questions:

Can non-expert consumers spot differences in wines’ quality in line with observed differences in selling prices?

More generally, are non-expert consumers really able to identify quality in wines?

We designed A simple wine-tasting experiment to test this!

We selected 3 quality wines having

- Similar but not identical oenological characteristics
- Not too distant experts’ evaluations: 3 most influential Italian wine guides (AIS, Veronelli, Slow Food-Gambero Rosso)
- Different prices

Subjects have 3 glasses of wines in front of them. Subjects can win a bottle of their preferred wine if correctly perform a simple task.

In stage 1, asked to blindly taste 3 glasses of wine and, for each wine, to
- describe perceived aromas
- express evaluations within a (0-100) grid
- declare their WTP for a bottle of that wine

In stage 2, told that 2 out of 3 glasses are actually containing exactly the same wine!

Task is to identify which glasses are containing the same wine

Also repeat exercise as in stage 1: aromas, evaluation, WTP

- Can subjects really identify aromas and wines?
- Can differences in evaluations/WTP among wines account for differences in prices?
Which experiments in Brescia?

A field experiment...

• In last July ESA World conference in Rome, we looked for 100 experimental economists to participate into a wine-tasting experiment after the last session of the conference. Thank you, John, Daniela and all LUISS staff!

• More than 130 people signed in: we only manage to accommodate up to 108 due to technical constraints (size of the classroom, glasses, bottles of wine).

• 84 subjects had proper glasses, 24 plastic glasses.

• Rome in August, about 35 degrees: definitely white wines!

The wines

3 prestigious Sauvignon from Sud Tirol, northern region by the Alps close to Austria, best for aromatic white wines (with Friuli)

Same village, close to Bozen, with identical hydro-geological and oenological characteristics, same (good) 2005 year, outstanding producers

• Similar but not identical aromatic profiles: mainly citrus, elder flowers, mineral

• Close enough evaluations by wine experts

• Very different prices

With an abuse of notation, G, M, B

G: St Michael-Eppan Sanct Valentin, one of the best Italian wine ever, 94/100 average evaluation, rich profile of 5 aromas, 24 €

M: Elena Walch-Castel Ringberg, 89/100, 4 aromas, 16 €

B: Terlaner-Winkl, 84/100, 3 aromas, 9 €

The experiment

• Subjects chose an envelope, to assess if they could win a bottle in case of a correct task (paid subjects, P), or not (unpaid subjects U)

  Approximately half the subjects (56.4%) were P

• Subjects indeed tasted only 2 wines

• Group GM of subjects had G+M, group MB had M+B, group GB G+B (highest variation of experts’ evaluations and prices)

  Put a hundred of (drunk) experimentalists together after a long day: chaos!

  We knew it and we allowed chats and noise...

To avoid influence from neighbours, used a perfect stranger matching

• 18 treatments, in a way that no subject had wines in same glass positions of his neighbours: e.g. left to right GBG, MMB, MGG, etc

• Also useful to control for effect of order/position of glasses

More on the experiment

• Subjects were wine lover from 12 nationalities: 68% males

• Wines kept in fridge and poured into glasses just slightly colder to be in perfect temperature conditions during experiment.

• Subjects had 3 glasses containing wine, a pen and two envelopes

• At stage 1, subjects tasted the wine and answered the questions

• They then opened another envelope, with instructions for stage 2

• Told that 2 out of 3 glasses just tasted were actually containing same wine

  The task to win a bottle was to identify which ones

As prize, paid subjects were given a bottle of the preferred wine.
Results

On average only 7% of “true” aromas were identified. Also, subjects describe aromatic profiles with vague and generic expressions: in stage 1 only 42,8% of terms were appropriate and specific from an expert’s perspective; in stage 2, 39,7%. Most adopted adjectives were “fresh, dry, sweet, fruity”.

2. Can subjects correctly evaluate the wines? No.
They underestimated quality: in stage 1, average evaluations were for G 48,1 points, for M 49,3, for B 47,6!

3. Are subjects’ WTP in line with actual prices? No.
Greatly underestimated prices: 6 € for G, 6,9 € for M, 6,2 for B!

4. Do subjects perceive the “true” quality differences in wines? No.
When comparing wines of different qualities by experts in stage 1, 21,8% of times subjects did not perceive any difference 44,4% of times gave to the wine of lower quality a higher evaluation!
On average 19,4 points, more than actual (absolute) difference!

5. Do subjects declare WTP in line with actual prices? No.
Same as above: in stage 1, 21,7% subjects gave same WTP for bottles of different qualities, and 45,4% gave higher WTP for wine with lower quality!
On average 3,84 € higher WTP, smaller difference than actual one

6. Do subjects review aromas/evaluation/WTP from stage 1 to 2? Yes.
In stage 2 only 67,8% of aromas perceived in stage 1; description even more imprecise; on average evaluations increased by 6,44 points, WTP increased by 3,78 €.

7. Now, can subjects correctly identify the two identical wines? No.
Only 43,5% of subjects identified correctly: 33,3% in group GM, 44,4% in group MB, 52,8% in group GB
closer the wines in terms of quality, lower the frequencies: All < *** 2/3, probability to correctly guess by randomly selecting any two glasses
Moreover, at stage 1, 59,3% of subjects evaluate differently the two glasses of the same wine with an average difference of 19,2 points (and 2,78 € in WTP).

8. Do plastic glasses matter? Yes.
Only 5 out of 24 with plastic glasses guessed correctly! (20,8%)
Sommelier are right: oxygenation matters
Even for sub-sample with proper glasses, only one in two subjects guessed correctly, still significantly < *** 2/3 (except in group GB)

13 subjects understood already in stage 1 that 2 wines were indeed the same!
Experienced or talented with good olfaction.
All but one with proper glasses!
They gave more specific description of aromatic profiles
Dropping them, frequencies (for proper glasses) are even lower! 35,78%: 22,2% in group GM, 25% in MB, 47,2% in GB.

Not significant differences between correct guesses by P and U
One exception is group GB: P identified significantly more than U.
Providing incentives does not help blind tasting… unless aromatic differences are evident!
11. Does the order in which glasses are presented/tasted matter?

Maybe.

Subjects gave guesses in approx. same proportions across all orderings of glasses:

- 31.5% said same wine in glasses 1&2, 37% in 1&3, 31.5% in 2&3.

However, correct guesses were actually 27.7% “1&2”, 29.7% “2&3” and 42.6% “1&3”.

Any effect of “distance” of glasses on likelihood of guesses?

12. What drives a correct guess?

A probit estimation of correct identification

Using all subjects who did not identify two glasses with identical wine already in stage 1

| Band   | Coef  | Std. Err | z    | P>|z|  | [95% Conf. Interval] |
|--------|-------|----------|------|-----|-------------------|
| Plastdum | -1.430 | 0.436    | -3.27 | 0.001      | -2.286 to -0.573 |
| GMdum   | -0.720 | 0.360    | -2.00 | 0.045      | -1.425 to -0.146 |
| MBdum   | -0.580 | 0.367    | -1.58 | 0.114      | -1.299 to 0.139  |
| Closedum| -0.603 | 0.315    | -1.91 | 0.056      | -1.222 to 0.015  |
| AvArID  | -0.023 | 0.018    | -1.26 | 0.208      | -0.059 to 0.013  |
| ArQualF | -0.001 | 0.005    | -0.05 | 0.960      | -0.011 to 0.009  |
| Paidum  | 0.186  | 0.308    | 0.61  | 0.544      | -0.417 to 0.790  |
| Const   | 0.764  | 0.496    | 1.54  | 0.124      | -0.208 to 1.737  |

And so, what?

- If you are an expert and trained wine lover, prepare to pay huge differences in money for slight differences in quality and for subtle differences in aromatic profiles
- If you are not, and you don’t know which bottle of quality wine to buy in a wine shop….
  
  ...just grab the cheapest bottle!
  
  You would never perceive the difference, anyway!