Hedonic Price for Italian wine

Rosella Levaggi and Paola Zuccolotto

Abstract The price formation mechanism in the domestic wine market has not been widely studied in spite of the importance of the Italian market as concerns sales. We use a unique dataset to estimate the hedonic price function for Italian wine sold on the Italian market in the period 2005-2011. For each bottle considered, the dataset allows us to know several characteristic such as the price by retail channel (price in large distribution and in wine shops), label characteristics, chemical analysis, sensory evaluations and experts' judgments. The aim of the analysis is to examine the price formation mechanism in the large distribution and in wine shops and to explore the differences in price formation for red and white wines.

1 Introduction

Wine producing is one of most important industries in Italy, in spite of the present recession, and Italy has overtaken France as the world's largest wine producer[2]. Although the importance of wine market is widely recognized, only few studies have been proposed to explain the main determinants of price formation in this market. Brentari and Levaggi (2010) show that pricing strategies for red wine are channel-dependent while Brentari, Levaggi, Zuccolotto (2011) show that sensory characteristics have a marginal role in price formation and this is especially true for the wine sold in the large distribution. In this paper we examine this relationship in more detail and we propose to use procedure combining two statistical methods, in order to overcome some of the problems related to the lack of theory behind hedonic pricing estimation techniques. In detail, we preliminarily carry out a dimensionality reduction by means of canonical correlation analysis, in order to replace the original chemical and sensory variables with some latent factors globally accounting for chemical and sensory quality.

---

[1] Rosella Levaggi, University of Brescia; email: levaggi@eco.unibs.it
Paola Zuccolotto, University of Brescia; email: zuk@eco.unibs.it

[2] http://www.telegraph.co.uk/foodanddrink/wine/8571222/Italy-overtakes-France-to-become-worlds-largest-wine-producer.html

Then we select the most important determinants of price and estimate the price formation mechanism by using step-wise regression.

2 Methods and data

Although various approaches may be used to estimate the price of wine, the vast majority of the literature adopts the hedonic price approach. The general specification of our hedonic price function is

$$p = g(L, C, S, D)$$

where L groups the characteristics of the wine that can be inferred from the label, C its chemical characteristics, S the sensory characteristics and D the variables that describe the distribution process. In this article we use the unique dataset that *Altroconsumo* uses for its yearly guide. Each year about 300 wines (red and white) in the low to medium/high price range are bought and their characteristics are evaluated using a panel of experts. Within this range, wines are chosen so as to represent the wide variety of Italian wines as regards vineyards, producers and region of origin. The sensory analysis is made using a detailed protocol and the price of each wine is estimated using a specific market analysis. We use data for red and white wines for the period 2005-2011: 1077 observations from 45 appellations for red wines and 724 observations from 48 appellations for white wines. The main descriptive statistics are presented in Table 1 and 2, for red wines and white wines, respectively. The main features of the dataset are more accurately described in Brentari et. al (2011).

Since the variables describing appellation and region of origin have a high number of categories, some of them have been grouped according to the following criterions. A specific ungrouped category was left to red wines appellations with frequency higher than 20, and to white wine appellations with frequency higher than 10. The other appellations have been grouped into four categories: *OsupN*, *OsupC*, *OsupS* - which comprises *DOC* and *DOCG* appellations divided by area (North, Centre, South) - and *Other* - which comprises less known and more generic appellations. The same has been done for Regions, with specific categories left only to Regions with frequency higher than 65 and 55, for red and white wines respectively. The others have been grouped into three categories, again according to area (*North*, *Centre*, *South*).

For what concerns chemical and sensory variables, following Brentari and Zuccolotto (2011) and Brentari et al. (2011), we have constructed sensory and chemical latent factors able to summarize the effects of the corresponding variables. Latent factors can be obtained using a variety of dimensionality reduction techniques. We decided to carry out Canonical Correlation Analysis (CCA) between the set C and S of chemical and sensory variables, a multivariate statistical technique introduced by Hotelling in 1936, aimed at defining the coordinate system that describes the maximum cross–covariance between two datasets. We argue that forcing chemical and sensory latent factors to be correlated to each other should hopefully result in latent factors globally correlated to the wine's quality. Let *LFC1*, *LFC2*, *LFS1*, *LFS2* be respectively the first and the second chemical and sensory latent factors. They will be added to the database and used as explanatory variables in place of the single variables composing the sets C and S.

The functional form of the hedonic price model is still a matter of empirical investigation. As in Brentari and Levaggi (2010), we restricted the choice to linear and log linear equations and, thanks to a RESET test, we show that a log-linear form should be preferred also in this case (Table 3).

3 Price formation in the large distribution

We performed a step-wise regression with the following starting model:

$$\ln p_{GDO} = k + aDOC + bDOCG + \sum_{i=1}^{z} c_i AP_i + \sum_{i=1}^{w} d_i REG_i + \sum_{i=1}^{2} k_i LFC_i + \sum_{i=1}^{2} l_i LFS_i + mENOp + nVINT + \varepsilon_i$$

where z=27 for red wines and 34 for white wines while w= 9 for red and 8 for white wines. Results are presented in Table 4. For red wines, about 80% of the total variance in price is explained by the model. The alcoholic content is highly significant as much as label characteristics. In particular the appellation tends to increase the price for wines with appreciable reputation. We interestingly note that both the sensory and the chemical quality latent factors enter the model. Results for white wines are similar: in this case the explanatory power of the model is lower (about 67%) and only the first chemical factor and the second sensory are considered. The premium for selling using both channels is slightly higher in this case and Regions of origin are never significant regressors. Also "pooled" appellations do not enter this model.

4 Price formation in wine shops

In this case the starting model for the step-wise regression was given by:

$$\ln p_{ENO} = k + aDOC + bDOCG + \sum_{i=1}^{z} c_i AP_i + \sum_{i=1}^{w} d_i REG_i + \sum_{i=1}^{2} k_i LFC_i + \sum_{i=1}^{2} l_i LFS_i + mGDOd + nVINT + \varepsilon_i$$

Results are presented in Table 4. Appellation are not significant explanatory variables for white wine. In general, price seems to be a little more homogeneous across appellations for red wines (only few and with a positive mark enter the model) while for white wines only a few appellations are significant for the price formation mechanism. The explanatory power of the model is lower than for GDO both for red and white wines. In both cases the Region of origin is important, especially for white wines where a North/Centre /South pattern seem to emerge as concerns the influence of the Region on price. The chemical quality indicators are not significant in both models, while sensory indicators are important, more for red than white wines. Finally it is interesting to note that selling using both channels has no impact on price for the red wine while it adds a small mark-up to price of white wines.

Table 1: Descriptive statistics of the sample for red wines

Binary variables (1 = presence of the specific characteristic)				**Other variables**			
AP	Sample composition				Mean	Min	Max
Chianti	0,0613			p_{ENO}	10,4320	3	45
Nero D'Avola	0,0604	Up	0,0288	p_{GDO}	6,1161	1,4	34
Montepulciano	0,0539	DOC	0,5557	Alcdic	12,7970	10	16,5
Barbera	0,0511	DOCG	0,2115	Alcver	12,8520	10.25	16.44
Siangiovese	0,0483	IGT	0,2180	Sugar	3,9348	1,1	27,6
Chianti Classico	0,0474	Riserva	0,0074	Acitot	5,5141	4,1	8,1
Dolcetto	0,0455			Acivol	0,4455	0,17	0,95
Merlot	0,0353	**REG**		RSO2	0,2062	0	1,4444
Rosso di Montalcino	0,0335	Toscana	0,1978	SO2	72,2840	3	166
Valtellina Superiore	0,0335	Piemonte	0.1578				
Aglianico	0,0316	Veneto	0,1049				
Rosso Toscano	0,0300	Sicilia	0,0854	Colour	7,0664	4	9
Grignolino	0,0232	Lombardia	0,0771	Violet	4,9225	0	8
Bardolino	0,0223	Puglia	0.0585	Orange	3,1783	0	8
Barolo	0,0214			Intolf	6,8928	5	8
Nebbiolo	0,0214	Nord	0,1133	Floral	3,9847	1	6
Refoscolo dal Peduncolo Rosso	0,0204	Centre	0,1300	Fruits	5,2298	3	7
Cabernet Sauvignon	0,0195	Sud	0,0743	Spicy	3,5130	1	6,5
Amarone	0,0186			Vegetables	2,7711	0	5
Castelli Romani	0,0186			Structure	6,7168	5	8
Primitivo	0,0186			Roundness	5,8946	4	8
Terre di Franciacorta	0,0186			Acidity	4,2191	2,5	6
Valpolicella	0,0186			Bitterness	2,0799	0	5
Ainf	0,0650	**Asupc**	0,0483	Astringency	4,3319	0	7
Vino Rosso		Sagrantino di Montefalco		AromRich	6,5042	4	8
Sicilia		Rosso del Conero		Persistency	6,4819	4	9
Negramaro		Rosso Piceno		Attraency	6,9875	5	8
Isola dei Nuraghi		Morellino di Scansiano		Clean	6,8185	4	8
Syrah		**Asups**	0,0576	Quality	6,7526	4	8
Asupn	0,0715	Ciro		CleanRet	6,4271	4	8
Barbaresco		Cannonau		QualityRet	6,8454	4	8
Bonarda		Monica		Giuglo	6,5956	4	8,5
Cabernet		San Severo		Zob	6,6305	4	8,5
Lagrein		Salice Salentino		IE	7,3808	4,7	8,8
Marzemino				IC	0,5494	0,4	0,7
Oltrepò				Vintage		2003	2010
Teroldego							
Valcalepio							

Table 2: Descriptive statistics of the sample for white wines

Binary variables (1 = presence of the specific characteristic)				**Other variables**			
AP	Sample composition				Mean	Min	Max
Trebbiano	0,0608	DOC	0,6989	p_{ENO}	7,9780	3,50	16,00
Vermentino	0,0608	DOCG	0,1133	p_{GDO}	4,8513	1,40	14,50
Chardonnay	0,0552	IGT	0,1796	Alcdic	12,2560	9,5	14,5
Pinot Grigio	0,0539			Alcver	12,3270	9,60	14,90
Verdicchio	0,0483	**REG**		Sugar	3,9849	0,00	16,90
Soave	0,0414	Trentino A.A.	0.1160	Acitot	5,3968	4,01	8,29
Orvieto	0,0401	Veneto	0.1064	Acivol	0,2301	0,00	0,58
Pinot Bianco	0,0331	Friuli VG	0.0925	RSO2	0,3257	0,02	80,00
Terre di Franciacorta	0,0331	Sicilia	0.0801	SO2	97,1270	0,09	176,0
Sauvignon	0,0318	Toscana	0.0746				
Sicilia	0,0318	Nord	0.1575	Colour	5,9551	2,5	8
Bianco di Toscana	0,0290	Centro	0.2017	Green	3,3833	1	6
Vernaccia	0,0290	Sud	0.1063	Gold	4,5953	1	7
Bianco di Custoza	0,0276	**Ainf**	0,0235	Intolf	6,6091	4,5	8
Muller Thurgau	0,0276	Salento		Floral	4,7169	2,5	7
Cirò	0,0249	Galestro		Fruits	4,7859	2,5	7
Alcamo	0,0235	Pecorino		Spicy	2,0829	0	5
Frascati	0,0221	Friulano		Vegetables	2,8267	0	6
Gavi	0,0221	Colline Pescaresi		Structure	6,3384	4	7
Gewurtztraminer	0,0221	**Asupn**	0,0469	Roundness	5,9185	4	7
Grillo	0,0221	Pigato		Acidity	4,6347	3	7
Greco di Tufo	0,0221	Lugana		Bitterness	1,8260	0	4
Isola di Nuragus	0,0166	Traminer		AromRich	6,2949	4	8
Roero	0,0166	Riesling		Persistency	6,2424	4	8
Castelli Romani	0,0152	Valcalepio		Attraency	6,7210	0	8
Est Est Est	0,0152	Erbaluce		Clean	6,6844	4	8
Locorotondo	0,0152	Cortese del Monferrato		Quality	6,5905	3,5	8
Tocai	0,0152	**Asupc**	0,0069	Harmony	6,4254	3	8
Falanghina	0,0138	Pomino Bianco		CleanRet	6,7459	4	8
Pinot Nero	0,0138	Pitigliano		QualityRet	6,4869	3	8
Sylvaner	0,0138	**Asups**	0,0249	IE	7,2620	4,6	8,41
		Fiano di Avellino		ZOB	0,4896	0,3	0,64
		Castel del Monte		IC	72,208	48,89	84,44
		San Severo		Vintage		2003	2010

Table 3: Choice of the functional form

		Red wine		*White wine*	
		Lin	*Loglin*	*Lin*	*Loglin*
p_{GDO}	RESET	252,57	0,108***	67,261***	1,48
	R^2	0,8035	0,8128	0,637	0,689
	BP	504,22***	51,44	294,131***	38,57
p_{ENO}	RESET	57,022***	0,776	17,503*	6,14
	R^2	0,786	0,710	0,648	0,652
	BP	516,22***	20,573	48,19	14,58

Note: RESET is the specification test of Ramesey (1969); BP is the Breush–Pagan (1979) test for heteroschedasticity; **means p-value $< 0{,}05$ ***means p-value $< 0{,}01$

Table 4: Censored Stepwise regression for red (left) and white (right) wines

	p_{GDO}		p_{ENO}	
	Estimate	t Statistic	Estimate	t Statistic
Constant	-1,2328	-4,100	-0,6853	-1,983
DOCG	0,3330	11,330	0,3446	10,800
Amarone	0,5771	7,568	0,4682	3,801
Bardolino	0,1298	2,266		
Barolo	0,4735	7,552	0,3439	3,866
Castel del Monte	-0,3347	-5,688		
Chianti	-0,2156	-5,137		
Dolcetto	0,2014	4,896		
Grignolino	0,3138	5,697		
Montepulciano	-0,2907	-7,682		
Nebbiolo	0,3266	4,863		
Primitivo di Manduria	-0,3161	-4,155	-0,2900	-2,700
Rosso di Montalcino	0,5247	10,260	0,5788	11,560
Sangiovese	-0,1392	-3,617		
Terre di Franciacorta	0,4146	6,252	0,5360	6,661
Valpolicella	0,2500	4,274		
Valtellina Superiore	0,2197	4,002		
OsupN	0,1290	3,851	0,2242	4,147
OsupC			0,1183	2,024
OsupS	-0,1643	-4,338		
Apulia	-0,0933	-2,159		
Piedmont			0,2419	5,940
Sicily	-0,1344	-4,290	-0,0860	-2,013
Veneto			0,2885	3,215
Alcdic	0,2162	9,145	0,2033	7,609
LFC1	0,0794	4,351		
LFC2	-0.0266	-2,236		
LFS1	0,0315	2,665	0,0890	5,174
LFS2	-0,0297	-3,196	-0,0480	-3,434
S_{ENO}	0,2795	14,160		
R^2	0,8000		0,7437	
BP				
N	973		416	
LL	-9.0535		3,2185	

	p_{GDO}		p_{ENO}	
	Estimate	t Statistic	Estimate	t Statistic
Constant	-1.0251	-3.444	0.2053	0.723
Alcamo	-0,2103	-3,290	-0,2140	-2,532
Bianco di Custoza	-0,1698	-2,829		
Castelli Romani	-0,4940	-6,037	-0,6099	-2,744
Est Est Est	-0,2689	-3,510	-0,6209	-5,572
Frascati	-0,1999	-3,100		
Gavi	0,2350	3,507		
Grillo	-0,3268	-4,783	-0,1910	-2,521
Greco di Tufo	0,4355	5,930	0,4388	5,115
Locorotondo	-0,2348	-3,034		
Isola di Nuragus	-0,3083	-3,859		
Orvieto	-0,2707	-5,452		
Pinot Grigio	0,0986	2,174		
Soave	-0,1801	-3,408	0,2676	2,588
Terre di Franciacorta	0,1896	2,857	0,2442	3,130
Tocai			-0,5960	-4,419
Toscano Bianco	-0,1405	-2,411		
Trebbiano	-0,4411	-10,110	-0,1956	-2,442
Verdicchio	-0,1236	-2,743		
OsupCS			0,3458	4,083
Friuli V.G			0,4037	7,817
Trentino Alto Adige			0,1200	2,957
Nord			0,4001	8,259
Sud			-0,1243	-2,143
Alcdic	0,2050	8,424	0,1350	6,078
LFC1	0,0464	3,135		
LFC2				
LFS1			0,0455	2,988
LFS2	0,0392	3,662		
S_{ENO}	0,3384	14,390		
S_{GDO}			0,0794	2,356
R^2	0.6658		0.5756	
BP				
N	652		280	
LL	-2.7432		41.3055	

References

1. Altroconsumo Edizioni, Guida Vini (2006-2012), Milano.
2. Brentari E., Levaggi R., Zuccolotto P. (2011), Pricing strategies for Italian red wines, Food Quality and Preference, 22, 725-732.
3. Brentari E., Zuccolotto P. (2011), The impact of chemical and sensory characteristics on the market price of Italian red wines, Electronic Journal of Applied Statistical Analysis, 4, 265-276.

Zironi R., Odello L., Brentari E. (2003), Un nuovo indice per misurare la qualità edonica del vino, L'assaggio, 1, 25-29; published also on: Il Sommelier, 19(3), 15-17.