Forecasting the Value of Fine Wines

Joseph L. Breeden, auctionforecast.com

ABSTRACT

Fine wines have gained attention globally as an investment opportunity with possible diversification benefits relative to more traditional investments and potentially high rates of return. Using a database from auctionforecast.com with over two million prices from auctions globally, we analyzed the price dynamics of fine wines to quantify predictive factors for investors. Brand, vintage, ratings, auction houses, bottle size, age of the wine, market trends, and more are considered in this analysis. Although data mining techniques are common in other applications, long-range forecasting requires a careful separation between current discrimination factors and long-term drivers.

In this article we demonstrate how one must handle these issues in model development. The models used are a combination of Age-Period-Cohort models and traditional scoring techniques. Expanding upon previous work, this study explores the ability to predict prices for infrequently traded wines using known lifecycle and market attributes along with measures of brand and vintage value for the specific wine.

This approach has direct applicability to consumer behavior in many industries and has specifically been used to great success in retail lending.

INTRODUCTION

Fine wines are both a luxury consumable and an investment category. Whether buying for pleasure or investment, data analysis can provide guidance on reasonable price ranges for purchase or sale. Recent improvements in data availability and data analysis methods have moved wine investment toward a quantitative activity, rather than a purely qualitative activity.

Infrequently traded wines are challenging for wine value forecasting. Auction houses need estimates for such wines when negotiating consignments, and wine buyers would like to have more than intuition for bidding on such wines. For frequently traded wines, age-period-cohort models and similar hedonic pricing models can make effective long-range forecasts of price. For rarely traded wines, the goal is to use a forecast model based upon brand value and vintage value instead.

Because we use the APC algorithm to model constituent drivers of price sensitivity, it falls within the class of hedonic regression (Rosen, 1974). The APC modeling technique is also similar to repeat sales regression (Bailey, Muth, and Nourse, 1963), which assumes that the log appreciation rate equals the log appreciation rate of an environment plus an error term. Log differences of repeat sales are regressed on time dummies. Although similar in concept to looking at recurring sales, the APC model in this case is applied directly to price values rather than to price differences. We make this choice when setting up the APC algorithm because of the nonuniformity of the time data.

Not all hedonic regressions of wine prices consider all three dimensions of age, vintage, and calendar date, nor do they all include the detailed treatment of the implicit specification error in vintage data. Therefore, we hope our employment of APC models provides a useful example of how to model age, vintage, and time effects while managing the linear specification error.

AGE-PERIOD-COHORT MODELS

Age-Period-Cohort (APC) models are nothing new. They evolved from the Lexis diagrams created in the 1890s as a way to view mortality trends for different cohorts. In mortality, a cohort is just defined as the year of one's birth. In the 1960s and 70s, techniques were developed to analyze aggregate time series of cohort performance in order to measure the main drivers of performance.

Although begun with mortality studies, APC models became most popular in studies of epidemiology and demography. This paper seeks to introduce APC models to a broader audience and demonstrate the

breadth of applications possible through specific examples from banking, fine wines, SETI@home, and tree rings. In addition, we describe how they could be applied to many other areas.

In the case of mortality modeling, the probability of death should follow a binomial distribution, so an APC model could be defined as

$$log\left(\frac{p(a,p,c)}{1-p(a,p,c)}\right) = F(a) + G(c) + H(p)$$

where the left side of the expression is the logit transformation of the probability of death $p(a,c,p)$, a is the age at death, c is the birth cohort, and p is the period (calendar date) at death. Three functions are estimated, one each for age, cohort, and period, $F(a)$, $G(c)$, and $H(p)$ respectively. Note also that $a = p - c$, the age of the entity can be computed by subtracting the cohort date from the period.

The reason is that these functions are estimated without explanatory factors. In cases where the root causes are unknown, the model can still measure the impact of the "environment" on a given calendar date, variation by cohort, and timing versus age.

Stated more generally, APC models model some rate by estimating functions of age, cohort, and time with a link function appropriate for the distribution of the event. In other areas of application, the terms age a, vintage v, and time t are used in place of age, cohort, and period.

$$r(a,v,t) = Link\big(F(a) + G(v) + H(t)\big)$$

Intuitively, the corresponding functions as estimated for some rate can be described as:

Lifecycle, $F(a)$: This quantifies the expected rate as a function of age at time of event. The magnitude is scaled to the rate being modeled. In later examples, this could be the probability of default on a loan versus the age of the loan or the probability of making an online purchase versus the age of the customer account.

Vintage quality, $G(v)$: A measure of variation by the origination date (a.k.a. cohort or vintage date) for a group of people, customers, loans, etc. For a consumer loan, it would measure the net credit risk by vintage after normalizing for lifecycle and environment. For online sales, it could be relative propensity to purchase versus the date or even time when the customer first registered. The magnitude is measured in terms of the relative change in the rate where the zero level is the average and the vintage estimates may be roughly normally distributed about that.

Environment, $H(t)$: A measure of the net impact of all factors that impact performance as a function of calendar date. In most business applications, one thinks of this including macroeconomic drivers and any management policy changes that affect all customers simultaneously. The magnitude is again scaled as the relative change in rate with zero representing the average.

SIMILARITY TO OTHER TECHNIQUES

The basic concept of APC models is similar to survival and hazard models. In all these methods, an entity is being monitored for the probability of some event. The age of the entity is a key determinant of the event rate. For example, part failures tend to increase versus the age of the part. This is referred to as the hazard or lifecycle function.

Survival models measure the survival function as the probability that an entity will survive to a certain age. The hazard function is the probability that an event will occur at a given age conditioned on the event not having already occurred. Thus, survival models capture one of the three dimensions (age) of the APC models. Survival or hazard functions are generally estimated nonparametrically. In SAS this can be done with

```
proc lifetest
```

Cox Proportional Hazard models (Cox PH) are an extension to survival models that include factors that adjust the probability of the event. The usual notation is

$$\lambda(a|\boldsymbol{X_i}) = \lambda_0(a)exp(\boldsymbol{X_i} \cdot \boldsymbol{\beta})$$

where $\lambda(t|X_i)$ is the event rate conditioned on scaling factors X_i specific to entity i and β are the coefficients to be estimated. Cox showed that the estimation of the coefficients is independent of the estimation of the hazard function, so he developed a partial likelihood optimization procedure for estimating the coefficients. In SAS this is found in

```
proc phreg
```

The Cox PH coefficients could be used to estimate the vintage or environment function parameters, but are more commonly applied to explanatory factors. The challenge is that Cox PH was not developed with the three dimensions of age, vintage, and time in mind, where a linear relationship exists between these dimensions. Therefore, the APC framework gives us more explicit control over the estimation of these functions, which becomes of critical important during forecasting.

Model Estimation

Before estimating the three APC functions, we must again consider the relationship $a = t - v$, age equals time minus vintage. If we start by assuming that the three functions of age, vintage, and time are completely general, then they must have constant, linear, and nonlinear parts. Without losing any generality, this can be expressed as

$$F(a) = \alpha_0 + \alpha_1 a + F'(a)$$
$$G(v) = \beta_0 + \beta_1 v + G'(v)$$
$$H(t) = \gamma_0 + \gamma_1 t + H'(t)$$

Substituting into our APC equation, we get

$$p(a, v, t) \sim \alpha_0 + \alpha_1 a + F'(a) + \beta_0 + \beta_1 v + G'(v) + \gamma_0 + \gamma_1 v + H'(t)$$

This equation has a few obvious problems. α_0, β_0, and γ_0 are the constant terms for the three respective functions, however, we cannot simultaneously estimate three constant terms in the same equation. Therefore we define a single constant term $\alpha_0' = \alpha_0 + \beta_0 + \gamma_0$. This means that the constant term is being included in the age function and all other functions are estimated relative to the age function. This is equivalent to what is done in Cox PH where the hazard function is scaled to the historic probability and the Cox PH regression coefficients are measured relative to it.

In addition, because $a = t - v$, α_1, β_1, and γ_1 cannot all be estimated independently. Only two of the three linear trends can be estimated independently. To resolve this, an assumption must be made about the allocation of the linear components among the three functions. In applications where the rate being modeled is stationary through time, we may assume that $\gamma_1 = 0$ so that the other two linear terms can be estimated uniquely.

Holford (1983) explains well that a decision must be made about the trend ambiguity in any vintage analysis, whereby only two linear components can be estimated from the three available dimensions. This decision is always domain-specific. Breeden, Bellotti, and Yablonski (2015) also explained that this problem occurs for any model of vintage data. Methods like Cox PH provide unique solutions only because of trend allocation assumptions embedded within the estimator. Breeden and Thomas (2016) provide specific solutions for models that incorporate macroeconomic factors.

Holford also proves that aside from the constant and linear term considerations, all other components of the functions are uniquely estimable. Therefore, the primary issue in using APC models is to determine the linear trend allocation. For the analyses shown here, sufficient history is available for us to assume that the environment function has no linear trend. Therefore, the following assignments are made for the functions.

$$F(a) = c_0 + c_1 a + F'(a)$$
$$G(v) = c_2 v + G'(v)$$
$$H(t) = H'(t)$$

ESTIMATORS

APC models come in many flavors (Yang and Land, 2013). In all implementations, the functions are estimated without reliance on external explanatory factors. Rather, the functions are estimated either parametrically via splines or some other basis functions, or nonparametrically.

In SAS, APC models may be estimated using the following procedures

- Spline estimation is available via `proc transreg`.
- Bayesian estimation (Schmid and Held, 2007) is available via `proc genmod`.
- Partial least squares estimation is available via `proc pls`.
- Ridge regression estimation is available via `proc reg (w/ridge= option)`.

These techniques are compared for a sample dataset in Figure 1.

Spline Estimation

The most common implementations use spline estimation to approximate the three functions, but several nonparametric estimation techniques exist as well. Splines are piecewise polynomial function approximations, where the user must specify the number and potentially location of the spline nodes. In the author's experience, fewer nodes are effective for the age function, since the prior assumption is usually that it should change smoothly with age. For the time and vintage functions, the nodes are best distributed according to the density of the data, and with as many nodes a supportable by the data. Both of these functions can have frequent discontinuities, because of sudden changes in the environment or in origination conditions for the vintages.

Bayesian Estimation

A Bayesian estimation algorithm was proposed by Schmid and Held (2007). Their algorithm estimates the functions nonparametrically, meaning that each point of the lifecycle, environment, and vintage functions is a separate parameter estimated without the smoothness constraint of the spline algorithm. It can use an initial prior for those functions, but then uses a Monte Carlo estimation procedure to refine the functions. Because the Monte Carlo simulation can be very slow, a quick spline estimation can serve as a useful prior to accelerate convergence.

The Bayesian result looks "noisier", but as a nonparametric estimation it has the ability to better estimate the complex structure that can arise in the functions. For example, sudden shocks may occur in the environment function due to operational failures. Similarly, the vintage function may have seeming discontinuities because of changes in underwriting policies.

For Bayesian estimation, the essential parameters are the number of Monte Carlo simulations used to create the estimate and the step size for each simulation.

Partial Least Squares

As an alternative to Bayesian estimation, one can also use partial least squares (PLS). PLS creates a set of principle components, or eigenvectors, on which the regression is performed. The primary control is in the number of components to be used in the estimation.
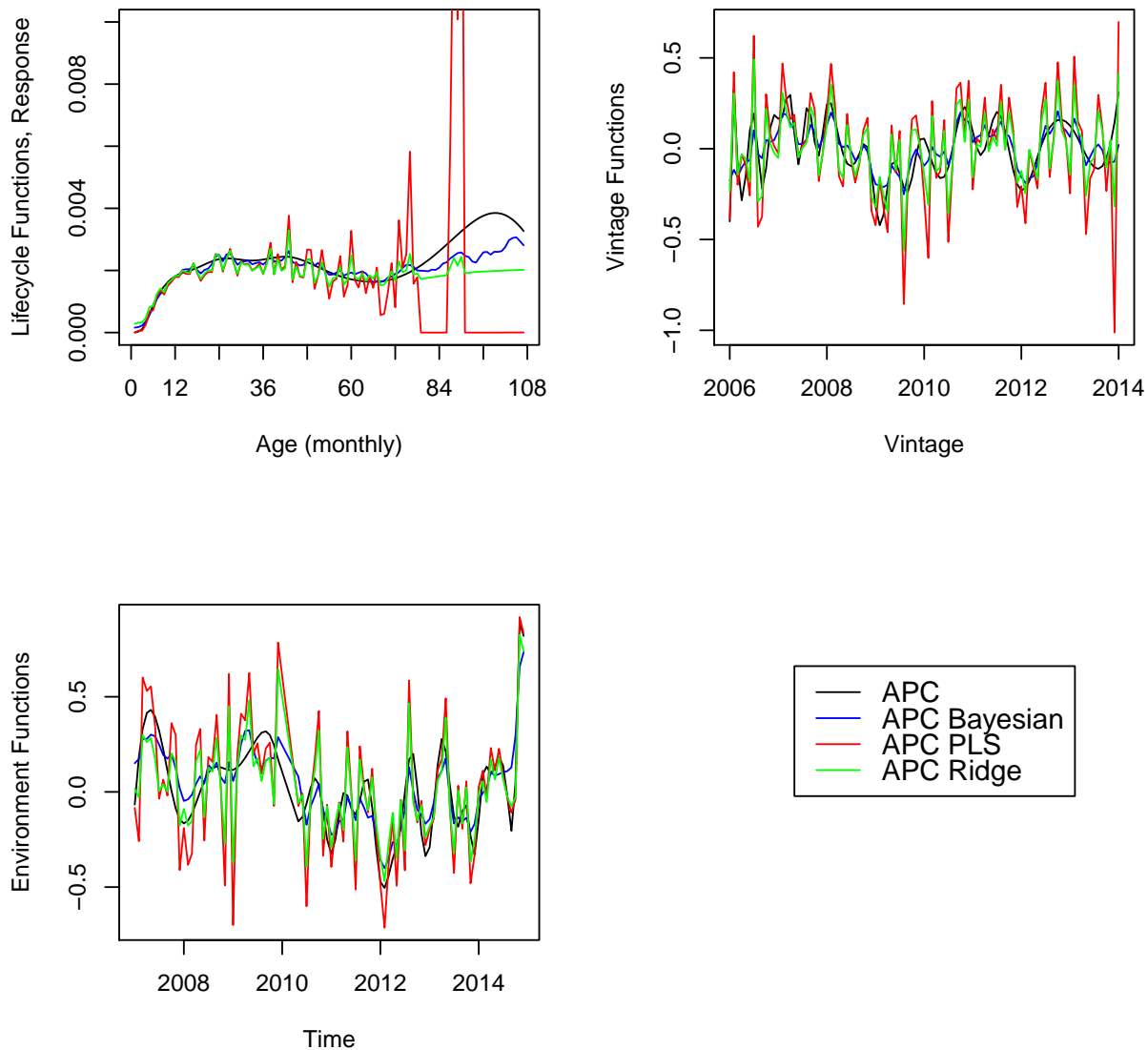
Figure 1: A comparison of APC estimation techniques as applied to modeling probability of default (PD) for a $400 million auto loan portfolio.

Ridge Regression

Ridge regression is the last technique we have used for APC estimation. It can also be an effective estimation method, although on some data sets it may have convergence problems depending on the step size used for the optimization.

In the above example, the different estimation techniques are compared. This example suggests that one could rank the smoothness of the result as APC (spline), Bayesian, Ridge regression, and PLS. However, on other data sets, we have observed a reordering of these.

One important point of comparison is how they perform for recent vintages with few observations. Spline estimation is generally unstable for these vintages whereas Bayesian estimation can use a starting prior of an average vintage and slowly diverge as observations accrue.

DATA

For the application to wine analysis, a dataset from auctionforecast.com was used to study auction prices for Bordeaux wines at 10 different auction houses: Acker Wines, Bidford, Bonhams, the Chicago Wine Company, Christies, Langton's, Sothebys, Spectrum, Veiling Sylvies, and Zachy's. 640,000 auction prices spanning a period from 2001 to 2016 and vintages from 1970 to 2011 were used for the analysis. To be considered for the test, a wine must have at least 10 auction results for other vintages, so that the needed models could be created.

AGE-PERIOD-COHORT ANALYSIS OF WINE

Leveraging previously published work, the auction prices were analyzed using an age-period-cohort model to measure the price dynamics by age of the wine, calendar date (market effects), and a specific wine factor. The lifecycle versus age and market versus date are common to all Bordeaux wines. The specific wine factor is discarded in favor of the comparative analysis done here.

Since the auction prices follow a lognormal distribution, the following for was used for APC analysis.

$$\log\big(price(a, v, t)\big) = F(a) + G(v) + H(t)$$

When applied to auction price data for fine wines, the lifecycle function, F(a), measures the expected average price for a wine in a segment as a function of the age of the wine. Thus, the lifecycle shows the expected rate of appreciation in a wine's value across different spans of age.

The vintage function, G(v), captures how much higher or lower a given wine is priced relative to the average lifecycle for the segment. This allows for the estimation of separate price scaling by vintage while maintaining a common market index (environment function) and common lifecycle function across all Lafite wines.

The environment function, H(t), measures how much auction prices are above or below the expected lifecycle values on a given calendar date.

Lifecycle

The lifecycle estimation for price versus age of the wine was first done with a Bayesian APC estimator (Figure 2A). The lifecycle function was then smoothed for business purposes, but the confidence intervals are preserved.

The lifecycle shows that the average auction price actually declines until the 5th or 6th year, at which point the prices stabilize and begin to rise again. The same effect is seen for all wines at auction with some variation in where the bottom occurs.

The most rapid price increases occur in the couple decades after that minimum before slowing their rate of appreciation throughout the remaining lives of the wines. The cumulative price appreciation between 5 and 25 years of age is 81% or 3.2% annually. The importance of these estimates is that they are cleaned of changes in market conditions and represent the performance of average Lafite wines cleaned of differences in specific vintage performance. Although 81% appreciation sounds impressive, 3.2% annual capital appreciation for the wine is less exciting when considering transaction costs, storage costs, inflation, etc.

Vintage

The vintage function (Figure 2B) captures the overall trend in increasing prices for newer Lafite vintages, and the exceptional prices for some vintages. The highest spikes in the graph correspond to the 1982, 2000, 2009 and 2010 vintages, which aligns well with industry expertise.
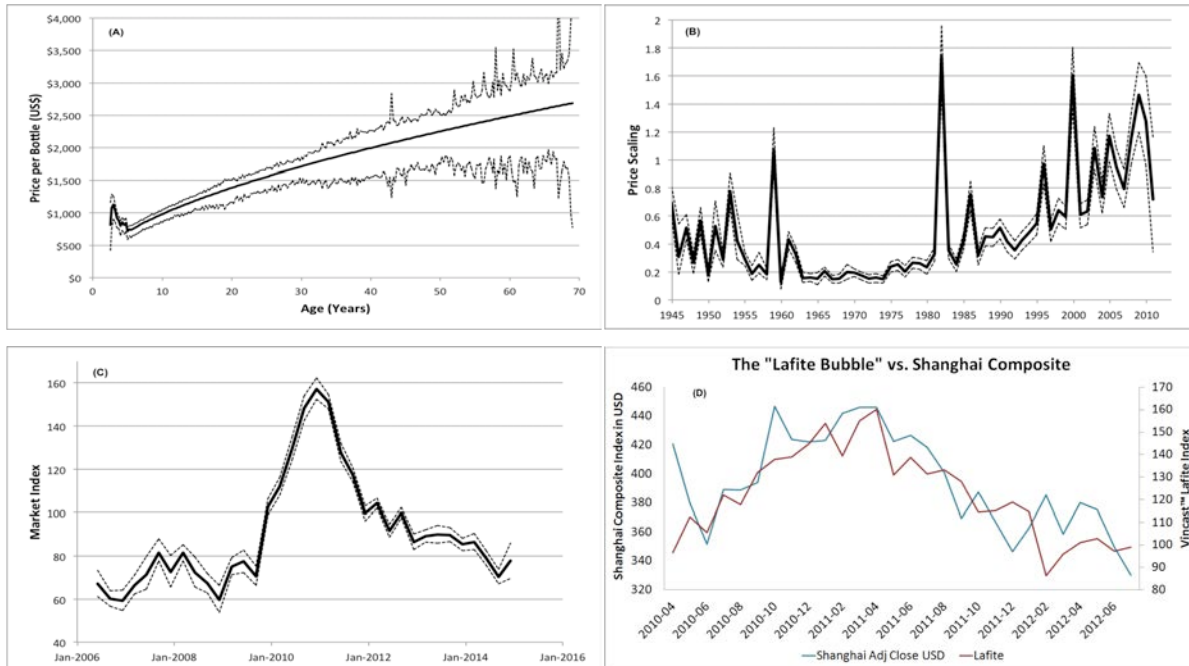
Figure 2: The APC decomposition for Château Lafite Rothschild wines: (A) lifecycle, (B) vintage, and (C) environment. Then an overlay of the environment versus the Shanghai Composite stock index

Environment

The environment function measures in the decomposition provide significant new insights. Many wine market indices are available, expressed as baskets of specific vintages. Like a stock market index, these baskets can be changed over time to swap in newer vintages, but wine-basket indices have are problematic as measures of the wine market. As seen in the lifecycle analysis, wines appreciate over time. Therefore, even if buyer interest is flat, a wine basket index will continue to rise unless manually readjusted. Further, these baskets generally have a small number of select wines and therefore do not capture the broader market conditions. Therefore, when annualized returns are quoted for specific wine portfolios, we cannot immediately conclude the causes of the appreciation, whether inherent to the wine or due to market trends.

A market index (Vincast Lafite Index, VLI) was created from the environment function for ease of interpretation by financial analysts as

$$VLI(t) = \exp\big(H(t)\big) * 100$$

A value of 100 represents the historic average for the market. The traditional approach to gauging the wine market is to create a basket of top wines and track their value. However, as the analysis here shows, that approach confounds the normal appreciation from the lifecycle with changes in market conditions. APC provides a market index that can leverage all wines auctioned, not just those in a select list, and is normalized for the natural appreciation in the value of the wines over time.

In contrast, the environmental functions in Figure 2C are also market indices, but with broad coverage (all vintages with a minimum number of auction results, set to 16 for robustness) and normalized for lifecycle and vintage effects. Normalizing by lifecycle means that the appreciation discussed in the previous section, including inflation is removed automatically. Normalization by vintage means that prices for highly valuable wines are adjusted to a measurement of changes that is comparable to price movements in less valuable vintages.

The market index for Lafite clearly shows the peak in early 2011 known as the "Lafite Bubble". From June 2010 to Feb 2011, prices for Lafite wines, adjusted for lifecycle and vintage effects, jumped by roughly

50%. By April 2013 the environment function shows a decline to levels below the June 2010 start of the bubble.

ADDING AUCTION FACTORS

After estimating the Age-Period-Cohort model, a secondary panel regression model was created that retained the lifecycle and environment from the APC model, but then added factors for auction house, location of auction (country), and bottle size. These results were reported in the previous study. This provided a baseline model for everything that could be known except the quality of the specific wine. That is the subject of the next analysis.

ILLIQUID WINE ANALYSIS

On top of the baseline model, two models were created. The first model (brand-vintage) estimates factor values by brand and vintage, excluding all data for the wine being forecast. Brand refers to the name of the wine ignoring vintage, so a specific producer may have multiple brands. Note that the brand-vintage model is not using any history for the specific wine, other for other vintages of the same brand and the same vintage for other brands. This approach required the brand-vintage models to be re-estimated every time the test wine was changed, because excluding that wine from the dataset would potentially affect the estimates of both the brand and vintage coefficients.

The second model (history-only) uses previous auctions prices for a specific wine to predict future auction prices for that wine. This is a simple regression model using the baseline forecast and the previous auction results for a given wine as inputs in order to estimate the price scaling for the wine. To create a realistic test, the "next auction price" was assumed to be no sooner than four weeks, allowing time for data gathering and analysis in a real-world setting. This filter specifically excludes cases where multiple lots for the same wine at the same auction will carry highly correlated prices. Although useful to know on that day, it has no bearing on the current goals of longer range forecasting.

RESULTS

The model forecast accuracies were compared as the difference between absolute percentage errors for each auction. Given the auction-level forecast errors, a regression model was created to predict the errors versus factors that should affect the accuracy of the two models: the amount of historic data used for modeling brand value, vintage value, price per bottle, popularity of the wine in terms of number of trades per year, and the number of previous auction values available for the model.

The brand-vintage model accuracy grows relative to the historic-data model as the log of the number of previous observations of brand value and vintage. However, the historic-data model accuracy grows with increasing price per bottle and increasing frequency of auctions for that wine. Looking at the standardized coefficients, the strongest error predictor was the log of the amount of auctions used to measure brand value. The second most important was the amount of training data for estimating the vintage value, followed by the price per bottle. Log transformations were used for all of these factors in order to maximize forecast accuracy. Other error measures were tested. No change in the overall results was found, but the difference of absolute percentage errors was found to be the most predictable measure, probably because of the reduced sensitivity to outliers.

CONCLUSION

The overall finding is that brand-vintage models are affective for predicting prices of illiquid (infrequently traded) wines, but historic-data models work well for the most popular and highest priced wines. The latter wines are assumed to be the ones that are so closely studied by the market that models are barely needed to predict price. All of these results make good intuitive sense. The results also suggest that an optimal approach would be to combine both models into a single forecast, which will also be tested later in order to assess how much potential benefit can be gained.

Future analysis will repeat the above study for Burgundy, California, and Australian wines to confirm that the same patterns hold throughout.

REFERENCES

Auctionforecast.com. 2018. "Auction Prices." Accessed January, 2018. http://www.auctionforecast.com.

Bailey, M. J., Muth, R. F., and Nourse, H. O. 1963. "A regression method for real estate price index construction". *Journal of the American Statistical Association*, 58, 933–942.

Breeden, J. L. 2010. *Reinventing Retail Lending Analytics: Forecasting, Stress Testing, Capital and Scoring for a World of Crises*. London: Risk Books.

Breeden, J.L., A. Bellotti, and A. Yablonski. 2015. "Instabilities using Cox PH for forecasting or stress testing loan portfolios". *Proceedings of the Credit Scoring and Credit Control Conference XIV, Edinburgh, 2015*.

Breeden, J. L. and S. Liang. 2017. "Auction-Price Dynamics for Fine Wines from Age-Period-Cohort Models". *Journal of Wine Economics*, 12:173-202.

Breeden, J.L. and L.C. Thomas. 2016. "Solutions to Specification Errors in Stress Testing Models", to appear in Journal of the Operational Research Society, 2016, doi:10.1057/jors.2015.97.

Holford, T.R. 1983. "The Estimation of Age, Period and Cohort Effects for Vital Rates", *Biometrics*, Vol. 39, No. 2 pp. 311-324

Rosen, S. 1974. "Hedonic prices and implicit markets: Product differentiation in pure competition". *Journal of Political Economy*, 82(1), 34–55.

Schmid V.J. and L. Held. 2007. *Journal of Statistical Software*, Volume 21, Issue 8. "Bayesian Age-Period-Cohort Modeling and Prediction – BAMP".

Yang, Y. and K. C. Land. 2013. *Age-Period-Cohort Analysis: New Models, Methods, and Empirical Applications.* New York: Chapman & Hall / CRC.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Joseph L. Breeden, PhD
Auctionforecast.com
breeden@auctionforecast.com
www.auctionforecast.com