**Experts, Reputation and Umbrella Effects:
Empirical Evidence from Wine Prices**

by

Dieter PENNERSTORFER
Christoph R. WEISS
Andreas HUBER

Working Paper No. 1908
March 2019

Experts, Reputation and Umbrella Effects: Empirical Evidence from Wine Prices

Abstract

We investigate the relationship between external quality evaluation via experts, firm reputation and product prices and extend the existing empirical literature in three dimensions. First, we empirically account for endogenous reputation effects. An increase in quality has an immediate positive impact on product prices but also improves the reputation of a firm, which contributes to higher prices in the future. Secondly, we analyse umbrella effects of reputation: investments in product quality of the ‚top product' are particularly profitable as they also generate a ‚reputational dividend' for other products with lower quality. And finally, we investigate selection effects in expert evaluations. Experts typically evaluate a selection of products only and we find endogenous selection effects to be important for analysing product quality empirically.

Keywords: experts evaluations, hedonic pricing, wine quality, endogenous reputation, sample selection, umbrella branding

JEL code: C33, L66, Q11

1 Introduction

Many markets for vertically differentiated products are characterized by information asymmetries, with suppliers usually being better informed about important product characteristics than consumers. The fact that consumers often are unable to ascertained quality before actual consumption (‚experience goods problem‘) can hinder the functioning of markets to the possible extreme of markets failing to operate despite obvious gains from trade. Economists have investigated different mechanisms for firms to overcome the experience good problem. Two mechanisms are particularly important: reputation and expert reviews.

While numerous empirical studies find a significant and positive impact of firms' reputation for high quality on performance, less research has focused on how firms may build a reputation (Basdeo et al., 2006). This lack of empirical work on the determinants of reputation contrasts sharply with the theoretical literature on this topic, which stresses the idea that a firm's reputation is a result of its market actions, i.e. the quality of its products delivered in the past. In Shapiro's (1983) classical paper, for example, individuals make consumption decisions on the basis of firms' reputation, which is again determined by product quality delivered in the previous period. Depending on the specific form in which reputation is accumulated from past actions (product quality delivered in previous periods), investments in product quality may not only have direct (contemporaneous) effects but may also trigger indirect (long-run) effects on firm performance and product prices. Dynamic effects of investment in quality (via endogenous reputation) are explicitly studied in Board and Meyer-ter Vehn (2013). The authors argue: ‚Once quality is established, it is persistent and generates a stream of reputational dividends until it becomes obsolete‘ (p. 2393). They also observe that most empirical studies are static, focus on quantifying the short-run value of reputation only and thus tend to ignore the long-run ‚stream of reputational dividends‘.

Our empirical analysis contributes to this literature by explicitly modelling reputation effects endogenously. In addition to short-run effects of quality on product prices, we also aim at quantifying the long-run ‚stream of reputational dividends‘. Empirical studies using cross-sectional data to investigate the contemporaneous relationship between reputation, product quality and market prices underestimate the returns to product quality since long-run (endogenous reputation) effects are ignored. Results from a structural econometric model estimated for a large number of individual wines produced between 2004 and 2007 by 488 wineries in Austria suggests that this bias is significant.

We further extend the existing literature by studying how reputation is established in a multi-product context. A growing literature in business and economics on brand extension and umbrella branding suggests that reputation can be leveraged across products. Rasmusen (2016), for example, argues that the quality of the ‚top product‘ of a multi-product firm (i.e. the product with the highest quality among all products of the firm) is particularly important for establishing a good firm reputation and also generates a ‚reputational dividend‘ for (other) products with lower quality (spill-over

or umbrella effect). Our empirical results for multi-product firms in the wine market support these arguments: investing in the quality of the top wine has a particularly strong effect on a winery's reputation and pays off in terms of higher prices for all wines of the winery.

Finally, our paper contributes to the growing literature in economics and marketing science on the effects of expert opinions and test reviews. A wide range of media provides information on the quality of products based on expert reviews. Empirical studies have been carried out for different markets such as visual and performing arts (Ginsburgh, 2003), movies (Reinstein and Snyder, 2005), books (Ginsburgh, 2003 and Sorensen, 2007), cars (Dewenter and Heimeshoff, 2014, 2015) and wine (Hilger et al., 2011 and Friberg and Grönqvist, 2012). The present paper addresses the issue of sample-selection for expert reviews. Product reviews often publish results (expert evaluations / grades) of a selection of products only. We argue that a non-random selection of products for grading can lead to biased parameter estimates in the quality-price relationship when the selection process is not explicitly accounted for – an issue largely ignored in the existing empirical literature (Caliendo et al., 2015). In our empirical analysis for the wine market, we find that selection effects in external certification (expert reviews) are indeed significant in the empirical analysis of product quality.

The wine market is particularly suited for this analysis for a number of reasons. First, vertical differentiation is key in this market[1] and experts' evaluations of product quality can be obtained easily.Secondly, a key component of vertical product differentiation in this market is product and/or firm reputation. Benfratello et al. (2009) suggest that ‚reputation acquired by wines and producers during the years is more important than taste in driving market prices' (p. 2206). The availability of a well established measure of winery reputation allows us to measure its impact on product prices as well as to identify factors determining winery reputation (such as product quality delivered in the past). Thirdly, ‚wine is an experience good' (Storchmann, 2012, p.7). Consumers face a huge variety of different products characterized by a large set of attributes and some of these attributes cannot be evaluated before consumption. As a result, consumers often rely heavily on ‚expert opinion' and the market for expert opinion on wine thus is large.[2] And finally, the impact of quality (as evaluated by experts) on wine prices has been the topic of a lively debate in economics, which started with first empirical studies in the 1990's (Nerlove (1995) and Combris et al. (1997), for example). Oczkowski and Doucouliagos (2015) survey more than 180 hedonic wine price models published over the last 20 years and covering many different countries. The present paper takes reference to these studies.

The remainder of the article is organized as follows. The next section 2 provides an overview of the theoretical and empirical literature. Section 3 describes the data and

---

[1]Fine wine can regularly fetch bottle prices that exceed several thousand dollars. In fact, the world's most expensive bottle, a 1787 Chateau Lafite, was auctioned off in 1985 for £105.000 (Storchmann, 2018).

[2]According to Storchmann (2012), the seven major U.S. wine magazines in 2010 had a combined subscriptionship of more than 500,000.

outlines the empirical estimation strategy. Section 4 presents and discusses estimation results and section 5 summarizes and concludes.

2 Related Literature

2.1 Reputation

The paper is most closely related to pricing models for product quality, for which theoretical foundations were developed in Rosen's (1974) hedonic model. Following first empirical applications (Golan and Shalit, 1993, Oczkowski, 1994, Nerlove, 1995, Combris et al., 1997), the last two decades have seen a large number of hedonic pricing models in the wine market. The existing empirical literature investigating the quality-price relationship in this market delivers conflicting results, however.[3]

One of the controversial issues in this discussion is related to the question of whether and how ‚true' quality can be observed and measured. Hedonic analyses assume that both demanders and suppliers have perfect knowledge about the quality attributes of a product, and that these characteristics are ‚objectively measureable' (Rosen, 1974): hedonic price functions may serve to reflect consumers' preferences for quality only to the extent that consumers actually have perfect knowledge about the relevant product characteristics. Early empirical studies (Combris et al., 1997, 2000, as well as Lecocq and Visser, 2006) include sensorial characteristics as explanatory variables in the price equation (in addition to objective attributes, such as vintage, denomination, grape variety, that usually appear on the label). Parameter estimates for some measures of a wine's aroma, body and finish would then indicate the consumers' willingness to pay for these quality attributes. In the case of Bordeaux wine, Combris et al. (1997) observe that many of these sensory characteristics of wines do not play a role in the determination of the market price. The authors interpret the statistical unimportance of individual sensory variables as an indication for the consumers' lack of information about wine quality attributes. Subsequent empirical studies found similar results and conclude that wine is an experience good: consumers are unable to value its current quality but make their consumption decisions on the basis of other factors and in particular on the basis of product and/or firm reputation.

A large number of empirical studies indeed observe a positive and significant impact of product and firm reputation on wine prices (Oczkowski and Doucouliagos, 2015). Landon and Smith (1997, 1998) estimate several cross-section models for French premium wines. They observe that reputation explains much more variation in the consumers' willingness to pay than does current product quality. Oczkowski (2001) also finds significant reputation effects for a cross-section of Australian premium wines. For the Italian premium wine market, Benfratello et al. (2009) suggest that ‚reputation ac-

---

[3]In the first issue of the *Journal of Wine Economics*, Lecocq and Visser (2006) summarize this literature and conclude that ‚it has always been animated and controversial' (p. 43). More recent surveys are provided in Estrella Orego et al. (2012) and Oczkowski and Doucouliagos (2015).

quired by wines and producers during the years is more important than taste in driving market prices' (p. 2206). They conclude that ,producers should aim at building a well established reputation – both at wine and at firm level – by promotional activities (e.g. participation to wine exhibitions) which facilitate citations in well-known guides' (p. 2207).

Empirical studies comparing the effects of reputation and measures of product quality by including both variables in hedonic cross-section price models can be misleading, however, since reputation and product quality are not independent. Good reputation comes from good quality delivered in the past. This interrelationship between reputation and product quality is a key element in theoretical reputation models and is the main reason why reputation-effects can eliminate the inefficiencies associated with asymmetric information (Klein and Leffler, 1981, Kreps et al., 1982, Shapiro, 1983, Allen, 1984, Hörner, 2002, Rob and Fishman, 2005, Board and Meyer-ter Vehn, 2013). In the classical paper of Shapiro (1983) entitled ,Premiums for High Quality Products as Returns to Reputation', for example, consumers' decisions are based on the reputation of a firm, which again is determined by product quality delivered in the past. Here, product quality does not pay off immediately and directly but only indirectly (in the next period) as higher quality feeds into the firm's reputation. In a related study by Kreps et al. (1982), reputation is modeled as a Bayesian updating process: based on the observation of past transactions, sellers form a belief about the type of seller they interact with.

The interrelated (and reinforcing) processes of investment in quality, reputation formation and firm performance are further investigated in Rob and Fishman (2005) who stress differences between large and small firms. According to the authors, reputation is more valuable to larger firms. As larger firms invest more in quality, their reputation further improves which contributes to a better firm performance and faster firm growth. As a result, quality and reputation are persistent over time.

The dynamics of quality and reputation are explicitly modelled in Board and Meyer-ter Vehn (2013). The authors adopt a capital-theoretic approach in which firms can invest in product quality. A firm's reputation is defined as the market's belief about this quality. They show that investment in quality pays off with a delay. In addition to any immediate (or short-run) impact of product quality on prices, an improvement in product quality will also have a positive (long-run) effect on reputation, which again commands a price premium. Ignoring this endogenous reputation effect implies underestimating the effects of quality on market prices. An empirical analysis of the price-quality relationship should thus take into account both, short-run and long-run effects of quality. This requires explicitly investigating the effects of current quality on firm reputation in a structural econometric model.[4]

An important extension of the literature on reputation investigates firms' incentives to deliver high quality in a multi-product context. Economies of scope in delivering high

---

[4]As an alternative, some authors have adopted a ,reduced form approach' by including lagged product quality in the price equation (Landon and Smith, 1997, 1998, and Ali and Nauges, 2007). The pros and cons of structural models are discussed in more detail in Reiss and Wolak (2007).

4

quality are already mentioned in the classical paper by Klein and Leffler (1981): ‚As long as consumers react to receiving unexpectedly low quality from a diversified firm by reducing purchases of the firm's entire product line, all the firm's nonsalvageable capital serves to assure the quality of each product it produces' (p. 636). The Klein-Leffler model has been extended to multiple products by Andersson (2002) and Cabral (2009). Andersson (2002) and Cabral (2009) consider a repeated game with moral hazard in which a multi-product firm must repeatedly invest in the quality of each of its products. They show that brand extension (‚umbrella branding') can support high quality equilibria, which are not feasible if products are sold as separate brands. One of the main advantages is that umbrella branding in markets for experience goods may provide for a harsher punishment on cheating sellers.[5] Hakenes and Peitz (2009) extend this literature by simultaneously investigating two signalling strategies: umbrella branding and third-party (expert) certification. They show that umbrella branding can be a cost-efficient way to signal product quality and identified parameter constellations under which umbrella branding can fully (or partially) substitute external certification. This paper highlights two issues that are important for the present empirical analysis: (a) the interrelationship between reputation in a multi-product setting (umbrella branding) and external quality evaluation (experts) as well as (b) the fact that not all products are externally evaluated and the decision to externally evaluate the quality of a firms' product is endogenous. The second issue will be discussed in more detail in section 2.2.

Meanwhile, there is a fairly sizeable theoretical literature on brand extension and umbrella branding (for recent studies see Rasmusen, 2016, and the literature mentioned there), suggesting that reputation can be leveraged across products. Investments in high quality for one particular product, which contributes to a good reputation of the firm, might generate a ‚stream of reputational dividends' not only for this, but also for the other products sold under the same brand name.[6]

---

[5]This is nicely illustrated in Cabral (2009), for example: ‚suppose that a low-quality product fails with probability $1 - \beta$. If a firm sells two products under different names and consumers punish each product failure separately, then the probability that cheating is detected in a given product is given by $1 - \beta$. If however the firm sells both products under the same name and the firm is punished in both products following any product failure, then the probability of punishment given that the firm shirks in both products is given by $1 - \beta^2$, which is greater than $1 - \beta$' (p. 206).

[6]Note that our analysis is also related to a substantial literature on collective reputation, in which otherwise autonomous firms share a common brand name. Example include regional agricultural products wherever country of origin labeling is salient and where brand names are protected by designation of origin and geographical indication status. Although the two concepts are interrelated (in Tirole's 1996 reputation theory, for example, collective reputation emerges as an aggregate of individual reputations), there are still important differences: regarding individual firms' incentives to invest in establishing a reputation. The main difference is that in an umbrella brand a central authority (the winery) makes investment decisions for each of the brand's products and internalizes the effect of each individual product's quality on the reputation of the entire brand. By contrast, in a collective brand, individual members are concerned only with the effect of their investment decisions on the value of their own product. Collective reputation and free-riding problems are studied in Fishman et al. (2018) and investigated empirically empirically in Castriota and Delmastro (2015) as well as Gergaud et al.

2.2 Experts

The second important mechanism for transmitting information about product quality to consumers are expert reviews. Experts typically are better at evaluating the quality attributes of products than novice consumers, and in some markets their judgment can have a vital impact on financial performance. For example, in urban centers, ‚theater and dance critics wield nearly live-or-death power over ticket demand' (Caves, 2000, p. 189). In the wine market, a number of empirical studies rely on wine ratings from experts and publicly accessible wine guides,[7] such as scores given by the Australian wine critic James Halliday (Schamel and Anderson, 2003) or the wine magazine *Winestate* (Schamel and Anderson, 2003), Robert Parker's *Wine Advocate* (Ali and Nauges, 2007, Dubois and Nauges, 2010), the *Shield and Meyer Guide* (Oczkowski, 1994), the *Wine Spectator* (Landon and Smith, 1997, 1998, and Benfratello et al., 2009), the Italian *Duemila Vini* guide (Benfratello et al., 2009), or from six major Swedish print media (Friberg and Grönqvist, 2012).

These studies find that expert grades exert a statistically significant and positive impact on prices, but the magnitude of these effects typically is rather small and short-lived. Ali and Nauges (2007), for example, conclude that ‚a one-point increase in [Parker's] grade has almost no effect on the price set by producers' (p. 96f). Examining weekly data on the demand for wine in Swedish state liquor stores over a time period of five years, Friberg and Grönqvist (2012) observe that the effects of favourable reviews generates a transitory, but quantitatively important increase in consumer demand. The effect of a positive review peaks in the week after publication and remains significant up to 22 weeks after the review. No significant effects of positive reviews can be found after half a year. The effect of negative reviews always is approximately zero.

As mentioned in section 2.1, one reason for the weak and short-lived effect of expert ratings is related to the fact that long-run (indirect) effects via endogenous reputation building are ignored. Selection effects could be another reason. When using expert evaluations published in wine guides in hedonic pricing models, the assumption is implicitly made that the selection of wines into these wine guides is random and in particular uncorrelated with the quality of wines.[8] If, however, the selection of wines is correlated

---

(2017).

[7]Exceptions are the studies carried out on the basis of data sets obtained from a random selections of wines in France (Combris et al., 1997, 2000, and Lecocq and Visser, 2006)

[8]Combris et al. (1997) identify five conditions that are required for using expert reviews in econometric quality-price models and conclude that the widely published and easily accessible wine guides do not in general verify these conditions. Their first condition is that ‚all the wines that are tasted should be included in the sample, regardless of whether the wine is considered good or bad. In wine guides the wines of inferior quality are often deliberately under-represented for commercial reasons' (p. 392). Note that sample selection effects are also mentioned in one of the first hedonic studies for wine (Nerlove, 1995). In contrast to the usual hedonic regression, in which unit variety prices are regressed on a vector of quality attributes, Nerlove (1995) estimates a regression on quantity sold on price and quality attributes. This is justified by the assumption that prices and attributes can be taken as exogenous to the Swedish consumer. The author argues: ‚Of course, to the extent that wines imported into Sweden are not a random sample of wines produced world-wide, there is a selectivity

with wine quality, parameter estimates on the quality-price relationship may be biased. In the present empirical application, each wine grower decides on the basis of his/her subjective quality evaluation whether to submit a particular wine for grading, and the resulting bias in the estimated price-quality equation will be discussed in more detail in Section 3.

Selection effects have also been studied in other markets, where critics are powerful market actors and their decision on which product to select for reviews might not be random but instead may be affected by selection effects related to different sources. In the market for movies, Reinstein and Snyder (2005) control for selection effects due to product quality in a difference-in-difference model. For the case of two popular movie critics, they show that selection effects related to inherent product quality lead to overestimation of the critics' effect. Investigating the impact of seller reputation on eBay auctions, Cabral and Hortacsu (2010) carefully investigate the possibility of sample section biases. Caliendo et al. (2015) propose a combined approach of propensity score matching with difference-in-differences methods for the German book market and identify substantial selection effects of individual critics on book sales.[9]

In the following section, we explicitly investigate two mechanisms for transmitting information about product quality to consumers: reputation and expert's product reviews. Specific emphasis will be given to endogenous reputation in a multi-product context (umbrella branding) as well as sample selection issues in expert reviews.

3 Data and Estimation Strategy

3.1 Data

In contrast to empirical studies investigating auction prices in the niche market of very exquisite wines, the present analysis focuses on price setting and product quality in the ‚regular wine market' i.e. wines sold directly from wineries to consumers. The sample used in the present paper consists of 24,547 wines which are produced between 2004 and 2007 by 488 wineries in Austria. These wineries cover about 35% of the annual production of quality wines. For each wine, data on numerous characteristics are available, such as the price per standard bottle, type (white, red, rosé and four types of sweet wine) and variety of grape, the year of harvest, age (the time span between harvest and supply to the market) as well as the size of the winery and the diversity of its assortment. Summary statistics can be found in Table A.1 in Appendix A.

This information is matched with data on the experts' evaluation of the specific

---

bias present even in the estimates of such confluent relationships' (p. 1707). Other potential problems associated with expert reviews are discussed in Ginsburgh (2003), Storchmann (2012) and Oczkowski and Doucouliagos (2015).

[9]Note that selection issues have also been studied in markets, where consumers (instead of experts) review products. As pointed out by Luca and Zervas (2016) ‚people who purchase a product on average like the product better than those who choose not to purchase it. Second, people who choose to leave a review may have different preferences than those who choose not to' (p. 3414).

bottle of wine, which is obtained from consulting the most influential 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 flavour and finish. This expert evaluation of wine quality will be denoted q^E in the following. For the time period 2004 – 2007, the average grade in our sample is 89.05. It is important to note that wineries usually publish price lists and start selling wines of the current vintage between February and June, while the wine guide (including experts' grades) is not published until July of the same year. Therefore, wieneries' pricing decisions are determined by their own evaluation of product quality (q^W), which is unobservable and not necessarily corresponds to the experts' evaluation in the wine guide (q^E).

Data from the Falstaff-Wine-Guide also include information on the reputation of a winery. This information is available for a longer period 1999 – 2007. Wineries are classified on a scale from 0 to 3 between 1999 and 2006 and from 0 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).[10]

Note that the Fallstaff-wine guide does not grade all wines produced by the 488 wineries in our sample. Expert grades are available for a subset of 9,914 bottles of wine only (about 40% of all wines in our data set). The number of wines graded per year also differs across wineries: for most wineries, six wines per vintage are graded (see Figure 1). While five (four) wines are graded for 20% (12%) of all wineries, only 14% of all wineries have more than six wines graded in the wine guide. Typically, these are wineries with a high reputation: reputation averages 0.68 for this group of wineries, whereas the average reputation is 0.20 for wineries where six or less wines are graded.

3.2 Estimation Strategy

Following the existing empirical literature, we aim at measuring the impact of expert evaluations of product quality q^E and firm reputation r on product prices p in a hedonic model: $p = f(q^E, r, \boldsymbol{X})$, where $\boldsymbol{X}$ denotes a vector of exogenous control variables. We extend this literature by considering (a) endogenous reputation and ‚umbrella effects' and (b) non-random selection of wines for grading.

As suggested by the theoretical reputation models mentioned in Section 2, reputation of a particular winery is determined by its actions in the past, i.e. the quality of its wines delivered in previous years. The empirical reputation model thus uses observations on the quality ratings of wines produced at time $t - k$ (with $k = 1, \ldots, 4$) as explanatory variables determining the winery's reputation at time t. We contribute to the small empirical literature on the determinants of reputation by testing different specifications of our reputation model. In addition to the average quality ($\bar{q}^E_{t-k}$) of all wines from a particular winery produced at $t - k$, we also investigate the impact of the wine with the highest quality ranking (i.e. a winery's top wine). We expect to find that

---

[10]For a detailed discussion on different concepts and measures of organizational reputation see Lange et al. (2011).

Figure 1: Number of Graded Wines per Winery (in percent)
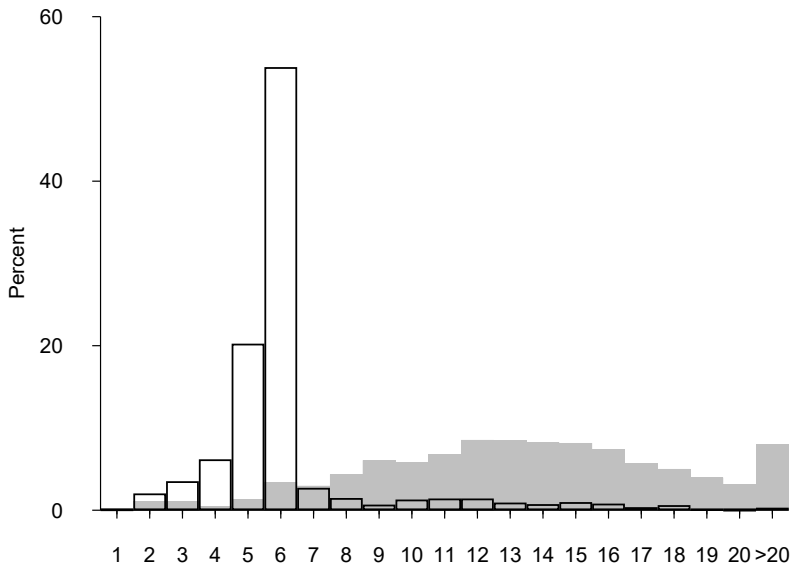
Notes: Histograms indicate the share of wineries with a particular number of graded wines (transparent bars) and a particular number of produced wines (gray bars) in one year. The figure is based on 1,615 observations.

(a) the average quality of wines in previous time periods has a positive impact on the reputation of the winery, (b) the impact of past wine quality on reputation becomes weaker as k increases, and (c) that the top wine of a winery has the strongest impact on reputation.

Secondly, the estimation strategy has to take into account the fact that our explanatory variable, expert evaluations q^E, can be observed only for wines that have been submitted for grading ($S = 1$). This, however, is not a random sample. As pointed out by Combris et al. (1997), ‚the wines of inferior quality are often deliberately under-represented for commercial reasons‘ in the wine guides (p. 392).

Assume that the wine producers' decision to submit wines for grading depends on his/her quality evaluation of the particular wine (q^W), which is unobservable to the econometrician, as well as on a number of other exogenous factors ($\boldsymbol{Z}$). While the estimation model for the full population of all wines can be written as $p = f\left(q^E, r, \boldsymbol{X}, S\left(q^W, \boldsymbol{Z}\right)\right)$, the existing literature typically estimates $p = f(q^E, r, \boldsymbol{X})$ for wines where data on expert evaluations are available, i.e. the sample of wines selected for grading ($S = 1$). The estimated relationship between prices p and experts' evaluations q^E in the selected sample will provide unbiased parameter estimates only if (a) $cov\left(p|\left(q^E, r, \boldsymbol{X}\right), S\left(q^W, \boldsymbol{Z}\right)\right) = 0$ or (b) $cov\left(q^E, S\left(q^W, \boldsymbol{Z}\right)\right) = 0$. The first condition (a) will be violated since the wine producer decides about both, wine prices (p) and submission for grading (S) on the basis of its own quality evaluation (q^W). If the wine producer is very optimistic about the quality of a particular wine, he/she will be more inclined to submit the wine for grading

9

and, at the same time, charge a higher price for this wine, ceteris paribus. Similarly, it is plausible to expect the second condition (according to which the evaluation of experts q^E and wine producers q^W is uncorrelated) to be violated in the wine market. While the subjective component in the judgement of a particular wine certainly plays an important role, systematic empirical evidence suggests that some objective components are shared by all experts which suggests a positive correlation between quality ratings among independent judges (between q^E and q^W and thus between q^E and $S\left(q^W, \boldsymbol{Z}\right)$).[11]

This non-random selection implies that (a) our measure of product quality (q^E) is endogenous and (b) the relationship between experts' grading (q^E) and wine prices (p) can be biased when sample selection is not taken into account. The appropriate estimation strategy is discussed in Wooldrigde (2002, Chapter 17.4.2) and can be summarized by the following structural model.

$$Prob\left(S_t = 1 | \boldsymbol{Z}\right) = \Phi\left(\boldsymbol{Z\gamma}\right) \tag{1}$$

$$E\left[q_t^E | \left(\bar{\boldsymbol{q}}_{\boldsymbol{t-k}}^{\boldsymbol{E}}, \boldsymbol{Y}, \hat{\lambda}\right), S_t = 1\right] = \sum\nolimits_{k=1}^{K} \mu_k \bar{q}_{t-k}^E + \boldsymbol{Y\mu} + \rho_2 \lambda\left(\boldsymbol{Z\hat{\gamma}}\right) \equiv \hat{q}_t^E \tag{2}$$

$$E\left[p_t | \left(\hat{q}_t^E, \hat{r}_t, \boldsymbol{X}, \hat{\lambda}\right), S_t = 1\right] = \alpha_1 \hat{q}_t^E + \alpha_2 \hat{r}_t + \boldsymbol{X\alpha} + \rho_4 \lambda\left(\boldsymbol{Z\hat{\gamma}}\right) \tag{3}$$

$$E\left[r_t | \left(\bar{\boldsymbol{q}}_{\boldsymbol{t-k}}^{\boldsymbol{E}}, \boldsymbol{Y}, \hat{\lambda}\right), S_t = 1\right] = \sum\nolimits_{k=1}^{K} \beta_k \bar{q}_{t-k}^E + \boldsymbol{Y\beta} + \rho_3 \lambda\left(\boldsymbol{Z\hat{\gamma}}\right) \equiv \hat{r}_t \tag{4}$$

$\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are vectors of explanatory variables, $\boldsymbol{\gamma}$, μ_k, β_k, ρ_2, ρ_3, ρ_4, α_1, and α_2 are parameters to be estimated, and Φ is the cumulative distribution function of the standard normal distribution. From the estimation of a first-stage probit-model on the selection of wines for grading (equation (1)), we compute the inverse Mills ratio $IMR \equiv \hat{\lambda} = \lambda\left(\boldsymbol{Z\hat{\gamma}}\right)$, which is then used as an additional explanatory variable in equations (2), (3) and (4).

In a second stage, the system of equations (2) - (4) is estimated by two-stage-least-squares (2SLS). Details on the identification (exclusion restrictions in the system of equations) are discussed in Appendix C (online). In estimating equations (1), (2), (3) and (4) we include winery fixed effects (η_w) to control for unobserved and time-invariant heterogeneity between wineries like differences in costs, soil conditions, consumer preferences or managerial skills. We further include time fixed-effects (a dummy variable ω_t for each year) to control for unobserved differences in the quality of wine between different years, as well as various dummy variables on the variety of the grape, the type and the age of the wine, to account for product-specific heterogeneity.

---

[11]Empirical evidence reveals a ‚moderately high level of consensus' (Stuen et al., 2015, p. 47) between the four most prominent wine quality ratings in the U.S. wine market, for example. ‚As measured by the correlation coefficient of wine quality ratings, the degree of consensus between pairs of publications ranges from 0.28 to 0.63' (p. 59). More details on selection effects are provided in Appendix B (online).

4 Results

4.1 Selection of Wines for Grading and Wine Quality

Table 1 reports parameter estimates of a probit model estimated to investigate the variation in the binary variable S in model [1] as well as estimates obtained from a two-stage least squares (2SLS) model estimated on the quality of wines (expert grades) in model [2].

The probit model [1] correctly classifies 74% of all observations, and 66% of wines that are actually submitted for grading are predicted correctly. Selection for grading is non-random but significantly related to observable characteristics of the winery. The number of times the particular wine has been graded between 1999 and 2002 has a significant and positive impact on the probability of the wine to be selected for grading in a particular year in the period between 2004 and 2007 (model [1]). Similarly, wines that have been graded more frequently in the past receive significantly higher grades (model [2]). The larger the number of different wines produced in the winery, the smaller is the probability of a specific wine to be considered for evaluation (model [1]). This variable does not contribute significantly to the explanatory power of model [2] (experts' evaluation of the quality of wines). A longer time span between harvest and selling (age of wine) is associated with a higher probability of a wine to be selected for evaluation (model [1]). This might be explained by the fact that these wines can be expected to receive higher grades, ceteris paribus. Results from the second econometric model reported in [2] indeed suggest that the age of the wine has a significant and positive effect on experts' evaluation of wine quality.

Table 1 further provides evidence on the impact of experts' quality evaluations in the past as well as winery reputation. The results reported in model [1] suggest that the average quality of all wines in the previous year $(t-1)$ has a positive impact on the likelihood of submission and grading at time t. This effect is significantly different from zero at the 10 %-level. The positive effect of expert evaluations in $t-1$ on the probability of submission is particularly pronounced for top wines. If the particular wine received the highest expert scores among all wines evaluated in the winery in $t-1$ (i.e., if the dummy variable ‚top wine‘ is set equal to 1), this wine has a significantly higher probability of being submitted for external evaluation again (at time t).

Model [2] of Table 1 suggests a regression towards the mean in experts' quality evaluations: if the average quality of wines produced in a winery was high in previous years (at time $t-1$ to $t-3$), quality evaluations tend to be significantly lower in period t, ceteris paribus. Further note that a winery's reputation in $t-1$ does not exert a significant impact on product quality (as evaluated by experts). These results do not support Rob and Fishman's (2005) conjecture according to which quality is persistent and firms with a good reputation invest more in product quality. Note, however, that results are different again for top wines. The wine with the highest expert scores within a particular winery in $t-1$ again receives significantly higher scores in the following year. The parameter estimate of the dummy variable ‚top wine‘ in model [2] suggests

Table 1: Regression Results Explaining the Selection and the Quality of Wines

Method	Probit			2SLS		
Model	[1]			[2]		
Dependent variable	Select (S)			Quality (q^E)		
Variables	Coeff.	Std. Dev.	Sign.	Coeff.	Std. Dev.	Sign.
# of grades between 1999 and 2002	0.3372	(0.0131)	***	0.3330	(0.0753)	***
# of different wines (in period t)	-0.0384	(0.0105)	***	0.0191	(0.0195)	
Wine age: 2 years	0.4020	(0.0682)	***	1.5721	(0.1055)	***
Wine age: 3 years	0.7020	(0.1628)	***	2.8944	(0.2564)	***
Average quality (at $t-1$)	0.0395	(0.0229)	*	-0.1961	(0.0406)	***
Average quality (at $t-2$)	0.0086	(0.0197)		-0.2361	(0.0345)	***
Average quality (at $t-3$)	-0.0253	(0.0170)		-0.1241	(0.0326)	***
Average quality (at $t-4$)	0.0057	(0.0179)		-0.0202	(0.0299)	
Reputation (at $t-1$)	0.1654	(0.1178)		-0.2833	(0.2579)	
Top Wine (at $t-1$)	0.5425	(0.0485)	***	1.4878	(0.1099)	***
IMR ($\hat{\lambda}$)				1.2003	(0.3537)	***
Constant	-2.8932	(4.6788)		134.7191	(8.6146)	***
Winery effects	Yes (345)			Yes (345)		
Vintage effects	Yes (3)			Yes (3)		
Variety of the grape	Yes (31)			Yes (31)		
Type of wine	Yes (6)			Yes (6)		
Crop year	Yes (4)			Yes (4)		
Number of observations	16,749			7,355		
Log-likelihood	-8,629					
Pseudo-R^2	0.249					
R^2				0.573		
Test on instruments	χ^2-test: 734.946			F-test: 18.531		
	$df = 11$			$df = (11, 345)$		
	$p = 0.0000$			$p = 0.0000$		

Notes: Standard errors are reported in parentheses and are clustered at the winery level. *** significant at 1 %, ** significant at 5 %, * significant at 10 % level. 7,656 observations are excluded because quality information of the respective wineries are not reported for all of the four previous years or information on the reputation of the winery of the previous year is not observed. 142 additional observations are excluded because winery fixed effects predict the selection perfectly. The χ^2-test and the F-test on instruments tests the hypothesis that the coefficients on the instruments (excluded from the regression on prices, reported in column [1] of Table 2) in the selection equation and in the regression on quality, respectively, are jointly zero.

a substantial effect of about 1.5 score points in experts' quality evaluations suggesting that quality is persistent for the top wine only.

A particularly interesting result from the quality model [2] in Table 1 is the parameter estimate for the inverse Mills ratio (calculated from the results of the selection equation). The parameter estimate of $IMR(\hat{\lambda})$ is positive and statistically significant from zero at the 1%-level. This suggests that the error terms in the selection equation [1] and the quality equation [2] are positively correlated. Unobserved factors that make the selection of wines for grading more likely (such as a positive subjective quality judgment of the wine producer, for example) tend to be associated with higher grades received from wine experts. This provide empirical support for our claim of a non-random sample of wines for which expert grades are available. Whether non-random selection process also introduces a bias in the price-quality relationship will be evaluated in the next step.

4.2   Wine Prices and Winery Reputation

The results of two-stage least squares (2SLS) estimation models on prices and reputation are summarized in Table 2.[12] To investigate the importance of umbrella effects of reputation, the results of two different specifications of the reputation model (specifications [a] and [b]) are reported.[13] All regressions explicitly consider sample selection effects by including the inverse Mills ratio $(IMR(\hat{\lambda}))$ computed from the selection equation. We find that the parameter estimates for IMR are not significantly different from zero at the 10%-level in the price and reputation models in Table 2; explicitly controlling for selection effects does not contribute significantly to the explanatory power of these models.[14]

The parameter estimates for wine quality are positive and significantly different from zero at the 1%-level in the price equation. An increase in quality of a particular

---

[12]The endogenous variables in the price equation (models [3a] and [3b]) are the quality of the wine and the reputation of the winery. Statistical tests reject the assumption that these variables are exogenous in the price equation. A Hansen-J-test of overidentifying restrictions is not rejected at the 10%-significance level, indicating that the explanatory variables for wine quality and winery reputation are valid (i.e. uncorrelated with the error term).

[13]The selection and the quality equation reported in Table 1 correspond to models [3a] and [4a] of Table 2. The regression results on the selection and the quality equation, which correspond to columns [3b] and [4b] of Table 2, are very similar and are thus not reported but available from the authors upon request. Parameter estimates on all variables also included in the previous model specification are hardly affected by this modification (both in size and statistical significance), and statistical tests again suggest that the variables excluded from the price equation are strong and valid instruments.

[14]Since 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 D.1 in Appendix D (online). The results support Wooldridge's (2002) observation that ,[i]n many cases, the adjustments do not lead to important differences' (p. 562).

wine by one index-point increases its by 6.9% (model [3a]). This quantitatively important positive effect is in contrast to some earlier empirical studies reporting very small effects of experts' grades ((Ali and Nauges, 2007)). Despite the difficulties in objectively and consistently assessing wine quality using expert ratings, these ratings seem to provide valuable information for consumers in the present context. Table Table 2 further suggests, that the price-quality relationship is not identical for all wines of the winery. Even after controlling for wine quality, wineries charge an additional price premium for their top wines (i.e. the wine with the highest quality ranking among all wines within the winery in a particular year). Columns [3a] and [3b] suggest that this ,top wine premium' is substantial and amounts to about 13%.

Consistent with existing empirical studies, winery reputation also has a significant and positive effect on product prices. An increase in reputation from 0 to 1 (i.e. from the lowest to the highest level of reputation) commands a substantial price premium; according to model [3a] of Table 2, prices increase by 39.9%. This price premium of reputation is even somewhat larger (41.2%) in model [3b] and significantly different from zero at the 1%-level.

If the time span between the harvest of the grapes and selling to consumers (age of wine) is two years instead of one year (the reference category) prices increase by 41%, and a time span of three or more years is associated with a price increase of 81%.

Endogenous repuation formation is investigated in models [4a] and [4b]. As suggested by Rob and Fishman (2005), we observe some degree of path dependence in reputation. A winery's reputation in the previous period $(t-1)$ is positively related to the contemporaneous reputation; the impact of this variable is significantly different from zero at the 5%-level in models [4a] and [4b]. Note, however, that the estimated persistence parameter of around 0.08 is only moderate in size.

Even after controlling for the persistence of reputation, both specifications (models [4a] and [4b]) suggest that the quality of wines delivered in the past contributes to the contemporaneous reputation of a winery. In specification [4a], we observe that the average quality of wines produced by the winery in the previous period $(t-1)$ exerts a positive impact on the winery's reputation, which is significantly different from zero at the 1%-level. As expected, the effect of average quality on reputation tends to decrease over time: the parameter estimates for the average quality produced in $(t-2)$ and $(t-3)$ are smaller compared to the average quality delivered in $(t-1)$, but still significantly different from zero (at the 10% level) and positive. The estimated effect of the average quality in $(t-4)$ is not significantly different from zero.

While model [4a] analyses the impact of the average quality of wines produced by a winery, model [4b] investigates umbrella effects by distinguishing between the quality of the top wine and the median quality of all graded wines. The results in column [4b] show that the quality of a top wine in $(t-1)$ has a strong influence on the contemporaneous reputation of a winery. The parameter estimate of this variable is positive and significantly different from zero at the 1% level. Again, the reputation-effect of the quality of a top wine declines over time, but is still positive and significantly different from zero at the 10%-level for two out of three time periods between $(t-2)$

and $(t-4)$. On the other hand, parameter estimates for the median quality of all graded wines produced in the previous years are not significantly different from zero. This suggests that a firm's reputation for high quality is primarily determined by the quality of its' top product. While consumer recognition for an increase in quality seems to be particularly strong for this product, a quality improvement for the other products does not generate the same positive reputation effect.

The number of times a wine was evaluated in the past (between 1999 and 2002) as well as the number of different wines produced by a winery does not contribute significantly to the explanatory power of the reputation model (columns [4a] and [4b]).

Our analysis provides empirical evidence for (dynamic) spill-over or umbrella effects of reputation: Consistent with theoretical models of Hakenes and Peitz (2009), Cabral (2009), and Rasmusen (2016), we find that reputation can be leveraged across products. A positive expert evaluation for one product generates a ‚stream of reputational dividends' (Board and Meyer-ter Vehn, 2013) not only for this product, but also for all other products of the winery. These (dynamic) umbrella effects will be illustrated in more detail in the following counterfactual analysis.

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

Method	2SLS						2SLS					
Dependent variable / Model	Price ($ln(p)$) / [3a]			Reputation (r) / [4a]			Price ($ln(p)$) / [3b]			Reputation (r) / [4b]		
Variables	Coeff.	Std. Dev.	Sign.	Coeff.	Std. Dev.	Sign.	Coeff.	Std. Dev.	Sign.	Coeff.	Std. Dev.	Sign.
Quality (in period t)	0.0697	(0.0142)	***				0.0680	(0.0144)	***			
Top wine (in period $t-1$)	0.1319	(0.0181)	***	-0.0008	(0.0032)		0.1334	(0.0184)	***	0.0039	(0.0034)	
Reputation (in period t)	0.3998	(0.1557)	**				0.4122	(0.1410)	***			
Reputation (in period $t-1$)				0.0789	(0.0393)	**				0.0890	(0.0385)	**
Wine age: 2 years	0.4114	(0.0249)	***	-0.0029	(0.0062)		0.4133	(0.0252)	***	-0.0027	(0.0061)	
Wine age: 3 years	0.8151	(0.0544)	***	-0.0090	(0.0112)		0.8188	(0.0542)	***	-0.0100	(0.0112)	
Average quality ($t-1$)				0.0301	(0.0070)	***						
Average quality ($t-2$)				0.0099	(0.0056)	*						
Average quality ($t-3$)				0.0194	(0.0052)	***						
Average quality ($t-4$)				0.0053	(0.0042)							
Quality of top wine ($t-1$)										0.0182	(0.0050)	***
Quality of top wine ($t-2$)										0.0063	(0.0047)	
Quality of top wine ($t-3$)										0.0130	(0.0043)	***
Quality of top wine ($t-4$)										0.0064	(0.0034)	*
Median quality ($t-1$)										0.0038	(0.0073)	
Median quality ($t-2$)										0.0022	(0.0066)	
Median quality ($t-3$)										0.0001	(0.0070)	
Median quality ($t-4$)										-0.0023	(0.0050)	
# of grades between 1999 and 2002				-0.0001	(0.0020)					0.0006	(0.0020)	
# of different wines (in period t)				-0.0011	(0.0032)					-0.0005	(0.0033)	
IMR ($\hat{\lambda}$)	-0.0153	(0.0184)		-0.0009	(0.0107)		-0.0174	(0.0187)		0.0038	(0.0109)	
Constant	-4.3021	(1.2787)	***	-5.3970	(1.3006)	***	-4.1595	(1.2913)	***	-3.9773	(1.2598)	***
Winery effects	Yes (345)			Yes (345)			Yes (345)			Yes (345)		
Vintage effects	Yes (3)			Yes (3)			Yes (3)			Yes (3)		
Variety of the grape	Yes (31)			Yes (31)			Yes (31)			Yes (31)		
Type of wine	Yes (6)			Yes (6)			Yes (6)			Yes (6)		
Crop year	No			Yes (4)			No			Yes (4)		
Instrumented variables	Reputation, Quality						Reputation, Quality					
Number of observations (R^2)	7,355 (0.829)			7,355 (0.940)			7,355 (0.828)			7,355 (0.940)		
Hansen-J χ^2-test	12.071	$df = (9)$	$p = 0.2093$				15.204	$df = (13)$	$p = 0.2948$			
F-test on exogeneity	25.035	$df = (2, 345)$	$p = 0.0000$				23.196	$df = (2, 345)$	$p = 0.0000$			
F-test on instruments	4.325	$df = (11, 345)$	$p = 0.0000$				3.703	$df = (15, 345)$	$p = 0.0000$			

Standard errors are in parentheses and are clustered at the winery level. *** significant at 1 %, ** significant at 5 %, * significant at 10 % level. Fixed effects for the winery, the vintage and the variety-of-the-grape level have been partialled out to calculate the Hansen-J-test statistic. The F-tests on endogeneity of instrumented variables (H0: instrumented variables are exogenous) are based on a robust regression tests developed by Wooldridge (1995). The F-test on instruments tests the hypothesis that the coefficients on the instruments in the regression on reputation are jointly zero.

4.3   Counterfactual Analysis: Does It Pay to Produce High Quality?

Summarizing the estimation results, we conclude that an increase in wine quality has (a) an immediate and positive effect on wine prices as well as (b) a significant and positive (long-run) effect on the reputation of the winery, which allows wineries to raise prices even further. In addition, we find (c) that the effect of quality on reputation is particularly strong for the top wine of a winery, which implies a positive spill-over effect of a quality increase of this wine on prices of other wines. The following counterfactual analysis aims at illustrating these results by simulating the price effects of an exogenous increase in the quality of specific wines. 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 regressions reported in Table 2.

The log-run ‚stream of reputational dividends' are shown in Figure 2a and 2b. Figure 2a illustrates price effects in % from a one-time increase in the experts' evaluation of all wines of a winery at time t. Based upon the results of models [3a] and [4a], an increase in experts' evaluation by one index-point raises wine prices in year t by 6.9% on average. In the year after the (one-time) quality improvement, wineries are still expected to charge 1.1% higher prices for the new wines. This additional (lagged) effect of quality is due to the fact that better expert evaluations improve the reputation of the winery, which again has a positive effect on wine prices (‚reputational dividend'). In the following years ($t + 2$ and $t + 3$), the price effect is still positive, but below 1%. Figure 2b reports effects of a permanent increase in the experts' evaluation of all wines of a winery by one index-point. The endogenous reputation effect of quality implies that the long run price effects (+9.8%) are substantially (42%) larger than the contemporaneous price increase (+6.9%). Our results thus put into perspective the findings from studies focusing on the quantity sold in the wine market suggesting that the effects of experts are short-lived (Friberg and Grönqvist (2012).

Figure 3 illustrates the importance of umbrella effect. Results from the bootstrap simulation are based on specifications [3b] and [4b] reported in Table 2. In contrast to Figure 2, price effects of an exogenous quality improvement of one specific wine only are simulated. Note that a quality improvement of a specific wine not only has a positive price effect for this wine but will also prompt wineries to raise prices of all wines (umbrella effect). The solid line represents the median price effects (over all wines in the winery) in % associated with an increase in the quality of the top wine by one index-point in year t. Figures 3a and 3b suggest that a top wine generates strong umbrella effects, i.e. a substantial ‚reputational dividend' for all other wines. In the long-run, the median price increase of other wines (i.e. wines with no quality improvement) is nearly two per cent.

In contrast, umbrella effects are negligible for wines of lower quality. The dash-dotted line shows that the umbrella effect is close to zero in the case of a quality improvement for a median quality wine (i.e. the wine with the median quality score

17

of all graded wines in the winery). These results strongly support the notion that the largest ‚reputational dividend' can be earned from investing in quality improvements in the top quality segment of a winery. Wine quality matters, but the quality of the top wine has by far the strongest long-term effect on wine prices.

5   Summary and Conclusions

The present paper investigates the relationship between experts' evaluation of product quality and wine prices. We extend the existing literature on this topic in three dimensions. First, in addition to the direct effect of quality on price, we also explicitly model indirect reputation effects. Secondly, we empirically investigate umbrella effects by modelling reputation in a multi-product context. And finally, we control for selection effects in the quality-price relationship.

The empirical analysis is based on 24,547 observations (bottles of wine) from 488 wineries in Austria in the period 2004 until 2007. Note that only a subset of all wines of a particular winery is graded in the Fallstaff-Wine-Guide. Explicitly taking into account the selection of wines for grading, we estimate a structural model for wine prices, the experts' evaluation of wine quality as well as the wineries' reputation. Modelling the wineries' selection of wines for grading in a sample-selection framework suggests that selection effects are important in the quality equation. Ignoring the fact that submission for grading as well as the decision on the price of a particular wine depends on winery's quality evaluation (which is unobservable and can deviate from the experts' evaluation) could introduce a bias in the estimated quality-price relationship. We find that this selection bias is not significant in the present context.

Consistent with the existing literature, we find that expert quality ratings exert a positive and significant impact on the price of wine. In addition to this direct effect of quality on price, we also observe a significant indirect effect: An increase in product quality improves the reputation of a winery in the following periods, which again contributes to higher wine prices in the next years (‚stream of reputational dividends'). These results are consistent with well established theoretical models of firm reputation according to which a firm's reputation is shaped by its past actions (i.e. the quality of its products delivered in the past). Simulation experiments indicate that this (endogenous) reputation effect of quality is important; the long-run price effects of a one-time quality increase are about 34% larger when the ‚stream of reputational dividends' are taken into account compared to a situation where (endogenous) reputation effects are ignored.

We find that reputation effects are particularly important for the top wines of a winery. The long run price effects of an increase in the quality of the top wine (via the improvement of a winery's reputation) are significantly larger then those of a quality improvement of any other wine. Moreover, these ‚stream of reputational dividends' from a quality increase of the top wine can be realized for all wines sold from the winery. Reputation generated from top wines has a significant spill-over effect (‚umbrella effect')

Figure 2: The (Long-Run) ‚Stream of Reputational Dividends'

| (a) transitory quality improvement | (b) permanent quality improvement |
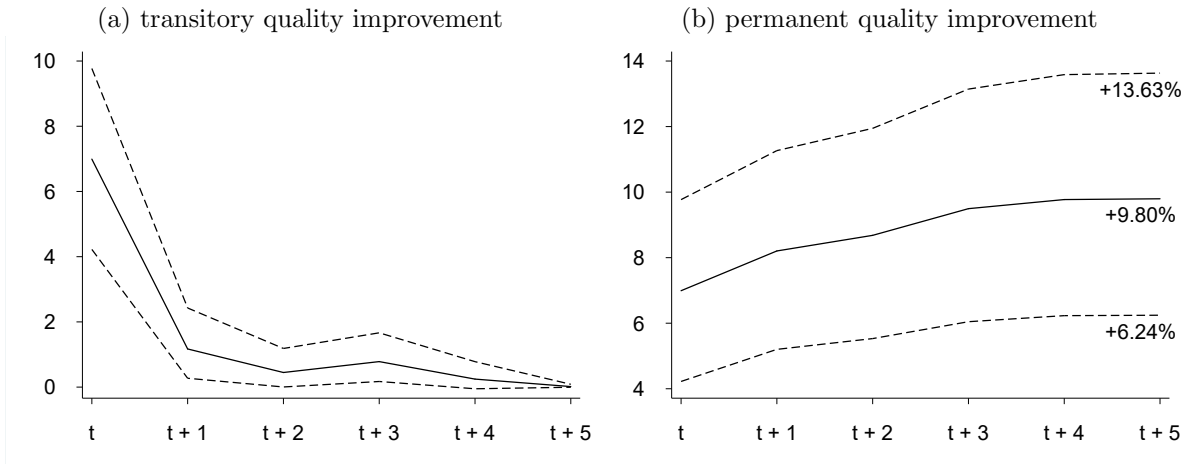
Notes: Figures illustrate price effects in % of an increase in quality of all wines of a winery by one index-point in year t. Figure (a) depicts the effect of a transitory quality increase in year t only, whereas Figure (b) shows the effects of a permanent quality improvement. Estimated effects are obtained by applying a bootstrap simulation technique with 10,000 replications. Each parameter is drawn randomly from a normal distribution with the mean and the standard deviation obtained from the regressions reported in columns [3a] and [4a] of Table 2. The solid lines denote the median of the price effects and the dotted lines denote the 2.5 and the 97.5 percentiles.

Figure 3: Umbrella Effects of Quality Improvements

| (a) transitory quality improvement | (b) permanent quality improvement |
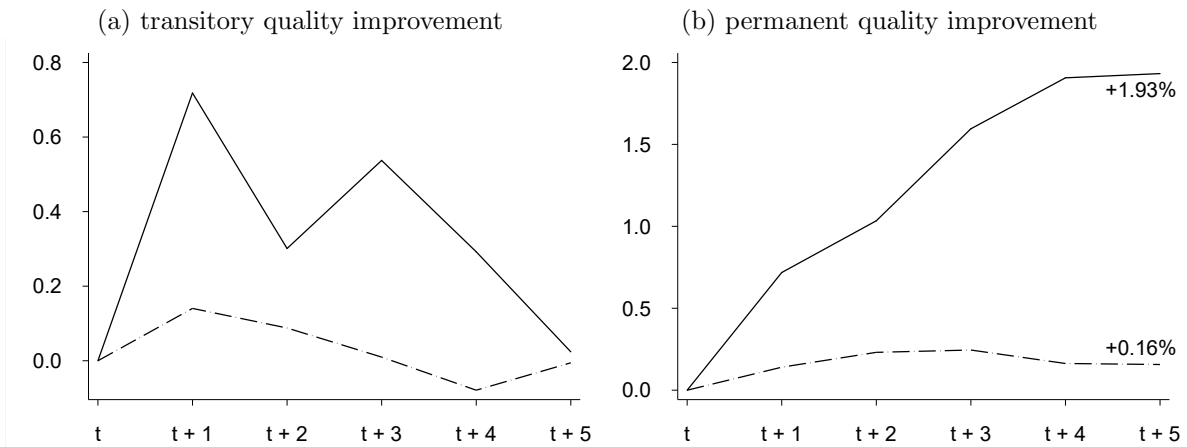
Notes: Figures illustrate the median (over all wines within a winery) price effects in % of an increase in reputation associated with a one index-point quality increase (i) of the top wine (solid line) or (ii) of a median quality wine (i.e. the wine with the median quality score of all graded wines in the winery). Figure (a) depicts the effect of a transitory quality increase in year t only, whereas Figure (b) shows the effects of a permanent quality improvement. Estimated effects are obtained by applying a bootstrap simulation technique with 10,000 replications. Each parameter is drawn randomly from a normal distribution with the mean and the standard deviation obtained from the regressions reported in columns [3b] and [4b] of Table 2.

on all other wines sold by the winery.

These results also have managerial implications. Form the observation that reputation is more important than taste in determining market prices (Benfratello et al., 2009), researchers suggest that firms should primarily focus at building an established reputation for their product by pursuing promotional activities. While we do not question the importance of promotional activities, our results suggest that it is the sustained sensory quality of a wine over time that leads to a winery's high reputation. Our analysis thus provides strong empirical support for theoretical models focusing on the interrelationship between quality and reputation (Board and Meyer-ter Vehn, 2013) in experience goods markets. We further provide a clear answer to the ongoing debate in the wine price and quality literature on the question: Does quality (as evaluated by experts) matter? Our answer is affirmative; quality matters and is particularly important for the top wine of a winery when the long-run consequences of improvements in quality, i.e. the ‚stream of reputational dividends' as well as umbrella effects, are explicitly taken into account.

References

Ali, H. H. and Nauges, C. (2007). The pricing of experience goods: The example of en primeur wine. *American Journal of Agricultural Economics* 89: 91–103.

Allen, F. (1984). Reputation and product quality. *The RAND Journal of Economics* 15: 311–327.

Andersson, F. (2002). Pooling reputations. *International Journal of Industrial Organization* 20: 715–730.

Basdeo, D. K., Smith, K. G., Grimm, C. M., Rindova, V. P. and Derfus, P. J. (2006). The impact of market actions on firm reputation. *Strategic Management Journal* 27: 1205–1219.

Benfratello, L., Piacenza, M. and 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: 2197–2209.

Board, S. and Vehn, M. Meyer-ter (2013). Reputation for quality. *Econometrica* 81: 2381–2462.

Cabral, L. and Hortacsu, A. (2010). The dynamics of seller reputation: Evidence from ebay. *The Journal of Industrial Economics* 58: 54–78.

Cabral, L. M. B. (2009). Umbrella branding with imperfect observability and moral hazard. *International Journal of Industrial Organization* 27: 206–213.

Caliendo, M., Michel, C. and Edlira, S. (2015). The effect of individual professional critics on books' sales: Capturing selection biases from observable and unobservable factors. *Marketing Letters* 26: 423–436.

Castriota, S. and Delmastro, M. (2015). The economics of collective reputation: Evidence from the wine industry. *American Journal of Agricultural Economics* 97: 469–489.

Caves, R. (2000). *Creative Industries*. MA: Harvard University Press.

Combris, P., Lecocq, S. and Michael, V. (1997). Estimation of a hedonic price equation for Bordeaux wine: Does quality matter? *Economic Journal* 107: 390–402.

Combris, P., Lecocq, S. and Michael, V. (2000). Estimation of a hedonic price equation for Burgundy wine. *Applied Economics* 32: 961–967.

Dewenter, R. and Heimeshoff, U. (2014). Media bias and advertising: Evidence from a German car magazine. *Review of Economics* 65: 77–94.

Dewenter, R. and Heimeshoff, U. (2015). Do expert reviews really drive demand? Evidence from a German car margazine. *Applied Economics Letters* 22: 1150–1153.

Dubois, P. and 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: 205–212.

Estrella Orego, M., Defrancesco, E. and Gennari, A. (2012). The wine hedonic price models in the 'old and new world': State of the art. *Revista de la Facultad de Ciencias Agrarias, UNCuyo* 44: 205–220.

Fishman, A., Finkelstein, I., Simhon, A. and Yacouel, N. (2018). Collective brands. *International Journal of Industrial Organization* 59: 316–339.

Friberg, R. and Grönqvist, E. (2012). Do expert reviews affect the demand for wine? *American Economic Journal: Applied Economics* 4: 193–211.

Gergaud, O., Florine, L., Rickard, B. and Warzynski, F. (2017). Evaluating the net benefits of collective reputation: The case of bordeaux wine. *Food Policy* 71: 8–16.

Ginsburgh, V. (2003). Awards, success and aesthetic quality in the arts. *The Journal of Economic Perspectives* 17: 99–111.

Golan, A. and Shalit, H. (1993). Wine quality differentials in hedonic grape pricing. *Journal of Agricultural Economics* 44: 311–321.

Hakenes, H. and Peitz, M. (2009). Umbrella branding and external certification. *European Economic Review* 53: 186–196.

Hilger, J., Rafert, G. and Villas-Boas, S. (2011). Expert opinion and the demand for experience goods: An experimental approach in the retail wine market. *The Review of Economics and Statistics* 93: 1289–1296.

Hörner, J. (2002). Reputation and competition. *The American Economic Review* 92: 644–663.

Klein, B. and Leffler, K. B. (1981). The role of market forces in assuring contractual performance. *Journal of Political Economy* 89: 615–641.

Kreps, D., Milgrom, P., Roberts, J. and Wilson, R. (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory* 27: 245–252.

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

Landon, S. and Smith, C. E. (1998). Quality expectations, reputation, and price. *Southern Economic Journal* 64: 628–647.

Lange, D., Lee, P. M. and Dai, Y. (2011). Organizational reputation: A review. *Journal of Management* 37: 153–184.

Lecocq, S. and Visser, M. (2006). What determines wine prices: Objective vs. sensory characteristic. *Journal of Wine Economics* 1: 42–56.

Luca, M. and Zervas, G. (2016). Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science* 62: 3393–3672.

Nerlove, M. (1995). Hedonic price functions and the measurement of preferences: The case of swedish wine consumers. *European Economic Review* 39: 1697–1716.

Oczkowski, E. (1994). A hedonic price function for Australian premium table wine. *Australian Journal of Agricultural and Resource Economics* 38: 93–110.

Oczkowski, E. (2001). Hedonic wine price functions and measurement error. *Economic Record* 77: 374–382.

Oczkowski, E. and Doucouliagos, H. (2015). Wine prices and quality ratings: A meta-regression analysis. *American Journal of Agricultural Economics* 97: 103–21.

Rasmusen, E. B. (2016). Leveraging of reputation through umbrella branding: The implications for market structure. *Journal of Economics and Management Strategy* 25: 261–273.

Reinstein, D. A. and 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: 27–51.

Reiss, P. C. and Wolak, F. A. (2007). Structural econometric modeling: Rationales and examples from Industrial Organization. In Heckman, J. J. and Leamer, E. E. (eds), *Handbook of Econometrics*. Elsevier, *6A*, chap. 64, 4280–4412.

Rob, R. and Fishman, A. (2005). Is bigger better? Customer base expansion through word-of-mouth reputation. *Journal of Political Economy* 113: 1146–1162.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *The Journal of Political Economy* 82: 34–55.

Schamel, G. and Anderson, K. (2003). Wine quality and varietal, regional and winery reputations: Hedonic prices for Australia and New Zealand. *Economic Record* 79: 357–369.

Shapiro, C. (1983). Premiums for high quality products as returns to reputations. *The Quarterly Journal of Economics* 98: 659–679.

Sorensen, A. T. (2007). Bestseller lists and product variety. *The Journal of Industrial Economics* 55: 715–738.

Storchmann, K. (2012). Wine economics. *Journal of Wine Economics* 7: 1–33.

Storchmann, K. (2018). Wine economics. In Ashenfelter, O., Gergaud, O., Storchmann, K. and Ziemba, W. (eds), *Handbook of the Economics of Wine*. World Scientific, 3–40.

Stuen, E. T., Miller, J. R. and Stone, R. W. (2015). An analysis of wine critic consensus: A study of Washington and California wines. *Journal of Wine Economics* 10: 47–61.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies* 63: 1–22.

Wooldridge, J. M. (1995). Score diagnostics for linear models estimated by two stage least squares. In Maddala, G. S., Phillips, P. C. B. and Srinivasan, T. N. (eds), *Advances in Econometrics and Quantitative Economics*. Oxford: Blackwell, 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 A.1: Descriptive Statistics of Variables used in the Empirical Analysis

Variable	# of Obs.	Mean	Std. Dev.	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
Top wine (at $t-1$)	24,547	0.08		0	1
Average quality	22,036	88.95	1.29	85.75	94
Maximum quality	22,036	91.20	1.87	87	98
Median quality	22,036	88.83	1.37	85.5	95
Wine age: 1 year	24,547	0.75		0	1
Wine age: 2 years	24,547	0.24		0	1
Wine age: 3 years	24,547	0.01		0	1
# of grades between 1999 and 2002	24,547	0.98	1.35	0	4
# of different wines of winery (in year t)	24,547	15.38	5.12	1	38
Type of Wine:					
White wine	24,547	0.55		0	1
Red wine	24,547	0.33		0	1
Rosé wine	24,547	0.03		0	1
Spaetlese (sweet wine)	24,547	0.02		0	1
Beerenauslese (sweet wine)	24,547	0.02		0	1
Trockenbeerenauslese (sweet wine)	24,547	0.03		0	1
Eiswein (sweet wine)	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
Rose	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

Appendix B (online): Selection Problem

To analyze the selection problem it is convenient to think about it as a missing variables bias. To illustrate the problem we start with the price equation (3):

$$E\left[p_t\vert\left(\hat{q}_t^E,\hat{r}_t,\boldsymbol{X},\hat{\lambda}\right),\ S_t=1\right]=\alpha_1\hat{q}_t^E+\alpha_2\hat{r}_t+\boldsymbol{X}\boldsymbol{\alpha}+\rho_4\lambda\left(\boldsymbol{Z}\hat{\boldsymbol{\gamma}}\right)$$

and simplify the regression to $p_t=\eta_w+\alpha_1 q_t^E+\epsilon_t$ for convenience, with η_w as winery specific fixed effects and ϵ_t as the error term. As q_t^E is not observed by the winery (because prices are chosen before the quality is assessed by the experts), the winery has to decide on prices based on its own quality evaluations q_t^W. Both quality ratings are related as follows: $q_t^W=q_t^E+m_t$. m_t is simply the difference in the quality assessments between the wineries and the experts.

As the prices are based on the wineries' quality assessments, the ,true' price equation is $p_t=\eta_w+\alpha_1 q_t^W+\epsilon_t=\eta_w+\alpha_1 q_t^E+\alpha_1 m_t+\epsilon_t$. The parameter of interest, α_1, can be estimated consistently if m_t and q_t^E are uncorrelated. Assume that $m_t=\nu_w+\beta_1 q_t^E+\upsilon_t$, with ν_w are winery-specific differences in quality evaluations relative to experts and υ_t is the (random) error term. Inserting m_t in the (simplified) price regression gives $p_t=\eta_w+\alpha_1 q_t^E+\alpha_1(\nu_w+\beta_1 q_t^E+\upsilon_t)+\epsilon_t$ or $p_t=(\eta_w+\alpha_1\nu_w)+\alpha_1(1+\beta_1)q_t^E+(\epsilon_t+\alpha_1\upsilon_t)$. While we are interested in estimating α_1, we indeed estimate $\alpha_1(1+\beta_1)$. Obviously, β_1 must be zero for $\hat{\alpha}_1$ to be an unbiased estimator of α_1.

Note that the estimated effect of quality on prices, $\hat{\alpha}_1$, is unbiased, even if the random component of the difference between wineries' and experts' grading, υ_t, is high, or if wineries systematically over- or underestimate the quality of all of their wines (captured by ν_w). The only assumption necessary to get a consistent estimator for α_1 is that the difference in wineries' and experts' quality evaluations does not systematically depend on the quality q_t^E, i.e. $\beta_1=0$.

The issue is less straightforward in case of sample selection. We describe two different situations to illustrate possible omitted variables biases introduced by the non-random sample selection. As the wineries select which wines are sent to the wine guide for grading, the selection depends on the wineries' evaluations q_t^W, i.e. it depends on both q_t^E and m_t. (i) Assume that the difference m_t is homoscedastic with respect to q_t^E, i.e. the absolute difference between wineries and experts in evaluating wine quality is independent of q_t^E. In this case, wines of mediocre quality q_t^E are only selected if m_t is high, while wines of very high quality are selected irrespective of m_t (because q_t^E is sufficiently high such that q_t^W is large enough). As a consequence of the sample selection, β_1 is negative and $\hat{\alpha}_1$ is biased downwards. (ii) Assume that the difference m_t is heteroscedastic with respect to q_t^E, i.e. the expected absolute difference between wineries and experts in evaluating wine quality increases with q_t^E. This means that experts and wineries have similar evaluations of mediocre wines, but disagree more strongly on high quality wines. If the variance of m_t is large relative to the dispersion of q_t^E, m_t is more important in the selection equation than q_t^E. In the limit the selection depends

solely on m_t. As only wines are selected if m_t passes a particular value (e.g. if $m_t > 0$), m_t will be positively correlated with q_t^E (i.e. $\beta_1 > 0$) and $\hat{\alpha}_1$ will be biased upwards.

Appendix C (online): Identification

To identify the selection equation (1) and the first-stage regressions on quality (2) and reputation (4) we need at least three exclusion restrictions in the price equation (3). These instrumental variables excluded from the price equation include three (sets of) variables in addition to the average quality provided by the winery in previous years ($\bar{q}_{t-k}^E$): (i) The number of times a particular wine is graded by the wine guide between 1999 and 2002 (i.e. before the sample period). We expect that the quality of a particular wine is correlated over time, as some variables affecting product quality – like soil or climatic conditions or producer-specific skills – do not change much over time and favour some wines produced by a winery more than others. If a wine is graded very often before the sample period, this can be interpreted that this wine is (likely to be) of high quality relative to other wines produced by this winery. This variable should thus be positively correlated with the quality of a wine and with the probability of a wine to be submitted to the wine guide to receive expert reviews within the sample period. (ii) The number of wines produced by a winery in a particular vintage. As space in the wine guide is limited, an increase in the number of wines should reduce the probability of a particular wine to be selected. (iii) Last, dummy variables indicating the year the grapes were harvested (crop year). These variables might influence product quality, as these variables also summarize overall weather conditions, while prices should be affected by the vintage rather than the crop year.[15]

Appendix D (online): Bootstrapped Standard Errors

To explicitly account for $IMR(\hat{\lambda})$ as a generated regressor rather than an observed variable we use bootstrapped standard errors in this sensitivity analysis. We report the point estimates of the price equation corresponding to column [3a] and column [3b] of Table 2, along with the bootstrapped and the non-adjusted ('non-bootstrapped') standard errors, in Table D.1.

Note that the number of observations is somewhat smaller compared to the main specifications reported in Table 2, because we exclude wineries with only a small number of selected wines for grading or with only a small number of non-selected wines. This has technical reasons: Bootstrapping is a random sampling method with replacement. If only a small number of wines of a particular winery is selected for grading in the full sample, then it may be the case that none of the wines of this winery in one particular randomly drawn sample is selected for evaluation. In this case the respective winery

---

[15]Note that all exogenous variables are included in the selection equation (1) and in the first stage regressions on quality (2) and reputation (4), as suggested by Wooldridge (2001).

fixed effect predicts the outcome of the selection equation perfectly. The respective fixed effect cannot be estimated in this particular replication (because the selection equation is estimated as a probit model), and thus the regression results of this replication cannot be used to compute bootstrapped standard errors. To reduce the probability of this to happen we reduce the sample accordingly. Nevertheless, out of 1,000 randomly selected samples only 450 and 470 replications can be used for calculating the bootstrapped standard errors in each specification, respectively, for the reasons outlined above. The point estimates and the (non-adjusted) standard errors reported in Table D.1 thus differ slightly from the regression results summarized in Table 2.

Comparing the bootstrapped with the non-adjusted standard errors in Table D.1 shows (i) that the difference is rather small and (ii) that the bootstrapped standard errors are always smaller compared to the non-adjusted standard errors. This suggests that we are conservative when evaluating the statistical significance of the parameter estimates reported in Table 2.

Table D.1: Regression Results Explaining Prices using Bootstrapped Standard Errors

Dependent Variable	ln(price)					ln(price)				
Method	2SLS					2SLS				
	Coeff.	Std. Dev. (bootstr. S.E.)	Sign.	Std. Dev. (non-bootstr. S.E.)	Sign.	Coeff.	Std. Dev. (bootstr. S.E.)	Sign.	Std. Dev. (non-bootstr. S.E.)	Sign.
Quality	0.0681	(0.0122)	***	(0.0150)	***	0.0648	(0.0111)	***	(0.0148)	***
Reputation	0.4320	(0.1484)	***	(0.1767)	**	0.3877	(0.1254)	***	(0.1499)	**
Wine age: 2 years	0.4149	(0.0213)	***	(0.0277)	***	0.4189	(0.0185)	***	(0.0276)	***
Wine age: 3 years	0.8223	(0.0431)	***	(0.0580)	***	0.8296	(0.0411)	***	(0.0571)	***
Top wine (at $t-1$)	0.1274	(0.0157)	***	(0.0184)	***	0.1303	(0.0144)	***	(0.0183)	***
$IMR(\hat{\lambda})$	-0.0126	(0.0122)		(0.0188)		-0.0155	(0.0119)		(0.0192)	
Constant	-4.1495	(1.1065)	***	(1.3500)	***	-3.8437	(0.9914)	***	(1.3230)	***
Winery effects	Yes (307)					Yes (307)				
Vintage effects	Yes (3)					Yes (3)				
Variety of the grape	Yes (31)					Yes (31)				
Type of wine	Yes (6)					Yes (6)				
Crop year	No					No				
Instrumented variables	Reputation, Quality					Reputation, Quality				
Number of observations	Selection equation: 21,340					Selection equation: 21,340				
Number of observations	Structural equation: 6,652					Structural equation: 6,652				
R^2	0.826					0.825				

Standard errors are reported in parentheses and are clustered at the winery level. Bootstrapped standard errors are based on 450 replications (first regression) and 470 replications (second regression), respectively. *** significant at 1 %, ** significant at 5 %, * significant at 10 % level. Wineries with 8 or less selected wines or with 8 or less non-selected wines over the entire sample period are excluded from the analysis. The first (second) regression summarized in this Table corresponds to specification [3a] ([3b]) reported in Table 2.