Wine Analytics

Burak Kazaz

bkazaz@syr.edu

Whitman School of Management

Syracuse University

Syracuse, NY 13244

May 21, 2020

This manuscript has been prepared in response to an invitation to contribute to the INFORMS TutORial in Operations Research.

Please cite this article as follows:

Kazaz, B. Wine Analytics. To appear in C. Druehl and W. Elmaghraby, eds. *Pushing the Boundaries: Frontiers in Impactful OR/OM Research, TutORials in Operations Research*, INFORMS, 2020.

Wine Analytics

Among agricultural goods, wine is a specialty product. Fine wine grapes require exceptional care and attention. After the harvest, grapes are crushed and the wine goes through a long aging process. In the case of Bordeaux-style wines, for example, the aging process in barrels can last 18 to 24 months. The aging continues for another 20 to 30 years in the bottle. This long aging process makes wine production a risky venture. Consumers follow the evolution of these fine wines closely, track their corresponding tasting reviews and scores, and are often informed about climatic conditions during the growing seasons. Thus, wine is one of the most heavily tracked and publicized agricultural products. Considering the long production time, winemakers can mitigate their operational and financial risks by selling their wines in advance in the form of wine futures.

In this tutorial, we describe predictive and prescriptive analytical methods that assist primary enterprises that produce and distribute wine in their decision-making processes. The tutorial begins with predictive models that estimate the true value of wine futures prices. These estimation models are essential to the financial exchange known as the London International Vintners Exchange (Liv-ex) where wine futures contracts are traded. Coined as “realistic prices” by Liv-ex, these predictive models assist buyers in their purchasing decisions as they can determine whether a futures contract is underpriced or overpriced. The tutorial then develops risk mitigation models to assist winemakers in mitigating uncertainty in weather conditions and tasting expert reviews. These prescriptive models rely on predictive analytics which help determine consumers’ utilities from buying the wine in advance, or later, or not purchasing it at all. Prescriptive models such as a multinomial logit model focus on determining how much of the wine should be sold in advance in order to reduce risk exposure and maximize the expected profits of the winemaker. On the buyer side, the tutorial introduces stochastic portfolio optimization models for both wine distributors and importers in their decision regarding how to allocate limited budgets between wine futures contracts and bottled wine. These prescriptive models are, once again, built on predictive analytics that estimate the evolution of futures and bottle prices over time under fluctuating market and weather conditions.

Wine is an exemplary agricultural product; its production and quality perceptions are widely tracked by businesses and consumer. The predictive and prescriptive models of this tutorial help create transparency in this largely opaque market. They assist the industry in its drive towards market efficiency. The tutorial also offers future research directions in wine analytics and describes how these techniques can be beneficial for the production and distribution of other agricultural products.

Keywords: *wine analytics, wine futures, price, weather, barrel score, Liv-ex 100, wine economics*

1. Introduction

Wine is an important agricultural product with a growing global interest. The global wine market is estimated to be greater than $385 billion in 2020 and is expected to grow annually by 3.7%. The US wine industry generates $52 billion annually. Italy produces the largest amount of wine, but France is the most celebrated wine producing country with the greatest amount of exports. The Bordeaux region of France, for example, produces the most sought-after wines in the world. Bordeaux wine prices often influence the prices of wines globally. Many countries try to emulate the Bordeaux-style winemaking by growing

Bordeaux varietal grapes. Thus, Bordeaux wines are perceived as the best wines and they set the pace in global wine supply chains.

Wine production departs from other agricultural products for several reasons. First, the time it takes to produce wine is substantially longer than most agricultural commodities. Second, consumers follow the reviews of well-known tasting experts and their publications wholeheartedly. While grape production also encounters risks stemming from climatic conditions, diseases (e.g., botrytis) and economic fluctuations, it has the added risks of long production times and potentially devastating reviews. Third, wine is an alternative investment mechanism. Financial institutions and investment banks follow the trends in the industry and provide reports that offer recommendations for industry participants and their clients. McMillon (2020), for example, offers a report from Silicon Valley Bank regarding the state of the wine industry in the US. The report demonstrates the need for rigorous analytical methods that will support the interests of the wine enterprises as well as the investors.

We focus on the risk mitigation aspects of analytics that help decision makers in the wine supply chain. Specifically, we focus on one element that combines the tools known in marketing as "advance selling," as "inventory financing" in supply chain finance and as "wine futures" in the wine industry. Because wine takes a long time to age, wine futures enable winemakers to sell the wine in the form of futures contracts. We illustrate in Section 2 through predictive analytics models how wine futures can be assessed, evaluated and estimated. In Section 3, we show how winemakers can determine the price and amount of wine futures to be sold in order to mitigate the risks stemming from uncertain tasting reviews and retail prices. Section 4 utilizes these instruments in designing a buyer's (e.g., wine distributor and/or importer) purchasing portfolio under limited budgets and varying risk profiles. Section 5 offers prescriptions regarding future research directions.

What is *wine analytics*? Wine analytics refers to employing analytical methodologies, be it predictive or prescriptive, to solve problems in the wine industry. One might wonder whether wine analytics constitutes an area of study, especially in the presence of a field called wine economics. It should be emphasized that the wine economics literature focusses on "explaining" the impact of factors such as climatic conditions and tasting reviews in wine problems. Wine analytics, on the other hand, develops methods that "predict" the evolution of agricultural, climatic, economic conditions and market conditions and build "prescriptive models" in order to guide executives and managers of wine businesses.

1.1. Literature Review

Wine draws attention from various fields including agricultural economics, marketing, tourism, hospitality and operations management. The emergence of wine economics as a field of study has led many scholars in economics to examine various problems. We provide some background on these areas.

Wine Economics. The research question associated with the pricing of wine has attracted numerous scholars to investigate the factors that influence the value of these agricultural goods. The wine economics literature dedicates econometric models in order to "explain" wine prices using weather conditions. However, there is a limited amount of work that "predict" the value of these goods. The two publications known for predictive analytics are Ashenfelter (2008) and Hekimoğlu and Kazaz (2020). The wine economics literature examines the pricing of aged wine using weather and expert opinions, however, these studies fail to estimate young wine prices, i.e., the prices that initiate the trade when the wine is in its early stages of aging. Young wine prices are important because leading Bordeaux wines are primarily sold before they are bottled; they are traded in the form of wine futures contracts. Ashenfelter et al. (1995) and Ashenfelter (2008) show that mature Bordeaux wine prices can be explained using weather and age. These two publications report significantly high errors for young wines. Hekimoğlu and Kazaz (2020) uses Ashenfelter (2008) as the benchmark study in order to demonstrate the effectiveness of their predictive models. Table 1.1 provides a comprehensive list of publications that explain wine prices using either weather information or tasting experts' reviews. Bazen and Cardebat (2018) build statistical models in order to predict the prices of generic wines in Bordeaux; these wines are often sold in bulk and are not traded in the form of wine futures. Gergaud et al. (2017) evaluate the collective economic benefits rather than the price. Wine analytics literature complements this wine economics literature by building predictive pricing models that provide transparency to the constituents of the wine supply chains.

Publication / Factors examined	Temperature	Rainfall	Tasting score	Liv-ex 100 index	Consecutive vintage comparison
Ashenfelter et al. (1995)	+	+			
Byron and Ashenfelter (1995)	+	+			
Combris et al. (1997)			+		
Jones and Storchmann (2001)	+	+	+		
Cardebat and Figuet (2004)			+		
Haeger and Storchmann (2006)	+	+	+		
Lecocq and Visser (2006)	+	+			
Wood and Anderson (2006)	+	+			
Ali and Nauges (2007)			+		
Ali et al. (2008)			+		
Ashenfelter (2008)	+	+			
Ashenfelter and Storchmann (2010)	+	+			
Dubois and Nauges (2010)			+		
Ashenfelter and Jones (2013)	+	+	+		
Dimson et al. (2015)	+	+			
Ashton (2016)			+		
Cardebat et al. (2017)			+		
Hekimoğlu and Kazaz (2020)	+	+	+	+	+

Table 1.1. The list of factors used in publications that examine wine prices.

The wine economics literature pays close attention to the influence of tasting expert opinions. There are publications (e.g., *Wine Advocate*, *Wine Spectator*, *Decanter*) that give tasting reviews and scores.

Avid consumers follow these numbers closely as they create a quality perception. Wine prices are influenced by the tasting scores. Quality perceptions are also created for other agricultural products. The quality of an olive oil, for example, is determined by an oleic acidity test (Ayvaz-Çavdaroğlu et al. 2020); thus, it is much less subjective than a tasting expert's score. There are futures for olive oil in global markets; however, the oils that get traded are almost always poorer in quality. The predictive and prescriptive analytical methods developed for wine can also be used in other agricultural products. In the case of an olive oil, the tasting score would be replaced by an oleic acidity level (a numerical value between 0.3% and 2%).

Fluctuations in economic conditions also influence wine prices. Traditional financial indices, e.g., The Standard & Poor's 500 index and the Dow Jones index, do not explain the movements in wine prices. Rather, the industry needs to follow new and different indices. London International Vintners' Exchange (Liv-ex) is the financial exchange where fine wines and their futures contracts are traded. Liv-ex has established the Liv-ex 100 index which constitutes the prices of the most sought-after 100 wines in the world, most of which are from Bordeaux. This index is used to describe the financial health and sustainability of the global wine industry; it is widely cited by Bloomberg and Reuters to inform investors. Hekimoğlu et al. (2017) show that the Liv-ex 100 index can be used to examine the influence of market fluctuations. Cardebat and Jiao (2018) conclude that the Liv-ex 100 serves as a proxy reflecting the consumers' willingness to pay for fines wines. Hekimoğlu and Kazaz (2020) show how the Liv-ex 100 index can be used in a predictive analytical model for Bordeaux wine futures. However, the literature in the creation of wine indices is sparse. Masset and Weisskopf (2018b) argue that indices that are developed for the wine industry are not as efficient as other commodity markets.

Finance. Wine is an alternative investment asset. Storchmann (2012) provides a comprehensive review about wine economics literature focusing on wine as an investment option. Dimson et al. (2015) find that young Bordeaux wines yield greater returns than mature ones. This finding further amplifies the importance of explaining the evolution of young wine prices; Hekimoğlu et al. (2017) provide the predictive models to develop functional forms of this price evolution. Jaeger (1981), Burton and Jacobsen (2001), Masset and Henderson (2018), Masset and Weisskopf (2016) and Masset et al. (2016) conclude that wine can be used as a long-term investment. Jaeger (1981), Burton and Jacobsen (2001), and Dimson et al. (2015) show that wines can yield greater returns than treasury bills, but less than equities. Masset and Weisskopf (2018a) demonstrate that fine wines often yield higher returns than equities especially when there is a financial crisis and when financial assets are highly correlated.

Marketing. This literature focuses on consumers' preferences, their willingness to pay functions, and their experiences in tasting rooms. Schmit et al. (2013) use sensory effects to explore customer valuation of environmentally friendly wines. Kelley et al. (2017) present a conjoint analysis to examine consumers' preferences for wines. Perla (2014) examines consumer preferences and sensory effects and provides guidance in the marketing of wine in restaurants. Kelley et al. (2020) extend this work to marketing in tasting rooms. Park et al. (2018) investigate what makes consumers revisit a winery. Back et al. (2019) relates these marketing preferences to margins and markups in fair trade wine supply chains.

Operations-Marketing Interface. This literature builds on the advance selling mechanisms originally developed in marketing. Xie and Shugan (2001), Boyacı and Özer (2010), Cho and Tang (2013), Tang and Lim (2013) and Yu et al. (2015a, 2015b) examine advance-selling quantity decisions under uncertainty and risk. Noparumpa et al. (2015) develop a prescriptive model for a winemaker (seller) to determine the proportion of wine to be sold in the form of wine futures – the remaining proportion is distributed after the wine is bottled. Hekimoğlu et al. (2017) develop a stochastic program to solve a wine distributor's (buyer) purchasing decision between wine futures and bottled wine.

Operations literature examines problems for other agricultural products in the context of uncertainty in climatic conditions and crop supply. Examples include Jones et al. (2001), Kazaz (2004), Kazaz and Webster (2011, 2015), Boyabatli et al. (2011), Boyabatli (2015) and Noparumpa et al. (2020).

In the remaining part of this tutorial, we build predictive and prescriptive models that estimate and employ wine futures. They can create a less risky operating environment for winemakers, offer transparency to the buyers, and build confidence in the efficiency of the wine market.

2. Predictive Analytics for Wine Futures

Wine futures is a term that refers to selling wine in advance of bottling. It is a mechanism to help winemakers mitigate the uncertainty in their future cash flows. By selling early, a winemaker can recuperate her cash investment in the wine that will be sitting in the barrel for an additional year before it is bottled. The wine may not even sell immediately after bottling and it may sit on retail shelves without earning money for some time. Selling wine in the form of wine futures has been practiced by French winemakers since the 17th century. Winemakers are often willing to sell their wine in the form of wine futures at a discounted price. These reduced prices make it attractive for buyers such as wine distributors and importers. In this section of the tutorial, we begin our discussion by describing the trade mechanisms that serve as the foundation for the most celebrated wines, those of the Bordeaux region. Then, we develop predictive analytics models that focus on the most critical and challenging question in the wine industry: What is the right price for these wine futures contracts?

Ashenfelter et al. (1995) and Ashenfelter (2008) use level panel data and provide successful regression results in estimating the impact of weather on aged Bordeaux wines. However, their models do not estimate young wine prices accurately. We present new predictive analytics approaches that lead to accurate estimations for young wine prices. We focus on the wines that are still aging in the barrel but traded in the form of France's *en primeur* system which is loosely translated into English as wine futures.

Let us begin our discussion with the winemaking process in the Bordeaux wine supply chain. Figure 2.1 presents a timeline of events in the winemaking process. Chateaus grow their grapes on their estate from May to August each summer. Grapes are harvested and pressed in early fall. The wine is referred to as vintage t and it begins its aging process in barrels at this time. In the spring of calendar year $t + 1$, tasting experts visit chateaus, taste samples, and establish what is known as "the barrel score." These barrel scores are often described out of 100 points and they create a perception of quality for buyers and consumers. In the French system, following the release of barrel scores, chateaus begin to sell their wines to middleman known as négociants. Négociants purchase the wine from chateaus at a price what the market refers to as the *ex-chateau* price. The *ex-chateau* prices are private information and are not known publicly. At the beginning of the summer of year $t + 1$, négociants sell the wine immediately in terms of futures contracts through a financial exchange – this price is referred to as the *ex-négociant* price and is known as the futures price. A vast majority of Bordeaux wines, more than 80% and often 100%, are traded in the form of wine futures in the summer of year $t + 1$.

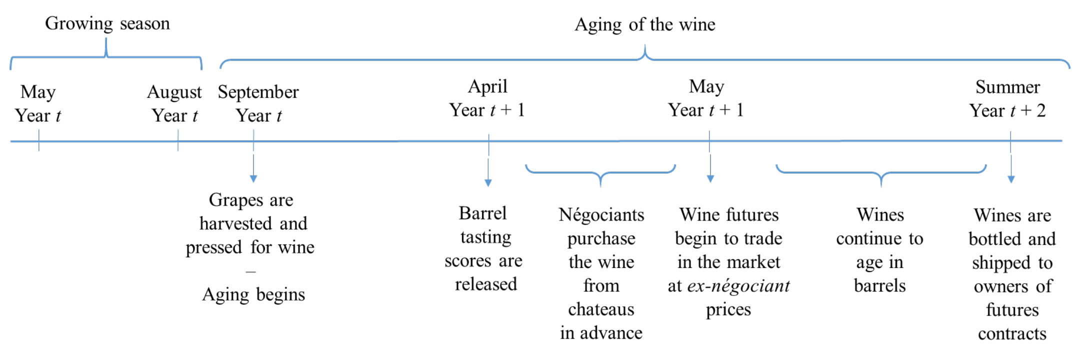

Figure 2.1. The sequence of events leading to the revelation of *ex-négociant* prices for vintage t wines.

It is important to highlight that the *ex-négociant* price is the "market price" for wine futures. The predictive analytics model presented in this section aims to estimate this market price; the estimate represents the true value of the wine futures price known as the *ex-négociant* price in Liv-ex. This price is important for the entire supply chain because it sets the pace in the downstream and influences the trade significantly. After being sold in the form of wine futures, the wine continues to age in barrels for another year. It gets bottled during the summer of year $t + 2$. As a result, the total aging process is 18 to 24 months. In the absence of wine futures, this is a long period of time to have cash tied up in this continually evolving good. After bottling, the wine gets delivered to the buyers of futures contracts – the shipment is made from the négociants to the buyers who are often distributors and importers.

Estimating the true value of the market price is a challenging task and even industry experts fail to provide meaningful estimates. A 2016 Liv-ex survey involving 440 of the world's leading wine merchants shows the estimation errors of industry experts regarding the futures prices of the 2015 vintage: https://www.liv-ex.com/2016/06/merchants-underestimated-bordeaux-2015-release-prices/. The survey relied on a basket of futures contracts from Bordeaux's premium wine producers Cheval Blanc, Cos d'Estournel, Leoville Las Cases, Mission Haut Brion, Montrose, Mouton Rothschild, Pavie, Pichon Lalande, Pontet Canet and Talbot. The 2016 survey results show that these 440 leading wine merchants under-estimated the total value of the basket by 21%. These experts estimated a 17.8% increase in the futures prices for the 2015 vintage from the futures prices of the same wines' 2014 vintage. The actual prices, however, increased by 45.8%. This is a great example of the need for a predictive analytics study intended to estimate wine futures prices accurately.

A reliable futures price estimate is essential for the wine industry. In the absence of good estimations, buyers would have the inertia to invest in wine futures. Conversations with the executives of the largest wine distributors in the US confirm this concern. These executives point to the necessity for establishing a benchmark price so that they can determine in confidence whether a wine is overpriced or underpriced, and thus, whether they should invest a smaller or large amount of money into futures contracts. One might wonder why we cannot decide this value by using just the tasting scores. We must remember that barrel scores established by the tasting experts do not immediately translate to prices. The predictive analytics approach in Hekimoğlu and Kazaz (2020) fills the necessary gap and it establishes the much-needed benchmark prices. Reliable estimations from predictive analytics bring transparency into this otherwise highly opaque market.

The adjustments in futures prices from the previous vintage exhibit a highly similar behavior in Figure 2.2. The reliance on the prior vintage's futures prices appear to be an operational planning phenomenon. Chateaus replace barrels each year, bottle the wine and push it downstream, and take in the new batch of grapes in order to produce the next vintage's wine. By selling their wine in the form of wine futures, they recover their cash investment in one year and use these funds to finance next year's operations. Négociants also operate on a yearly basis; they move the inventory of the wine futures contracts they purchased from chateaus to the market so that they can recover their cash investment in these futures agreements. In sum, this one-year planning phenomenon creates a coordinated financial and physical flow that influences adjustments from the prior vintage's futures prices. It turns out that, when this one-year operational emphasis is incorporated into the predictive model, the estimations improve significantly. Hekimoğlu and Kazaz (2020) make use of the one-year operational planning phenomenon displayed in Figure 2.2. They define a unique variable definition that compares two consecutive vintages and use this "change" variable rather than level panel data. This variable definition is shown to improve accuracy.

Hekimoğlu and Kazaz (2020) develop a comprehensive predictive analytics model in order to estimate wine futures prices. Their study examines a total of 33 factors but identify that there are only five key factors that estimate futures prices accurately: (1) Temperature; (2) precipitation; (3) the Liv-ex 100 index; (4) barrel scores of the most influential tasting experts, and (5) a positive interaction term that captures the improvement in temperatures and market conditions from the previous vintage.

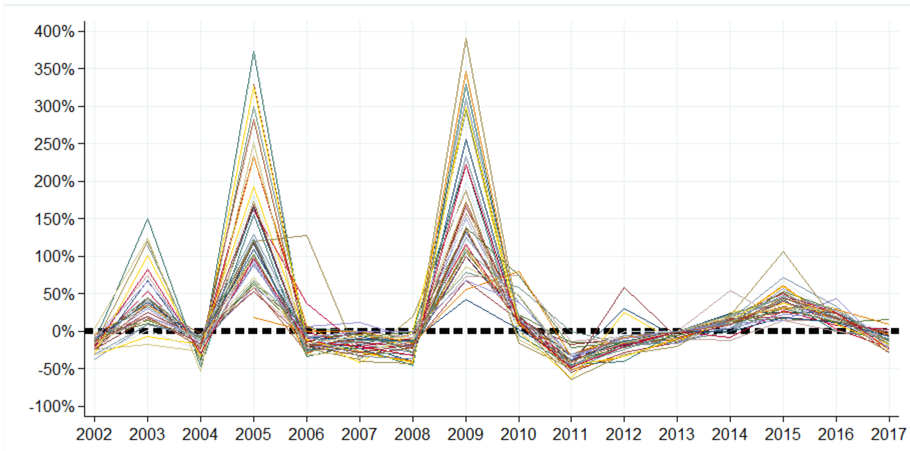

Figure 2.2. Percentage change in the futures prices from the previous vintage between 2002 and 2017 for 40 Bordeaux chateaus examined in Hekimoğlu and Kazaz (2020).

The online supplement provides the details of the data used in the analysis of Hekimoğlu and Kazaz (2020). In summary, futures prices are collected from Liv-ex (www.liv-ex.com); Table A1 of the online supplement provides the details of the chateaus involved in the analysis. Météo-France is the primary source for local weather information. The Merignac and Saint-Emilion weather stations are the main weather locations for the Left Bank and Right Bank chateaus, respectively. The weather data consists of daily maximum temperatures (in °C) and daily total rainfall (in cm) during the growing season (May 1 – August 31). Figure A1 of the online supplement illustrates the weather data between 2001 and 2017. Barrel tasting scores are provided by Liv-ex. They rely on the most influential publication: *Wine Advocate* and RobertParker.com[1]. The Liv-ex Fine Wine 100 Index captures market-wide fluctuations in this industry. Even though the index includes only bottled wines of older vintages, the value of the index is an excellent proxy for reflecting consumers' willingness to pay for fine wines. Figure A2 of the online supplement illustrates the values of Liv-ex 100 since its inception in July 2001.

[1] The late President François Mitterrand recognized Robert Parker with the Chevalier de l'Ordre National du Mérite in 1993, and President Chirac awarded Robert Parker with France's Legion of Honor, an extremely rare distinction, in 2005 for his contributions to the quality and education of French wines. The industry perceives Robert Parker as the most influential tasting expert.

2.1. Predictive Analytics Model

The predictive model in Hekimoğlu and Kazaz (2020) uses the following dependent and independent variable definitions:

Change in futures prices. The dependent variable is the logarithmic change across the futures prices of two consecutive vintages from the same chateau, i.e., $\Delta p_{i,t} = \log(p_{i,t}/p_{i,t\text{-}1})$ where $p_{i,t}$ is the futures price of vintage t of chateau i.

Change in average temperature. The temperature variable is the logarithmic change across the average growing season temperatures of two consecutive vintages, i.e., $\Delta m_{i,t} = \log(m_{i,t}/m_{i,t\text{-}1})$ where $m_{i,t}$ is the average of daily maximum temperatures during the growing season (May 1 – August 31) of year t in the region where chateau i is located. Because warmer temperatures are expected to produce a higher quality wine, a positive change is expected to result in higher prices from the previous vintage.

Change in total rainfall. The precipitation variable is the logarithmic change across the total growing season rainfall of two consecutive vintages, i.e., $\Delta r_{i,t} = \log(r_{i,t}/r_{i,t\text{-}1})$ where $r_{i,t}$ is the total rainfall during the growing season period of year t in the region where chateau i is located. A rainier growing season is expected to have a negative impact on the *ex-négociant* price.

Change in barrel tasting score. The barrel score variable is the difference between the barrel tasting scores of two consecutive vintages of the same chateau, i.e., $\Delta s_{i,t} = s_{i,t} - s_{i,t\text{-}1}$ where $s_{i,t}$ is the barrel tasting score of vintage t of chateau i. A higher score is expected to have a positive impact on the *ex-négociant* price.

Change in Liv-ex 100. The Liv-ex 100 index variable is the logarithmic change in the value of Liv-ex 100 index between the *en primeur* campaign of the previous vintage and shortly before the *en primeur* campaign of the new vintage. It is expressed as $\Delta l_t = \log(l_t^{mar}/l_{t\text{-}1}^{may})$ where $l_{t\text{-}1}^{may}$ is the value of Liv-ex 100 around the *en primeur* campaign of vintage $t - 1$ (corresponding to May of year t), and l_t^{mar} is the value of Liv-ex 100 in March prior to the *en primeur* campaign of vintage t (corresponding to March of year $t + 1$). The index value at the end of March is used in order to predict prices, rather than to explain prices. A positive change in the Liv-ex 100 index is expected to have a positive impact on the *ex-négociant* price.

Table A2 in the online supplement shows that the correlation coefficients among these variables are not strong enough to suggest any collinearity issues. Figure 2.2 highlights the hype effect that can be seen in vintages that are tagged as phenomenal, e.g., 2003, 2005 and 2009. To capture this hype effect, Hekimoğlu and Kazaz (2020) define additional positive interaction variables.

Positive Interaction variables. The following six interaction variables are defined in order to combine the pairwise positive effects of temperature and rainfall $(mr_{i,t})^+$, temperature and Liv-ex 100 $(ml_{i,t})^+$, temperature and barrel score $(ms_{i,t})^+$, rainfall and Liv-ex 100 $(rl_{i,t})^+$, rainfall and barrel score $(rs_{i,t})^+$, and Liv-ex 100 and barrel score $(ls_{i,t})^+$:

$(mr_{i,t})^+ = \Delta m_{i,t} \times |\Delta r_{i,t}|$ if $m_{i,t} > m_{i,t\text{-}1}$ and $r_{i,t} < r_{i,t\text{-}1}$, $(mr_{i,t})^+ = 0$ if otherwise;

$(ml_{i,t})^+ = \Delta m_{i,t} \times \Delta l_t$ if $m_{i,t} > m_{i,t\text{-}1}$ and $l_t^{mar} > l_{t\text{-}1}^{may}$, $(ml_{i,t})^+ = 0$ if otherwise;

$(ms_{i,t})^+ = \Delta m_{i,t} \times \Delta s_{i,t}$ if $m_{i,t} > m_{i,t\text{-}1}$ and $s_{i,t} > s_{i,t\text{-}1}$, $(ms_{i,t})^+ = 0$ if otherwise;

$(rl_{i,t})^+ = |\Delta r_{i,t}| \times \Delta l_t$ if $r_{i,t} < r_{i,t\text{-}1}$ and $l_t^{mar} > l_{t\text{-}1}^{may}$, $(rl_{i,t})^+ = 0$ if otherwise;

$(rs_{i,t})^+ = |\Delta r_{i,t}| \times \Delta s_{i,t}$ if $r_{i,t} < r_{i,t\text{-}1}$ and $s_{i,t} > s_{i,t\text{-}1}$, $(rs_{i,t})^+ = 0$ if otherwise;

$(ls_{i,t})^+ = \Delta l_t \times \Delta s_{i,t}$ if $l_t^{mar} > l_{t\text{-}1}^{may}$ and $s_{i,t} > s_{i,t\text{-}1}$, $(ls_{i,t})^+ = 0$ if otherwise.

Other explanatory variables. As will be seen later, Hekimoğlu and Kazaz (2020) report that they test a total of 33 explanatory variables; however, as presented in the online supplement, the remaining variables are not selected by the Lasso analysis.

2.2. Analysis and Results

The results of the ordinary least squares (OLS) regression of various models is presented in Table 2.1 with cluster-robust standard errors (using classical standard errors leads to the same statistical inferences). Models 1 through 4 show that each independent variable (temperature, rainfall, barrel score, and Liv-ex 100) influence futures prices independently at the highest statistical significance. Moreover, their coefficients produce the positive and negative values as expected. Model 5 uses only weather information: temperatures and rainfall. Model 6 adds barrel scores, and Model 7 uses all four independent variables. All variables in Models 5, 6, and 7 continue to be significant at 1%. Variance inflation factors (VIF) for the variables in Model 7 are 1.06 for Δl_t, 1.24 for $\Delta s_{i,t}$, 1.41 for $\Delta m_{i,t}$, and 1.62 for $\Delta r_{i,}$; they are substantially lower than the well-established threshold of 5.00 for collinearity (Studenmund 2016). Models 8 through 13 add the positive interaction terms to Model 7; they reveal positive coefficients as expected in accordance with their definition and are statistically significant at 1%. Models 14 – 16 study the combined effects of two interaction terms. Model 9 prevails as the predictive model to be used. This conclusion comes from the Akaike information criterion (AIC) that determines the best estimation model. The preferred model, Model 9, has the minimum AIC value -133.56 and a relatively high R^2 of 74.62% with all statistically significant coefficients. Model 9 can be described as:

$$\Delta p_{i,t} = \alpha_0 + \alpha_1 \Delta m_{i,t} + \alpha_2 \Delta r_{i,t} + \alpha_3 \Delta s_{i,t} + \alpha_4 \Delta l_t + \alpha_5 (ml_{i,t})^+ + \varepsilon_{i,t}. \quad (1)$$

VIF for Model 9 variables are 1.24 for $\Delta s_{i,t}$, 1.41 for Δl_t, 1.63 for $\Delta m_{i,t}$, 2.01 for $\Delta r_{i,}$, and 2.43 for $(ml_{i,t})^+$ which are all less than the commonly used threshold of 5.00 for collinearity. Thus, there is no collinearity issue in the predictive analytics model. It is also important to observe that Model 15 has a higher R^2 than Model 9, however, the additional interaction term $(rs_{i,t})^+$ is not significant and it has a lower AIC score indicating less desirability for predictive purposes.

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**	**Model 9**	**Model 10**	**Model 11**	**Model 12**	**Model 13**	**Model 14**	**Model 15**	**Model 16**
Int.	0.060 *15.64****	0.077 *19.59****	0.056 *12.35****	-0.007 *-1.28*	0.072 *19.33****	0.066 *14.95****	-0.004 *-0.64*	-0.056 *-6.65****	-0.086 *-8.84****	-0.022 *-2.60***	-0.049 *-5.80****	-0.029 *-3.30****	-0.024 *-2.57***	-0.061 *-6.85****	-0.083 *-8.29****	-0.061 *-6.36****
$\boldsymbol{\Delta m_{i,t}}$	2.931 *13.66****				1.205 *6.40****	1.026 *5.68****	1.594 *8.21****	1.027 *5.84****	0.568 *3.59****	1.189 *5.58****	1.775 *8.59****	1.568 *8.02****	1.538 *7.99****	1.063 *6.22****	0.536 *3.54****	1.465 *6.44****
$\boldsymbol{\Delta r_{i,t}}$		-0.507 *-17.48****			-0.427 *-16.87****	-0.352 *-14.57****	-0.241 *-11.67****	-0.114 *-5.38****	-0.060 *-2.99****	-0.253 *-11.80****	-0.054 *-2.34***	-0.174 *-6.79****	-0.218 *-8.28****	-0.117 *-5.60****	-0.068 *-3.17****	-0.071 *-3.16****
$\boldsymbol{\Delta s_{i,t}}$			0.066 *11.76****			0.035 *6.88****	0.035 *8.14****	0.033 *8.35****	0.032 *8.76****	0.028 *5.66****	0.037 *8.95****	0.029 *6.65****	0.026 *6.91****	0.028 *7.72****	0.034 *8.39****	0.031 *6.72****
$\boldsymbol{\Delta l_t}$				1.556 *18.23****			1.405 *16.61****	1.168 *14.31****	0.835 *12.37****	1.391 *16.68****	1.084 *14.09****	1.303 *16.32****	1.196 *14.86****	1.076 *14.03****	0.836 *12.34****	1.087 *14.14****
$\boldsymbol{(mr_{i,t})^+}$								7.626 *8.67****						6.712 *6.31****		
$\boldsymbol{(ml_{i,t})^+}$									54.318 *14.50****						56.319 *13.95****	
$\boldsymbol{(ms_{i,t})^+}$										0.393 *3.15****						0.293 *2.81****
$\boldsymbol{(rl_{i,t})^+}$											2.109 *9.70****					2.021 *9.32****
$\boldsymbol{(rs_{i,t})^+}$												0.079 *4.38****			-0.017 *-0.98*	
$\boldsymbol{(ls_{i,t})^+}$													0.280 *3.74****	0.162 *2.07***		
$\boldsymbol{R^2}$	19.00%	34.48%	21.41%	28.52%	36.82%	41.83%	63.76%	67.32%	74.62%	64.50%	67.26%	65.29%	65.30%	67.78%	74.68%	67.66%
$\boldsymbol{N}$	626	626	623	626	626	623	623	623	623	623	623	623	623	623	623	623
AIC	581.41	448.63	562.69	503.15	427.81	379.20	86.40	24.05	-133.56	75.52	25.23	61.57	61.40	17.24	-133.05	19.44

Table 2.1. Regression results for the dependent variable $\Delta p_{i,t}$ in Hekimoğlu and Kazaz (2020).
T-statistics using cluster-robust standard errors are given in italics below the coefficients.
*, **, *** denote statistical significance at 10%, 5%, 1%, respectively.

Using the equation (1) of Model 9, the predictive analysis populates the estimated futures prices $\hat{p}_{i,t}$ by converting the realized future prices of the previous vintage, denoted $p_{i,t-1}$:

$$\hat{p}_{i,t} = \exp\left(\hat{\Delta} p_{i,t}\right) p_{i,t-1}. \tag{2}$$

Figure 2.3 shows the accuracy of Model 9 by plotting the estimated and actual futures prices and fitting a line. The fitted line $y = 1.0002x$ has a slope extremely close to 1 and has an R^2 value of 94.87%. The figure reveals a remarkable success in prediction accuracy.

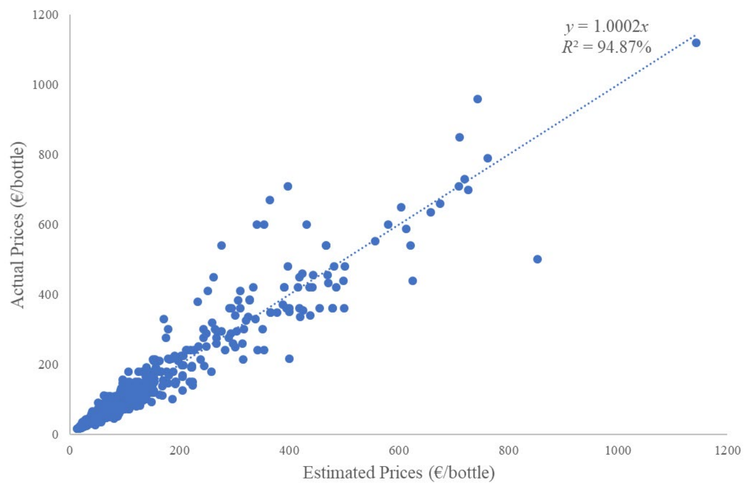

Figure 2.3. The fit between the actual and estimated futures prices in Hekimoğlu and Kazaz (2020) for the vintages between 2002 and 2017 using Model 9 where $N = 623$.

2.4. Prediction Accuracy with Out-of-Sample Testing

Ashenfelter (2008) reports that its predictive model has a mean absolute percentage error of 36.14% for vintages between 1967 and 1972. How does Model 9 perform in an out-of-sample test? Hekimoğlu and Kazaz (2020) employ an out-of-sample test by holding out the data for vintages between 2015 and 2017. Their approach trains the model by using the data for vintages up until 2014 and makes predictions for the 2015 vintage. The test is then replicated by including the data for 2015 and 2016 in re-estimating the coefficients in order to predict the 2016 and 2017 vintage prices, respectively. Defining the percentage error for the 2017 vintage as $e_{i,t=2017} = \left(\hat{p}_{i,t=2017} - p_{i,t=2017}\right) / p_{i,t=2017}$, Table 2.2 shows the estimated and actual futures prices for the 2017 vintage and their percentage errors. Model 9 has a mean absolute percentage error of 9.19% with a standard deviation of 7.17%. More than two thirds of the chateaus have prediction errors with less than 10%.

To see the prediction accuracy, Model 9 is compared with three benchmark models:

Model B0: $\log(p_{i,t}) = \alpha_0 + \alpha_1 \log(m_{i,t}) + \alpha_2 \log(r_{i,t}) + \mu_i + \varepsilon_{i,t}$

Model B1: $\log(p_{i,t}) = \alpha_0 + \alpha_1\log(m_{i,t}) + \alpha_2\log(r_{i,t}) + \alpha_3\log(l_t^{mar}) + \mu_i + \varepsilon_{i,t}$;

Model B2: $\log(p_{i,t}) = \alpha_0 + \alpha_1\log(m_{i,t}) + \alpha_2\log(r_{i,t}) + \alpha_3\log(l_t^{mar}) + \alpha_4 s_{i,t} + \mu_i + \varepsilon_{i,t}$

where μ_i represents the time-invariant chateau characteristics. Models B0-B2 rely on the traditional level panel data and do not employ the definition of change variables. Model B0 is equivalent to the prediction model of Ashenfelter (2008) which relies solely on weather information. Model B1 incorporates the Liv-ex 100 index to Model B0, and Model B2 adds the barrel scores to Model B1. Note that Model 7 uses the same four factors but employs the change variable definition.

Chateau (i)	Estimated Price (€) $\hat{p}_{i,t=2017}$	Actual Price (€) $p_{i,t=2017}$	Error (%) $e_{i,t=2017}$	Chateau (i)	Estimated Price (€) $\hat{p}_{i,t=2017}$	Actual Price (€) $p_{i,t=2017}$	Error (%) $e_{i,t=2017}$
Angelus	269.38	276.00	-2.40%	Lafleur	425.56	460.00	-7.49%
Ausone	505.77	480.00	5.37%	Leoville Barton	63.33	52.80	19.94%
Beychevelle	46.62	52.80	-11.70%	Leoville Las Cases	153.04	144.00	6.27%
Calon Segur	56.51	60.00	-5.81%	Leoville Poyferre	57.91	54.00	7.25%
Carruades Lafite	134.42	135.00	-0.43%	Lynch Bages	79.08	75.00	5.44%
Cheval Blanc	474.80	432.00	9.91%	Margaux	368.55	348.00	5.90%
Clarence (Bahans) Haut Brion	101.56	102.00	-0.43%	Mission Haut Brion	285.67	240.00	19.03%
Clinet	63.92	56.00	14.14%	Montrose	92.38	96.00	-3.77%
Clos Fourtet	73.51	72.00	2.09%	Mouton Rothschild	368.55	348.00	5.90%
Conseillante	133.17	120.00	10.97%	Palmer	224.34	192.00	16.84%
Cos d'Estournel	105.30	108.00	-2.50%	Pape Clement	59.77	61.20	-2.33%
Ducru Beaucaillou	126.07	120.00	5.06%	Pavie	245.02	276.00	-11.23%
Duhart Milon	49.81	48.00	3.77%	Pavillon Rouge	103.25	132.00	-21.78%
Eglise Clinet	206.16	168.00	22.72%	Pichon Baron	103.25	96.00	7.55%
Evangile	154.83	180.00	-13.98%	Pichon Lalande	108.68	90.00	20.76%
Grand Puy Lacoste	52.65	52.80	-0.28%	Pontet Canet	97.81	80.00	22.27%
Gruaud Larose	47.82	51.75	-7.60%	Smith Haut Lafitte	69.56	67.20	3.51%
Haut Bailly	73.71	72.00	2.37%	Troplong Mondot	91.99	72.00	27.77%
Haut Brion	380.38	348.00	9.30%	Vieux Chateau Certan	181.57	168.00	8.08%
Lafite Rothschild	438.97	420.00	4.52%				
Mean Absolute % Error = 9.19%				Min. of Absolute % Error = 0.28%			
Std. Dev. of Absolute % Errors = 7.17%				Max. of Absolute % Error = 27.77%			

Table 2.2. The estimated and actual futures prices for the 2017 vintage in Hekimoğlu and Kazaz (2020).

Table 2.3 provides the comparison of models B0, B1, B2, 7, and 9 in the same out-sample-test and makes the following conclusions: First, Model 9 makes significantly more accurate predictions than the widely practiced benchmark B0. Incorporating market information (Liv-ex 100 index) improves the performance by 20.97% and barrel scores by another 3.17%. Changing the variable definition prompts an improvement of 3.65% and the interaction term adds another 0.55% performance boost.

Vintage	Mean Absolute % Error				
	Model B0	**Model B1**	**Model B2**	**Model 7**	**Model 9**
2015	38.20%	18.57%	16.10%	10.12%	12.74%
2016	45.79%	21.24%	16.20%	12.29%	11.42%
2017	34.38%	15.66%	13.66%	12.61%	9.19%
Average	**39.46%**	**18.49%**	**15.32%**	**11.67%**	**11.12%**
Δ Liv-ex 100 (B0 – B1)	**20.97%**				
Δ barrel score (B1 – B2)		**3.17%**			
Δ variable definition (B2 – 7)			**3.65%**		
Δ positive interaction (7 – 9)				**0.55%**	

Table 2.3. Out-of-sample testing results of models B0, B1, B2, 7 and 9 in Hekimoğlu and Kazaz (2020).

A comprehensive robustness analysis is presented in the online supplement. The Lasso analysis is a popular machine learning methodology for variable selection. Analyzing 33 independent variables leading to 2^{33} = 8,589,934,592 independent models, the analysis confirms that the optimal set of variables is the set of variables featured in Model 9. Using quantile regression, robust regression, hierarchical linear modeling, the hierarchical Bayes model, parametric and non-parametric variable transformations, robustness tests show the extraordinary accuracy of this predictive model. As a result, Liv-ex has decided to publish the estimated prices of Model 9 as "realistic prices." Realistic prices provide transparency in the wine industry: Buyers can now determine whether a wine is underpriced or overpriced.

Improving the performance of price estimations would build additional investor confidence in wine futures. A higher number of participants in the futures market would improve the efficiency of the financial exchange. Future research should construct additional predictive analytics for other wine producing regions. This is particularly important in the US wine industry where there is no financial exchange to trade wine futures even though many offer it through wine club memberships (e.g., Harlan Estate, Hawkes, Far Niente Group's Nickel & Nickel). As a result, winemakers rely primarily on their direct-to-consumer channels. This can be seen in McMillan (2020) where a winery's financial sustainability is measured in terms of the size of its direct-to-consumer channel. In the absence of a financial exchange, when a winemaker goes bankrupt, the financial institution receives priority in taking possession of the existing inventory of wines. Thus, club members who paid the winemaker in advance may not receive their wines. In sum, establishing a financial exchange would be beneficial for protecting their investments in futures; it would incentivize more people to participate in wine futures trade.

3. The Supplier: Winemaker

Winemakers who focus on producing premium quality wines take a substantial risk with the long production times with 18 to 24 months before the wine is bottled. It is easy to see that monetary investment is tied up in this liquid for a long time. To combat the risks in the value during the production process, French winemakers have always sold premium wines in the form of wine futures. This risk-mitigation technique corresponds to the popular advance selling mechanism in marketing and to the inventory financing mechanism in supply chain finance. Section 2 developed a predictive analytics model that determined the value of the wine futures contracts. In this section of the tutorial, we develop a prescriptive model that helps winemakers endogenize the pricing decision and determine the amount of wine to be sold in advance in the form of futures contracts.

Noparumpa et al. (2015) develop predictive and prescriptive analytics approaches in order to manage the risks in the long production process. To help mitigate risks, their study develops a prescriptive analytical method that relies on a multinomial logit (MNL) model. The winemaker makes two critical decisions at the half point of the aging process, i.e., when the wine is still in the barrel: (1) What proportion of the wine should be sold in advance in the form of wine futures and what proportion of the wine should be kept for retail sales? And, (2) what should be the price of wine futures contracts? The latter decision sets an expectation about the retail price upon bottling. In this MNL model, consumers can purchase the wine at two different points in time: Either as a wine futures contract while the wine continues to age in the barrel or at retail after the wine is bottled.

Tasting experts review premium wines at two points in time: first when the wine is still aging in the barrel (and before it is offered to the public in the form of wine futures), and second when the wine is bottled and distributed for retail sales. Each time, the review will be a score given out of 100 points.[2] The first review is the "barrel score" and is denoted s_1; the wine is still aging in the barrel at this time. The second review is called the "bottle score" and is denoted by $\tilde{s}_2$ representing the randomness as it takes place in a future time period, i.e., after the wine is bottled. The barrel scores are published before the wine is offered in the form of wine futures and we already established the fact these barrel scores influence the price of wine futures. For example, when the most influential tasting expert Robert Parker provided his almost perfect score for the 2008 Lafite Rothschild, the price of wine futures jumped approximately 50 to 75% in value.

The MNL model describes consumers' valuation in terms of three decision options:
(1) Purchase the wine in advance of bottling and in the form of wine futures (at time epoch t_1);
(2) Wait for a year and purchase the wine after it is bottled at an unknown retail price (at time epoch t_2);
(3) Not purchase the wine.

The average valuation of a wine future at the time it is offered to the public is influenced by three factors: (a) the expectation from the unknown bottle score which is often assumed to be equal to the barrel score, i.e., $E\left[\tilde{s}_2\right] = s_1$; (b) the variation in the bottle score σ; and, (c) the risk-free rate r_f that indicates the

[2] There are a few tasting experts who provide scores out of 20, e.g., Jancis Robinson.

time-value-of-money from paying today and receiving the product a year later. Note that the weather information during the growing season is already revealed at this time; it is not random anymore.

We denote the average consumer value from a future under bottle-score uncertainty with v_f and employ a Conditional Value at Risk (CVaR) framework to determine its value. For a given $\xi \in (0, 1]$, let $z(\xi) = G^{-1}(\xi)$ describe the ξ^{th} percentile of the bottle score, i.e., $s_2(\xi) = s_1 + z(\xi)\sigma$. The valuation of wine futures by an average consumer is equal to the conditional expected value of the bottle score discounted to time t_1 at the risk-free rate, i.e.,

$$v_f = \left(1+r_f\right)^{-1} E\left[\tilde{s}_2 \mid s_1, \tilde{s}_2 \le s_2\left(\xi\right)\right] = \left(1+r_f\right)^{-1}\left(s_1 - E\left[-\tilde{z} \mid \tilde{z} \le z\left(\xi\right)\right]\sigma\right) = \left(1+r_f\right)^{-1}\left(s_1 - \gamma\sigma\right)$$

where $\gamma = E\left[-\tilde{z} \mid \tilde{z} \le z\left(\xi\right)\right]$ represents a measure of sensitivity to uncertainty in bottle score. Note that γ is decreasing in ξ, $\gamma \ge 0$ (due to $E[\tilde{z}] = 0$), and $\gamma = 0$ when $\xi = 1$. Then, v_f can be written as $v_f = \theta s_1$ where $\theta = (1 + r_f)^{-1}(1 - \gamma(\sigma/s_1))$ is the risk-adjusted discount factor. Note that in this definition of consumers' risk-adjusted discount factor, the risk-free discount factor ($(1 + r_f)^{-1}$) is reduced by the uncertainty in the bottle score (γ) and the coefficient of variation of the bottle score (σ/s_1). The valuation of a wine future by a random consumer is then $V_f = v_f + \varepsilon_f = \theta s_1 + \varepsilon_f$ where ε_f is a random variable with $E[\varepsilon_f] = 0$. The utility of a future is equal to the consumer surplus, i.e., the difference between valuation and price, $U_f = V_f - p_f = \theta s_1 + \varepsilon_f - p_f$. The average utility of a future among consumers can be written as $u_f = E[U_f] = \theta s_1 - p_f$.

One can observe that the utility from purchasing the wine in the form of a wine futures contracts is increasing in the expected bottle score (s_1) and is decreasing in price (p_f), the uncertainty in bottle score (σ), the risk-free discount rate (r_f), and risk aversion (γ).

A consumer who does not purchase the wine in the form of a wine future has two choices at the time that the wine is bottled: (1) purchase the wine at the unknown retail price $p_r\left(\tilde{s}_2 \mid s_1\right) = \tilde{s}_2$ at time t_2; or, (2) do not purchase it. The average utility of a retail purchase choice among consumers is the difference between the expected valuation and the expected price discounted by the risk-adjusted discount rate, i.e., is $u_r = \theta\left(E\left[\tilde{s}_2 \mid s_1\right] - E\left[p_r\left(\tilde{s}_2 \mid s_1\right)\right]\right) = 0$ and the random utility is $U_r = \varepsilon_r$ where ε_r is a random variable with $E[\varepsilon_r] = 0$. The average utility of the no purchase option is zero and the random utility is $U_0 = \varepsilon_0$. The utility of not purchasing a wine future in the first time epoch is the maximum utility among the two no-purchase alternatives, i.e., $\max\{U_r, U_0\} = \max\{\varepsilon_r, \varepsilon_0\}$.

We next derive the purchase probability for wine futures considering that a consumer selects the alternative with the highest utility. The fraction of consumers who purchase the wine in the form of wine futures is $P\left[U_f > \max\left\{U_r, U_0\right\}\right] = P\left[\max\left\{\varepsilon_r, \varepsilon_0\right\} - \varepsilon_f < \theta s_1 - p_f\right]$. Assuming that ε_f, ε_r, and ε_0 are i.i.d.

Gumbel random variables with zero mean and scale parameter β, $\max\{\varepsilon_r, \varepsilon_0\}$ is a Gumbel random variable with $E[\max\{\varepsilon_r, \varepsilon_0\}] = \beta \ln 2$ and scale parameter β (because the Gumbel distribution is closed under maximization), $\max\{\varepsilon_r, \varepsilon_0\} - \varepsilon_f$ is a logistic random variable (because the difference between two independent Gumbel random variables with the same scale parameter is a logistic random variable). Thus, the futures purchase probability conforms to the multinomial logit (MNL) model:

$$P\left[U_f > \max\{U_r, U_0\}\right] = \frac{e^{(\theta s_1 - p_f)/\beta}}{2 + e^{(\theta s_1 - p_f)/\beta}} = \frac{e^{\left((1+r_f)^{-1}(s_1 - \gamma\sigma) - p_f\right)/\beta}}{2 + e^{\left((1+r_f)^{-1}(s_1 - \gamma\sigma) - p_f\right)/\beta}}.$$

MNL models are commonly used in retail settings. What sets wine apart from a traditional retail good is that the market size increases/decreases with higher/smaller barrel scores. Let us describe this market size as a non-decreasing function of the barrel score s_1 and denote it with $M(s_1)$ where $M'(s_1) \geq 0$. Individuals will purchase a wine futures contract only when it provides the highest utility, i.e.,

$$q_f(p_f) = M(s_1) P\left[U_f > \max\{U_r, U_0\}\right] = M(s_1)\left[e^{(\theta s_1 - p_f)/\beta} / \left(2 + e^{(\theta s_1 - p_f)/\beta}\right)\right]. \tag{3}$$

Inverting (3) allows us to write the futures price as a function of quantity:

$$p_f(q_f) = \theta s_1 + \beta \ln\left[\left(M(s_1) - q_f\right)/2q_f\right]. \tag{4}$$

The winemaker's risk-adjusted discount factor is denoted with ϕ and its value reflects the risk of selling a bottle of wine at an uncertain retail price in the future. The higher the uncertainty of the bottle price and the more risk-averse the winemaker, the lower the value of ϕ.

A predictive analytics method relying on a Capital Asset Pricing Model (CAPM) can be employed in order to estimate the risk-adjusted discount factor for a winemaker. The estimate of the value of ϕ can then be used to solve the prescriptive analytical model for the winemaker. In this approach, let us describe the winemaker's risk-adjusted discount factor with $\phi = (1 + r_f + \gamma(r_m - r_f))^{-1}$ where r_m is the market return. For the international wine companies, it is appropriate to use the average annual percentage change in the Liv-ex 100 index as the market return r_m. Then, $r_m - r_f$ is the risk premium. The value of γ representing the winemaker's risk measure can be estimated through a CAPM approach using $\gamma = COV(r_j, r_m)/VAR(r_m)$ which represents the covariance between the returns of the specific winemaker (r_j) and the market returns (defined as $COV(r_j, r_m)$) divided by the variance in market returns (defined as $VAR(r_m)$).

The winemaker's risk-adjusted expected profit can then be expressed as follows:

$$\Pi(q_f) = q_f p_f(q_f) + \phi E\left[p_r(\tilde{s}_2 \mid s_1)\right](Q - q_f) = q_f\left((\theta - \phi)s_1 + \beta \ln\left[\frac{M(s_1) - q_f}{2q_f}\right]\right) + \phi s_1 Q \tag{5}$$

and the winemaker's objective function maximizes the above risk-adjusted expected profit function:

$$\rho^* = \max_{q_f \le Q} \Pi\left(q_f\right). \qquad (6)$$

The above derivations help in developing the expressions for the optimal values of the expected profit ρ^*, the optimal amount of wine to be sold in the form of wine futures q_f^*, and the optimal futures price p_f^*. These expressions rely on the Lambert W function $W(z)$ (Corless et al. 1996); $W(z)$ is the value of w satisfying $z = we^w$. Let us describe the winemaker's total production amount with Q. The optimal fraction of the market that will buy the wine in the form of wine futures is then equal to $\alpha^{o} = e^{(\theta-\phi)s_1/\beta - W\left(\frac{e^{(\theta-\phi)s_1/\beta}}{2e}\right)} \Big/ \left(2e + e^{(\theta-\phi)s_1/\beta - W\left(\frac{e^{(\theta-\phi)s_1/\beta}}{2e}\right)}\right)$ when the supply constraint is nonbinding, i.e., when $Q \ge M(s_1)\alpha^{o}$. The optimal solution for a MNL model using consumers' CVaR perspective and winemakers' risk-adjusted profits is characterized in Table 3.1:

Production is nonbinding, i.e., $Q \ge M(s_1)\alpha^{o}$	**Production is binding,** i.e., $Q < M(s_1)\alpha^{o}$
$q_f^* = M(s_1)\alpha^{o}$ $p_f^* = \phi s_1 + \beta\left[1 + W\left(\frac{e^{(\theta-\phi)s_1/\beta}}{2e}\right)\right]$ $\rho^* = M\left(s_1\right)\left[\beta W\left(\frac{e^{(\theta-\phi)s_1/\beta}}{2e}\right) + \phi s_1 \frac{Q}{M\left(s_1\right)}\right]$	$q_f^* = Q$ $p_f^* = \theta s_1 + \beta \ln\left[\frac{M\left(s_1\right) - Q}{2Q}\right]$ $\rho^* = Q\left(\theta s_1 + \beta \ln\left[\frac{M\left(s_1\right) - Q}{2Q}\right]\right)$

Table 3.1. The optimal solution of the MNL model based on the winemaker's production amount.

Noparumpa et al. (2015) estimate empirically that Bordeaux chateaus increase their futures allocation by an average of 27.65% due to the risks stemming from uncertainty in the bottle reviews and retail prices. It is also estimated that the model improves their expected profits by an average of 10.10%.

In Section 2, we estimated the realistic price, i.e., the *ex-négociant* price, at which the futures contract should be released to the financial exchange Liv-ex. Let $\hat{p}_f$ denote the *ex- négociant* price estimate. Let δ represent the négociant's margin; thus, the winemaker's realistic earning is $\hat{p}_f - \delta$. Consider the case when production is not binding, i.e., $Q \ge M(s_1)\alpha^{o}$ and $\phi s_1 + \beta\left[1 + W\left(e^{(\theta-\phi)s_1/\beta}/2e\right)\right] > \hat{p}_f - \delta$. This case implies that the winemaker's optimal price choice is greater than the estimated realistic earning. It tells the winemaker that this price can potentially create a problem in terms of the winemaker's estimate of the demand function derived from the MNL model. In this case, the winemaker is recommended to reduce the futures price to $p_f^* = \hat{p}_f - \delta$; the amount of futures contracts $q_f(p_f)$ that the winemaker should expect to sell can be obtained by substituting $p_f^* = \hat{p}_f - \delta$ in (3).

In the US, winemakers are not required to grow their grapes in their own estates like the French winemakers. The US winemakers lease vineyards in order to secure the supply of grapes and control the quality of their agricultural input. The above MNL model can easily be extended in order to determine the optimal amount of a vineyard lease for such winemakers. Incorporating the MNL-based prescriptive model into the leasing decisions enables winemakers to grow their businesses without requiring a substantial amount of capital, and thus, in a financially healthier manner.

Boutique and artisanal winemakers in the US exhibit significantly more risk averse behavior than Bordeaux chateaus. These small winemakers often possess a smaller amount of cash and operate under higher financial uncertainty. Thus, the estimates for the risk-adjusted discount factors for these boutique and artisanal winemakers in the US are substantially greater than those of Bordeaux chateaus. This implies that the value from selling wine in advance in the form of wine futures is significantly more valuable to boutique and artisanal winemakers of the US than the Bordeaux chateaus. Noparumpa et al. (2015) estimates empirically that one such winemaker, Heart and Hands Wine Company, would increase its wine futures allocation by an average of 55.03% and improve its expected profits by an average of 13.87%.

Noparumpa et al. (2015) show that winemakers benefit from selling wine in the form of wine futures. Will the buyers of wine futures, distributors and importers, lose if they engage in wine futures?

4. The Buyer: Distributor and Importer

This section develops a prescriptive analytics model that helps wine distributors determine the investment allocation between wine futures and bottled wine. Hekimoğlu et al. (2017) examine how the prices of wine futures and bottled wine are influenced by randomness in weather and market conditions. The predictive analytics framework in their publication provides the necessary empirical foundation for the functional forms that will be used in the prescriptive model. Every May of calendar year t, a distributor can buy wine futures of vintage $t - 1$ and bottled wines of vintage $t - 2$ from each winemaker. How much money should the wine distributor invest in the wine futures of vintage $t - 1$ and how much money should she allocate to the bottled wines from vintage $t - 2$? In order to decide, the wine distributor needs to predict the price evolution of wine futures and bottled wine from May of calendar year t to May of calendar year $t + 1$. The online supplement presents the details of the predictive analytics models that estimate the evolution of prices for wine futures and bottled wine. Figure 4.1 illustrates the evolution of realized futures and bottle prices over time.

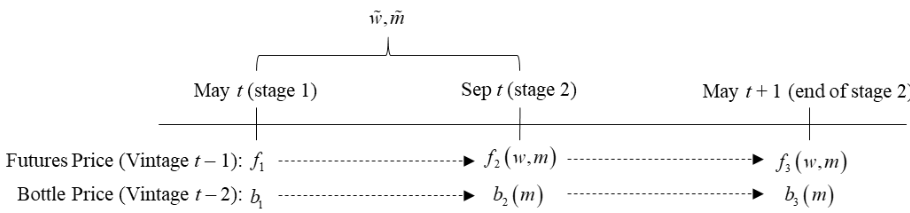

Figure 4.1. The evolution of futures and bottled wine prices under weather and market uncertainty.

We denote weather fluctuations with random variable $\tilde{w}$ and its realization with w, and market fluctuations with random variable $\tilde{m}$ and its realization with m. The predictive analytics study establishes the functional forms of futures and bottled wine prices in our analytical model as functions of w and m.

The predictive analysis presented in the online supplement shows that futures prices are negatively influenced by the new weather information: The value of f_2 decreases from f_1 with better weather information designated by a positive value of w; and the value of f_3 further decreases (with a larger coefficient) from f_2 with a positive value of w. Futures prices improve with better market conditions: The value of f_2 increases from f_1 with better market information designated by a positive value of m; and, the value of f_3 further increases (with a larger coefficient) from f_2 with a positive value of m. The predictive analysis pertaining to the bottled wine prices shows that the next vintage's weather information does not influence the evolution of bottled wine prices because the coefficient of w is not statistically significant in the price evolution from b_1 to b_2 or from b_2 to b_3. Market information positively impacts the evolution of bottled wine prices: The value of b_2 increases from b_1 with a positive value of m; and, the value of b_3 further increases (with a larger coefficient) from b_2 with a positive value of m. From the analysis of the predictive analytics model, Hekimoğlu et al. (2017) define the following:

$\partial f_2(w, m)/\partial w < 0$ and $\partial f_3(w, m)/\partial w < 0$ (because of the negative coefficients), with

$\partial f_3(w, m)/\partial w < \partial f_2(w, m)/\partial w < 0$ (because of the greater value of coefficients in later periods); and

$\partial f_2(w, m)/\partial m > 0$ and $\partial f_3(w, m)/\partial m > 0$ (because of the positive coefficients), with

$\partial f_3(w, m)/\partial m > \partial f_2(w, m)/\partial w > 0$ (because of the greater value of coefficients in later periods); and,

$\partial b_2(m)/\partial m > 0$ and $\partial b_3(m)/\partial m > 0$ (because of the positive coefficient), with

$\partial b_3(m)/\partial m > \partial b_2(m)/\partial m > 0$ (because of the greater value of coefficients in later periods).

4.1. The Prescriptive Model for the Buyer

Distributors often have a dedicated budget, denoted B, for each winemaker. The question is: How much of that money should be allocated into wine futures (of vintage $t-1$) versus bottled wine (of vintage $t-2$)? A two-stage stochastic program with recourse provides the optimal decisions.

<u>Stage 1 (May of year t):</u> The distributor determines the optimal values of

x_1: the amount of money to be invested in the wine futures of vintage $t-1$, and

y_1: the amount of money to be invested in the bottled wine from vintage $t-2$

within the limited budget of B and under risk aversion. Wine distributors exhibit risk aversion that conforms to the Value at Risk (VaR) measure where they try to limit the amount of loss with an associated probability. We denote the unit prices of wine futures and bottled wine with f_1 and b_1 and normalize their values to $f_1 = b_1 = 1$. We normalize the means of the two random variables $\tilde{w}$ and $\tilde{m}$ to zero, i.e., $E[\tilde{w}] = E[\tilde{m}] = 0$. Their probability density functions (pdf) are denoted $\phi_w(w)$ and $\phi_m(m)$ on respective support $[w_L, w_H]$ and $[m_L, m_H]$. We define the set $\Omega = [w_L, w_H] \times [m_L, m_H]$.

Stage 2 (September of year t): After observing the realizations for weather and market fluctuations (w, m) in September of year t, the distributor determines the optimal values for

x_2: the amount of additional money to be invested in the wine futures of vintage $t-1$, and

y_2: the amount of additional money to be invested in the bottled wine from vintage $t-2$.

Note that the distributor can sell some of the futures purchased in stage 1 in May, and thus, the value of x_2 can be negative. The unit price for wine futures is $f_2(w, m)$ and for bottled wine is $b_2(m)$.

The returns from wine futures and bottled wine at the end of stage 2 (May of year $t+1$) is also uncertain; they are described by random variables $(\tilde{z}_f, \tilde{z}_b)$ and their mean values are $E[\tilde{z}_f] = E[\tilde{z}_b] = 0$. The realized prices at the end of stage 2 are described as $f_3(w, m) + z_f$ and $b_3(m) + z_b$. Thus, $E[f_3(w, m) + \tilde{z}_f] = f_3(w, m)$ and $E[b_3(m) + \tilde{z}_b] = b_3(m)$.

From the analysis of the predictive models, note that when the futures and bottle prices move in one direction in the evolution from f_1 to $f_2(w, m)$ and from b_1 to $b_2(m)$, then often they also evolve in the same direction from $f_2(w, m)$ to $f_3(w, m) + z_f$ and from $b_2(m)$ to $b_3(m) + z_b$. As a result, we assume:

If $f_2(w, m) \lozenge f_1$, then $E[f_3(w, m) + \tilde{z}_f] \lozenge f_2(w, m)$ for all $\lozenge \in \{>, =, <\}$ and for all (w, m). (7)

If $b_2(m) \lozenge b_1$, then $E[b_3(m) + \tilde{z}_b] \lozenge b_2(m)$ for all $\lozenge \in \{>, =, <\}$ and for all m. (8)

The realized profit at the end of stage 2 can be written as:

$$\Pi(x_1, y_1, w, m, x_2, y_2, z_f, z_b)$$
$$= -x_1 - y_1 - f_2(w, m)x_2 - b_2(m)y_2 + [f_3(w, m) + z_f](x_1 + x_2) + [b_3(m) + z_b](y_1 + y_2). \quad (9)$$

The stage 2 model can be expressed as follows:

$$\max_{x_2, y_2} E\left[\Pi\left(x_1, y_1, w, m, x_2, y_2, \tilde{z}_f, \tilde{z}_b\right)\right] \quad (10)$$

subject to

$$f_2(w, m)x_2 + b_2(m)y_2 \leq B - x_1 - y_1 \quad (11)$$

$$P\left[\Pi\left(x_1, y_1, w, m, x_2, y_2, \tilde{z}_f, \tilde{z}_b\right) < -\beta\right] \leq \alpha \quad (12)$$

$$x_2 \geq -x_1 \quad (13)$$

$y_2 \geq 0$ (14)

Inequality (11) is the second-stage budget constraint; the distributor can use the remaining budget from stage 1 in addition to the money generated through the sale of futures in stage 2 (when $x_2 < 0$). Inequality (12) is the second-stage VaR constraint; the probability of realized profit less than $-\beta$ should not exceed α. Inequality (13) indicates that the distributor cannot sell more futures in stage 2 than the amount purchased in stage 1. For given x_1, y_1, w, m, we let (x_2^*, y_2^*) denote the optimal solution, i.e.,

$$\left(x_2^*\left(x_1, y_1, w, m,\right), y_2^*\left(x_1, y_1, w, m,\right)\right) = \underset{x_2, y_2}{\arg\max}\, E\left[\Pi\left(x_1, y_1, w, m, x_2, y_2, \tilde{z}_f, \tilde{z}_b\right)\right] \text{ s.t. (11) – (14).}$$

We can define $z_{f\alpha}$ and $z_{b\alpha}$ as the realizations of $\tilde{z}_f$ and $\tilde{z}_b$ at fractile α, i.e., $P\left[\tilde{z}_f \leq z_{f\alpha}\right] = P\left[\tilde{z}_b \leq z_{b\alpha}\right] = \alpha$ and assume that $z_{f\alpha} < 0$ and $z_{b\alpha} < 0$ implying that the distributor is worried about the negative returns at the end of stage 2. Distributors often invest only in bottled wine; therefore, it is practical to assume that the VaR constraint is satisfied in the event the distributor invests the entire budget on bottles, i.e.,

$$(1 - b_3(m_L) - z_{b\alpha})B < \beta. \quad (15)$$

The stage 1 model selects x_1 and y_1 that maximize the expected profit:

$$\max_{x_1, y_1 \geq 0} E\left[\Pi\left(x_1, y_1, \tilde{w}, \tilde{m}, x_2^*\left(x_1, y_1, \tilde{w}, \tilde{m}\right), y_2^*\left(x_1, y_1, \tilde{w}, \tilde{m}\right), \tilde{z}_f, \tilde{z}_b\right)\right] \quad (16)$$

subject to

$$x_1 + y_1 \leq B \quad (17)$$

$$P\left[\Pi\left(x_1, y_1, w, m, x_2^*\left(x_1, y_1, w, m\right), y_2^*\left(x_1, y_1, w, m\right), \tilde{z}_f, \tilde{z}_b\right) < -\beta\right] \leq \alpha \text{ for all } (w, m) \in \Omega \quad (18)$$

Inequality (17) is the budget constraint. Inequality (18) is the VaR constraint under a time-consistent risk measure (e.g., see Boda and Filar 2006). Some (x_1, y_1) decisions may satisfy the VaR constraint in stage 1 but violate it in stage 2; these decisions become infeasible in the model. To examine whether a distributor should invest in futures, it is appropriate to assume equal and positive returns at the end of stage 2:

$$E[f_3(\tilde{w}, \tilde{m}) + \tilde{z}_f] = E[b_3(\tilde{m}) + \tilde{z}_b] > 1. \quad (19)$$

4.2. Analysis of the Prescriptive Model

Partitioning the state space for $\left(\tilde{w}, \tilde{m}\right)$ allows us to identify when the distributor would (1) sell futures, (2) buy futures, and (3) sell futures in order to buy bottles.

$\Omega 0 = \{(w, m) \in \Omega : f_3(w, m)/f_2(w, m) = b_3(m)/b_2(m) = 1\}$

$\Omega 1 = \{(w, m) \in \Omega : f_3(w, m)/f_2(w, m) < 1 \text{ and } b_3(m)/b_2(m) < 1\}$

$\Omega 2 = \{(w, m) \in \Omega : f_3(w, m)/f_2(w, m) \geq \max\{b_3(m)/b_2(m), 1\} \setminus \Omega 0\}$

$\Omega 3 = \{(w, m) \in \Omega : b_3(m)/b_2(m) \geq \max\{f_3(w, m)/f_2(w, m), 1\} \cup \Omega 0\}$.

Let us define the equilibrium points as m_τ where $b_3(m_\tau)/b_2(m_\tau) = 1$ and $f_3(0, m_\tau)/f_2(0, m_\tau) = 1$, $w_\tau(m)$ where $f_3(w_\tau(m), m)/f_2(w_\tau(m), m) = 1$ for $m \le m_\tau$, and $w_\tau^- = w_\tau(m_L)$. Note that

$m_\tau < 0$, $w_\tau(m) < 0$ for all $m < m_\tau$, and $w_\tau(m_\tau) = 0$ (follows from (7), (8), (19)). (20)

We assume that

$m_\tau > m_L$ and $w_\tau(m_L) > w_L$. (21)

The set Ω0 identifies the set of realizations where the distributor is indifferent between wine futures and bottled wine. The set Ω1 defines realizations where it is best to sell the futures purchased in stage 1 and without buying bottled wine. The set Ω2 defines realizations where it is best to buy additional futures if there is any money left from the stage 1 investments. The set Ω3 defines realizations where it is best to sell the futures and buy bottles instead; this set of realizations shows the benefits of the liquidity from purchasing wine futures in stage 1.

It is useful to solve the risk-neutral version of stage 2 before analyzing risk aversion. This is accomplished by relaxing the VaR constraint in (12); we denote its solution as $\left(x_2^0, y_2^0\right)$ where

$$\left(x_2^0\left(x_1, y_1, w, m\right), y_2^0\left(x_1, y_1, w, m\right)\right) = \underset{x_2, y_2}{\arg\max}\, E\left[\Pi\left(x_1, y_1, w, m, x_2, y_2, \tilde{z}_f, \tilde{z}_b\right)\right] \text{ s.t. (11), (13), (14).}$$

Then, it can be seen that $\left(x_2^0, y_2^0\right)$ is given as follows:

$$\left(x_2^0, y_2^0\right) = \begin{cases} \left(-x_1, 0\right) & \text{if } \left(w, m\right) \in \Omega 1 \\ \left(\left(B - x_1 - y_1\right)/f_2\left(w, m\right), 0\right) & \text{if } \left(w, m\right) \in \Omega 2 \\ \left(-x_1, \left(B - x_1 - y_1 + f_2\left(w, m\right) x_1\right)/b_2\left(m\right)\right) & \text{if } \left(w, m\right) \in \Omega 3 \end{cases}, \tag{22}$$

Assuming that the bottled wine is more profitable than holding cash in stage 1, i.e.,

$$\partial E\left[\Pi\left(x_1, y_1, \tilde{w}, \tilde{m}, x_2^0, y_2^0, \tilde{z}_f, \tilde{z}_b\right)\right] / \partial y_1 \Big|_{(x_1, y_1) = (0,0)} > 0 \tag{23}$$

the following proposition shows that buying futures is even more profitable. Thus, for any (x_1, y_1),

$$\frac{\partial E\left[\Pi\left(x_1, y_1, \tilde{w}, \tilde{m}, x_2^0, y_2^0, \tilde{z}_f, \tilde{z}_b\right)\right]}{\partial x_1} \ge \frac{\partial E\left[\Pi\left(x_1, y_1, \tilde{w}, \tilde{m}, x_2^0, y_2^0, \tilde{z}_f, \tilde{z}_b\right)\right]}{\partial y_1} > 0. \tag{24}$$

Wine futures are advantageous because of the additional flexibility that they offer liquidity, i.e., being able to sell futures after observing weather and market random variables, and even swap them for bottles. Thus, a risk-neutral buyer would invest all budget in futures. The expected profit increases: (a) in the variation in market fluctuations in all cases; and, (b) in the variation of weather only when it increases the value from liquidity and swapping. Let us define the following for the risk-averse buyer:

$x_1^+ = \beta/[1 - f_2(w_H, m_L)]$; $x_1^V = [\beta + z_{b\alpha} B]/([1 - f_2(w_H, m_\tau)][1 + z_{b\alpha}])$; $x_1^s = (\beta - B[1 - b_3(m_L) - z_{b\alpha}])/[b_3(m_L) + z_{b\alpha} - f_2(w_H, m_L)]$; $y_1^V = [\beta - [1 - f_2(w_H, m_L)]x_1^V]/[1 - b_3(m_L) - z_{b\alpha}]$ and $y_1^s = (B[1 - f_2(w_H, m_L)] - \beta)/[b_3(m_L) + z_{b\alpha} - f_2(w_H, m_L)]$. Moreover, let

$$- z_{f\alpha} < \beta/B \tag{25}$$

$$\frac{\partial E\left[\Pi\left(x_1, y_1, \tilde{w}, \tilde{m}, x_2^0, y_2^0, \tilde{z}_f, \tilde{z}_b\right)\right] / \partial y_1\Big|_{(x_1,y_1)=(0,0)}}{\partial E\left[\Pi\left(x_1, y_1, \tilde{w}, \tilde{m}, x_2^0, y_2^0, \tilde{z}_f, \tilde{z}_b\right)\right] / \partial x_1\Big|_{(x_1,y_1)=(0,0)}} < \frac{1 - b_3\left(m_L\right) - z_{b\alpha}}{1 - f_2\left(w_H, m_L\right)}. \tag{26}$$

The values of x_1^+, x_1^V and y_1^V are determined at various (w, m) pairs that bind the VaR constraint in (18). The values of x_1^s and y_1^s are obtained when the VaR constraint (18) intersects with the budget constraint (17); note that x_1^s is strictly smaller than x_1^+ when $x_1^+ < B$. Inequality (25) states that investing the entire budget in futures in stage 2 at point (w_τ^-, m_L), the point at which the distributor switches to buying additional futures, does not violate the VaR constraint in (12). Inequality (25) is a weaker assumption than (15). Unlike (15), inequality (25) allows the possibility that investing the entire budget in futures in stage 1 can violate the VaR constraint in (18). In sum, the comparison of (25) with (15) reveals that there is greater uncertainty in the randomness in futures prices than that in bottle prices. Inequality (26) states that wine futures is preferred at the worst realizations of weather and market random variables.

It can be shown that it is always optimal to invest in some futures because $x_1^* > 0$ in all conditions. When (25) holds and $\left(\tilde{z}_f, \tilde{z}_b\right)$ follow a bivariate normal distribution, the optimal solution is as follows:

(a) If $\{x_1^+, x_1^V\} \geq B$, then $(x_1^*, y_1^*) = (B, 0)$ and $(x_2^*, y_2^*) = (x_2^0, y_2^0)$;

(b) If $x_1^V < B \leq x_1^+$, then $(x_1^*, y_1^*) = (x_1^V, B - x_1^V)$ and $(x_2^*, y_2^*) = (x_2^0, y_2^0)$;

(c) If $x_1^+ < \{x_1^V, B\}$, then

  (*i*) if (26) holds, then $(x_1^*, y_1^*) = (x_1^+, 0)$ and $(x_2^*, y_2^*) = (x_2^0, y_2^0)$;

  (*ii*) if (26) does not hold, then $(x_1^*, y_1^*) = (x_1^s, y_1^s)$ and $(x_2^*, y_2^*) = (x_2^0, y_2^0)$;

(d) If $x_1^s < x_1^V \leq x_1^+ < B$, then

  (*i*) if (26) holds, then $(x_1^*, y_1^*) = (x_1^V, y_1^V)$ and $(x_2^*, y_2^*) = (x_2^0, y_2^0)$;

  (*ii*) if (26) does not hold, then $(x_1^*, y_1^*) = (x_1^s, y_1^s)$ and $(x_2^*, y_2^*) = (x_2^0, y_2^0)$;

(e) If $x_1^V \leq x_1^s < x_1^+ < B$, then $(x_1^*, y_1^*) = (x_1^V, B - x_1^V)$ and $(x_2^*, y_2^*) = (x_2^0, y_2^0)$.

While it is optimal to invest in some futures all the time, it is not necessarily optimal to invest in bottles all the time as shown in (a) and (c)(*i*) above. The conclusion to always invest in some futures holds true even in the presence of a higher degree of uncertainty in futures and bottled wine and even if $\left(\tilde{z}_f, \tilde{z}_b\right)$ do not follow a bivariate normal distribution as specified in (25). When $\phi_w(w)$ follows a symmetric pdf and $\left(\tilde{z}_f, \tilde{z}_b\right)$ follow a bivariate normal distribution,

$$\frac{\partial E\left[\Pi\left(x_1, y_1, \tilde{w}, \tilde{m}, x_2^*, y_2^*, \tilde{z}_f, \tilde{z}_b\right)\right]}{\partial x_1} \geq \frac{\partial E\left[\Pi\left(x_1, y_1, \tilde{w}, \tilde{m}, x_2^*, y_2^*, \tilde{z}_f, \tilde{z}_b\right)\right]}{\partial y_1} > 0. \quad (27)$$

In conclusion, the distributors are recommended to invest in wine futures despite the fact that wine futures exhibit a higher degree of risk than purchasing bottled wine.

4.3. Financial Benefits from the Predictive and Prescriptive Models

We next present the financial benefits of the distributor's model using the Bordeaux chateaus employed in the analysis in Section 2. The evolution of futures prices for the 2010 and 2011 vintages are estimated from vintages 2007, 2008 and 2009. Similarly, the evolution of bottled wine prices for the 2009 and 2011 vintages are estimated from vintages 2006, 2007 and 2008. In May of 2011, the distributor can purchase the futures of the 2010 vintage or bottled wine of the 2009 vintage. At this time point, the distributor knows the actual futures and bottle prices (f_1 and b_1, respectively) for each winemaker. Similarly, in May of 2012, the distributor can purchase the futures of the 2011 vintage or bottled wine of the 2010 vintage. Using the coefficient estimates from the predictive analytics models, the evolution of futures and bottled wine prices are estimated for September of 2011 and 2012 (i.e., $f_2(w, m)$ and $b_2(m)$) and May of 2012 and 2013 (i.e., $f_3(w, m) + z_f$ and $b_3(m) + z_b$) for given realizations of all four random variables. The analysis here relaxes the earlier assumption that wine futures and bottled wine have equal expected returns as in (19). The distributor's tolerable loss is designated as 20% of the budget, i.e., $\beta = 0.2B$. Risk aversion is captured at $\alpha \in \{1, 0.20, 0.10\}$; $\alpha = 1$ corresponds to a risk-neutral distributor, whereas $\alpha = 0.20$ and $\alpha = 0.10$ represent low and high risk aversion, respectively. The results are independent of the choice of B which is set to $B = 10{,}000$. $E[\Pi_I^{j,t}(x_1^*, y_1^*)]$ denotes the optimal profit coming from winemaker j with futures and bottled wine in year t and $E[\Pi_I^{j,t}(0, y_1^{**})]$ is the expected profit from the current practice of investing only in bottled wine with no investment in futures, i.e., $(x_1, x_2) = (0, 0)$. The financial benefit from the prescriptive model is:

$$\Delta^{j,t} = (E[\Pi_I^{j,t}(x_1^*, y_1^*)] - E[\Pi_I^{j,t}(0, y_1^{**})])/E[\Pi_I^{j,t}(0, y_1^{**})]. \quad (28)$$

Table 4.1 summarizes the financial benefits with the average benefit $\bar{\Delta}^j = (1/2)\sum_t(\Delta^{j,t})$ for each winemaker at different levels of risk aversion. The average financial benefit for all chateaus is $\bar{\Delta} = \sum_j \bar{\Delta}^j / 35$.

Even the largest risk-neutral distributors would benefit significantly from investing in wine futures. The average expected profit improvement is estimated to be 17.83% with the largest improvement at 55.74% at Lafleur. The financial benefits increase in the presence of risk aversion: The average profit improvement is 19.46% for low risk aversion and 20.53% for high risk aversion; however, the improvement is not always monotone in risk aversion (see the decrease in Conseillante as an example).

Let f_I^{jr} denote the estimated realistic price for wine futures, corresponding to the initial release price, for chateau j from the analysis described in Section 2. Recall that the actual release price f_I^j can be greater than the estimated release price f_1^r; we argued earlier that these futures contracts are not desirable for the buyer. Among the 35 chateaus and two vintages, 39 of the initial release prices out of 70 futures prices (for vintages 2010 and 2011) are overpriced; this means that the actual release price is greater than the estimated realistic price, i.e., $f_1^j > f_1^{jr}$. For some of these overpriced futures prices, the appreciation over time can still make it an attractive investment especially when f_1^j evolves to f_2^j and f_3^j. For the 39 overpriced futures releases (with $f_1^j > f_1^{jr}$), the average benefit is only 5.24%. The benefit is not equal to zero because some futures contracts continue to appreciate at greater levels over time. Let us now compare this gain with the truly attractive – and underpriced – futures contracts where the release price is below the estimated realistic price, i.e., $f_1^j < f_1^{jr}$. From the 34 underpriced futures contracts in the sample, the average financial benefit is 33.05%. The analysis confirms our earlier assertion that underpriced futures are substantially more beneficial for the wine distributor than the overpriced futures. Realistic prices provide valuable insights for the wine distributor's purchasing decisions.

Winemaker (j)	Risk Neutral $\bar{\Delta}^j$	Low Risk Aversion $\bar{\Delta}^j$	High Risk Aversion $\bar{\Delta}^j$	Winemaker (j)	Risk Neutral $\bar{\Delta}^j$	Low Risk Aversion $\bar{\Delta}^j$	High Risk Aversion $\bar{\Delta}^j$
Angelus	4.45%	7.40%	10.00%	Leoville Barton	18.63%	18.63%	21.58%
Ausone	48.33%	53.18%	54.32%	Leoville Las Cases	28.20%	24.78%	25.92%
Beychevelle	0.00%	0.00%	0.00%	Leoville Poyferre	36.72%	23.82%	23.39%
Calon Segur	1.88%	1.88%	1.88%	Lynch Bages	20.97%	20.97%	20.97%
Cheval Blanc	29.71%	34.44%	36.89%	Margaux	31.84%	50.52%	53.81%
Clos Fourtet	38.92%	38.96%	39.30%	Mission Haut Brion	4.75%	12.99%	12.62%
Conseillante	10.69%	5.95%	5.35%	Montrose	14.90%	14.07%	17.98%
Cos d'Estournel	36.04%	31.53%	31.99%	Mouton Rothschild	10.93%	20.65%	22.62%
Ducru Beaucaillou	0.00%	2.30%	4.33%	Palmer	0.00%	0.00%	0.00%
Duhart Milon	10.35%	8.94%	12.74%	Pavie	24.46%	25.99%	28.53%
Eglise Clinet	6.64%	21.90%	21.71%	Pavillon Rouge	5.00%	5.00%	5.00%
Evangile	14.48%	33.16%	34.81%	Petit Mouton	3.69%	3.69%	3.69%
Grand Puy Lacoste	25.13%	26.18%	27.41%	Pichon Baron	17.06%	17.06%	17.06%
Gruaud Larose	7.34%	7.34%	7.34%	Pichon Lalande	10.29%	5.85%	7.49%
Haut Bailly	1.38%	1.38%	1.38%	Pontet Canet	10.44%	10.44%	10.44%
Haut Brion	9.91%	11.94%	14.32%	Troplong Mondot	32.24%	31.29%	31.21%
Lafite Rothschild	22.06%	43.32%	47.28%	Vieux Chateau Certan	21.33%	29.73%	31.83%
Lafleur	55.74%	35.73%	33.29%				

	Risk Neutral $\bar{\Delta}$	**Low Risk Aversion** $\bar{\Delta}$	**High Risk Aversion** $\bar{\Delta}$
Average	**17.83%**	**19.46%**	**20.53%**

Table 4.1. The average financial benefit to the wine distributor from the model.

A distributor may hold on to a greater amount of cash in under risk aversion. But, the presence of futures and its attractiveness with liquidity causes the distributor to increase early spending. In the absence of futures, the firm may hold excess cash, i.e., $y_1^{**} < B/b_1$. When futures are incorporated into the prescriptive model, the firm may invest more in stage 1 (i.e., $f_1x_1^* + b_1y_1^* > b_1y_1^{**}$) which leads to a greater average improvement than that for a risk-neutral distributor where $f_1x_1^* + b_1y_1^* = b_1y_1^{**} = B$. Collectively, it can be concluded that wine futures is increasingly beneficial for risk-averse distributors. Note that the distributor's profit can be further improved by considering a single budget for all chateaus; this adjustment can recommend carrying only a few chateaus. However, this approach is not always desirable for distributors because the business model requires that they carry wine from almost all chateaus.

5. Future Research Directions

Research in wine analytics will continue to contribute to the practice of winemaking and distribution. We offer a list of future research directions where analytical methods would benefit wine enterprises.

Climate Change. Wine is especially sensitive to climate change. The intensely warmer summers during the growing seasons cause the grapes to ripen faster with higher levels of sugar that converts to a higher level of alcohol. Wines from Central California often feature alcohol levels close to 16% in comparison to 13.5% often seen in Bordeaux. Investigating climate change impact on wine production might require a multi-faceted study on the choice of grape varietals that are resistant to hot summer temperatures to adopting better-suited locations for grape growth.

One of the consequences of climate change is the frequency of wildfires, such as those that took place in the summers of 2017 and 2019 in Northern California (the winemaking capital of the US), those in Australia and New Zealand as well as the spring frosts in Bordeaux. Wildfires are not only detrimental to grapes and vineyards; the smoke can taint the wines that are already aging in barrels. Frosts can be equally detrimental; the frost of 2017 reduced the overall crop yield substantially in Bordeaux.

Wine indices. Similar to the Liv-ex 100 index that reflects the consumers' willingness to pay for fine wines in a global context, new indices need to be developed in order to provide information about the local wine markets, e.g., California, Argentina. These new indices should reveal whether a vintage will be a success in a specific geography.

Digital transformation. Analytical methods are heavily employed in smart agricultural techniques. Weather forecasts tell winemakers whether there will be precipitation, and thus, whether they should irrigate the vineyard or not. Technological advancements need to be coupled with analytical methods in digital transformation. There is technology that identifies the minerals in the soil; important missing minerals can be supplemented in the irrigation systems for optimal soil composition necessary for grape growth.

Tasting scores. There are various tasting experts and their scores are demonstrated to be important for the price and quality perceptions. Future research should develop composite tasting scores using machine learning algorithms. These composite scores are likely to improve the predictions in futures prices.

Acknowledgments

The author is grateful for the contributions of Drs. M. Hakan Hekimoğlu of Rensselaer Polytechnic Institute, Tim Noparumpa of Chulalongkorn University, Thailand and Scott Webster of Arizona State University. Thanks go to Liv-ex executives James Miles, Neil Taylor, Sarah Philips, Theodora Moscalu, Andreas Neder, Giannis Kerestetzis for their continuous feedback; Southern Glazers' Wine & Spirits executives Steven Becker and Steve Slater, and the Silicon Valley Bank executive Rob McMillan for their insights.

References

A. Ahrens, C. B. Hansen and M. E. Schaffer. LASSOPACK: Stata module for lasso, square-root lasso, elastic net, ridge, adaptive lasso estimation and cross-validation. Statistical software components S458458, Boston College Department of Economics, revised 03 Nov 2019, 2019a.

A. Ahrens, C. B. Hansen and M. E. Schaffer. lassopack: Model selection and prediction with regularized regression in Stata, 2019b. https://arxiv.org/abs/1901.05397.

H. H. Ali, S. Lecocq and M. Visser. The impact of gurus: Parker grades and *en primeur* wine prices. *Economic Journal*, **118:**158-173, 2008.

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

O. Ashenfelter. Predicting the prices and quality of Bordeaux wines. *Economic Journal*, **118**(529):F174-F184, 2008.

O. Ashenfelter, D. Ashmore and R. Lalonde. Bordeaux wine vintage quality and the weather. *Chance*, **8**(4):7-13, 1995.

O. Ashenfelter and G. V. Jones. The demand for expert opinion: Bordeaux wine. *Journal of Wine Economics*, **8**(3):285-293, 2013.

O. Ashenfelter and K. Storchmann. Using a hedonic model of solar radiation to assess the economic effect of climate change: The case of Mosel valley vineyards. *Review of Economics and Statistics*, **92**(2):333-349, 2010.

R. H. Ashton. The value of expert opinion in the pricing of Bordeaux wine futures. *Journal of Wine Economics*, **11**(2):261-288, 2016.

N. Ayvaz-Çavdaroğlu, B. Kazaz and S. Webster. Incentivizing farmers to invest in quality through quality-based payment. 2020. In review.

R. M. Back, X. Liu, B. Niklas, K. Storchmann and N. Vick. Margins and mark-ups for fair-trade wine along the supply chain: Evidence from South African wine sold in the US. *Journal of Wine Economics*, **14**(3):274-297, 2019.

S. Bazen and J. M. Cardebat. Forecasting Bordeaux wine prices using state space methods. *Applied Economics*, **50**(47):5110-5121, 2018.

A. Belloni, D. Chen, V. Chernozhukov and C. Hansen. Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, **80**(6):2369–2429, 2012.

A. Belloni, V. Chernozhukov, C. Hansen and D. Kozbur. Inference in high-dimensional panel models with an application to gun control. *Journal of Business & Economic Statistics*, **34**(4):590–605, 2016.

A. Belloni, V. Chernozhukov and L. Wang. Square-root lasso: Pivotal recovery of sparse signals via conic programming. *Biometrika*, **98**(4):791–806, 2011.

A. Belloni, V. Chernozhukov and L. Wang. Pivotal estimation via square-root Lasso in nonparametric regression. *The Annals of Statistics*, **42**(2):757–788, 2014.

K. Boda and J. A. Filar. Time consistent dynamic risk measures. *Mathematical Methods in Operations Research*, **63**(1):169-186, 2006.

O. Boyabatlı. Supply management in multi-product firms with fixed proportions technology. *Management Science*, **61**(12):3013-3031, 2015.
O. Boyabatlı, P. Kleindorfer and S. Koontz. Integrating long-term and short-term contracting in beef supply chains. *Management Science*, **57**(10):1771-1787, 2011.
T. Boyacı and Ö. Özer. Information acquisition for capacity planning via pricing and advance selling: When to stop and act? *Operations Research*, **58**(5):1328–1349, 2010.
B. J. Burton and J. P. Jacobsen. The rate of return on investment in wine. *Economic Inquiry*, **39**:337-350, 2001.
R. P. Byron and O. Ashenfelter. Predicting the quality of an unborn Grange. *Economic Record*, **71**:40-53, 1995.
J. M. Cardebat, B. Faye and E. Le Fur, K. Storchmann. The law of one price? Price dispersion on the auction market for fine wine. *Journal of Wine Economics*, **12**(3):302-331, 2017.
J. M. Cardebat and J. M. Figuet. What explains Bordeaux wine prices? *Applied Economics Letters*, **11**(5):293-296, 2004.
J. M. Cardebat and L. Jiao. The long-term financial derivatives of fine wine prices: The role of emerging markets. *The Quarterly Review of Economics and Finance*, **67**:347-361, 2018.
S. H. Cho and C. S. Tang. Advance selling in a supply chain under uncertain supply and demand. *Manufacturing & Service Operations Management*, **15**(2):305–319, 2013.
P. Combris, S. Lecocq and M. Visser. Estimation of a hedonic price equation for Bordeaux wine: Does quality matter? *The Economic Journal*, **107**(441):390-402, 1997.
E. Dimson, P. Rousseau and J. Spaenjers. The price of wine. *Journal of Financial Economics*, **118**(2):431-449, 2015.
N. Duan. Smearing estimate: A nonparametric retransformation method. *Journal of the American Statistical Association,* **78**(383):605-610, 1983.
P. Dubois and C. Nauges. Identifying the effect of unobserved quality and expert reviews in the pricing of experience goods: Empirical application on Bordeaux wine. *International Journal of Industrial Organization*, **28**(3):205-212, 2010.
O. Gergaud, F. Livat, B. Rickard and F. Warzynski. Evaluating the net benefits of collective reputation: The case of Bordeaux wine. *Food Policy*, **71**:8–16, 2017.
J. W. Haeger and K. Storchmann. Prices of American pinot noir wines: Climate, craftsmanship, critics. *Agricultural Economics*, **35**:67-78, 2006.
M. H. Hekimoğlu, B. Kazaz and S. Webster. Wine Analytics: Fine wine pricing and selection under weather and market uncertainty. *Manufacturing & Service Operations Management*, **19**(2):202-215, 2017.
M. H. Hekimoğlu and B. Kazaz. Analytics for wine futures: Realistic prices. Forthcoming in *Production and Operations Management*, 2020. https://doi.org/10.1111/poms.13193.
M. Hu, Y. Liu and W. Wang. Socially beneficial rationality: The value of strategic farmers, social entrepreneurs, and for-profit firms in crop planting decisions. *Management Science*, **65**(8):3654-3672, 2019.
E. Jaeger. To save or savor: The rate of return to storing wine. *Journal of Political Economy*, **89**(3):584-592, 1981.
G. Jones and K. Storchmann. Wine market prices and investment under uncertainty: an econometric model for Bordeaux Cru Classes. *Agricultural Economics*, **26:**114-33, 2001.
P. C. Jones, T. Lowe, R. D. Traub and G. Keller. Matching supply and demand: The value of a second chance in producing hybrid seed corn. *Manufacturing & Service Operations Management*, **3**(2):122-137, 2001.
B. Kazaz. Production planning under yield and demand uncertainty. *Manufacturing & Service Operations Management*, **6**(3):209-224, 2004.
B. Kazaz and S. Webster. The impact of yield dependent trading costs on pricing and production planning under supply uncertainty. *Manufacturing & Service Operations Management*, **13**(3):404-417, 2011.

B. Kazaz and S. Webster. Technical note – Price-setting newsvendor problems with uncertain supply and risk aversion. *Operations Researc,h* **63**(4):807-811, 2015.
K. Kelley, J. Zelinskie, M. Centinari, D. M. Gardner, R. Govindasamy, J. Hyde, B. Rickard and K. Storchmann. Consumer preferences for sustainable wine attributes: A conjoint analysis. *Journal of Wine Economics*, **12**(4):416–425, 2017.
K. M. Kelley, J. Bruwer, J. Zelinskie, D. Gardner, R. Govindasamy, J. Hyde and B. Rickard. Wine consumers' willingness to adopt environmentally friendly packaging practices at tasting rooms. *British Food Journal*, **122**(1):309–327, 2020.
S. Lecocq and M. Visser. Spatial variations in weather conditions and wine prices in Bordeaux. *Journal of Wine Economics*, **1**(2):114-124, 2006.
G. Li. Robust regression. In *Exploring Data Tables, Trends, and Shapes*, ed. D.C. Hoaglin, C.F. Mosteller, and J.W. Tukey, pages 281-340. Wiley, New York, NY 1985.
P. Masset and C. Henderson. Wine as an alternative asset class. In *Handbook of the Economics of Wine* edited by O. Ashenfelter et al., pages 297-335. World Scientific Publishing Co., Singapore, Singapore, 2018.
P. Masset and J. P. Weisskopf. A study of the performance of fine wine on the Swiss market. *International Journal of Entrepreneurship and Small Business*, **26**(4):566-582, 2016.
P. Masset and J. P. Weisskopf. Raise your glass: Wine investment and the financial crisis. In *Handbook of the Economics of Wine* edited by O. Ashenfelter et al., pages 271-295. World Scientific Publishing Co., Singapore, Singapore, 2018a.
P. Masset and J.P. Weisskopf. Wine indices in practice: Nicely labeled but slightly corked. *Economic Modelling*, **68**:555-569, 2018b.
P. Masset, J. P. Weisskopf, B. Faye and E. Le Fur. Red obsession: The ascent of red wine in China. *Emerging Markets Review*, **29**:200-225, 2016.
R. McMillan. State of the wine US industry. Silicon Valley Bank Report, 2020.
T. Noparumpa, B. Kazaz and S. Webster. Wine futures and advance selling under quality uncertainty. *Manufacturing & Service Operations Management*, **17**(3):411-426, 2015.
T. Noparumpa, B. Kazaz and S. Webster. Production planning under supply and quality uncertainty with two customer segments and downward substitution. Working paper, 2020.
J. Y. Park, D. Bufquin and R. M. Back. Why do they become satiated? An examination of the relationship among winery tourists' satisfaction, repeat visits and revisit intentions. *Journal of Destination Marketing and Management*, **11:**231-239, 2018.
J. Perla, B. Rickard and T. Schmit. Looking for locapours: Using Zagat Survey® data to examine restaurant demand for local wine. *Agricultural and Resource Economics Review*, **43**(1):69–86, 2014.
T. M. Schmit, B. J. Rickard and J. Taber. Consumer valuation of environmentally friendly production practices in wines considering asymmetric information and sensory effects. *Journal of Agricultural Economics*, **64**(2):483–504, 2013.
K. Storchmann. Wine economics. *Journal of Wine Economics*, **7**(1):1-33, 2012.
J. Street, R. Carroll and D. Ruppert. A note on computing robust regression estimates via iteratively reweighted least squares. *The American Statistician*, **42**(2):152-154, 1988.
A. H. Studenmund. Using Econometrics: A Practical Guide. Pearson, Boston, MA, 2016.
C. S. Tang and W. S. Lim. Advance selling in the presence of speculators and forward looking consumers. *Production and Operations Management*, **22**(3):571-587, 2013.
D. Wood and K. Anderson. What determines the future value of an icon wine? New evidence from Australia. *Journal of Wine Economics*, **1**(2):141-161, 2006.
J. Xie and S. M. Shugan. Electronic tickets, smart cards, and online payments: When and how to advance sell. *Marketing Science*, **20**(3):219–243, 2001.
M. Yu, H. S. Ahn and R. Kapuscinski. Rationing capacity in advance selling to signal quality. *Management Science*, **61**(3):560-577, 2015.
M. Yu, R. Kapuscinski and H. S. Ahn. Advance selling: Effects of interdependent consumer valuations and seller's capacity. *Management Science*, **61**(9):2100-2117, 2015.

Online Supplement

for

Wine Analytics

Burak Kazaz

bkazaz@syr.edu

Whitman School of Management

Syracuse University

Part A

This section of the online supplement provides the supplemental information about the predictive analytics model presented in Section 2 of the publication.

Data Used in the Predictive Analytics Model:

The data in Hekimoğlu and Kazaz (2020) consists of the leading 40 chateaus that are listed in the Bordeaux 500 index. Their futures prices (in €/bottle) are collected from Liv-ex (www.liv-ex.com). Five Sauternes wine producers (Yquem, Climens, Coutet, Suduiraut, and Rieussec) are excluded because the production process and timeline of these wines are different than the traditional Bordeaux wines. Latour and Forts Latour wines are not offered in the form of wine futures and Petrus, Fleur Petrus, and Pin have missing futures prices in the Liv-ex database. Table A1 provides the list of these chateaus and their appellations; it also presents the average futures prices and the standard deviation for vintages between 2001 and 2017.

Chateau	Region	Price (€/bottle)		Chateau	Region	Price (€/bottle)	
		Average	Std. Dev.			Average	Std. Dev.
Angelus	Right Bank	154.12	78.58	Lafleur	Right Bank	400.71	145.57
Ausone	Right Bank	521.29	254.63	Leoville Barton	Left Bank	43.11	14.58
Beychevelle	Left Bank	35.18	14.13	Leoville Las Cases	Left Bank	115.18	52.43
Calon Segur	Left Bank	39.07	14.09	Leoville Poyferre	Left Bank	44.09	18.54
Carruades Lafite	Left Bank	69.40	42.36	Lynch Bages	Left Bank	55.19	24.43
Cheval Blanc	Right Bank	398.35	196.69	Margaux	Left Bank	310.75	165.41
Clarence (Bahans) Haut Brion	Left Bank	56.50	30.81	Mission Haut Brion	Left Bank	216.79	160.99
Clinet	Right Bank	51.35	15.66	Montrose	Left Bank	67.79	31.64
Clos Fourtet	Right Bank	46.11	18.93	Mouton Rothschild	Left Bank	297.82	169.15
Conseillante	Right Bank	78.83	39.76	Palmer	Left Bank	141.88	60.92
Cos d'Estournel	Left Bank	97.85	48.02	Pape Clement	Left Bank	62.14	18.88
Ducru Beaucaillou	Left Bank	86.85	42.75	Pavie	Right Bank	164.76	69.84
Duhart Milon	Left Bank	36.29	16.90	Pavillon Rouge	Left Bank	67.19	35.92
Eglise Clinet	Right Bank	131.35	72.28	Petit Mouton	Left Bank	71.76	34.81
Evangile	Right Bank	112.94	44.44	Pichon Baron	Left Bank	66.69	30.43
Grand Puy Lacoste	Left Bank	36.92	13.10	Pichon Lalande	Left Bank	73.55	31.42
Gruaud Larose	Left Bank	34.45	10.34	Pontet Canet	Left Bank	56.39	26.05
Haut Bailly	Left Bank	47.33	22.39	Smith Haut Lafitte	Left Bank	42.89	18.07
Haut Brion	Left Bank	299.00	184.99	Troplong Mondot	Right Bank	60.19	26.18
Lafite Rothschild	Left Bank	351.31	196.34	Vieux Chateau Certan	Right Bank	99.74	51.50

Table A1. List of chateaus in Hekimoğlu and Kazaz (2020), their region, average futures prices and standard deviation.

Météo-France, the national meteorological service organization, is the primary source for the local weather information; the weather information from Météo-France is complemented by Wolfram Mathematica. The Bordeaux appellation is divided into two main regions: Left Bank and Right Bank. The Merignac weather station is the main weather location for the Left Bank chateaus and Saint-Emilion weather station is the main weather location for the Right Bank chateaus. The weather data consists of the daily maximum temperatures (in °C) and the daily total rainfall (in cm) during the growing season (May 1 – August 31). Figure A1 illustrates the weather data between 2001 and 2017. Barrel tasting scores are provided by Liv-ex. They rely on the most influential publication RobertParker.com.

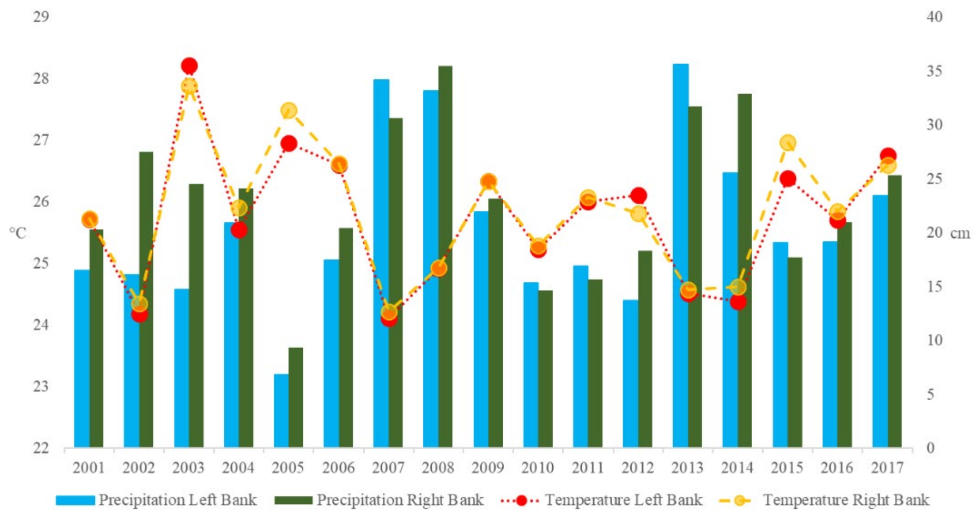

Figure A1. Average of the daily maximum temperatures and the total rainfall observed in the Left Bank and Right Bank of Bordeaux as reported in Hekimoğlu and Kazaz (2020) during growing season between years 2001 and 2017.

The Liv-ex Fine Wine 100 Index (shortly Liv-ex 100) is a monthly index that represents the price movements of the world's most sought-after 100 wines and is frequently quoted by Bloomberg and Reuters as the wine industry benchmark. It captures market-wide fluctuations in this industry. Even though the index includes only bottled wines of older vintages, the value of the index is an excellent proxy for reflecting consumers' willingness to pay for fine wines. Figure A2 illustrates the values of Liv-ex 100 index since its inception in July 2001.

Figure A2. Historical values of the Liv-ex Fine Wine 100 index since its inception.

Table A2 presents the correlation coefficients between the main variables used in the predictive analytics study. As can be seen from this table, the correlation coefficients between the four independent variables and the dependent variable are not strong enough to suggest collinearity.

Correlation Coefficients	Change in futures prices ($\Delta p_{i,t}$)	Change in temperature ($\Delta m_{i,t}$)	Change in rainfall ($\Delta r_{i,t}$)	Change in barrel score ($\Delta s_{i,t}$)	Change in Liv-ex 100 (Δl_t)
Change in futures prices ($\Delta p_{i,t}$)	1				
Change in temperature ($\Delta m_{i,t}$)	0.4358	1			
Change in rainfall ($\Delta r_{i,t}$)	-0.5891	-0.5188	1		
Change in barrel score ($\Delta s_{i,t}$)	0.4627	0.2988	-0.4271	1	
Change in Liv-ex 100 (Δl_t)	0.5343	-0.0381	-0.1742	0.0563	1

Table A2. The correlation coefficients among variables as reported in Hekimoğlu and Kazaz (2020).

Robustness Analysis

We next present a comprehensive robustness analysis confirming the predictive accuracy of Model 9 presented in Hekimoğlu and Kazaz (2020).

Quantile Regression Analysis

Does Model 9 predict consistently well at different price levels? Using the quantile regression approach, one can see from Table A3 that Model 9 exhibits consistent performance at all price levels.

Regression Percentiles	Price Thresholds (€)	Fitted Line	R^2
Below 25th percentile	price ≤ 45.86	$y = 1.0600x$	97.36%
Between 25th and 50th percentile	45.86 < price ≤ 72.18	$y = 1.0041x$	95.00%
Between 50th and 75th percentile	72.18 < price ≤ 143.90	$y = 1.0091x$	95.42%
Above 75th percentile	price > 143.90	$y = 0.9989x$	94.81%
All observations		$y = 1.0002x$	94.87%

Table A3. Results of regression in quantiles in Hekimoğlu and Kazaz (2020) using Model 9.

Lasso Analysis for Variable Selection

Hekimoğlu and Kazaz (2020) provide a comprehensive robustness check on the proposed predictive model. Their study begins with a Lasso analysis which is a popular machine learning methodology for variable selection. Lasso analysis makes the selection with the aim to balance the in-sample fit and the out-of-sample prediction accuracy. It departs from traditional approaches like OLS regression which can yield good in-sample performance (e.g., high R^2 values) with a relatively poorer prediction performance in an out-of-sample. Hekimoğlu and Kazaz (2020) report that the optimal variable selection according to the Lasso analysis features the exact same variables used in Model 9.

A comprehensive discussion of various Lasso applications is provided in Ahrens et al. (2019a, 2019b). Hekimoğlu and Kazaz (2020) employ the square-root Lasso with theory-driven rigorous penalization which is the most appropriate method (see Belloni et al. 2011, 2012, 2014, 2016) in order to control overfitting and to guarantee consistent out-of-sample prediction performance (Ahrens et al. 2019b). The square-root Lasso estimates the coefficients that minimize the following expression:

$$\hat{\beta}_{\sqrt{lasso}} = \arg\min \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i^T\beta\right)^2} + \frac{\lambda}{n}\sum_{j=1}^{p}\psi_j\left|\beta_j\right|$$

where the first term is the square root of mean squared error and the second term is a penalty term with the overall penalty level λ and individual penalty loading ψ_j for each regressor under n observations and p regressors. The square-root Lasso with theory-driven rigorous penalization approach yields the optimal overall penalty level $\lambda = (1.1)(n)^{1/2}\Phi^{-1}(1 - 0.05/(\log(n)p))$, where $\Phi^{-1}(\cdot)$ is the inverse of the standard normal cumulative distribution, and computes the individual penalty loading ψ_j using an iterative algorithm (see Ahrens et al. 2019b for details).

Hekimoğlu and Kazaz (2020) include a total of 33 independent variables in their analysis. Complementing the independent variables defined earlier, their study employs the following additional variables:

Quadratic variables. Quadratic terms of the changes in temperature, rainfall, barrel score and Liv-ex 100 variables account for potential nonlinearity in the form of convexity/concavity and are denoted by $(\Delta m_{i,t})^2$, $(\Delta r_{i,t})^2$, $(\Delta s_{i,t})^2$ and $(\Delta l_t)^2$, respectively.

Negative interaction variables. The negative interaction variables account for possible negative synergies stemming from multiple negative news. They are described as:

$(mr_{i,t})^- = |\Delta m_{i,t}| \times \Delta r_{i,t}$ if $m_{i,t} < m_{i,t-1}$ and $r_{i,t} > r_{i,t-1}$, $(mr_{i,t})^- = 0$ if otherwise;

$(ml_{i,t})^- = \Delta m_{i,t} \times \Delta l_t$ if $m_{i,t} < m_{i,t-1}$ and $l_t^{mar} < l_{t-1}^{may}$, $(ml_{i,t})^- = 0$ if otherwise;

$(ms_{i,t})^- = \Delta m_{i,t} \times \Delta s_{i,t}$ if $m_{i,t} < m_{i,t-1}$ and $s_{i,t} < s_{i,t-1}$, $(ms_{i,t})^- = 0$ if otherwise;

$(rl_{i,t})^- = \Delta r_{i,t} \times |\Delta l_t|$ if $r_{i,t} > r_{i,t-1}$ and $l_t^{mar} < l_{t-1}^{may}$, $(rl_{i,t})^- = 0$ if otherwise;

$(rs_{i,t})^- = \Delta r_{i,t} \times |\Delta s_{i,t}|$ if $r_{i,t} > r_{i,t-1}$ and $s_{i,t} < s_{i,t-1}$, $(rs_{i,t})^- = 0$ if otherwise;

$(ls_{i,t})^- = \Delta l_t \times \Delta s_{i,t}$ if $l_t^{mar} < l_{t-1}^{may}$ and $s_{i,t} < s_{i,t-1}$, $(ls_{i,t})^- = 0$ if otherwise.

Change in exchange rates. The variable $\Delta fx\_\$_t$ is the logarithmic change in the USD (\$) to Euro (€) exchange rate between the futures price of the previous vintage and shortly before the futures price of the new vintage, i.e., $\Delta fx\_\$_t = \log(fx\_\$_t^{mar}/fx\_\$_{t-1}^{may})$ where $fx\_\$_{t-1}^{may}$ is the value of \$/€ rate around the *en primeur* campaign of vintage $t – 1$ (corresponding to May of year t), and $fx\_\$_t^{mar}$ is the value of \$/€ rate in March prior to the *en primeur* campaign of vintage t (corresponding to March of year $t + 1$). The variable $\Delta fx\_£_t = \log(fx\_£_t^{mar}/fx\_£_{t-1}^{may})$ is the logarithmic change in the Pound Sterling (£) to Euro (€) exchange rate. These exchange rate variables are intended to capture the influence of currency fluctuations.

Left Bank dummy. The binary variable lb_i takes a value of 1 if chateau i is located in the Left Bank region and a value of 0 if it is in the Right Bank region. It is important to remember that the Left Bank chateaus produce wines that are dominated by Cabernet Sauvignon grapes and the Right Bank chateaus focus on wines that employ Merlot grapes; thus, this classification can be a potentially important factor.

The 1855 Bordeaux Classification (for Left Bank) variables. The 1855 Bordeaux Classification is an official ranking system established by Napoleon III in 1855. It is still in effect. It classifies chateaus according to their reputation into five groups from first to fifth growth where the first growth represent the most reputable chateaus. Five binary variables are defined, $first_i$, $second_i$, $third_i$, $fouth_i$, $fifth_i$, and they take the value of 1 if chateau i belongs to the corresponding category and 0 if it does not. These five binary variables aim to capture the reputation effects on prices.

The Saint-Emilion Classification (for Right Bank) variables. The Saint-Emilion Classification categorizes the notable Right Bank chateaus based on their reputation. Binary variables are defined for each category: Premier Grand Cru Classe A (pga_i) and Premier Grand Cru Classe B (pgb_i) take the value of 1 if chateau i belongs to the corresponding category and 0 if it does not. Table A4 provides the comprehensive set of classifications used for the Left Bank and the Right Bank chateaus.

Other chateau variables. Three additional time-variant chateau variables are employed in the analysis: (1) Annual trade volume $\Delta vol_{i,t} = vol_{i,t} - vol_{i,t\text{-}1}$ where $vol_{i,t}$ describes the percentage of the total trade volume that belongs to chateau i in year t; (2) annual trade value $\Delta val_{i,t} = val_{i,t} - val_{i,t\text{-}1}$ where $val_{i,t}$ describes the percentage of the total trade value that belongs to chateau i in year t; and, (3) the number of unique wines produced by chateau i between two consecutive vintages, i.e., $\Delta unq_{i,t} = unq_{i,t} - unq_{i,t\text{-}1}$ where $unq_{i,t}$ is the number of unique wines produced by chateau i in year t.

The 1855 Bordeaux Classification (for Left Bank)		**The Saint-Emilion Classification (for Right Bank)**
First Growth	**Fourth Growth**	**Premier Grand Cru Classe A**
Haut Brion	Beychevelle	Angelus
Lafite Rothschild	Duhart Milon	Ausone
Margaux	**Fifth Growth**	Cheval Blanc
Mouton Rothschild	Grand Puy Lacoste	Pavie
Second Growth	Lynch Bages	**Premier Grand Cru Classe B**
Cos d'Estournel	Pontet Canet	Clos Fourtet
Ducru Beaucaillou	**Unclassified**	Troplong Mondot
Gruaud Larose	Carruades Lafite	**Unclassified**
Leoville Barton	Clarence (Bahans) Haut Brion	Clinet
Leoville Las Cases	Haut Bailly	Conseillante
Leoville Poyferre	Mission Haut Brion	Eglise Clinet
Montrose	Pape Clement	Evangile
Pichon Baron	Pavillon Rouge	Lafleur
Pichon Lalande	Petit Mouton	Vieux Chateau Certain
Third Growth	Smith Haut Lafitte	
Calon Segur		
Palmer		

Table A4. Chateaus according to the 1855 Bordeaux Classification and the Saint-Emilion Classification.

The Lasso method identifies the exact same set of variables identified by the predictive analytics model in the study, i.e., Model 9: $\Delta m_{i,t}$, $\Delta r_{i,t}$, $\Delta s_{i,t}$, Δl_t, and $(ml_{i,t})^+$. It is important to note that the Lasso analysis confirms this predictive model by testing 2^{33} (= 8,589,934,592) unique model combinations.

Robust Regression

Robust regression is employed in order to ensure that the findings are robust to outlying observations in the sample. The analysis shows that no outlier observations need to be dropped from consideration according to Cook's *D*. Robust regression replicates regressions in an iterative manner using Huber weights followed by biweights until the weights of observations converge (Li 1985). It assigns smaller weights to outliers in order to reduce their impact. Replicating the analysis using robust regression, Hekimoğlu and Kazaz (2020) report that all statistical inferences remain intact, yielding the conclusion that the findings are robust to outliers.

Hierarchical Linear Modeling

Hierarchical linear modeling is employed in order to see whether the data features group effects. This methodology introduces chateau-specific intercept and coefficients for all of the variables in the predictive model. The revised version of Model 9 is as follows:

$$\Delta p_{i,t} = \alpha_0 + \beta_{0,i} + (\alpha_1 + \beta_{1,i})\Delta m_{i,t} + (\alpha_2 + \beta_{2,i})\Delta r_{i,t} + (\alpha_3 + \beta_{3,i})\Delta s_{i,t} + (\alpha_4 + \beta_{4,i})\Delta l_t + (\alpha_5 + \beta_{5,i})(ml_{i,t})^+ + \varepsilon_{i,t}$$

where α_k represents the fixed (average) effect and $\beta_{k,i}$ represents the random effect for chateau *i* such that $\beta_{k,i} \sim N(0, \sigma_k^2)$ for $k \in \{0,\ldots, 5\}$. Maximum likelihood is used in order to estimate the parameters α_k and σ_k^2 for each *k*, along with σ_ε^2 denoting the variance for $\varepsilon_{i,t}$. From the variance estimates, it is concluded that the effects of temperature, Liv-ex 100 index and their positive interaction vary across chateaus. The consequence of this observation is that the futures prices across chateaus respond differently to these three independent variables. Therefore, the analysis is replicated with the mixed-effects regression that features chateau-specific intercept and coefficients. The results of a mixed-effect regression for the hierarchical linear modeling approach is tabulated in Table A5. It demonstrates that the results do not improve using the hierarchical linear model.

2.5.4. Hierarchical Bayes Modeling

This section presents a Bayesian alternative to hierarchical modeling of the group-level effects which treats α_k, σ_k^2, and σ_ε^2 as fixed unknown parameters estimated through maximum likelihood. The Bayesian approach treats these parameters as random variables. A normal likelihood model is employed that considers

$$\Delta p_{i,t} \sim N(\alpha_0 + \beta_{0,i} + (\alpha_1 + \beta_{1,i})\Delta m_{i,t} + (\alpha_2 + \beta_{2,i})\Delta r_{i,t} + (\alpha_3 + \beta_{3,i})\Delta s_{i,t} + (\alpha_4 + \beta_{4,i})\Delta l_t + (\alpha_5 + \beta_{5,i})\, (ml_{i,t})^+, \sigma_\varepsilon^2)$$

where the prior and hyperprior distributions are $\alpha_k \sim N(0, 10000)$, $\beta_{k,i} \sim N(0, \sigma_k^2)$, $\sigma_k^2 \sim$ Inv-Gamma(0.01, 0.01) for $k \in \{0,\ldots, 5\}$, and $\sigma_\varepsilon^2 \sim$ Inv-Gamma(0.01, 0.01). Metropolis-Hastings and Gibbs sampling methods are used in order to simulate the posterior distributions. Table A5 presents the results of the analysis pertaining to hierarchical Bayes. It demonstrates that the results do not improve using hierarchical Bayes modeling.

2.5.5. Dependent Variable Retransformation

The predictive model in Hekimoğlu and Kazaz (2020) requires retransforming a logarithmic dependent variable back to its untransformed scale. Two alternative methods can be used to remove potential bias due to this retransformation: The normal theory estimation and the smearing estimation in Duan (1983). Normal theory estimation assumes that errors in (1) are normally distributed. In light of this assumption, it is appropriate to revise (2) as

$$\hat{p}_{i,t}^{N} = \exp\left(\hat{\Delta}p_{i,t} + \hat{\sigma}^2/2\right) p_{i,t-1} \quad (29)$$

where $\hat{\sigma}^2$ denotes the mean squared error in (1). Duan's smearing estimation is an alternative and nonparametric method which does not require any knowledge on the error distribution. For the smearing estimates, it is appropriate to revise (2) as follows:

$$\hat{p}_{i,t}^{D} = \left[N^{-1} \sum_{i} \sum_{t} \exp\left(\hat{\varepsilon}_{i,t}\right) \right] \exp\left(\hat{\Delta}p_{i,t}\right) p_{i,t-1} \quad (30)$$

where $\hat{\varepsilon}_{i,t}$ denotes the residual for vintage *t* of chateau *i* in (1).

Table A5 compares the out-of-sample performance of the original estimates with that of the normal theory and the non-parametric approach. The analysis concludes that the original estimates do not feature a systematic bias.

Vintage	Mean Absolute % Error					
	OLS	**Robust Regression**	**Hierarchical Linear Model**	**Hierarchical Bayes**	**Normal Theory Retransformation**	**Smearing Method Retransformation**
2015	12.74%	13.39%	12.88%	12.55%	11.31%	11.31%
2016	11.42%	11.69%	11.39%	12.22%	10.22%	10.23%
2017	9.19%	9.29%	9.15%	9.58%	10.45%	10.44%
Average	**11.12%**	**11.46%**	**11.14%**	**11.45%**	**10.66%**	**10.66%**

Table A5. Summary of out-of-sample testing results in Hekimoğlu and Kazaz (2020). The comparison of Model 9 using OLS of versus Robust Regression, Hierarchical Linear Model, Hierarchical Bayes Model, Normal Theory-based Retransformation and Duan's Smearing Method in Retransformation.

Part B

This section of the online supplement provides the supplemental information about the predictive analytics model presented in Section 4 of the publication.

Predictive Analytics for the Evolution of Prices

We begin our discussion with realized values of futures and bottled wine prices. In May of calendar year t, futures for vintage $t-1$ are released at the futures price of $f_1^{j,t-1}$ for winemaker j. We express the futures price of the same vintage for winemaker j in September of calendar year t as $f_2^{j,t-1}$, and in May of calendar year $t+1$ as $f_3^{j,t-1}$. In May of calendar year t, bottled wine of winemaker j from vintage $t-2$ is also released, and we express this bottle price as $b_1^{j,t-2}$. We denote the bottle price of vintage $t-2$ from winemaker j in September of calendar year t with $b_2^{j,t-2}$, and in May of calendar year $t+1$ with $b_3^{j,t-2}$.

After the wine distributor makes investments in futures of vintage $t-1$ and bottled wine of vintage $t-2$ from winemaker j in May of calendar year t, new summer weather information becomes available in calendar year t. This new summer weather information, which is fully observed by September of calendar year t, provides a relative comparison for the wines that are from vintages $t-1$ and $t-2$. For the case of wine futures of vintage $t-1$, the new weather information from May–September period of year t compared to the growing season of grapes (i.e., May–September period of year $t-1$) can play a role. Thus, both $f_2^{j,t-1}$ and $f_3^{j,t-1}$ can be influenced by the new weather information. For the case of bottled wine of vintage $t-2$, the new weather information from May–September period of year t compared to the growing season of grapes (i.e., May–September period of year $t-2$) can also influence the values of $b_2^{j,t-2}$ and $b_3^{j,t-2}$. Similarly, market conditions change from May to September of year t. As a consequence, the weather and market information observed at the end of summer in calendar year t can have an impact of the values of $f_2^{j,t-1}$, $f_3^{j,t-1}$, $b_2^{j,t-2}$, and $b_3^{j,t-2}$.

Evolution of Futures Prices

For the futures of vintage $t-1$, we denote the average temperature difference between the new growing season (of calendar year t) and the wine's own growing season by w_t. A positive (negative) w_t implies that the new growing season is relatively warmer (colder) than the growing season of the futures. We denote the percentage change in Liv-ex 100 index over the new growing season (of calendar year t) by m_t. A positive (negative) m_t implies that the market conditions improved (worsened) over the new growing season.

We develop the following linear regression models designated as Model 1A and Model 1B, respectively, where $t = \{2008, 2009, 2010, 2011, 2012\}$ and $j = \{1, 2, \ldots, 44\}$:

$$(f_2^{j,t-1} - f_1^{j,t-1}) = \gamma_0 + \gamma_1 w_t + \gamma_2 m_t + \varepsilon_{j,t}, \tag{31}$$

$$(f_3^{j,t\text{-}1} - f_2^{j,t\text{-}1}) = \eta_0 + \eta_1 w_t + \eta_2 m_t + \varepsilon_{j,t}. \tag{32}$$

Table B1 provides the regression results of the impact of new summer weather and market information on the price evolution of futures with $f_2^{j,t\text{-}1}$ (in Model 1A) and $f_3^{j,t\text{-}1}$ (in Model 1B). Four conclusions can be made from this empirical analysis. First, better weather of the upcoming vintage (i.e., higher value of w_t) has a negative impact on the evolution of futures price from $f_1^{j,t\text{-}1}$ to $f_2^{j,t\text{-}1}$. This weather effect is statistically significant at 1% level. This can be easily understood: If the upcoming vintage had better weather conditions, then the futures price for the current vintage would decrease. Moreover, better weather of the upcoming vintage (i.e., higher value of w_t) has a continued negative impact (statistically significant at 1%) on the evolution of futures prices from $f_2^{j,t\text{-}1}$ to $f_3^{j,t\text{-}1}$. This implies that the new weather information is not completely priced in the futures prices as of September of calendar year t. A similar observation is made in Ashenfelter (2008). Second, the negative coefficient representing the impact of weather in the evolution of futures prices from $f_2^{j,t\text{-}1}$ to $f_3^{j,t\text{-}1}$ is greater in absolute value than that of $f_1^{j,t\text{-}1}$ to $f_2^{j,t\text{-}1}$. Third, improving market conditions during the summer of calendar year t (with a higher value of m_t) has a positive impact on the evolution of futures prices both from $f_1^{j,t\text{-}1}$ to $f_2^{j,t\text{-}1}$ and from $f_2^{j,t\text{-}1}$ to $f_3^{j,t\text{-}1}$. This market effect is statistically significant at 1% level. Fourth, the positive coefficient representing the impact of market conditions in the evolution of futures prices from $f_2^{j,t\text{-}1}$ to $f_3^{j,t\text{-}1}$ is greater than that of $f_1^{j,t\text{-}1}$ to $f_2^{j,t\text{-}1}$.

	Model 1A: $f_2^{j,t\text{-}1} - f_1^{j,t\text{-}1}$		**Model 1B:** $f_3^{j,t\text{-}1} - f_2^{j,t\text{-}1}$	
Parameter	Coefficient	*t-stat*	Coefficient	*t-stat*
Intercept	0.0296	*2.85****	0.0788	*4.45****
w_t	-0.0501	*-4.58****	-0.1281	*-6.88****
m_t	0.0079	*5.47****	0.0223	*9.01****
Adjusted R^2	0.19		0.37	
Observations	220		220	

Table B1. Linear regression results demonstrating the impact of weather and market conditions on the evolution of futures prices. *** denotes statistical significance at 1%.

Evolution of Bottled Wine Prices

Using the same notation, we develop predictive analytics models in order to examine the impact of the upcoming vintage's weather condition and market fluctuations on the price evolution of bottled wine. We develop the following linear regression models designated as Model 2A and Model 2B, respectively, where $t = \{2008, 2009, 2010, 2011, 2012\}$ and $j = \{1, 2, \ldots, 44\}$:

$$(b_2^{j,t-2} - b_1^{j,t-2}) = \theta_0 + \theta_1 w_t + \theta_2 m_t + \varepsilon_{j,t} \tag{33}$$

$$(b_3^{j,t-2} - b_2^{j,t-2}) = \lambda_0 + \lambda_1 w_t + \lambda_2 m_t + \varepsilon_{j,t}. \tag{34}$$

Table B2 provides the regression results of the impact of new summer weather and market information on the evolution of bottle prices described as $b_2^{j,t-2}$ (in Model 2A) and $b_3^{j,t-2}$ (in Model 2B). Three conclusions can be made from the analysis. First, weather conditions of the upcoming vintage (i.e., the value of w_t) does not have a statistically significant effect on the evolution of bottle prices. This holds true when prices evolve from $b_1^{j,t-2}$ to $b_2^{j,t-2}$ and from $b_2^{j,t-2}$ to $b_3^{j,t-2}$. Second, improving market conditions during the summer of calendar year t (with a higher value of m_t) has a positive impact on the evolution of bottle prices both from $b_1^{j,t-2}$ to $b_2^{j,t-2}$ and from $b_2^{j,t-2}$ to $b_3^{j,t-2}$. Third, the positive coefficient representing the impact of market conditions in the evolution of bottle prices from $b_2^{j,t-2}$ to $b_3^{j,t-2}$ is greater than that of $b_1^{j,t-2}$ to $b_2^{j,t-2}$.

	Model 2A: $b_2^{j,t-2} - b_1^{j,t-2}$		**Model 2B:** $b_3^{j,t-2} - b_2^{j,t-2}$	
Parameter	Coefficient	*t-stat*	Coefficient	*t-stat*
Intercept	0.0248	*1.52*	0.0187	*0.53*
w_t	-0.0082	*-0.59*	0.0245	*0.82*
m_t	0.0059	*2.19***	0.0255	*4.43****
Adjusted R^2	0.01		0.12	
Observations	220		220	

Table B2. Linear regression results demonstrating the impact of weather and market conditions on the evolution of bottle prices. ** and *** denote statistical significance at 5% and 1%, respectively.