2015·BRNO·-·CZ¶
XXII·ENOMETRICS·CONFERENCE

Quality, Reputation and the Price of Wine

Andreas Huber[1]; Dieter Pennerstorfer[2]; Christoph R. Weiss[3]; Dániel Békési[3]

Preliminary draft version, please do not cite!

Abstract

The existing empirical literature on the determinants of wine prices emphasizes the importance of dynamic interactions between quality and reputation, i.e. an increase in reputation may be the consequence of high product quality in the past. On the other hand, the potential danger of endogeneity and especially sample selection issues are usually not addressed. Based on a unique panel data set for more than 300 wineries over the period of 2004-2007, our hedonic analysis decouples the direct (short-run) and the indirect (long-run) effects of product quality on prices in the Austrian market for bottled domestic premium wines. By explicitly accounting for the endogeneity of wine quality (based on the grades given by the Falstaff wine guide) and reputation, we apply an estimation procedure that corrects for a potential sample selection bias. Our results suggest that the proposed approach is superior to specifications ignoring self-selection and assuming exogenous quality and winery reputation. While influencing reputation positively over time, wine quality is found to have a significant effect on prices simultaneously. Consequently, short- and long-run price effects of quality can be quantified, which might be of great relevance for stakeholders in the wine industry.

Keywords: hedonic pricing, wine quality, reputation, endogeneity, sample selection
JEL-Code: C33, L66, Q11

[1] andreas@weingut-huber.at
[2] Austrian Institute of Economic Research (WIFO)
[3] Department of Economics - Vienna University of Economics and Business

Introduction

The interrelated nature of product price, quality and reputation bears great relevance for potential consumers and hence, for producers as well. This is especially true for the wine market, where consumers face a huge variety of different alternatives characterized by a large set of attributes. It is rather plausible that in many cases, consumers lack information about the true quality of a given wine (Oczkowski & Doucouliagos, 2014). Some of these product traits cannot be evaluated until consumption (i.e. wine is considered to be an experience good). Some aspects of quality, on the other hand, may readily be proxied at the purchase occasion by diverse indicators like vintage, type of grape, the identity of the winegrower, etc. Very prominent sources of additional information on the actual value of the to-be-consumed product are third-party (expert) reviews[4]. These guides attempt to reduce the gap between buyers and producers caused by asymmetric information (Dubois & Nauges, 2010) and can serve to reflect demand side influences on prices (Oczkowski, 2001). According to Johnson & Bruwer (2004) wine guides could be considered as means of risk reduction in the absence of perfect information used by consumers. Moreover, recently, Friberg & Gronqvist (2012) presented positive effects on wine demand that occur after the release of favorable reviews. Thus wine experts and guides effectively influence consumer behavior and as such must be considered an important vehicle of capturing potential buyers, which in the meantime strengthen shoppers´ loyalty as well. Therefore a better understanding of these determining factors of wine prices seems to be valuable not only to market researches but also to every stakeholder in the wine industry.

In line with the theoretical model of Shapiro (1983), consumers´ decisions are driven mainly by the reputation of the firm, which again is determined by prior quality. Firms with a good reputation for high quality products have a stronger incentive to deliver high quality products in the future too. It implies that a firm’s investment in quality improvement not only has an immediate and direct positive effect on market prices, but may be responsible for an even stronger (indirect) long-run positive effect. From an empirical standpoint, these two must be separately accounted for.

The paper at hand aims at quantifying the price effects of Austrian bottled wine quality defined as grades given by a well-established wine guide (Falstaff). Simultaneously, we disentangle short-run effects from long-run influences of quality that are operating thorough the channel of the overall reputation of wine makers. Our contribution to the existing literature is manifold. By using a unique, very detailed and extensive data set of Austrian wines, besides accounting for the endogeneity of wine quality and reputation, we explicitly attempt to detect and correct for the potential bias introduced by the sample selection and which is linked to the mechanism by which individual wines are selected to be graded by the guide. This aspect has so far been neglected by similar works. As a further appealing feature of the current paper, to the best of our knowledge, the number of empirical studies using wine data from the German-speaking region and a hedonic pricing model is really scarce indeed.

[4] The influence of expert reviews has been found to be a key factor in other areas like art (Ginsburgh, 2003), movies markets (Reinstein & Snyder, 2005), and “primeur” wine sales (Héla Hadj, et al., 2008) too.

Literature review

The foundation of economic theory concerned with the crucial relationship between product price and quality can be traced back to the seminal paper by Rosen (1974), which posits that goods are valued for their utility-generating attributes. The model predicts that a higher sensory quality results in higher prices. This relationship for experience goods has been studied extensively in theoretical works (e.g. Shapiro (1983), Tirole (1996), and Mahenc & Meunier (2003)) dealing with quality signaling from the sellers´ perspective. These and other extensions of the original framework of Rosen focus on the role of reputation, individual sensory quality variables – later – endogeneity and measurement error (see the meta-analysis of Oczkowski & Doucouliagos (2014)).

Shapiro (1983) presents a theoretical framework to examine the effects of individual producer reputation on prices. He demonstrates that reputation makes it possible for high-quality producers to sell their goods at a higher price that may be interpreted as a return from investments in building reputation. By analyzing the long term dynamics of such investments in reputation, Rob & Fishman (2005) show that these can initiate a virtuous circle where reputation becomes increasingly important. On the demand side, it is costly for consumers to improve their information about product quality. Given imperfect information, using the reputation of a producer in the decision making process can be an effective way for consumers to become better informed (Schamel & Kym, 2003). Based on the theoretical work of Tirole (1996), Landon & Smith (1997) report that in case of the Bordeaux wine, overall reputation has a considerably stronger impact on consumers´ decisions than short-term quality variations. The fact that reputation plays a very important role in determining wine prices and that it is linked to prior quality reviews, is also underpinned by the studies of Oczkowski (2001) and Benfratello, et al. (2009) on Australian and Italian premium wines, respectively.

Providing a fruitful domain for empirical research, wine has been a subject of many works that attempt to measure the effects of additional sources of information regarding quality on product prices. Hedonic price analysis aims at measuring the implicit price of a specific product attribute. Being concerned with quality factors influencing vegetable prices, in the realm of agricultural economics, Waugh (1928) was the first to apply the hedonic price model. By using quality ratings of wine guides, quality was linked to prices in many further studies. More recently, it was first Ashenfelter (1989) who applied this method to the wine industry to quantify the impact of expert ratings of product quality on the auction prices of Bordeaux wines. Oczkowski (1994) estimates (mostly objective) attribute effects of Australian table wines. In his analysis, objective attributes like cellaring potential, grape variety, grape vintage, region, and producer size are found to be significant. The only nonobjective variable, the rating of an Australian wine guide, is a very important factor influencing prices. However, the lack of modelling the role of winemakers´ reputation in consumer decisions does not allow for a comparison of effects in the short and long run. Later, Schamel & Kym (2003) applied the hedonic price model to Australian and New Zealand data from the Winestate magazine and James Halliday's ratings differentiating implicit prices for sensory quality ratings, wine varieties and – this time – for regional and winery reputation as well. Investigating Bordeaux wine, Landon & Smith (1997) and Landon

& Smith (1998) find that, as consumers react slowly to changes in product quality, reputation has an even more important effect on consumers relative to current quality ratings. Their finding underpins the fact that neglecting reputation effects might lead to an over-estimated impact of short-run changes in quality. Costanigro, et al. (2007), Roberts & Reagans (2007), and Cuellar & Claps (2013) proxy wine quality by Wine Spectator magazine rating indices and find significant influence of expert reviews on New World wine prices. Roma, et al. (2013) suggest that, besides quality indicators, wine price strongly depends on objective features such as vintage, alcoholic content, geographical origin, grape variety, producer size and cellaring potential as well. All in all, evidence from a growing number of empirical studies is consistent with the expectation that quality is positively linked to wine prices and that reputation is a factor that has to be taken into consideration while estimating the price effect of quality (see the very comprehensive meta-analysis of Oczkowski & Doucouliagos (2014) conduted by investigating many separate studies).

By recognizing the deficiency of earlier approaches regarding the potential endogeneity of quality and reputation, Oczkowski (2001) extends the analysis on Australian wines and accounts for measurement errors of these product attributes. While, his findings reveal significant reputation effects, the author warns that "standard OLS procedures may seriously distort the statistical significance of attributes". In this vain, San Martin, et al. (2008) report the same outcome regarding Argentinean wine export data. As experts may be wrong in assessing wine quality at the time of primeur sales, Dubois & Nauges (2010) deal with the potential endogeneity of quality (modelled by using expert opinions on Bordeaux wine) caused by omitted variable bias (i.e. unobserved quality). Their empirical results also reinforce the assumption that failing to control for endogeneity may lead to an over-estimated influence of experts´ grades on the "primeur" price. Haeger & Storchmann (2005) and Blackman Bicknell & MacDonald (2012), on the other hand, use U.S. and New Zealand data sets, respectively and come to the conclusion that the rating variable is indeed exogenous and OLS estimates are consistent. Consequently, it seems to be an empirical question, if endogeneity is indeed a problem, but if it is the case, it has to be accounted for appropriately. To highlight the potential complications caused by endogeneity, by reviewing 36 influential studies from the field, Oczkowski & Doucouliagos, (2014) report that it is indeed an important issue, even if the number of works that account for it is relatively low.

The source of potential endogeneity might also be related to the sample selection process of graded wines by the wine guide in question – an aspect omitted in similar works so far. Moreover, the fact that reputation is endogenous and is determined by high quality produced over prior time periods is ignored in many existing studies. Hence, our effort is to fill these gaps regarding the role played by wine quality and reputation in price formation.

The rest of the paper is structured as follows. Section 2 describes the unique data set and explains how the quality and reputation variables in the analysis are constructed. Section 3 presents the empirical model and the estimation strategy. Section 4 contains the estimation results. Finally, Section 5 concludes the paper.

Data

The sample used in the present paper consists of more than 7,000 single wines which are produced by Austria's most prestigious wineries and covers about 35% of the annual national production of quality wines. For each wine, numerous characteristics are available, such as its price per standard bottle, type (white/red/rose/sweet), the year of harvest, the temporal distance between the harvest and bottling, the variety of grape the wine has been made of, as well as the size of the winery and the size of its assortment. The summary statistics are to be found in Appendix A. This information is matched with data on the sensory quality of the specific bottle of wine, which is obtained from consulting the most influential written guide on Austrian wine – the annually published Falstaff-Wine-Guide. Experts grade on a scale from 1 to 100 on color and appearance, aroma and bouquet, as well as flavor and finish. The average quality of wines in our sample is 89.05 and only wines on the scale between 82 and 100 are included in the Falstaff publication. The number of wines graded per year differs across wineries (see next section). The reputation of a winery is classified on a scale from 1 to 3 between 2004 and 2006 and from 1 to 5 in 2007. To avoid the different scaling of this variable to affect our estimation results, we use relative reputation in the empirical analysis (defined as the level of reputation relative to the maximum level of reputation in that particular year). All the data were collected by one of the authors (Andreas Huber), a more detailed description is available in Huber (2013).

Empirical Model and Estimation Strategy

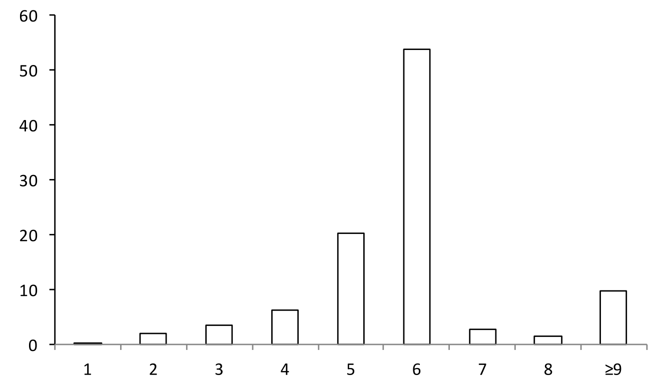
In order to estimate the short- and long-run price effects of quality we have to deal with concerns on endogeneity and sample selection.

First, we consider the reputation as an endogenous variable and use the winery's (average) quality levels in previous years as exogenous instruments. We have to instrument the variable on reputation for two reasons: (i) As we are interested in the long-run price effects of quality, we have to link both quality in previous years with reputation and reputation with price. (ii) The variable on reputation does not (only) measure what it is expected to do, namely summarizing past performance, but is also influenced by the contemporaneous quality levels produced by a winery (see Table B1 in Appendix B).

Second, not all wines produced by the wineries considered in the Falstaff wine guide are graded, but only a sample of 40% in the sample period between 2004 and 2007. The sample selection causes the variable on product quality – although probably exogenous in the full sample – to be endogenous in the selected sample. The sample selection therefore has to be modeled explicitly to correct for the potential bias introduced by the sample selection and we have to instrument the quality evaluation of the wine guide. The mechanism of the sample selection needs some explanation: For most wineries, space in the wine guide is limited to six graded wines per vintage. Figure 1 below shows that for more than half of all wineries exactly six wines are graded per year. For 20% of all wineries five wines are graded, and for 12% four or less wines are evaluated. Only for 14% of all wineries more than six wines are assessed by the wine guide. These are usually wineries with high reputation: Reputation

averages 0.68 for this group of wineries, whereas the average reputation is 0.20 for wineries with six or less graded wines. Each winery decides on which wines to send to the wine guide for evaluation. While the wine guide need not publish the results of all wines received by the wineries, a rejection to publish the grade is a rare event. So, basically, the winegrowers decide which wines are selected for grading.

Figure 1: Number of Graded Wines per Winery (in percent)



The selection of wines will therefore depend on the quality of the wine, and the missing values can be described as "non missing at random" ("NMAR").[5] However, the wineries usually publish their price lists and start selling the wines of the current vintage between February and June, while the wine guide is not published until July of the same year. The price of a wine and the probability of a wine to be selected for evaluation can therefore not directly depend on the grade a wine receives from the wine guide, but depends on the quality perception of the winery (the winegrower), which is unobservable. The evaluation of the wine guide will be correlated with the quality perceived by the winegrower and can therefore serve as a proxy variable. In the full sample the difference in the quality evaluation of all wines of a winery between the wine guide and the winegrower might not be systematically and therefore not lead to biased parameter estimates. The endogeneity of the quality evaluation of the wine guide comes from the sample selection: If a winegrower overstates the quality of a particular wine, then the price and the probability of the wine to get selected for grading will be higher than expected (based on the quality evaluation of the wine guide). Therefore, the sample selection is correlated with the dependent variable of the structural equation of interest (the price), even if we control for the quality evaluation of the wine guide. As a consequence, the product quality is correlated with the error term in the structural equation of interest, which causes product quality to be endogenous in the selected sample and listwise deletion to produce inconsistent parameter estimates (see Allison (2001) and Little (1992)). The missing data mechanism is therefore nonignorable and the sample selection has to be modeled along with the overall model to gain consistent parameter estimates (Cameron & Trivedi, 2005, p. 927 f.).

[5] See e.g. Little & Rubin (2002) for an overview on missing data.

The relationship between price, quality and reputation can be stated as follows: [6]

(1) $Ln(Price_{iwt}) = \alpha_1 Qual_{iwt} + \alpha_2 Rep_{wt} + \boldsymbol{X}_{iwt}\boldsymbol{\beta} + IMR_{iwt}\gamma + \varepsilon_{iwt}$

(2) $Qual_{iwt} = (\boldsymbol{X}_{iwt}, IMR_{iwt}, \boldsymbol{Z}_{iwt})\boldsymbol{\delta}_1 + u_{1,iwt}$

(3) $Rep_{wt} = (\boldsymbol{X}_{iwt}, IMR_{iwt}, \boldsymbol{Z}_{iwt})\boldsymbol{\delta}_2 + u_{2,iwt}$

(4) $Select_{iwt} = 1\left((\boldsymbol{X}_{iwt}, \boldsymbol{Z}_{iwt})\boldsymbol{\delta}_3 + u_{3,iwt} > 0\right)$

The logarithm of the price of wine i of winery w at time t, $\ln(Price_{iwt})$, depends on the quality of the wine, $Qual_{iwt}$, on the reputation of the winery Rep_{wt} and on other exogenous variables (summarized in $\boldsymbol{X}_{iwt}$). $Select_{iwt}$ equals 1 if a wine is selected for evaluation, and zero otherwise. The matrix $\boldsymbol{Z}_{iwt}$ summarizes the additional instrumental variables necessary for identification and is included in the regression on product quality (2), on reputation (3), and in the selection equation (4). From the selection equation we calculate the inverse Mills ratio (*IMR*), which is then included as an additional regressor in the main regression (1). α_1, α_2, $\boldsymbol{\beta}$, γ, $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_2$ and $\boldsymbol{\delta}_3$ are the respective (vectors of) parameter estimates, and ε_{iwt}, $u_{1,iwt}$, $u_{2,iwt}$ and $u_{3,iwt}$ are the error terms.

The variables summarized in $\boldsymbol{X}_{iwt}$ include dummy variables indicating the vintage, the variety of the grape, the type of wine, the type of sweet wine (if it is a sweet wine), and variables indicating the time span between the harvest of the grapes and the year the winery starts to sell the wine at the market. Most importantly, $\boldsymbol{X}_{iwt}$ includes fixed winery effects. These dummy variables control for unobserved and time-invariant heterogeneity between wineries like differences in costs, consumer preferences or managerial skills.

Technically, matrix $\boldsymbol{Z}_{iwt}$ must contain at least two variables (due to the nonlinearity of the Mills ratio), but should contain at least three variables to allow us to convincingly identify product quality, firm reputation and the sample selection. While all exogenous variables should be included in the selection equation (4) and in the auxiliary regressions (2) and (3) we have clear expectations which variables can determine the selection and which variables serve as instruments for the endogenous variables reputation and quality, an issue to which we will turn now:

- The average quality of a winery at year $t - k$, with $k = 1,\ldots,4$ is calculated as the average grade of all graded wines of a particular winery at year $t - k$. These variables are used to explain reputation. We expect to find a positive effect on past average quality that diminishes over time.
- Dummy variables indicating the year the grapes were harvested. These variables should influence quality, as these variables also summarize overall weather conditions. Prices should be affected by the vintage, but not by the year the grapes are harvested.

[6] The following methodological considerations are based on Wooldridge (2001), who considers a very similar setting (see p. 567 ff.).

- The number of times a particular wine is selected for evaluation between 1999 and 2002 (i.e. before the time period under investigation, which is between 2004 and 2007). This variable indicates that this particular wine is of relative high quality (compared to other wines of the winery) – e.g. due to favorable soil conditions, a good climate for a particular type of wine or specific human skills – and that the winegrower views the grade received by the wine guide as a positive market signal. This variable should be positively correlated with the quality of a wine and with the probability of a wine to be considered for evaluation.
- The number of produced wines by the winery in a particular vintage. As space in the wine guide is scarce (at least for most wineries), an increase in the number of wines should reduce the probability of a single wine to be considered for publication.

We follow Wooldridge (2001) and start estimating the system of equation by estimating a probit model of $Select_{iwt}$ on all exogenous variables to obtain the inverse Mills ratio: IMR_{iwt}. We then apply a 2-stage least square (2SLS) procedure on equation (1) for the sub-population of all wines selected for evaluation using $\boldsymbol{X}_{iwt}$, $\boldsymbol{Z}_{iwt}$ and IMR_{iwt} as instruments for Rep_{wt} and $Qual_{iwt}$.

Table 1 below reports the results of the probit model explaining the selection of wines. The number of times the particular wine received a grade between 1999 and 2002 has a positive impact on the probability of the wine to be selected for evaluation in a particular year in the sample period (2004-2007). If a winery increases the number of wines produced, the probability of a single wine to be considered for evaluation is reduced. Both coefficients take the expected signs and are significantly different from zero at the 1% level.

The year the grapes are harvested does not have a systematic effect on the selection of wines. An increase in past quality tends to increase the probability of a wine to be considered for grading. Conditioned on the number of produced wines this indicates that an increase in average quality tends to increase the number of graded wines in the following years. However, the result is not very robust and the parameter estimates are only significantly different from zero in the previous year, but insignificant in the years before. A longer time span between harvest and bottling is associated with a higher probability of a wine to be selected for evaluation. This might be explained by the result that these wines can be expected to receive in general higher grades (see Table C1 in Appendix C).

Table 1: Regression Results Explaining the Selection of Wines for Evaluation

Dependent Variable Method Model	Select Probit [1]
CONST	−4.5457 (−1.21)
Number of times Wine received Grade between 1999 and 2002	0.3513 (39.95)
Number of different Wines of Winery (in Year t)	−0.0406 (−4.75)
Harvest 2002	0.0108 (0.06)
Harvest 2003	−0.0088 (−0.07)
Harvest 2004	−0.0684 (−0.70)
Harvest 2005	−0.0767 (−1.28)
Average Quality (at t-1)	0.0408 (1.98)
Average Quality (at t-2)	0.0122 (0.66)
Average Quality (at t-3)	−0.0242 (−1.50)
Average Quality (at t-4)	0.0095 (0.62)
Time between Harvest and Bottling: 2 Years	0.4609 (7.94)
Time between Harvest and Bottling: 3 Years	0.8468 (6.87)
Winery Effects	Yes (346)
Vintage Effects	Yes (3)
Variety of the Grape	Yes (31)
Type of Wine	Yes (3)
Type of Sweet Wine	Yes (3)
N	16,758
log-likelihood	−8,721.90
Pseudo-R²	0.241
Chi²-test statistic on instruments (df) [p-value]	1,638.52 (10) [0.0000]
	1,632.91 (6) [0.0000]

Notes: z-values are reported in parentheses. Chi²-test rejects the hypothesis that all instruments are jointly zero. The number of instrumental variables depends on the specification of the last stage regression reported in Table 2 below and equals 10 in specification [2] and 6 in specification [3].

Results

As we are interested in the price effects of quality, but not in explaining quality itself, the auxiliary regression on quality is summarized in Table C1 in Appendix C. The results on the regression on prices and on reputation are summarized in Table 2.

In model [1] we ignore the sample selection problem and the possible endogeneity of quality. An increase in quality by one point increases prices by 10.5% and an increase in reputation from 0 to 1 increases prices by 67.7%. If the time span between the harvest of the grapes and bottling is two years instead of one year (the reference category) prices increase by 38.3%, a time span of three or more years is associated with a mark-up of 76.8%. The average quality the winery produced in previous years has positive effects on reputation. The effect in general decreases over time, with the exception that the effect of average quality at $t-3$ is slightly larger than at $t-2$. The current quality of a particular wine (which is included as it is not considered to be an endogenous variable in this specification) has a positive effect on reputation, casting doubts on whether reputation truly summarizes only past performance. Surprisingly, the time between harvest and bottling has a negative impact on reputation. All described coefficients are significantly different from zero (at least) at the 5% level.

Statistical tests on the exogeneity of the reputation and on the joint significance of the instruments support our perception of the endogeneity of reputation and that the instruments significantly contribute to explain this variable. A Chi²-test of overidentifying restrictions indicates that the instruments may not be valid, casting doubts on the overall model specification.

In specification [2] we control for the sample selection and the endogeneity of product quality due to the sample selection and (i) include the inverse Mills ratio (*IMR*), based on the selection equation, and (ii) instrument product quality. The *IMR* is significantly different from zero at the 1% level, indicating the need to correct for the (non-random) sample selection.[7] Accounting for the sample selection reduces the impact of an increase in one quality point on prices to 7.1% while an increase in reputation from 0 to 1 causes prices to increase by 38.8%. The impact of quality therefore drops by roughly one third and the influence of reputation by nearly one half. The *IMR* is not significantly different from zero in the auxiliary regression on reputation. Consequently, correcting for the sample selection hardly affects these regression results. Statistical tests support the endogeneity of both quality and reputation and that the instruments are jointly significant. A Chi² test that the instruments are uncorrelated with the error term is not rejected at the 5% significance level.

As the hypothesis that the excluded instruments are valid instruments is not rejected at the 5%, but at the 10% level, we include the average quality of previous years (instead of

[7] As the *IMR* is a so-called *generated regressor* rather than an observed variable the standard errors and the test statistics of the structural equation should be corrected (Wooldridge, 2001). However, adjusting the variance of the parameter estimates of the structural equation "because of the two-step estimation is cumbersome" (p. 564). We circumvent this problem by using bootstrapped standard errors. As the standard deviations of the parameter estimates are hardly affected when using bootstrapped standard errors we report these results in Table D1 in Appendix D only. The results support (Wooldridge, 2002) supposition that "[i]n many cases, the adjustments do not lead to important differences" (p. 562).

reputation) directly in our regression analysis (specification [3]). Again, the *IMR* is significantly different from zero and statistical tests indicate the endogeneity of quality in our sample. In this specification there are no concerns on the validity of the instruments for product quality, as the respective test is not rejected at the 10% significance level. The impacts of average quality levels of the wineries in the previous year (in the year before the previous year) are positive and statistically different from zero at the 1% (10%) significance level. The effect of the average quality of older vintages is still positive, but smaller and not statistically different from zero. A non-recurring increase in the quality of all wines of a winery by one point in a particular vintage causes the winegrower to increase prices by nearly 8% in this year, more than 2% in the next year, and by less than 1% in years afterwards.

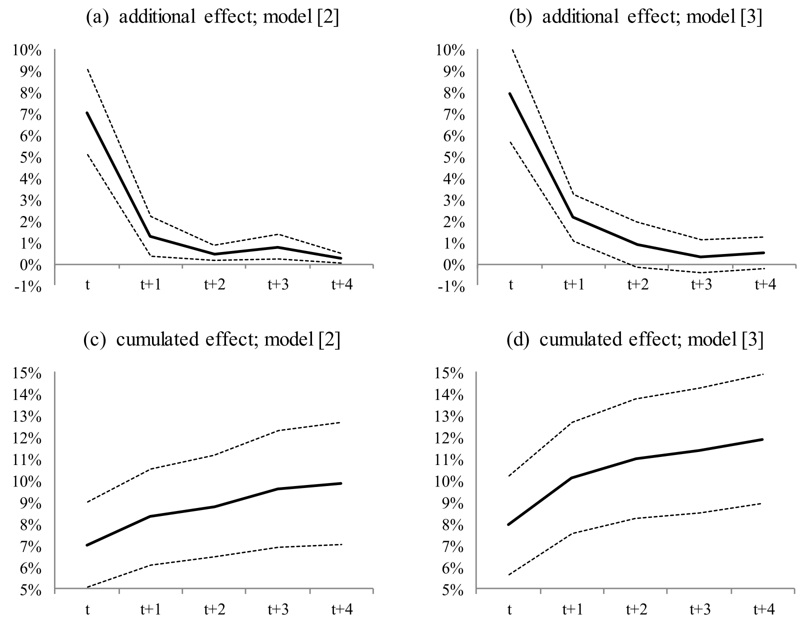
Table 2: Regression Results Explaining the Effects of Quality and Reputation on Prices

Dependent Variable	ln(Price)	Reputation	ln(Price)	Reputation	ln(Price)
Method	2SLS		2SLS		2SLS
Model	[1]		[2]		[3]
CONST	–7.6367	–6.7727	–4.4488	–6.0788	–8.6281
	(–42.95)	(–16.87)	(–4.78)	(–15.43)	(–4.44)
Quality	0.1049	0.0058	0.0705		0.0793
	(54.41)	(8.51)	(6.90)		(6.70)
Reputation	0.6770		0.3884		
	(5.48)		(2.85)		
Average Quality (at t – 1)		0.0338		0.0330	0.0217
		(16.51)		(15.60)	(3.87)
Average Quality (at t – 2)		0.0134		0.0125	0.0089
		(7.42)		(6.79)	(1.65)
Average Quality (at t – 3)		0.0209		0.0205	0.0036
		(13.95)		(13.53)	(0.90)
Average Quality (at t – 4)		0.0061		0.0063	0.0053
		(4.42)		(4.48)	(1.46)
Time between Harvest and	0.3830	–0.0098	0.4293	–0.0026	0.4161
Bottling: 2 Years	(36.69)	(–2.45)	(22.94)	(–0.43)	(20.01)
Time between Harvest and	0.7678	–0.0249	0.8561	–0.0077	0.8294
Bottling: 3 Years	(24.70)	(–1.98)	(20.97)	(–0.53)	(18.66)
IMR			–0.0330	0.0005	–0.0271
			(–2.73)	(0.04)	(–2.22)
Winery Effects	Yes (353)	Yes (353)	Yes (346)	Yes (346)	Yes (346)
Vintage Effects	Yes (3)	Yes (3)	Yes (3)	Yes (3)	Yes (3)
Variety of the Grape	Yes (31)	Yes (31)	Yes (31)	Yes (31)	Yes (31)
Type of Wine	Yes (3)	Yes (3)	Yes (3)	Yes (3)	Yes (3)
Type of Sweet Wine	Yes (3)	Yes (3)	Yes (3)	Yes (3)	Yes (3)
Instrumented Variables	Reputation		Reputation Quality		Quality
N	7,403	7,403	7,358	7,358	7,361
R^2	0.823	0.940	0.822	0.940	0.833
Overidentification restriction: Chi² test statistic (df) [p-value]		9.246 (3) [0.0262]		14.796 (8) [0.0632]	8.334 (5) [0.1388]
Exogeneity: Chi² test statistic (df) [p-value]		51.514 (1) [0.0000]		60.487 (2) [0.0000]	6.68462 (1) [0.0097]
Exogeneity: F-test statistic (df) [p-value]		49.733 (1,7004) [0.0000]		28.978 (2,6964) [0.0000]	6.394 [0.0115] (1,6965)
F-test statistic on instruments (df) [p-value]		95.630 (4,7002) [0.0000]		37.940 (10,6958) [0.0000]	

Notes: z-values (t-values) for the regressions on prices (reputation) are reported in parentheses and are based on robust standard errors. Tests on overidentifying restrictions (H0: Instruments are valid) and endogeneity of instrumented variables (H0: Instrumented variables are exogenous) are based on (Wooldridge, 1995) robust score tests and (Wooldridge, 1995) robust regression tests. The regression on reputation in specification [2] and [3] includes all exogenous variables of the regression on quality (in model [2]) and the selection equation. Parameter estimates of the respective variables are not reported for convenience. The F-test on instruments tests the hypothesis that the coefficients on the instruments in the regression on reputation are jointly zero.

Figure 2 below illustrates the size of short and long run price effects of quality improvements. We apply a bootstrap simulation technique to account for the uncertainty of the estimated parameter values. Each parameter is drawn randomly from a normal distribution with the mean and the standard deviation obtained from the regression in model [2] and [3] of Table 2. Figure 2 (a) and (b) show the price effects if a winery increases the quality of each wine in a particular year by one point. In the year after the (non-recurring) quality improvement wineries are expected to charge 1.3% (model [2]) or 2.2% (model [3]) higher prices. In the following years the price mark-ups are still positive, but below 1%. The price increases in the four years following the quality improvement are however sizable and aggregate to 40% (model [2]) or 50% (model [3]) of the contemporaneous price increase. Figure 2 (c) and (d) show the aggregate price effects or – put differently – the price increases if winegrowers succeed in increasing the wine quality permanently by one quality point.

Figure 2: Short- and Long-Run Price Effects of Quality



Notes: The results are based on the point estimates and on the standard deviation of the respective parameter estimates of model [2] (figure (a) and (c)) and model [3] (figure (b) and (d)). The solid line denotes the median of the price effects and the dotted lines denote the 2.5% and the 97.5% percentiles.

Summary and Conclusions

Given our findings, it can be stated that the proposed approach of accounting for self-selection regarding the selection procedure of wines to be graded by experts and considering endogenous wine quality and winery reputation was justified. Based on an impressive data set, compared to a model variant that ignores such selection bias and seems to overestimate the influence of quality and reputation, the proposed 2SLS specification decouples short- and long-run effects and leads to plausible results. A one point increase in the Falstaff guide grade leads to a 7.1% price increase in the same year. Similarly, reputation exhibits a strong positive effect as well, whereas, importantly, it is also influenced by prior average quality scores of the given winery. It is important to recognize that long-run effects of quality might be indication for the relative (to supply) importance of demand influences on prices (Oczkowski & Doucouliagos, 2014). Consumers seem rely on the wine's reputation, which is reflected by previous vintages.

These findings may bear importance for all the stakeholders in the wine industry. The results from the present analysis provide useful insights on the evaluation of high quality and reputation in the market. This information can be valuable for producers of premium wine in Austria in deciding whether and to what extent to invest in quality improvements and also for policy makers in designing adequate measures to improve the competitiveness of Austrian wines.

By differentiating their products through improving quality, winemakers might be in a position to capture consumers and increase their market share, and at any rate, accumulate consumer goodwill and recognition. Eventually, wine prices can be used to signal quality, hence, consumers´ uncertainty regarding purchases might be reduced by price information that is associated with a given level of quality (Oczkowski & Doucouliagos, 2014).

As far as further research is concerned, it might be interesting to take, on the one hand, potential geographic differences and/or spatial interactions of price effects into consideration by comparing different countries or smaller regional territories in regard of the impact of quality and reputation. On the other hand, consumer characteristics and habits may be seen as a further possible extension that could provide us an even more detailed picture and deeper understanding of how wines are or should be priced.

References

Allison, P. D., 2001. Missing Data. In: *Sage University Papers Series: Quantitative Applications in the Social Sciences.* Thousand Oaks, CA: Sage, pp. 107-136.

Ashenfelter, O., 1989. How Auctions Work for Wine and Art. *Journal of Economic Perspectives*, Volume 3, pp. 23-26.

Benfratello, L., Piacenza, M. & Sacchetto, S., 2009. Taste or Reputation: What Drives Market Prices in the Wine Industry? Estimation of a Hedonic Model for Italian Premium Wines. *Applied Economics*, 41(17), p. 2197–2209.

Blackman Bicknell, K. & MacDonald, I. A., 2012. Regional reputation and expert opinion in the domestic market for New Zealand wine. *Journal of Wine Research*, 23(2), p. 172–184.

Cameron, C. A. & Trivedi, P. K., 2005. *Microeconometrics: Methods and Applications.* New York; Melbourne; Madrid; Cape Town; Singapore; Sâo Paulo: Cambridge Universtiy Press.

Costanigro, M., McCluskey, J. & Mittelhammer, R. C., 2007. Segmenting the wine market based on price: Hedonic regression when different prices mean different products. *Journal of Agricultural Economics*, 58(3), pp. 454-466.

Cuellar, S. S. & Claps, M., 2013. Differential effects of brand, ratings and region on willingness to pay: a hedonic price approach. *Journal of Wine Research*, 24(2), pp. 138-155.

Dubois, P. & Nauges, C., 2010. Identifying the effect of unobserved quality and expert reviews in the pricing of experience goods: empirical application on Bordeaux wine. *International Journal of Industrial Organization*, 28(3), pp. 205-212.

Estrella Orrego, M. J. E., Defrancesco, E. & Gennari, 2012. A The Wine Hedonic Price Models in the “Old and New World”: State of the Art. *Journal of the Faculty of Agricultural Sciences, UNCuyo*, 44(1), p. 205–220..

Friberg, R. & Gronqvist, E., 2012. Do Expert Reviews Affect Demand forWine?. *American Economic Journal – Applied Economics*, 4(1), p. 193–211.

Ginsburgh, V., 2003. Awards, Success and Aesthetic Quality. *Journal of Economic Perspectives*, 17(2), pp. 99-111.

Haeger, J. W. & Storchmann, K., 2005. Prices of American Pinot Noir wines: Climate, craftsmanship, critics. *Agricultural Economics*, Volume 35, p. 67–78..

Héla Hadj, A., Lecocq, S. & Visser, M., 2008. The Impact of Gurus: Parker Grades and En Primeur Wine Prices. *The Economic Journal*, 118(529), pp. 158-173.

Huber, A., 2013. *Qualität und Preis auf dem österreichischen Weinmarkt: Eine hedonische Analyse.* Dissertation: Vienna University of Economics and Business.

Johnson, T. & Bruwer, J., 2004. Generic Consumer Risk-reduction Strategies (RRS) in Wine-related Lifestyle Segments of the Australian Wine Market. *International Journal of Wine Marketing*, 16(1), pp. 5-35.

Landon, S. & Smith, C. E., 1997. The use of quality and reputation indicates by consumers: The case of Bordeaux wine. *Journal of Consumer Policy*, 20(3), p. 289–323.

Landon, S. & Smith, C. E., 1998. Quality expectations, reputation, and price. *Southern Economic Journal*, 64(3), pp. 628-647.

Little, R. J., 1992. Regression with Missing X’s: A Review. *Journal of the American Statistical Association*, 87(420), p. 1227–1237.

Little, R. J. A. & Rubin, D. B., 2002. *Statistical Analysis with Missing Data. 2.* Hoboken, New Jersey: John Wiley & Sons.

Mahenc, P. & Meunier, V., 2003. Market and Signals of Quality. *The RAND Journal of Economics*, 34(3), pp. 478-494.

Oczkowski, E., 1994. A Hedonic Price Function for Australian Table Wine. *Australian Journal of Agricultural Economics*, Volume 38, pp. 93-110.

Oczkowski, E., 2001. Hedonic Wine Price Functions and Measurement Error. *Economic Record*, 77(239), pp. 374-382..

Oczkowski, E. & Doucouliagos, H., 2014. Wine Prices and Quality Ratings: A Meta-regression Analysis. *American Journal of Agricultural Economics*, 97(1), pp. 103-121.

Reinstein, D. A. & Snyder, C. M., 2005. The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics. *The Journal of Industrial Economics*, 53(1), pp. 27-51.

Roberts, P. W. & Reagans, R., 2007. Critical exposure and price-quality relationships for New World wines in the US market. *Journal of Wine Economics*, 2(1), pp. 84-97.

Rob, R. & Fishman, A., 2005. Is Bigger Better? Customer Base Expansion through Word-of-Mouth Reputation. *Journal of Political Economy*, 113(5), pp. 1146-1175.

Roma, P., Di Martino, G. & Perrone, G., 2013. What to show on the wine labels: a hedonic analysis of price drivers of Sicilian wines. *Applied Economics*, 45(19), pp. 2765-2778.

Rosen, S., 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economics*, Volume 82, pp. 34- 55.

San Martin, G. J., Troncoso, J. L. & Brümmer, B., 2008. Determinants of Argentinean wine prices in the U.S. *Journal of Wine Economics*, 3(1), p. 72–84.

Schamel, G. & Kym, A., 2003. Wine quality and varietal, regional and winery reputations: hedonic prices for Australia and New Zealand. *Economic Record*, 79(246), pp. 357-369.

Shapiro, C., 1983. Premiums for high quality products as rents to reputation. *Quarterly Journal of Economics*, Volume 98, pp. 659-680.

Tirole, J., 1996. A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality). *Review of Economic Studies*, Volume 63, pp. 1-22.

Waugh, F. V., 1928. Quality Factors Influencing Vegetable Prices. *Journal of Farm Economics*, Volume 10, pp. 185-196.

Wooldridge, J. M., 1995. Score diagnostics for linear models estimated by two stage least squares. In: Maddala, G. S., Phillips, P. C. B., Srinivasan, T. N. (eds), *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C. R. Rao.* Oxford: Blackwell, pp. 66-87.

Wooldridge, J. M., 2001. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

Wooldridge, J. M., 2002. *Introductory Econometrics: A Modern Approach.* Mason, OH: South Western Educational Publishing.

Appendix A

Table A1: Descriptive Statistics of Variables used in the Empirical Analysis

Variable	Number of Observations	Mean	Standard Deviation	Min	Max
Price	24,547	9.80	6.44	2.20	100.00
Quality	9,914	89.05	2.05	83	98
Reputation	22,210	0.29	0.31	0	1
Select	24,547	0.40		0	1
Average Quality (at t – 1)	21,705	88.87	1.31	85.40	94.00
Average Quality (at t – 2)	20,792	88.94	1.43	85.33	94.38
Average Quality (at t – 3)	19,762	88.88	1.56	84.50	95.00
Average Quality (at t – 4)	18,370	88.70	1.69	84.00	95.00
Time between Harvest and Bottling: 1 Year	24,547	0.75		0	1
Time between Harvest and Bottling: 2 Years	24,547	0.24		0	1
Time between Harvest and Bottling: 3 Years	24,547	0.01		0	1
Number of times Wine received Grade between 1999 and 2002	24,547	0.98	1.35	0	4
Number of different Wines of Winery (in Year t)	24,547	15.38	5.12	1	38
Harvest 2002	24,547	0.05		0	1
Harvest 2003	24,547	0.23		0	1
Harvest 2004	24,547	0.25		0	1
Harvest 2005	24,547	0.26		0	1
Harvest 2006	24,547	0.20			
Type of Wine:					
White wine	24,547	0.55		0	1
Red wine	24,547	0.33		0	1
Sweet wine ('Süßwein')	24,547	0.08		0	1
Rosé wine	24,547	0.03		0	1
Type of Sweet Wine:					
Spaetlese	24,547	0.02		0	1
Beerenauslese	24,547	0.02		0	1
Trockenbeerenauslese	24,547	0.03		0	1
Eiswein	24,547	0.02		0	1
Variety of Grape:					
Blauburger	24,547	0.01		0	1
Blaufränkisch	24,547	0.05		0	1
Blauer Portugieser	24,547	0.01		0	1
Blauer Wildbacher	24,547	0.00		0	1
Chardonnay	24,547	0.08		0	1
Cabernet Sauvignon	24,547	0.02		0	1
Cuvee Rot	24,547	0.09		0	1
Cuvee Weiss	24,547	0.04		0	1
Frühroter Veltliner	24,547	0.00		0	1
Gemischter Satz	24,547	0.01		0	1
Gelber Muskateller	24,547	0.02		0	1
Grüner Veltliner	24,547	0.15		0	1
Merlot	24,547	0.01		0	1
Muskat Ottonel	24,547	0.01		0	1
Müller Thurgau	24,547	0.01		0	1
Neuburger	24,547	0.01		0	1
Pinot Gris / Grauburgunder	24,547	0.01		0	1
Pinot Noir / Blauburgunder	24,547	0.03		0	1
Rotgipfler	24,547	0.01		0	1

Riesling	24,547	0.09	0	1
Rosé	24,547	0.02	0	1
Roter Veltliner	24,547	0.01	0	1
Sämling 88 / Scheurebe	24,547	0.01	0	1
Sauvignon Blanc	24,547	0.05	0	1
Schilcher	24,547	0.01	0	1
Sankt Laurent	24,547	0.02	0	1
Sortenvielfalt Weiss	24,547	0.01	0	1
Syrah	24,547	0.01	0	1
Traminer	24,547	0.02	0	1
Weissburgunder / Pinot Blanc	24,547	0.05	0	1
Welschriesling	24,547	0.04	0	1
Zierfandler	24,547	0.01	0	1
Zweigelt	24,547	0.10	0	1

Notes: 0.00 denotes that the value is rounded and not exactly zero. Standard deviations for dummy variables are not reported.

Appendix B

Table B1: Regression Explaining the Reputation of Wineries

Dependent Variable	Reputation	Reputation
Method	OLS	OLS
CONST	–5.4304	–10.7028
	(–5.54)	(–8.89)
Average Quality (at t)		0.0374
		(7.44)
Average Quality (at t – 1)	0.0307	0.0404
	(5.67)	(7.44)
Average Quality (at t – 2)	0.0111	0.0209
	(2.32)	(4.29)
Average Quality (at t – 3)	0.0167	0.0203
	(3.95)	(4.89)
Average Quality (at t – 4)	0.0060	0.0048
	(1.52)	(1.23)
Winery Effects	Yes (355)	Yes (353)
Vintage Effects	Yes (3)	Yes (3)
N	1,155	1,146
R^2	0.65	0.69

Notes: Regression is evaluated at the winery level. t-values are reported in parentheses.

Appendix C

Table C1: Regression Results Explaining Product Quality

Dependent Variable	Quality	Quality
Method	2SLS	2SLS
Model	[2]	[3]
CONST	140.7277	140.7326
	(22.54)	(22.55)
Number of times Wine received Grade between 1999 and 2002	0.3709	0.3712
	(7.92)	(7.93)
Number of different Wines of Winery (in Year t)	0.0149	0.0149
	(0.98)	(0.97)
Harvest 2002	0.1302	0.1304
	(0.44)	(0.44)
Harvest 2003	0.2512	0.2505
	(1.12)	(1.12)
Harvest 2004	–0.3236	–0.3236
	(–2.00)	(–2.00)
Harvest 2005	–0.3926	–0.3927
	(–4.06)	(–4.06)
Average Quality (at t – 1)	–0.2243	–0.2237
	(–6.59)	(–6.57)
Average Quality (at t – 2)	–0.2496	–0.2493
	(–8.20)	(–8.19)
Average Quality (at t – 3)	–0.1298	–0.1302
	(–4.77)	(–4.78)
Average Quality (at t – 4)	–0.0213	–0.0220
	(–0.85)	(–0.88)
Time between Harvest and Bottling: 2 Years	1.8117	1.8106
	(16.85)	(16.84)
Time between Harvest and Bottling: 3 Years	3.4322	3.4316
	(15.72)	(15.72)
IMR	1.1697	1.1703
	(5.18)	(5.18)
Winery Effects	Yes (346)	Yes (346)
Vintage Effects	Yes (3)	Yes (3)
Variety of the Grape	Yes (31)	Yes (31)
Type of Wine	Yes (3)	Yes (3)
Type of Sweet Wine	Yes (3)	Yes (3)
N	7,358	7,361
R^2	0.5355	0.5088
F-test statistic on instruments (df) [p-value]	25.84 (10, 6958) [0.000]	24.40 (6, 6961) [0.000]

Notes: t-values are reported in parentheses and are based on robust standard errors. The F-test on instruments tests the hypothesis that the coefficients on the instruments are jointly zero. The numbering of models [2]and [3] indicates the corresponding specification in Table 2.

Appendix D

Table D1 reports the regression results explaining prices when using bootstrapped standard errors. Note that the number of observations and the coefficients change slightly compared to the main results, summarized in Table 2, as we exclude all wineries with (on average) less than one graded wine per year or less than one nongraded wine per year. We exclude these wineries as the method of bootstrapping is a resampling method that draws a random sample with replacement of the same size as our data sample (see Wooldridge, 2001, 378 ff.) and as we need variation in the variable $Select_{iwt}$ within each winery (in each resampled data set) to apply this method. The rationale behind this is the following: If a particular winery produces only a small number of graded (ungraded) wines, it is likely that the bootstrapping method samples only ungraded (graded) wines of this winery. In this case the particular winery fixed effect explains the selection of wines of this winery perfectly. These observations therefore drop out of the regression explaining the selection, as this regression is estimated by a probit model. The results cannot be used for the method of bootstrapping if the number of observations differs between replications. To reduce the probability that some observations are predicted perfectly and are therefore excluded we decided to discard the wineries with a low number of graded or a low number of ungraded wines. These wineries account for 141 wines in the sample period and excluding these firms reduces the size of the sample by less than one percent.

We re-estimated model [2] and [3] of Table 2 with the reduced data sample and report bootstrapped and non-bootstrapped standard errors in Table D1. The size of the standard errors of the coefficients and (therefore) the z-statistics hardly change due to using bootstrapped standard errors.

Table D1: Regression Results Explaining Prices using Bootstrapped Standard Errors

Dependent Variable		ln(price)			ln(price)	
Method		2SLS			2SLS	
Model		[2]			[3]	
	Param. Est.	Boots-trapped S.E.	S.E.	Param. Est.	Boots-trapped S.E.	S.E.
CONST	–4.5317	1.0547	0.9273	–9.1587	2.2828	1.9407
Quality	0.0713	0.0115	0.0102	0.0812	0.0140	0.0117
Reputation	0.4261	0.1472	0.1383			
Average Quality (at t – 1)				0.0227	0.0058	0.0056
Average Quality (at t – 2)				0.0105	0.0060	0.0054
Average Quality (at t – 3)				0.0045	0.0041	0.0040
Average Quality (at t – 4)				0.0059	0.0038	0.0037
Time between Harvest and Bottling: 2 Years	0.4281	0.0202	0.0189	0.4132	0.0231	0.0209
Time between Harvest and Bottling: 3 Years	0.8537	0.0447	0.0409	0.8235	0.0506	0.0444
IMR	–0.0325	0.0133	0.0121	–0.0262	0.0138	0.0122
Winery Effects		Yes (334)			Yes (334)	
Vintage Effects		Yes (3)			Yes (3)	
Variety of the Grape		Yes (31)			Yes (31)	
Type of Wine		Yes (3)			Yes (3)	
Type of Sweet Wine		Yes (3)			Yes (3)	
Instrumented Variables		Reputation			Quality	
		Quality				
N (Selection equation)		16,617			16,617	
N (Structural Equation)		7,248			7,248	
R2		0.8203			0.8330	
# of Replications		586			586	

Notes: Robust standard errors are used within each replication. The numbering of models [2] and [3] indicates the corresponding specification in Table 2.

Appendix E

Table E1: Regression Results Explaining the Effects of Quality and Reputation on Prices using Reputation of the Previous Year

Dependent Variable	ln(price)		Reputation (at t – 1)	
Method		2SLS		
Model		[2]		
CONST	–5.5857	(–4.96)	–8.0796	(–22.29)
Quality	0.0828	(6.77)		
Reputation (at t – 1)	0.3963	(2.97)		
Average Quality (at t – 1)			0.0354	(18.72)
Average Quality (at t – 2)			0.0327	(19.11)
Average Quality (at t – 3)			0.0148	(10.91)
Average Quality (at t – 4)			0.0123	(9.69)
Time between Harvest and Bottling: 2 Years	0.4104	(19.20)	0.0019	(0.41)
Time between Harvest and Bottling: 3 Years	0.8165	(17.76)	0.0132	(1.28)
IMR	–0.0258	(–2.07)	0.0010	(0.09)
Winery Effects	Yes (346)		Yes (346)	
Vintage Effects	Yes (3)		Yes (3)	
Variety of the Grape	Yes (31)		Yes (31)	
Type of Wine	Yes (3)		Yes (3)	
Type of Sweet Wine	Yes (3)		Yes (3)	
Instrumented Variables		Reputation (at t – 1)		
		Quality		
N	7,361		7,361	
R2	0.8320		0.959	
Overidentification restriction: Chi² test statistic (df) [p-value]			14.458 [0.0706]	(8)
Exogeneity: Chi² test statistic (df) [p-value]			36.041 [0.0000]	(2)
Exogeneity: F-test statistic (df) [p-value]			17.079 [0.0000]	(2, 6967)
F-test statistic on instruments (df) [p-value]			62.580 [0.0000]	(10, 6961)

Notes: Notes: z-values (t-values) for the regressions on prices (reputation) are reported in parentheses and are based on robust standard errors. Tests on overidentifying restrictions (H0: Instruments are valid) and endogeneity of instrumented variables (H0: Instrumented variables are exogenous) are based on Wooldridge's (1995) robust score tests and Wooldridge's (1995) robust regression tests. The regression on reputation includes all exogenous variables of the regression on quality (in model [2]) and the selection equation. Parameter estimates of the respective variables are not reported for convenience. The F-test on instruments tests the hypothesis that the coefficients on the instruments in the regression on reputation are jointly zero.