Journal of Wine Economics, Volume 15, Number 3, 2020, Pages 284–311
doi:10.1017/jwe.2020.16

Pricing Models for German Wine: Hedonic Regression vs. Machine Learning

Britta Niklas [a] **and Wolfram Rinke** [b]

Abstract

This article examines whether there are different hedonic price models for different German wines by grape variety, and identifies influential factors that focus on weather variables and direct and indirect quality measures for wine prices. A log linear regression model is first applied only for Riesling, and then machine learning is used to find hedonic price models for Riesling, Silvaner, Pinot Blanc, and Pinot Noir. Machine learning exhibits slightly greater explanatory power, suggests adding additional variables, and allows for a more detailed interpretation of results. Gault&Millau points are shown to have a significant positive impact on German wine prices. The log linear approach suggests a huge effect of different quality categories on the wine prices for Riesling with the highest price premiums for Auslese and "Beerenauslese/Trockenbeerenauslese/Eiswein (Batbaice)," while the machine learning model shows, that additionally the alcohol level has a positive effect on wines in the quality categories "QbA," "Kabinett," and "Spätlese," and a mostly negative one in the categories "Auslese" and "Batbaice." Weather variables exert different affects per grape variety, but all grape varieties have problems coping with rising maximum temperatures in the winter and with rising minimum and maximum temperatures in the harvest season. (JEL Classifications: C45, L11, Q11)

Keywords: German wine prices, hedonic pricing, machine learning, weather changes.

I. Introduction and Literature

In the hedonic price approach, the price of a good or service is split into several implied prices that relate to specific characteristics of a good or service for which consumers are willing to pay. In the most common application of the hedonic

We are indebted to an anonymous referee and the participants at the 11th Annual AAWE Conference in Padua for many helpful comments.
[a] Institute of Development Research and Development Policy, Ruhr-University Bochum, Universitätsstr.105, 44789 Bochum, Germany; e-mail: britta.niklas@rub.de, corresponding author.
[b] Department of Information-Technology and Information-Management, Fachhochschule Burgenland GmbH, Campus 1, A-7000 Eisenstadt, Austria; e-mail: wolfram.rinke@fh-burgenland.at.

approach to analyzing wine prices, a log-linear regression is used to estimate the price

$$ln(P) = X\beta + \varepsilon, \tag{1}$$

where P is a price vector of a 0.75 liter bottle of wine, X denotes a matrix of characteristics which are supposed to have an influence on the wine price, β is the vector of parameters that are associated to these characteristics, and ε is the error term (Thrane, 2004).

There is a growing number of studies dealing with the impact of weather changes on wine quality and wine prices.[1] Most of these studies suggest that temperature and precipitation during the growing and harvesting seasons may have an important impact on wine prices (e.g., Ashenfelter, 2010; Ashenfelter and Storchmann, 2010; Byron and Ashenfelter, 1995; Haeger and Storchmann, 2006; Jones and Storchmann, 2001; Lecocq and Visser, 2006; Oczkowski, 2001, 2014, 2016; Ramirez, 2008; Storchmann, 2005, 2012). As weather information is not given on the wine label, visible quality indicators include the quality categories or the alcohol level as direct measures of quality that are affected by weather (Niklas, 2017), and wine guide scores as indirect measures of wine quality (Schamel, 2000, 2002, 2003; Shapiro, 1983; Tirole, 1996).

Most of the extant studies use either time series data or cross-sectional data and apply (log)linear regression approaches to identify the impact of the weather, while other disciplines such as engineering or stock exchange trading use machine learning (Shavlik and Diettrich, 1990) as the core technology for their model building capabilities (Stone et al., 2016).[2] There is only one paper to date, that deals with the prediction of wine prices applying machine learning. This paper focuses on time-series analyses required for stock exchange trading (Yeo, Fletcher, and Shawe-Taylor, 2015).

The current study focuses on separate equations for different German grape varieties, assuming that these grape varieties react differently to weather variables. The results of the log linear regression (and squared forms) and the machine learning are first compared for the Riesling grape variety, then the machine learning is applied for the other grape varieties.

The article is organized as follows. Section II presents the data. Section III describes the methodological approach. Section IV reports the results and Section V draws conclusions.

[1] Ashenfelter and Storchmann (2016) provide a survey of the relevant literature.
[2] Autonomous driving, speech recognition, or face recognition technology in modern photographic equipment or video monitoring are other examples for the application of machine learning (National Science and Technology Council, Committee on Technology, 2016).

II. Data on German Weather and Wine Prices

We base our analysis on a dataset that covers farm gate prices for 0.75 liter bottles of wine,[3] reported by 177 wine producers that we sampled at random for the Riesling, Silvaner, Pinot Blanc, and Pinot Noir grape varieties for the vintages 1998 to 2013, for all 13 German wine regions. We obtained farm gate prices[4] from these producers with data on various control variables from "Gault&Millau Weineguide Deutschland" (Diel and Payne, 2002–2009; Payne, 2010–2015). Table 1 summarizes all collected variables for the Riesling grape variety.[5]

We use German quality categories as direct measures of quality. German wines are categorized by the degree of ripeness of the grapes, measured as the content of natural sugar in the must (grape juice) at harvest. However, these so called "quality categories" do not measure, per se, whether a wine is of good or bad quality.[6] The German term is "degree Oechsle"[7] and the following categories are distinguished: Qualitätswein (QbA),[8] Kabinett, Spätlese, Auslese, Beerenauslese, Trockenbeerenauslese, and Eiswein[9] (Ashenfelter and Storchmann, 2010).

Instead of collecting data from a single weather station, as shown in the literature (Lecocq and Visser, 2006; Haeger and Storchmann, 2006), this study uses daily data (daily average temperature,[10] average of the daily maximum and daily minimum temperature all in degrees Celsius, sum of precipitation in mm, and daily average humidity in percent) from 13 different local weather stations,[11] as especially since precipitation varies between German wine regions, which therefore differ in their

[3] In the Gault&Millau Weinguide Deutschland (Diel and Payne, 2002–2009; Payne, 2010–2015), only prices of 0.75 liter bottles are published.

[4] Farm gate prices are those prices that consumers pay when visiting the farm and buying in the "farm shop," but could be related to recommended retail prices.

[5] Not all variables are included in all models; further explanations can be found in the methodology Chapter III. The descriptive statistics for Silvaner, Pinot Blanc, and Pinot Noir can be found in the online supplementary material.

[6] Hence there is no multicollinearity issue between quality levels and Gault&Millau points. Please see the correlation matrix between Gault&Millau points and quality levels in Table A1 of the Appendix.

[7] It denotes the specific weight of the must compared to the weight of the water at a temperature of 20 degrees. One liter of water weighs 1,000 g, which equals 0 degrees Oechsle. Grape must, with a mass of 1,084g per liter, has 84 degrees Oechsle. The mass difference is almost entirely due to the sugar dissolved in the must, so that Oechsle measures the sweetness of the grape juice (Ashenfelter and Storchmann, 2010).

[8] Although their production levels are insignificant (less than 1% of total production), there are two categories below the QbA, that is, Tafelwein and Landwein.

[9] The Oechsle thresholds vary from region to region. In the region Pfalz (Palatinate), for example, the quality categories fall into the following brackets: Qualitätswein (QbA) (60–72° Oechsle), Kabinett (73–84° Oechsle), Spätlese (85–91° Oechsle), Auslese (92–119° Oechsle), Beerenauslese and Eiswein (120–149° Oechsle), and Trockenbeerenauslese (>150° Oechsle) (Ashenfelter and Storchmann, 2010).

[10] The average is calculated from at least 21 hourly measurements each day.

[11] The German Weather Service provided the regional weather data from the following weather stations: Ahr: Bad Neuenahr-Ahrweiler – Baden: Karlsruhe/Rheinstetten and Freiburg, Franken: Würzburg as closest station, Hessische Bergstraße: Mannheim – Mittelrhein: Montabaur and Nastätten, Mosel:

Table 1
Descriptive Statistics for Riesling

Variable	*Obs.*	*Mean*	*Std. Dev.*	*Min.*	*Max.*	*Skewness*	*Kurtosis*
Price 0.75l bottle	8,334	19.42	32.76	2.10	550	5.74	47.94
Gault&Millau points	8,334	86.39	3.2	72	100	0.27	3.09
Age	8,334	2.06	0.21	2	9	9.74	195.59
Trend	8,334	10.07	3.94	3	16	–0.16	1,84
Alcohol	8,334	10.96	2.04	5.5	14.5	–0.68	2.16
Precipitation Growing Season (sum)[c]	8,334	266.23	79.16	70.4	584	0.92	4.64
Precipitation Harvest Season (sum)[c]	8,334	156.02	65.38	40.2	500.9	1.10	5.94
Air Temperature Growing Season (avg.)[c]	8,334	17.36	0.96	15.25	21.14	0.58	3.96
Air Temperature Harvest Season (avg.)[c]	8,334	10.34	1.11	7.76	13.78	0.49	3.69
Minimum Air Temperature Growing Season (avg.)[c]	8,334	12.12	0.72	10.62	15,07	0.79	3.83
Minimum Air Temperature Harvest Season (avg.)[c]	8,334	6.65	1.04	4.45	9.74	0.65	3.25
Maximum Air Temperature Winter Season (avg.)[c]	8,334	13.22	2.11	7.36	17.61	–0.10	2.51
Sunshine Hours Winter Season (sum)[c]	8,334	330.55	65.02	122.4	538.05	–0.20	3.29
Humidity Winter Season (avg.)[c]	8,334	7.62	0.61	6.03	8.99	0.01	2.35
Humidity Growing Season (avg.)[c]	8,334	13.59	0.66	12.01	15.80	–0.13	2.98
Humidity Harvest Season (avg.)[c]	8,334	10.57	0.91	8.78	13.69	0.86	4.70
QbA[a]	8,334	0.355	0.48	0	1	0.61	1.37
Kabinett	8,334	0.18	0.39	0	1	1.64	3.71
Spätlese	8,334	0.285	0.45	0	1	0.96	1.91
Auslese	8,334	0.12	0.33	0	1	2.34	6.50
Batbaice[b]	8,334	0.06	0.24	0	1	3.68	14.57
Ahr	8,334	0.003	0.058	0	1	17.16	295.68
Baden	8,334	0.041	0.020	0	1	4.61	22.27
Franken	8,334	0.049	0.22	0	1	4.16	18.33
Hessische Bergstr.	8,334	0.007	0.85	0	1	11.66	136.92
Mittelrhein	8,334	0.057	0.23	0	1	3.81	15.57
Mosel	8,334	0.329	0.47	0	1	0.73	1.53
Nahe	8,334	0.091	0.29	0	1	2.84	9.07
Pfalz	8,334	0.123	0.33	0	1	2.29	6.24
Rheingau	8,334	0.129	0.33	0	1	2.22	5.92
Rheinhessen	8,334	0.130	0.34	0	1	2.20	5.85
Saale-Unstrut	8,334	0.005	0.68	0	1	14.51	211.72
Sachsen	8,334	0.004	0.61	0	1	16.31	266.87
Württemberg	8,334	0.031	0.17	0	1	5.38	29.97

[a]QbA: Qualitätswein

[b]Batbaice: Beerenauslese/Trockenbeerenauslese/Eiswein

[c]Winter Season: 12/01–02/28; Growing Season: 03/01–09/15; Harvest Season: 09/16–10/31

Source: Authors' calculations.

suitability for grape growing, while one weather station would have been suitable in the case of the temperature due to spatial correlation (Ashenfelter and Storchmann, 2010; Haeger and Storchmann, 2006). Overall, our sample comprises 8,334 observations for Riesling, 2,004 for Pinot Noir, 1,294 for Pinot Blanc, and 917 for Silvaner.

Mainly the regions Mosel (32.9%), Pfalz (12.3%), Rheingau (12.9%), and Rheinhessen (13%) produce Riesling and the quality categories are QbA (35.5%), Spätlese (28.5%), Kabinett (18%), Auslese (12%), and Beerenauslese/Trockenbeerenauslese/Eiswein (Batbaice) (6%). The price of Riesling varies tremendously, from €2.10 to €550.00 per bottle, while Gault&Millau quality points vary between 72 and 100.

III. Methods

In the first step, we run a log linear regression for Riesling and follow the variables employed in the previous literature. The hedonic equation is

$$\begin{aligned} ln(P_i) = \beta_0 + \beta_1 GMP_{it} + \beta_2 age_{it} + \beta_3 trend_{it} + \beta_4 W_{4it} + \ldots + \beta_k W_{kit} \\ + \gamma_1(Quality) + \gamma_2(Region) + \gamma_3(Producer) + \varepsilon_{it}, \end{aligned} \tag{2}$$

where *GMP* denotes Gault&Millau points (as an indirect measure of quality), *age* is the age of the wine at the time of purchase, *trend* is an annual trend variable to capture inflationary trends, $W_4 \ldots W_k$ are various weather variables (average air temperature, squared temperature, sum of precipitation in mm) which we split into the growing and harvest seasons,[12] *Quality* is a series of dummy variables for the German quality categories (as a direct measure of quality), *Region* is a series of dummy variables depicting the region from where the grapes are sourced, and *Producer* is a series of dummy variables reflecting price policies and other farm specific factors (production cost, capital cost, owner structure, etc.).

In the second step, we substitute the log linear regression algorithm with machine learning (Witten, Eibe, and Hall, 2017). We build a non-linear regression model including all variables, which uses a classic feed forward artificial neural network (ANN) algorithm (Witten, Eibe, and Hall, 2017; Hornik, Stichcombe, and White, 1990; Rumelhart, Hinton, and Williams, 1986) to generate the functional model between environmental and other variables related to logarithmic bottle prices.

Bernkastel and Trier Petrisberg, Nahe: Bad Kreuznach, Pfalz: Bad Dürkheim, Rheingau: Geisenheim, Rheinhessen: Alzey, Saale-Unstrut: Osterfeld, Sachsen: Dresden-Hosterwitz, Württemberg: Sachsenheim.

[12] The whole season comprises the winter (Dec. 1 to Feb. 28), growing (Mar. 1 to Sep. 15.), and harvest (Sep. 16 to Oct. 31). The growing phase now starts in March, while the majority of grapes are picked from the middle of September to the end of October (Fecke, 2014; Kriener and Mortsiefer, 2017; Jones and Storchmann, 2001).

An ANN is a mathematical simulation of the biological nervous cell system and consists of many regression units, which are typically non-linear, but could also be linear for scaling purposes. These units are also known in the literature as perceptrons (Rosenblatt, 1958), neurons, or nodes and are organized in layers (Witten, Eibe, and Hall, 2017; Hornik, Stichcombe, and White, 1990; Rumelhart, Hinton, and Williams, 1986). Each layer is stacked on top of the other and all nodes of the underlying layer have a direct weighted link to each unit in the following layer. Several kinds of architecture exist, where the layers can be skipped or where there are feedback loops between previous layers (Witten, Eibe, and Hall, 2017). Feed forward networks, those without any feedback loops or recursions, typically have three types of layers. The input layer type that uses the values of the independent variables as input, the hidden layer type that is responsible for the non-linear functional regression, and the output layer type that performs the final transformation to the dependent variables. Equation (3) describes the general form of the network architecture as

$$ln(P_i) = f_o\left(\sum\nolimits_{k\rightarrow h} w_{kh} f_h\left(\sum\nolimits_{j\rightarrow k} w_{jk} f_i\left(\sum\nolimits_{i\rightarrow j} w_{ij} x_i\right)\right)\right), \tag{3}$$

where x_i is an independent variable and f_i is the transfer function on the input layer, f_h is the transfer function of the hidden layer, and f_o is the transfer function of the output layer. The transfer function in this case is the logistic sigmoidal function as described in Equation (4).

$$f(x) = (1 + e^{-x})^{-1} \tag{4}$$

In the third step, we calculate the dependency matrix (Rinke, 2015) for the previously generated ANN model. We derive the dependency matrix from a sensitivity analyses of the ANN model, which Hashem (1992) originally presented and Yeh and Cheng (2010) further investigated. It represents a normalized, accumulated Jacobi matrix (Rudin, 1976) over the observed data samples and expresses the relative importance of an independent variable with respect to the dependent variable of the ANN model.

There are two other methods of describing the "importance" of an independent variable with respect to the dependent variable, but these are based only on the learned internal network weights w_{ij} (Olden and Jackson, 2002; Olden, Joy, and Death, 2004; Garson, 1991; Goh, 1995; Giam and Olden, 2015). Yeh and Cheng (2010) show that calculating importance as a value, based on the first order partial derivative and further on, the second order partial derivative is more accurate than other common methods.

We calculate the dependency factor (Rinke, 2015) for each independent variable of the model with respect to the dependent variable separately using Equation (5). Yeh and Cheng (2010) call this dependency factor the "average linear importance factor"

$$DF(y(x)) = \sqrt{\frac{1}{n}\sum\left(\frac{\partial y}{\partial x}\right)^2} \tag{5}$$

https://doi.org/10.1017/jwe.2020.16 Published online by Cambridge University Press

where $\frac{\partial y}{\partial x}$ represents the partial derivative of the function $y(x)$ and n the number of samples.

In addition to the dependency factors, we examine the semi-elasticity ε of $y(x)$ (Owen, 2012). Semi-elasticity measures the percentage change in the dependent variable caused by a unit change in the independent variable. In the log-linear-regression model $log(y) = \beta_0 + \beta_{1x} + u$. The variable β_1 measures the semi-elasticity of the dependent variable in relation to the respective independent variable. Since we use a log-linear specification between the log-price and an independent variable x, a change of one unit of x results in an $\varepsilon \times 100$ percent change in y.[13] We calculate the semi-elasticity following Equation (6) and the average semi-elasticity following Equation (7).

$$\varepsilon_{y,x} = \frac{\partial y}{\partial x}\frac{1}{y} \tag{6}$$

$$\overline{\varepsilon_{y,xi}} = \frac{1}{n}\sum_{i=1..n} \varepsilon_{y,xi}, \tag{7}$$

for n observations

In the fourth step, we calculate the average semi-elasticity for each independent variable of the model with respect to the dependent variable separately for each German quality category (QbA, Kabinett, Spätlese, Auslese, and Batbaice) and discuss the results. For the remaining grape varieties, Silvaner, Pinot Blanc, and Pinot Noir, we repeat steps two to four and summarize the results (only available online).

IV. Results

A. Riesling Models

(1) Log Linear Regression

We apply the hedonic Equation (2) using average temperature functions for the growing and harvest seasons in *Model 1*, squared temperature functions in *Model 2*, and omitting Gault&Millau points in *Model 3* applying robust standard errors.[14]

The reason for using squared temperatures in *Model 2* is to check for non-linear temperature effects. The Gault&Millau points are omitted in *Model 3* to check whether the weather variables can cover most quality aspects. The results are shown in Table 2 with an R^2 of 0.7974 (*Model 1*), 0.7979 (*Model 2*), and 0.6168 (*Model 3*).[15]

[13] See Greene (2008) for a more detailed discussion.

[14] Since the Breusch–Pagan/Cook–Weisberg test finds heteroscedasticity for all three models, see Table A5 of the Appendix.

[15] There is no multicollinearity issue except for the temperature variables and its squared version in Model 2, the respective variance inflation factor test for *Models 1*, *2*, and *3*, and the correlation matrix can be found in Tables A6 and A7 of the Appendix.

Table 2
Comparison of Model 1, 2, and 3 for Riesling (without Regional and Producer Fixed Effects)

	ln price *Model (1)*	*ln price* *Model (2)*	*ln price* *Model (3)*
Gault&Millau points	0.133***	0.133***	
	(76.53)	(76.28)	
Age	0.0543*	0.0495	0.111**
	(2.08)	(1.93)	(2.68)
Trend	0.0196***	0.0190***	0.0376***
	(17.87)	(17.05)	(25.57)
Precipitation Growing Season (sum)[b]	0.000364***	0.000375***	0.000187*
	(6.22)	(6.31)	(2.29)
Precipitation Harvest Season (sum)[b]	0.000280***	0.000248***	0.000135
	(4.06)	(3.58)	(1.43)
Air Temperature Growing Season (avg.)[b]	0.0455***	0.390***	0.0428***
	(9.42)	(3.87)	(6.28)
Air Temperature Growing Season (avg.) squared[b]		–0.00976***	
		(–3.42)	
Air Temperature Harvest Season (avg.)[b]	0.0226***	0.124**	0.0127*
	(5.96)	(2.71)	(2.42)
Air Temperature Harvest Season (avg.) squared[b]		–0.00499*	
		(–2.33)	
Kabinett	–0.170***	–0.168***	–0.250***
	(–17.68)	(–17.48)	(–18.81)
Spätlese	–0.0454***	–0.0444***	0.127***
	(–4.74)	(–4.64)	(9.23)
Auslese	0.470***	0.470***	1.010***
	(25.81)	(25.79)	(42.95)
Batbaice[a]	1.508***	1.508***	2.308***
	(58.86)	(58.83)	(80.31)
Constant	–10.56***	–14.06***	0.694***
	(–57.36)	(–16.36)	(4.35)
Observations	8,335	8,335	8,335
R2	0.7974	0.7979	0.6168

Note: Robust *t* statistics are in parentheses; significance levels are *$p < 0.05$, **$p < 0.01$, and ***$p < 0.00$.

[a]Batbaice: Beerenauslese/Trockenbeerenauslese/Eiswein

[b]Winter Season: 12/01–02/28; Growing Season: 03/01–09/15; Harvest Season: 09/16–10/31

Source: Authors' calculations.

All significant weather variables show a positive impact on wine prices in all three models. *Model 3* suggests that average temperatures during the growing and harvest seasons have a significantly positive but decreasing effect, because the squared functions are negative. Based on these results, we calculate the price-maximizing temperature, which is 19.98 degrees Celsius for the growing season and 12.42 degrees Celsius for the harvest season.

https://doi.org/10.1017/jwe.2020.16 Published online by Cambridge University Press

Gault&Millau points have the greatest influence on the Riesling price, as the price increases by 13.3% with each additional point in *Models 1* and *2*. The exclusion of Gault&Millau points leads to less accurate results and a much lower R^2.

The price trend has a significantly positive impact in all three models (between 1.9% and 3.76%) and the wine price increases with the age of the wine by 5.43% in *Model 1* and 11.1% in *Model 3*, but is insignificant in *Model 2*.

Quality categories also have a significant impact. Compared to QbA, the Kabinett category leads to a decline in prices for all three models and Spätlese leads to a decline in prices for *Models 1* and *2*, while the higher-quality categories lead to much higher prices. The quality Auslese category has a positive price impact from 47% (*Models 1* and *2*) to 101% (*Model 3*), and Batbaice from 150.8% (*Models 1* and *2*) to 230.8% (*Model 3*).

We then estimate the above models applying regional fixed effects to control regional unobserved heterogeneity. Table 3 (left side) shows the results with an R^2 of 0.8020 (*Model 4*), 0.8020 (*Model 5*), and 0.6348 (*Model 6*).

While most of the results are confirmed, *Model 5* shows the most obvious change in results. Including squared temperature variables now makes all temperature variables insignificant. One reason might be that non-linear temperature effects are only observed in some regions, so the effects are no longer significant when we apply regional fixed effects. We confirm this assumption when running regressions for regional subsamples: squared temperature variables for both the growing and harvest seasons are only significant for the regions of Baden (in the very south of Germany) and for Mittelrhein (a sunny wine region), while for the harvest season squared temperature variables are only significant for the regions of Nahe and Württemberg (again, in the very south of Germany). The majority of observations (6,504 out of 8,335), however, cover German wine regions with linear temperature effects. *Model 6* again has less accurate results than the other two models.

Finally, we estimate the same models applying regional and producer fixed effects that can reflect pricing policies and other farm-specific factors such as production cost, capital cost, or owner structure. Table 3 (right side) shows the results, with an R^2 of 0.8517 (*Model 7*), 0.8518 (*Model 8*), and 0.7812 (*Model 9*).

The application of producer fixed effects generally leads to a higher explanatory power of all models. The effect of Gault&Millau points is less strong than before (10.7% instead of 13.1%), while age and trend have stronger effects. Only in *Model 7* do all weather variables show a significant and positive effect on wine prices, while the inclusion of squared temperature effects still leads to insignificant results. Similar to the previous explanation, this may be due to the majority of producers in this sample (74.1%) are in regions with linear temperature effects according to the regional subsamples.

Table 3
Models for Riesling

	ln price *Model (4)*	*ln price* *Model (5)*	*ln price* *Model (6)*	*ln price* *Model (7)*	*ln price* *Model (8)*	*ln price* *Model (9)*
Gault&Millau points	0.131***	0.131***		0.107***	0.107***	
	(73.81)	(73.78)		(53.40)	(53.35)	
Age	0.0537*	0.0544*	0.109**	0.0716**	0.0728**	0.110***
	(2.05)	(2.07)	(2.94)	(2.85)	(2.89)	(3.60)
Trend	0.0181***	0.0182***	0.0341***	0.0259***	0.0260***	0.0424***
	(16.43)	(16.34)	(23.44)	(24.55)	(24.45)	(34.38)
Precipitation Growing Season (sum)[b]	0.000336***	0.000320***	0.000153	0.000247***	0.000218**	0.0000481
	(4.48)	(3.98)	(1.54)	(3.72)	(3.06)	(0.61)
Precipitation Harvest Season (sum)[b]	0.000247***	0.000255***	–0.0000108	0.000167**	0.000181**	–0.0000166
	(3.39)	(3.46)	(–0.11)	(2.64)	(2.82)	(–0.21)
Air Temperature Growing Season (avg.)[b]	0.0232***	–0.00865	0.0170*	0.0178**	–0.0208	0.0129
	(3.54)	(–0.07)	(1.98)	(3.06)	(–0.20)	(1.87)
Air Temperature Growing Season (avg.) squared[b]		0.000852			0.00101	
		(0.26)			(0.36)	
Air Temperature Harvest Season (avg.)[b]	0.00868*	–0.0226	0.0114	0.00827*	–0.0491	0.00162
	(2.05)	(–0.42)	(0.20)	(2.24)	(–1.08)	(0.37)
Air Temperature Harvest Season (avg.) squared[b]		0.00145			0.00266	
		(0.59)			(1.26)	
Kabinett	–0.178***	–0.178***	–0.250***	–0.168***	–0.168***	–0.205***
	(–18.05)	(–18.03)	(–18.66)	(–17.61)	(–17.61)	(–16.89)
Spätlese	–0.0502***	–0.0503***	0.113***	0.0319***	0.0318***	0.209***
	(–5.22)	(–5.23)	(8.28)	(3.36)	(3.35)	(18.39)
Auslese	0.461***	0.461***	0.976***	0.565***	0.565***	0.966***
	(25.24)	(25.22)	(42.39)	(32.98)	(32.95)	(50.06)

Continued

Table 3
Continued

	ln price *Model (4)*	*ln price* *Model (5)*	*ln price* *Model (6)*	*ln price* *Model (7)*	*ln price* *Model (8)*	*ln price* *Model (9)*
Batbaice[a]	1.504***	1.504***	2.276***	1.673***	1.673***	2.303***
	(58.74)	(58.75)	(80.25)	(67.29)	(67.32)	(94.73)
Regional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Producer fixed effects	No	No	No	Yes	Yes	Yes
Constant	–9.691***	–9.225***	1.265***	–8.116***	–7.437***	0.825***
	(–45.07)	(–8.09)	(6.18)	(–36.72)	(–7.65)	(4.97)
Observations	8,335	8,335	8,335	8,335	8,335	8,335
R2	0.8020	0.8020	0.6348	0.8517	0.8518	0.7812

Note: Robust t statistics are in parentheses; significance levels are $*p < 0.05$, $**p < 0.01$, and $***p < 0.00$.

[a]Batbaice: Beerenauslese/Trockenbeerenauslese/Eiswein

[b]Winter Season: 12/01–02/28; Growing Season: 03/01–09/15; Harvest Season: 09/16–10/31

Source: Authors' calculations.

https://doi.org/10.1017/jwe.2020.16 Published online by Cambridge University Press

Due to the high degree of skewness (5.74) and kurtosis (47.94) in the price data (see Table 1), we assume that marginal effects might vary in different price brackets. Therefore, we finally run a quantile regression for *Model 4* (with regional fixed effects without squared temperatures[16]) and report the results and the corresponding ordinary least square (OLS) results in Table 4 (Column 1), allowing for a direct comparison of results.

Similar to the OLS results, Gault&Millau points have a positive affect on the price. The effect peaks in the 0.75 quantile and then decreases slightly afterwards.

The age of a wine bottle, which is statistically significant in the OLS regression, only exerts a significantly positive price effect in the 0.5 quantile and remains insignificant for all others.

The quantile regression shows significant trend effects like the OLS, but the trend effect decreases with the price quantiles.

Like in the OLS results, wine prices in the growing season are significantly positively affected by precipitation, with a peak in the 0.25 quantile. Precipitation in the harvest season only has a significant positive price effect in the 0.25 quantile and then again in the 0.75 and 0.9 quantiles, the effect increasing with the price quantiles.

The average temperature in the growing season shows a significantly positive price effect only for the 0.25 and 0.5 quantiles, with a peak for the 0.25 quantile. The same holds to the average temperature in the harvest season, which is significantly positive for the 0.9 quantile.

The highest German quality categories, Auslese and Batbaice, show a highly significant positive effect on wine prices. This effect increases with the price quantiles and accounts for a price increase of up to 61% for Auslese and up to 171.1% for Batbaice.

The Kabinett category shows no significant effect for the 0.25 quantile, but for all other quantiles there is an increasing, significantly negative effect on wine prices. For Spätlese, the effect is significantly positive for the 0.25 quantile, but significantly negative for the quantiles 0.5, 0.75, and 0.9.

(2) Machine Learning Approach

All available variables (see Table 1) can be used for the application of ANNs[17] without a negative impact on the model accuracy, since ANN models are

[16] Quantile regressions for models with producer fixed effects cannot be run, since these are underdetermined due to the large number of different producers (177).

[17] We use the "Orange: Data Mining Toolbox" version 3.13 for the model building and analyses process (Demsar et al., 2013). Orange is developed by the Bioinformatics Laboratory at the University of

Table 4
OLS vs. Quantile Regressions of Model 4 for Riesling

	ln price Model (4)	*ln price (Quantile 0.1)*	*ln price (Quantile 0.25)*	*ln price (Quantile 0.5)*	*ln price (Quantile 0.75)*	*ln price (Quantile 0.9)*
Gault&Millau points	0.131***	0.111***	0.123***	0.129***	0.132***	0.126***
	(73.81)	(35.78)	(63.80)	(69.82)	(44.30)	(34.82)
Age	0.0537*	0.0280	0.0495	0.0621**	0.0602	0.0339
	(2.05)	(0.93)	(1.56)	(2.61)	(1.59)	(1.19)
Trend	0.0181***	0.0246***	0.0239***	0.0171***	0.0155***	0.0137***
	(16.43)	(12.94)	(20.80)	(13.53)	(8.57)	(6.41)
Precipitation Growing Season (sum)[b]	0.000336***	0.000286*	0.000369***	0.000275*	0.000201*	0.000300*
	(4.48)	(2.34)	(5.42)	(2.57)	(2.07)	(2.43)
Precipitation Harvest Season (sum)[b]	0.000247***	0.000123	0.000158*	0.000132	0.000240**	0.000402***
	(3.39)	(1.18)	(2.32)	(1.53)	(2.58)	(4.00)
Air Temperature Growing Season (avg.)[b]	0.0232***	0.0126	0.0337***	0.0224***	0.0163	0.0131
	(3.54)	(1.78)	(4.10)	(4.43)	(1.82)	(1.20)
Air Temperature Harvest Season (avg.)[b]	0.00868*	0.00127	0.00862*	0.0101*	0.00830	0.0122*
	(2.05)	(0.26)	(2.01)	(2.37)	(1.76)	(1.97)
Kabinett	–0.178***	0.0211	–0.0961***	–0.173***	–0.264***	–0.367***
	(–18.05)	(1.53)	(–8.13)	(–15.359)	(–17.59)	(–15.40)
Spätlese	–0.0502***	0.104***	0.00846	–0.0541***	–0.105***	–0.171***
	(–5.22)	(6.38)	(0.57)	(–3.88)	(–6.83)	(–7.10)
Auslese	0.461***	0.460***	0.388***	0.403***	0.528***	0.610***
	(25.24)	(19.96)	(16.71)	(18.06)	(20.27)	(23.32)
Batbaice[a]	1.504***	1.474***	1.450***	1.520***	1.610***	1.711***
	(58.74)	(30.55)	(41.56)	(36.19)	(43.09)	(58.86)
Constant	–9.691***	–8.174***	–9.457***	–9.445***	–9.398***	–8.710***
	(–45.07)	(–24.19)	(–30.77)	(–43.11)	(–30.36)	(–21.58)

Continued

Table 4
Continued

	ln price *Model (4)*	*ln price* *(Quantile 0.1)*	*ln price* *(Quantile 0.25)*	*ln price* *(Quantile 0.5)*	*ln price* *(Quantile 0.75)*	*ln price* *(Quantile 0.9)*
Observations	8,335	8,335	8,335	8,335	8,335	8,335
R2	0.8020	0.4675	0.4872	0.5325	0.5888	0.6405

Note: All equations include a full set of regional fixed effects. Robust t statistics are in parentheses; significance levels are $*p < 0.05$, $**p < 0.01$, and $***p < 0.00$.

[a]Batbaice: Beerenauslese/Trockenbeerenauslese/Eiswein

[b]Winter Season: 12/01–02/28; Growing Season: 03/01–09/15; Harvest Season: 09/16–10/31

Source: Authors' calculations.

Table 5
Significant and Influential Variables for Riesling

Riesling (ANN)	*Dependency Factor*	*Average Semi-Elasticity*
Gault&Millau points	0.8941	0.0526
Alcohol	0.4703	0.0093
Trend	0.2285	0.0126
Air Temperature Growing Season (avg.)[a]	0.1886	0.0090
Minimum Air Temperature Growing Season(avg.)[a]	0.1409	0.0043
Minimum Air Temperature Harvest Season (avg.)[a]	0.2213	–0.0092
Maximum Air Temperature Winter Season (avg.)[a]	0.2874	–0.0170
Sunshine Hours Winter Season (sum)[a]	0.1751	0.0103
Humidity Winter Season (avg.)[a]	0.0790	0.0031
Humidity Growing Season (avg.)[a]	0.1002	–0.0054
Humidity Harvest Season (avg.)[a]	0.1845	0.0087
Observations	8,335	
R2	0.870	

[a]Winter Season: 12/01–02/28; Growing Season: 03/01–09/15; Harvest Season: 09/16–10/31

Source: Authors' calculations.

robust with regard to multicollinearity (Dumancas and A Bello, 2015; Garg and Tai, 2013).

First, we investigate a feasible architecture for the ANN model that matches the Riesling dataset. As a result of several experiments with different architectures and the intention to maximize the R^2, to minimize root mean squared error (RMSE) and to avoid overfitting, we found a four-layer architecture for the Riesling model. Thus, all 44 input variables, including quality fixed effects and regional fixed effects and, therefore, 44 nodes in the input-layer, 15 nodes in the first hidden layer and one node in the second hidden layer, and finally one output node for our dependent variable. In the next step, we train the Riesling model and calculate the dependency matrix and the semi-elasticity, which gives a sextuple for each independent variable, applying quality fixed effects.

We interpret a variable as important (significant) and influential if it has a high dependency value and a high semi-elasticity value.[18] Table 5 shows the significant and influential variables for Riesling.[19] We additionally include Humidity as an independent variable because it has rarely been used in the literature.

Ljubljana, Slovenia, in collaboration with an open source community. Orange is implemented in the Python programming language, with extra Python code developed by the authors to calculate the dependency and semi-elasticity required for proper analysis.

[18] Here, variables with a dependency factor > 0.5 are kept independent of the respective semi-elasticity values due to their importance for the whole model.

[19] Unless the variable is a fixed effects variable, as due to the definition of fixed effects, their semi-elasticity equals zero.

The ANN model is of higher explanatory power as it is possible to also capture non-linear relationships,[20] and the resulting R^2 is now 0.876. Furthermore, the RMSE with a value of 0.293 shows a better model accuracy (about 18.7% better) than the log linear regression model.

Again, we identify Gault&Millau points as the most important (nonfixed effect) variable with a dependency factor of 0.894, which also shows the highest semi-elasticity, with a price increase of 5.26% for each additional Gault&Millau point.

Like in the log linear regression model, there is a positive price trend, while the age of the wine does not seem to be important for the Riesling prices. The average (and also minimum) temperatures in the growing season again have a positive influence on wine prices, while precipitation does not seem to be relevant.

Some additional variables exert a price effect. The alcohol level has a positive effect on the price, while the rise in the maximum temperatures in the winter season and the minimum temperatures in the harvest season have negative prices effects, so that an increase in extreme weather conditions is, therefore, unfavorable for Riesling.

The ANN model allows the semi-elasticity to be split for each independent variable into a semi-elasticity per fixed effect (here the respective quality category), which improves the interpretation of the results. Table 6 shows the splitting into these semi-elasticities per quality category for Riesling.

The very left column of Table 6 shows the average coefficient (semi-elasticity), while the other 5 columns show the semi-elasticity per quality category. Here, the influence of additional Gault&Millau points on the price for QbA wines (+7.41%) is very high, but only small for wines of the Kabinett quality (+2.67%) and medium for Spätlese (+4.28%), Auslese (+5.85%), and Batbaice (+4.05%). Table 6 shows the positive influence of the alcohol level on the prices of QbA wines (+3.68%), Kabinett (+0.17%), and Spätlese (+0.63%), while the impact is negative regarding the prices of wines of the quality categories Auslese (–3.29%) and Batbaice (–3.02%). A rise in maximum temperatures during the winter season is not favorable for all quality categories of Riesling.

With the scatter plots in Figures 1 and 2, we further split the semi-elasticities and distinguish between the influence on low- and high-priced wines within a single quality category.

Figure 1 shows that wines with a lower QbA benefit from additional Gault&Millau points, while wines with a higher price do not. This is the opposite

[20] Therefore we do not apply squared functions. The same holds to producer fixed effects that would lead to an underdetermination of the equation system.

Table 6

Average Semi-Elasticity of Significant and Influential Variables for Riesling by Quality Category

	Average Semi-Elasticity	*Semi-Elasticity QbA*[b]	*Semi-Elasticity Kabinett*	*Semi-Elasticity Spätlese*	*Semi-Elasticity Auslese*	*Semi-Elasticity Batbaice*[c]
Gault&Millau points	0.0526	0.0741	0.0267	0.0428	0.0585	0.0405
Alcohol	0.0093	0.0368	0.0017	0.0063	–0.0329	–0.0302
Trend	0.0126	0.0169	0.0137	0.0100	0.0057	0.0102
Air Temperature Growing Season (avg.)[a]	0.0090	0.0051	0.0027	0.0079	0.0251	0.0238
Minimum Air Temperature Growing Season (avg.)[a]	0.0043	0.0061	0.0077	0.0060	–0.0023	–0.0121
Minimum Air Temperature Harvest Season (avg.)[a]	–0.0092	0.0012	–0.0090	–0.0137	–0.0282	–0.0123
Maximum Air Temperature Winter Season (avg.)[a]	–0.0170	–0.0156	–0.0132	–0.0179	–0.0232	–0.0199
Sunshine Hours Winter Season (sum)[a]	0.0103	0.0091	0.0081	0.0097	0.0167	0.0139
Humidity Winter Season (avg.)[a]	0.0031	0.0028	0.0015	0.0014	0.0092	0.0050
Humidity Growing Season (avg.)[a]	–0.0054	–0.0048	–0.0043	–0.0056	–0.0089	–0.0043
Humidity Harvest Season (avg.)[a]	0.0087	0.0061	0.0043	0.0086	0.0223	0.0111

[a]Winter Season: 12/01–02/28; Growing Season: 03/01–09/15; Harvest Season: 09/16–10/31

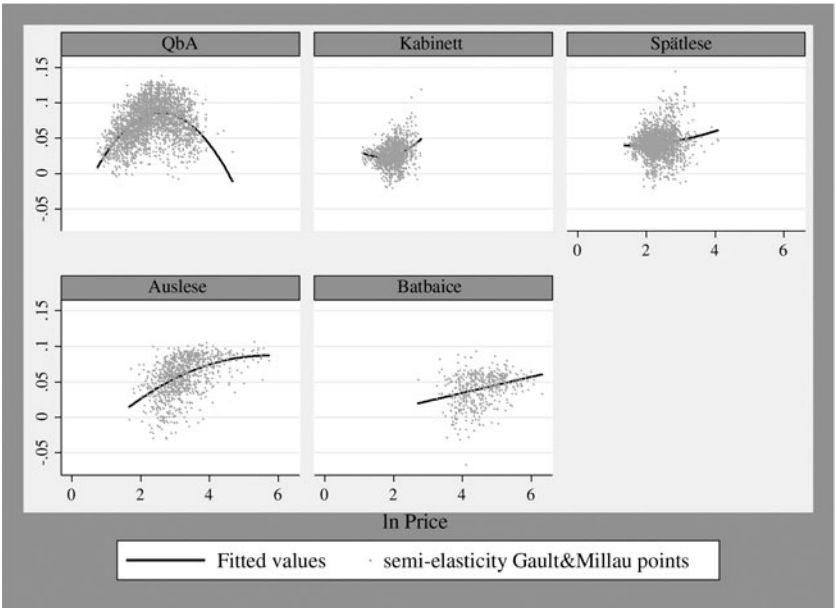
[b]Qualitätswein

[c]Batbaice: Beerenauslese/Trockenbeerenauslese/Eiswein

Source: Authors' calculations.

Figure 1

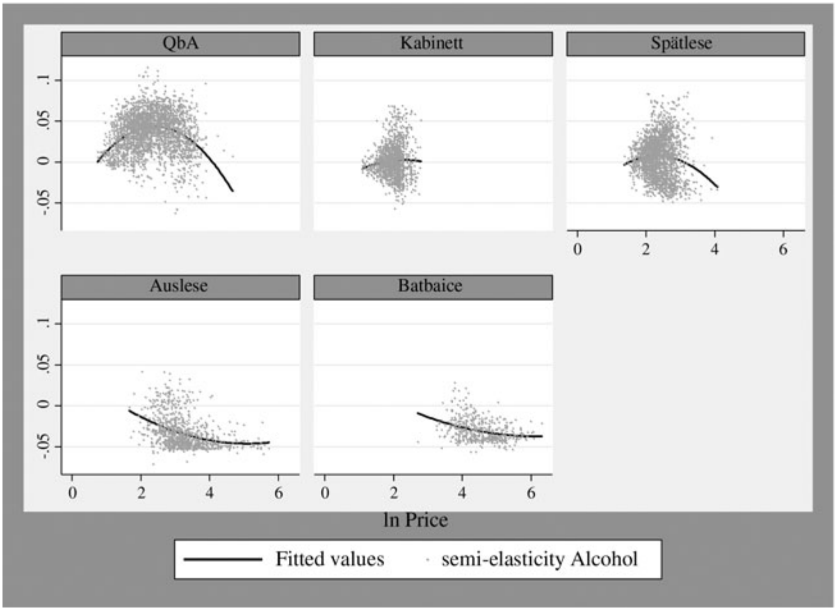
Average Semi-Elasticity of Gault&Millau Points for Each Riesling Quality Category



Source: Authors' calculations.

Figure 2

Average Semi-Elasticity of the Alcohol Level for Each Riesling Quality Category



Source: Authors' calculations.

https://doi.org/10.1017/jwe.2020.16 Published online by Cambridge University Press

for all other quality categories, since the higher the price, the more positive the price effect.

Figure 2 shows that the influence of the alcohol level is higher for QbA wines with lower prices than for QbA wines with higher prices, and that there is a slightly positive effect on higher-priced Kabinett wines. The scatterplot clearly confirms the negative influence of higher alcohol percentages on wine prices for the quality categories Spätlese, Auslese, and Batbaice.

B. ANN Regression Models for Silvaner, Pinot Blanc, and Pinot Noir

We also conduct the analysis for the Silvaner, Pinot Blanc, and Pinot Noir grape varieties, but as mentioned in Section III, we exclusively apply the machine learning approach due to its better performance. We again build an ANN and apply the same architecture as used for the Riesling model in order to compare the results.

Table 7 shows the dependency matrix and average semi-elasticities for the most important and influential variables for the other three grape varieties.[21]

Figure 3 (dependency factor) and Figure 4 (average semi-elasticity) visualize the differences between grape varieties regarding dependency factors and average semi-elasticities more easily for all grape varieties, including Riesling.

Regarding the dependency factors, Figure 3 clearly shows that the "Gault&Millau points" is the most important independent variable for the prices of all grape varieties and that the various weather variables have differing levels of importance for these prices with the variable "maximum air temperature in the winter season" being the most important one for Pinot Noir prices, and the variable "minimum air temperature in the growing season" for Pinot Blanc prices.

The same applies to the average semi-elasticities (Figure 4) with "Gault&Millau points" and on a lower level "Alcohol" as variables with a highly positive influence on the prices of all grape varieties, the variable "minimum air temperature in the growing season" having a highly positive effect on Pinot Blanc prices, the "air temperature in the winter season" having a highly negative effect on Silvaner prices, and "maximum air temperature in the winter season" on Pinot Noir prices.

We calculate the semi-elasticities per quality category of important (significant) and influential variables for Silvaner, Pinot Blanc, and Pinot Noir as we did for Riesling. The respective scatter plots show whether there is a difference between

[21] Numbers in bold, we only include non-bold numbers to allow for comparisons.

Table 7
Dependency Matrix and Average Semi-Elasticity for Silvaner, Pinot Blanc, and Pinot Noir

	Dependency Factor	*Silvaner Average Semi-Elasticity*	*Dependency Factor*	*Pinot Blanc Average Semi-Elasticity*	*Dependency Factor*	*Pinot Noir Average Semi-Elasticity*
Gault&Millau points	**0.347**	**0.0506**	**1.000**	**0.0533**	**1.000**	**0.0519**
Alcohol	**0.355**	**0.0369**	**0.638**	**0.0125**	**0.585**	**0.0225**
Trend	0.230	0.0386	**0.473**	**0.0184**	**0.505**	**–0.0022**
Precipitation Winter Season (sum)[a]	**0.237**	**0.0308**	**0.548**	**–0.0166**	0.386	–0.0045
Precipitation Growing Season (sum)[a]	0.121	–0.0009	**0.402**	**0.0102**	**0.403**	**–0.0003**
Precipitation Harvest Season (sum)[a]	0.148	–0.0051	0.375	0.0023	**0.423**	**0.0190**
Sunshine Hours Winter Season (sum)[a]	0.185	0.0266	0.358	–0.0140	**0.439**	**0.0176**
Air Temperature Winter Season (avg.)[a]	**0.262**	**–0.0531**	0.345	0.0070	**0.413**	**0.0182**
Air Temperature Growing Season (avg.)[a]	0.076	–0.0028	**0.457**	**–0.0273**	0.319	–0.0126
Minimum Air Temperature Growing Season (avg.)[a]	0.090	–0.0012	**0.708**	**0.0361**	0.182	0.0051
Minimum Air Temperature Harvest Season (avg.)[a]	0.177	–0.0156	**0.492**	**–0.0148**	**0.508**	**–0.0222**
Maximum Air Temperature Winter Season (avg.)[a]	0.097	–0.0055	0.359	–0.0210	**0.872**	**–0.0489**
Humidity Winter Season (avg.)[a]	***0.241***	***–0.0377***	***0.401***	***0.0006***	*0.378*	*0.0066*
Humidity Growing Season (avg.)[a]	*0.159*	*–0.0109*	*0.382*	*–0.0024*	***0.402***	***–0.0027***
Humidity Harvest Season (avg.)[a]	*0.149*	*–0.0015*	***0.404***	***–0.0223***	*0.362*	*0.0066*
Frost[b]	0.166	–0.0237	0.340	0.0074	**0.463**	**0.0137**
Observations	917		1,294		2,004	
R2	0.943		0.895		0.825	

Note: Significant and influential variables for each grape variety in bold letters.

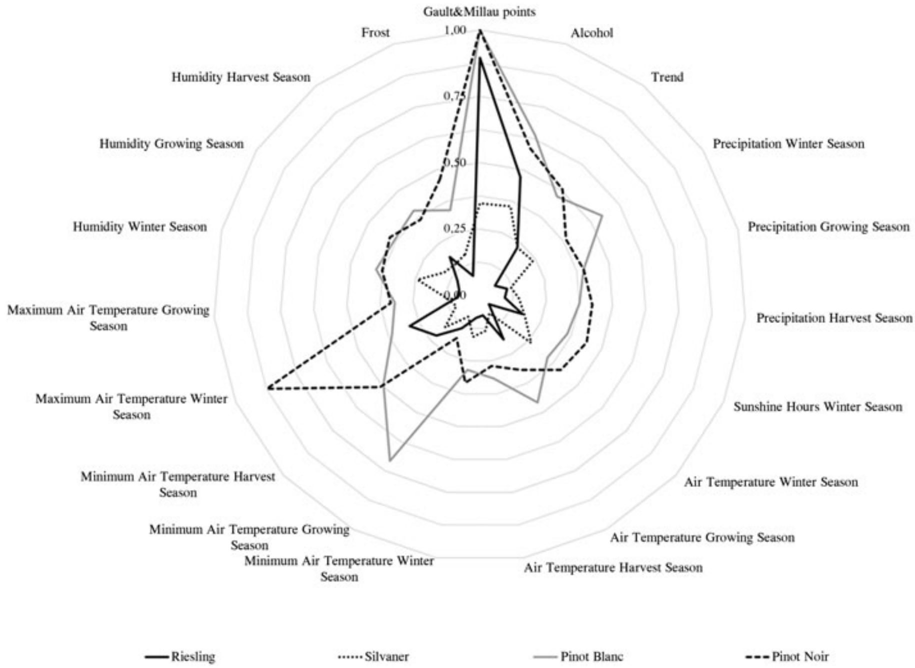
[a]Winter Season: 12/01–02/28; Growing Season: 03/01–09/15; Harvest Season: 09/15–10/31

[b]Frost: Sum of days of frost, with soil temperatures < 0 during winter, growing, and harvest seasons.

Source: Authors' calculations.

Figure 3

Dependency Factors for Riesling, Silvaner, Pinot Blanc, and Pinot Noir



Source: Authors' calculations.

lower- and higher-priced wines in each quality category for each independent variable.

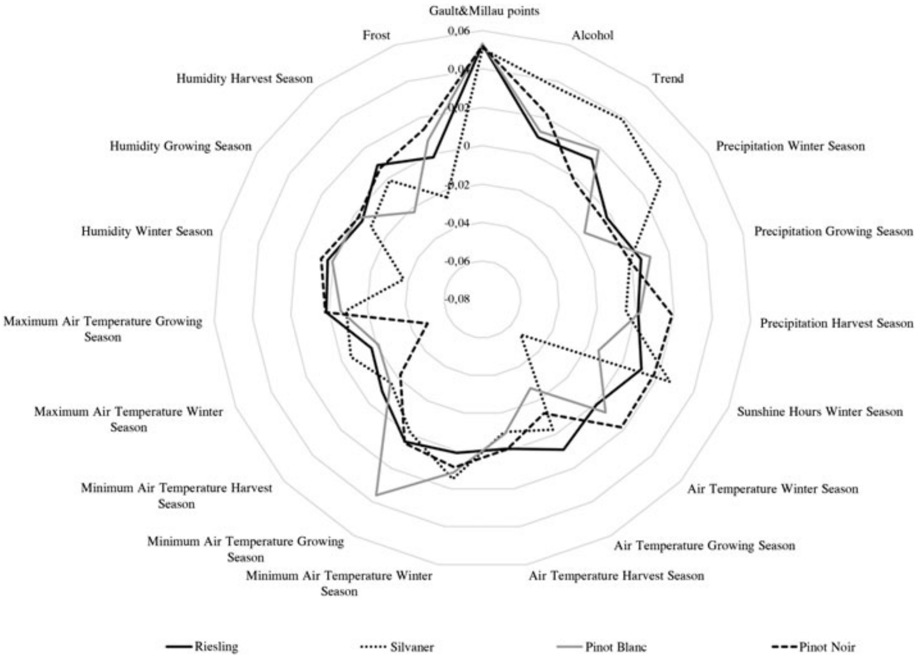
These results are only addressed in the online publication.

V. Conclusion

The results suggest that a simple hedonic price equation does not exist for all grape varieties. It is more appropriate to use different price equations for each grape variety. The log linear regression model for Riesling finds positive effects especially of Gault&Millau points, but also of age, trend, and average temperatures as well as precipitation on German wine prices. The quality categories Auslese and Batbaice lead to very high-price premiums for Riesling. The non-linear regression using ANNs performs slightly better than the log linear regression model because it delivers better results with respect to R^2 and RMSE, suggests some additional explanatory variables (such as alcohol level or minimum and maximum temperatures), and gives a more detailed insight into the interpretation of explanatory variables.

Figure 4

Average Semi-Elasticity for Riesling, Silvaner, Pinot Blanc, and Pinot Noir



Source: Authors' calculations.

Therefore, the results for Silvaner, Pinot Blanc, and Pinot Noir are based on the machine learning approach. It is shown that the influence of an independent variable on the wine price of a certain grape variety cannot be estimated by a single coefficient, since the influence clearly differs for each quality category (fixed effect) of the respective wine.

The results of the machine learning model also suggest that Gault&Millau points have a significant and high influence on German wine prices, as one additional point increases wine prices by around 5% regarding the average coefficients, but with differences between quality categories. The split analysis shows that the lowest German quality category QbA has the highest premiums, while the highest quality category Batbaice only has high premiums for Pinot Noir. Additional scatter plots show that the higher-priced QbA wines benefit more from additional Gault&Millau points than the lower-priced QbA wines.

The influence of the alcohol level on wine prices is positive with regard to the average coefficient, but the split analysis shows that this influence holds especially to the quality categories QbA, Kabinett, and Spätlese and that the influence is even negative for the quality categories Auslese (except for Pinot Noir) and Batbaice (except for Pinot Blanc).

There are essential differences between grape varieties regarding influential weather variables and their ability to cope with rising temperatures or extreme weather (minimum and maximum temperatures) in different seasons of the ripening process.

Rising average air temperatures during the winter season lead to a decrease in prices for Riesling and Silvaner, while Pinot Blanc copes better with rising winter temperatures and Pinot Noir, the only red variety in the sample, even shows a positive effect of both the rising minimum and average temperatures, but all varieties cannot cope with rising maximum temperatures in the winter season.

During the harvest season, especially higher minimum and maximum temperatures lead to a negative price effect for wines of all grape varieties and quality categories, so that earlier harvest is recommended for all grape varieties in cases of warmer harvest seasons, such as was the case in 2018.

Precipitation does not seem to have as high an impact but is different for each grape variety and quality category.

In the analysis, ripeness levels are used as proxies for quality levels according to the German classification system. Future research may include investigating the effects of climate change on wine prices from regions with distinct quality levels, such as Bordeaux.

Supplementary Material

To view supplementary material for this article, please visit https://doi.org/10.1017/jwe.2020.16.

References

Ashenfelter, O. (2010). Predicting the quality and prices of Bordeaux wines. *Journal of Wine Economics*, 5(1), 40–52.

Ashenfelter, O., and Storchmann K. (2010). Measuring the economic effect of global warming on viticulture using auction, retail and wholesale prices. *Review of Industrial Organization*, 37(1), 51–64.

Ashenfelter, O., and Storchmann K. (2016). Climate change and wine: A review of the economic implications. *Journal of Wine Economics*, 11(1), 105–138.

Byron, R. P., and Ashenfelter, O. (1995). Predicting the quality of an unborn Grange. *Economic Record*, 71(212), 400–414.

Demsar, J., Curk, T., Erjavec, A., Gorup, C., Hocevar, T., Milutinovic M., Mozina, M., Polajnar, M., Toplak, M., Staric, A., Stajdohar, M., Umek, L., Zagar, L., Zbontar, J., Zitnik, M., and Zupan, B. (2013). Orange: Data mining toolbox in Python. *Journal of Machine Learning Research*, 14(Aug.), 2349–2353.

Diel, A., and Payne, J. (eds.). (2002–2009). *Gault&Millau Weinguide Deutschland*. München: Christian Verlag.

Dumancas, G., and A Bello, G. (2015). Comparison of machine-learning techniques for handling multicollinearity in big data analytics and high performance data mining. *Supercomputing* 2015: The International Conference for High Performance Computing, Networking, Storage, and Analysis (Austin, TX). doi: 10.13140/RG.2.1.1579.4641.

Fecke, B. (2014). Klimawandel stellt Winzer vor neue Probleme. *Deutschlandfunk*, 03.11.2014. Available from: http://www.deutschlandfunk.de/weinanbau-klimawandel-stelltwinzer-vor-neue-probleme.697.de.html?dram:article_id=302145 (accessed on November 8, 2017).

Garg, A., and Tai, K. (2013). Comparison of statistical and machine learning methods in modelling of data with multicollinearity. *Inter. J. Modelling, Identification and Control*, 18(4), 295–312.

Garson, G. D. (1991). Interpreting neural network connection weights. *Artificial Intelligence Expert*, 6(4), 46–51.

Giam, X., and Olden, J. D. (2015). A new R^2-based metric to shed greater insight on variable importance in artificial neural networks. *Elsevier Ecological Modelling* 313, 307–313.

Goh, A. T. C. (1995). Back-propagation neural networks for modelling complex systems. *Artificial Intelligence in Engineering*, 9(3),143–151.

Greene, W. H. (2008). *Econometric Analysis*. 6th ed. Upper Saddle River, NJ: Pearson.

Haeger, J. W., and Storchmann, K. (2006). Prices of American pinot noir wines: Climate, craftsmanship, critics. *Agricultural Economics*, 35(1), 67–78.

Hashem, S. (1992). Sensitivity analysis for feedforward artificial neural networks with differentiable activation functions. *Proceedings of the International Joint Conference on Neural Networks* (Baltimore, MD), 419–424. New York: IEEE Press.

Hornik, K., Stichcombe, M., and White, H. (1990). Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 3, 551–560.

Jones, G. V., and Storchmann, K. (2001). Wine market prices and investment under uncertainty. An econometric model for Bordeaux Crus Classés. *Agricultural Economics*, 26(2), 115–133.

Kriener, M., and Mortsiefer, H. (2017). Weinproduktion sinkt auf 50-Jahres-Tief. Der Tagesspiegel, 2017-10-24. Available from https://www.tagesspiegel.de/wirtschaft/ernte-2017-weinproduktion-sinkt-auf-50-jahres-tief/20497304.html.

Lecocq, S., and Visser, M. (2006). Spatial variations in weather conditions and wines prices in Bordeaux. *Journal of Wine Economics*, 1(2), 114–124.

National Science and Technology Council, Committee on Technology (2016). *Preparing for the Future of Artificial Intelligence*. Available from: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

Niklas, B. (2017). Impact of annual weather fluctuations on wine production in Germany. *Journal of Wine Economics*, 12(4), 436–445.

Oczkowski, E. (2001). Hedonic wine price functions and measurement error. *Economic Record*, 77(239), 374–382.

Oczkowski, E. (2014). Wine prices and quality ratings: A meta regression analysis. *American Journal of Agricultural Economics*, 97(1), 103–121.

Oczkowski, E. (2016). The effect of weather on wine quality and prices: An Australian spatial analysis. *Journal of Wine Economics*, 11(1), 48–65.

Olden, J. D., and Jackson, D. A. (2002). Illuminating the “black-box”: A randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, 154, 135–150.

Olden, J. D., Joy, M. K., and Death, R. G. (2004). An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling*, 178, 389–397.

Owen, G. W. (2012). Applying point elasticity of demand principles to optimal pricing in management accounting. *International Journal of Applied Economics and Finance*, 6, 89–99.

Payne, J. (eds.). (2010–2015). *Gault&Millau Weinguide Deutschland*. München: Christian Verlag.

Ramirez, C. D. (2008). Wine quality, wine prices and the weather: Is Napa "different"? *Journal of Wine Economics*, 3(2), 114–131.

Rinke, W. (2015). Calculating the dependency of components of observable nonlinear systems using artificial neural networks. *MakeLearn & TIIM Conference Proceedings*, 367–374. Available from https://EconPapers.repec.org/RePEc:tkp:mklp15:367-374.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408.

Rudin, W. (1976). *Principles of Mathematical Analysis*, 3rd edition. New York: McGraw-Hill.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533–536.

Schamel, G. (2000). Individual and collective reputation indicators of wine quality. Centre for International Economic Studies, Discussion Paper No. 0009, March. Available from https://www.researchgate.net/publication/228286730_Individual_and_Collective_Repuatation_Indicators_of_Wine_Quality.

Schamel, G. (2002). California wine winners: A hedonic analysis of regional and winery reputation indicators. Paper presented at the AAEA Meeting, July 28–31, Long Beach, CA.

Schamel, G. (2003). A hedonic pricing model for German wine. *Agrarwirtschaft*, 52(5), 247–254.

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

Shavlik, J. W., and Diettrich, T. G. (eds.). (1990). *Reading in Machine Learning*. San Mateo, CA: Morgan Kaufmann.

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M., and Teller, A. (2016). Artificial intelligence and life in 2030. One hundred year study on artificial intelligence: Report of the 2015-2016 Study Panel, 8. Available from https://ai100.stanford.edu/sites/default/files/ai100report10032016fnl_singles.pdf.

Storchmann, K. (2005). English weather and Rhine wine quality: An ordered probit approach. *Journal of Wine Research*, 16(2), 105–119.

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

Thrane, C. (2004). In defence of the price hedonic model in wine research. *Journal of Wine Research*, 15(2), 123–134.

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

Witten, I., Eibe, F., and Hall, M. (2017). *Data Mining – Practical Machine Learning Tools and Techniques*, 4th edition, 261–269. New York: Morgan Kaufmann.

Yeh, I.-C., and Cheng, W.-L. (2010). First and second order sensitivity analysis of MLP. *Neurocomputing*, 73, 2225–2233.

Yeo, M., Fletcher, T., and Shawe-Taylor, J. (2015). Machine learning in fine wine price prediction. *Journal of Wine Economics*, 10(2), 151–172.

Appendix

Table A1

Correlation Matrix Between Gault&Millau Points and German Quality Categories

Correlation Coefficients	*QbA*[a]	*Kabinett*	*Spätlese*	*Auslese*	*Batbaice*[b]
G&M[c] points Riesling	–0.217	–0.262	0.020	0.324	0.379
G&M[c] points Silvaner	–0.133	–0.245	0.070	0.151	0.483
G&M[c] points Pinot Blanc	–0.164	–0.237	0.271	0.155	0.146
G&M[c] points Pinot Noir	0.011	–0.148	–0.046	0.082	0.107

[a]QbA: Qualitätswein

[b]Batbaice: Beerenauslese/Trockenbeerenauslese/Eiswein

[c]Gault&Millau

Table A2

Breusch–Pagan/Cook–Weisberg Test for Heteroskedasticity – Model 1, 2, and 3
Breusch–Pagan/Cook–Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fittes values of logprice075

Model 1	*Model 2*	*Model 3*
Chi2 (1) = 594.62 Prob > chi2 = 0.0000	Chi2 (1) = 608.75 Prob > chi2 = 0.0000	Chi2 (1) = 262.28 Prob > chi2 = 0.0000

Table A3
Variance Inflation Tests for Multicollinearity for Model 1, 2, and 3

Variable	*VIF Model 1*	*VIF Model 2*	*VIF Model 3*
Gault&Millau points	1.53	1.53	
Age	1.01	1.02	1.01
Trend	1.21	1.24	1.17
Precipitation Growing Season (sum)	1.49	1.56	1.49
Precipitation Harvest Season (sum)	1.33	1.35	1.32
Air Temperature Growing Season (avg.)	1.35	642.44	1.35
Air Temperature Growing Season (avg.)[b]		640.46	
Air Temperature Harvest Season (avg.)	1.08	170.15	1.08
Air Temperature Harvest Season (avg.)[b]		167.02	
Kabinett	1.28	1.28	1.27
Spätlese	1.39	1.40	1.34
Auslese	1.47	1.47	1.22
Batbaice[a]	1.44	1.44	1.13
Mean VIF	1.33	125.57	1.24

[a]Batbaice: Beerenauslese/Trockenbeerenauslese/Eiswein

[b]Growing Season: 03/01–09/15; Harvest Season: 09/16–10/31

Table A4

Correlation Matrix for Model 1

	Gault&Millau Points	*Age*	*Trend*	*Air Temp. Growing Season*	*Air Temp. Harvest Season*	*Precip. Growing Season*	*Precip. Harvest Season*	*QbA*[a]	*Kabinett*	*Spätlese*	*Auslese*	*Batbaice*[a]
Gault&Millau points	—											
Age	0.0749	—										
Trend	0.0885	0.0109	—									
Air Temp. Growing Season[c]	0.0003	0.0177	–0.2072	—								
Air Temp. Harvest Season[c]	–0.0011	–0.0024	–0.0875	0.2574	—							
Precipitation Growing Season[c]	–0.0375	0.0006	0.0535	–0.4285	–0.1227	—						
Precipitation Harvest Season[c]	–0.0656	–0.0515	–0.1819	–0.1740	–0.0651	0.4459	—					
QbA[a]	–0.2173	0.0059	0.2273	0.0110	–0.0160	0.0034	–0.0387	—				
Kabinett	–0.2617	–0.0527	–0.0477	–0.0435	–0.0246	0.0144	0.0298	–0.3489	—			
Spätlese	0.0205	–0.0192	–0.0981	–0.0171	–0.0237	–0.0068	0.0124	–0.4658	–0.2976	—		
Auslese	0.3242	0.0268	–0.0858	0.0360	0.0714	–0.0030	0.0132	–0.2724	–0.1740	–0.2323	—	
Batbaice[b]	0.3792	0.0734	–0.0760	0.0316	0.0194	–0.0132	–0.0121	–0.1877	–0.1199	–0.1601	–0.0936	—

[a]QbA: Qualitätswein

[b]Batbaice: Beerenauslese/Trockenbeerenauslese/Eiswein

[c]Growing Season: 03/01–09/15; Harvest Season: 09/16–10/31

Source: Authors' calculations.