The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1751-1062.htm

How heterogeneous wine lots result in a price discount relative to homogeneous lots

Joseph L. Breeden
AuctionForecast.com LLC, Santa Fe, New Mexico, USA

Received 10 December 2022
Revised 10 May 2023
18 July 2023
Accepted 22 July 2023

Abstract

Purpose – The purpose of this study is to determine whether the fine wine market is efficient between homogeneous lots and heterogeneous lots.

Design/methodology/approach – Auction price data for homogeneous (or solid) lots of fine wines was analyzed to create price prediction models. Those models were used to predict the expected auction price for the bottles within heterogeneous lots. Lastly, models were created to explain and predict the differences between expected and realized prices for heterogenous wine lots.

Findings – The results show that large inefficiencies exist. The more complex and expensive the heterogeneous lot, the greater the discount relative to what would have been realized if the bottles had been sold individually. This discount can exceed 50% of the expected auction price.

Practical implications – Heterogeneous lots may arise as a practical requirement from the auction house. Restaurant buyers probably have little interest in such lots because of the inclusion of wines the restaurant will be unable to sell. Collectors may be uniquely positioned to benefit from this price discount.

Originality/value – These results are unique in the literature, because the price dynamics of heterogeneous (or mixed) lots of fine wines have not previously been studied.

Keywords Wine investment, Auction prices, Age-period-cohort models, Market efficiency

Paper type Research paper

1. Introduction

Market prices for specific wine labels and wine market indices are always based upon data gathered from the sale of homogeneous (or solid) lots. A lot may have multiple bottles of wine, but for a homogeneous (or solid) lot, they are all bottles of the same wine label: same brand, vintage, size and preferably the same condition. Homogeneous lots are statistically easy to work with, but they do not reflect all lots sold at auction. Heterogeneous lots contain a mixture of labels. These can be a "vertical" lot where multiple vintages of the same wine are sold together, a "horizontal" lot where multiple wines of the same vintage are combined or what we will term a "diverse" lot where no intrinsic relationship is apparent.

This article seeks to answer the following questions. Do heterogeneous lots increase or decrease the total auction price as compared to what would have been expected were they sold individually? Can we say that markets are efficient if a bottle is sold individually or as

JEL classification – C23, D44, G11, G12, Q11

International Journal of Wine Business Research
Vol. 36 No. 1, 2024
pp. 1-13
Emerald Publishing Limited
1751-1062
DOI 10.1108/IJWBR-12-2022-0047

part of a heterogeneous lot? Can we identify possible drivers of such price disparities? Specifically, we look at the impact of anchoring a heterogeneous lot with one or a few high-value wines accompanied by less value wines to quantify if this has a net benefit to the auction price.

Section 2 provides a literature review on wine analysis. Section 3 describes the data used in the study. Section 4 describes the forecasting approach used to predict single-bottle prices in-sample and apply that to heterogeneous lots. Section 5 reviews the results of the study. Section 6 provides conclusions on market efficiency with heterogeneous lots.

2. Literature review

The earliest studies of the prediction of fine wine prices were conducted by Ashenfelter (1989, 2010). In a recent survey of wine price prediction, Outreville and Le Fur (2020) described over 100 studies.

Collectible wine prices have been studied to determine the suitability of wine as an investment (Burton and Jacobsen, 2001; Sanning *et al.*, 2008; Fogarty, 2010; Masset and Henderson, 2010). For portfolios only containing fine wines, the investment return on wine portfolios is compared with other investments (Burton and Jacobsen, 2001; Fogarty, 2010; Jaeger, 1981; Krasker, 1979). More recently, Sanning *et al.* (2008) found large excess returns; however, any study immediately prior to the 2011 market correction is likely to be overly optimistic. As with most markets, one must consider long periods to capture both rallies and corrections. Fogarty and Sadler (2014) nicely summarize previous studies in their Table 1, appropriately concluding that they have obtained a wide variety of results depending upon the time period examined and method of analysis applied.

When a diversified portfolio is created with multiple asset classes, fine wines may be included to provide diversification (Bouri *et al.*, 2018; Chu, 2014). A lower return series that is somewhat uncorrelated to other assets is also beneficial. So, any return modestly greater than zero could be beneficial if uncorrelated or, better yet, anti-correlated to other assets. Initial studies did find low or negative correlation to typical investment vehicles, especially stocks and bonds (Fogarty, 2010; Masset and Henderson, 2010; Masset and Weisskopf, 2018; Sanning *et al.*, 2008).

All of this preceding wine auction analysis has focused on realized prices for homogeneous lots. No comparable literature is available for heterogeneous lots of fine wines. Optimal pricing of product bundles has long been studied in the economics literature, including Stigler (1968), McAfee *et al.* (1989) and Chen and Riordan (2013), to name a few. Whereas early work focused on bundling by monopolies, auctions create interesting dynamics not found in the study of monopolies. Looking to the auction literature, Palfrey (1983) studied bundling of inequivalent

Metric	Lots size (bottles)	Label dominance	Lot price (US$)
Min.	1	0.094	120
First quart.	4	0.404	700
Median	6	0.524	1,053
Mean	7.2	0.512	1,420
Third quart.	10	0.605	1,600
Max.	72	1	143,175

Note: Label dominance is the ratio of the highest value sublot to the total lot, where a sublot is a group of bottles of the same label
Source: Table by the author

Table 1. Summary statistics for the continuous variables on heterogeneous (mixed) lots

items, showing that auctioning multiple distinct items in a single lot generated superior returns as compared to auctioning the items separately if there were only two bidders. Chakraborty (1999) generalized that result and found a critical number of bidders above which inefficiencies were created, which lowered the final auction price. Interestingly, a report by Brilliant (1999) on the auction of a mixed lot of music manuscript sheets commended that the lot, "[. . .] suffered the usual fate of mixed lots by being bought in by the auction house, in spite of modest estimates [. . .]" The phrase "bought in" means that it did not meet the reserve price and was not sold. This comment about the "usual fate" of mixed lots seems to agree with the findings of Chakraborty, assuming that auctions usually have enough potential bidders to be above the critical number.

To estimate any auction price discount for heterogeneous lots relative to homogeneous lots, the first step is to build a model for the homogeneous lots. Age-period-cohort (APC) models (Mason and Fienberg, 1985; Holford, 1983; Holford, 2005) are being used here as the foundation for the homogeneous lot analysis, because they directly estimate the dependence of wine prices on the age of the wine, market conditions by calendar date and variations by vintage. In fact, a simple extension of APC to wine forecasting is to make the vintage (cohort) function a set of coefficients by wine label, incorporating both vintage and brand. Wood and Anderson (2006) previously demonstrated that the age of a wine was predictive of price, although age has been used as a linear or polynomial factor in most wine price studies. APC models provide more resolution by offering nonparametric estimation of price versus the age of the wine.

APC modeling and the subsequent incorporation of auction details to predict price can be considered a type of hedonic regression [Rosen74]. APC modeling is also similar to repeat sales regression (Bailey *et al.*, 1963), with the nuance that we are not tracking repeat sales of the exact same bottles but treating bottles of a certain label approximately as a commodity. Repeat sales regression 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. Because repeat sales of specific labels can be spread randomly over many years, the APC model is being used to model price directly, rather than as a rate of change.

Dimson, Rousseau and Spaenjers examined the impact of aging on wine prices and, by using a value-weighted arithmetic repeat-sales regression over 1900–2012, examined the long-term return for high-end wines. They created annual price histories for vintages of Haut-Brion, Lafite-Rothschild, Latour, Margaux and Mouton-Rothschild and analyzed those to discover the price appreciation lifecycle with wine age, environmental impacts with calendar date and included production and quality as explanatory variables to explain vintage variation.

3. Data

The training data for homogeneous lot prices was taken from 2,450,000 auction prices from AuctionForecast.com spanning 14 auction houses (Acker Wines, Bidforwine, Bonhams, Christie's, Chicago Wine Company, Hart Davis Hart, Langton's, Skinner, Sotheby's, Spectrum, Veiling Sylvie's and Zachy's) from February 2003 through July 2022. For this analysis, a subset of 609,000 Bordeaux and 293,000 Burgundy auction prices were modeled. Prices are not adjusted for inflation, but they are adjusted for exchange rates at the time of the auction to convert all prices to US dollars. Figure 1 summarizes the number of homogeneous lots by the main categorical variables. Figure 2 is a box plot of price per bottle for lots aggregated by auction year.

The data for heterogeneous lots was also sourced from AuctionForecast.com, which contains 6,946 heterogeneous lots from Acker, Bonhams, Hart-Davis-Hart and Zachy's from March 2005 through June 2022. Within the heterogeneous lots, 30,886 sublots with a single

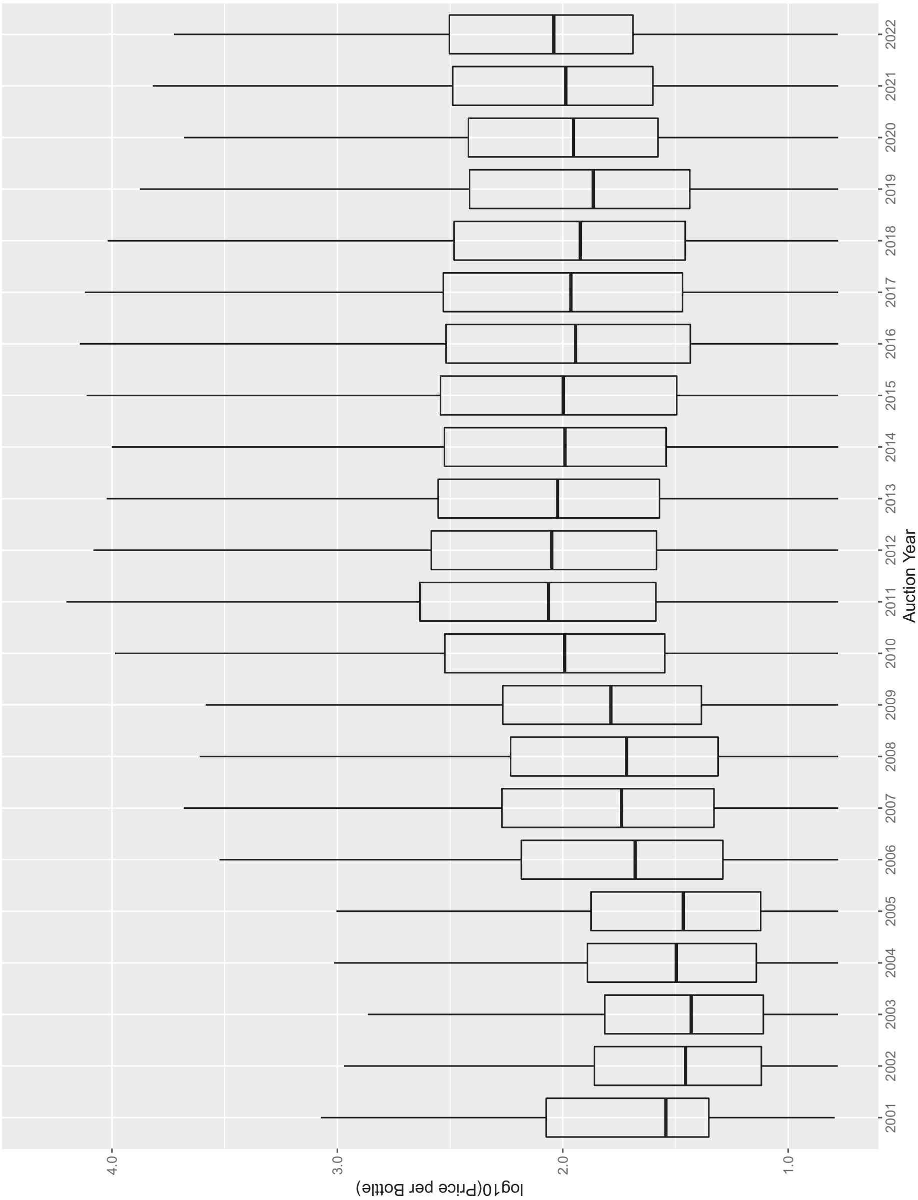
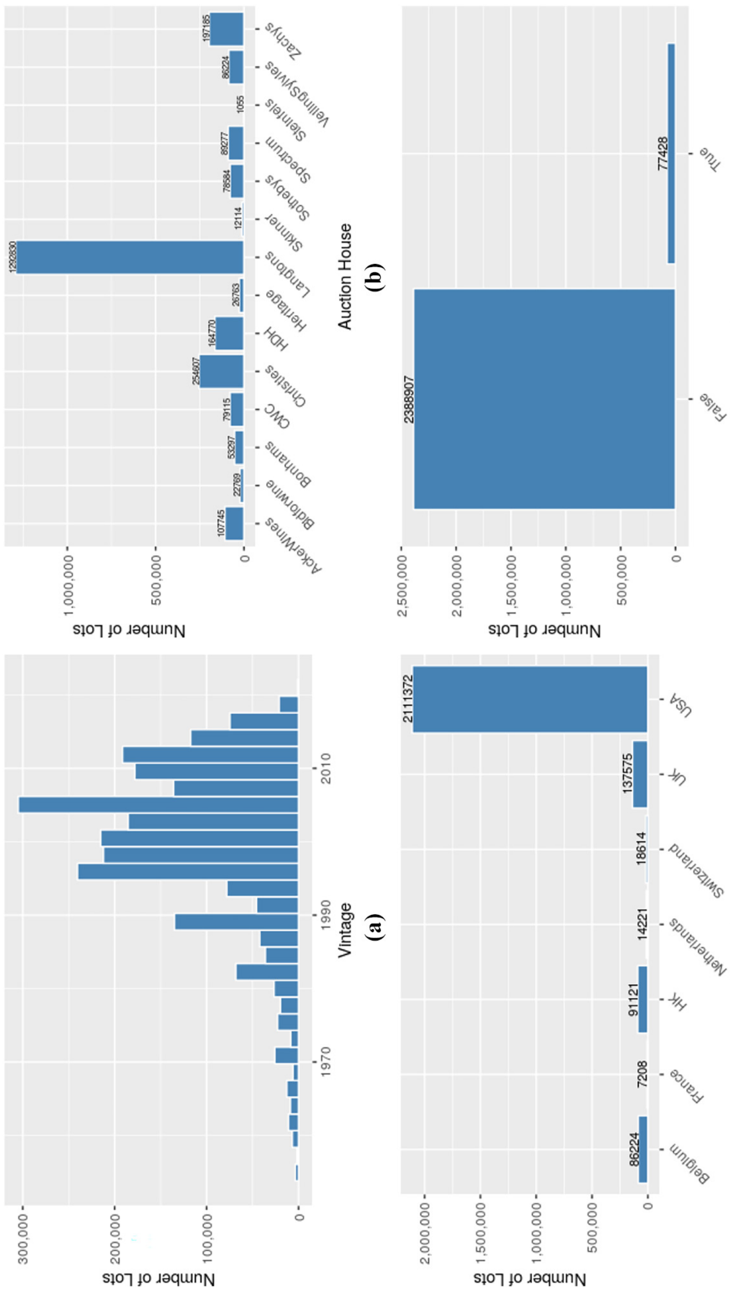


Figure 1.
Summary statistics for the price per bottle for homogeneous lots

Source: Figure by author



Notes: (a) Vintage year; (b) auction house; (c) auction location; (d) online auction

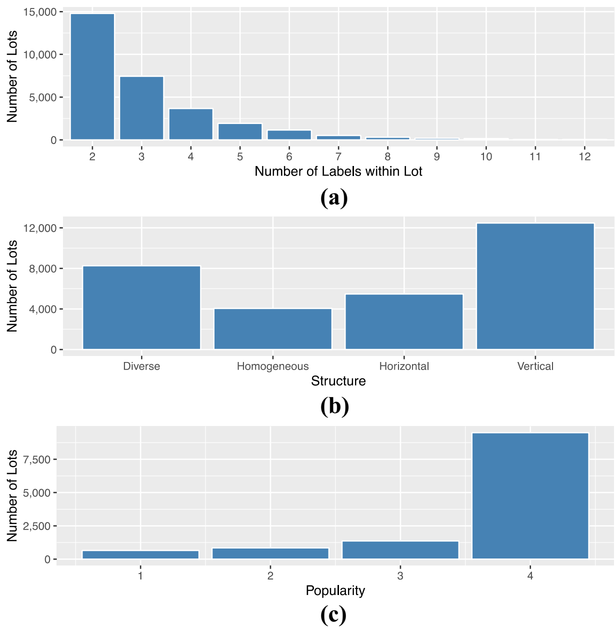
Figure 2.
Summary statistics on categorical variables for the homogeneous lot data

label and bottle size were identified. A sublot is defined as a collection of one or more bottles of a label within a larger heterogeneous lot. The forecast models will be applied at the sublot level. Figure 3 shows data density plots for the categorical variables for the heterogeneous lots. Table 1 provides summary statistics for the continuous variables.

For the heterogeneous lots, some explanatory factors were derived from the details of the lot. Label dominance was defined as the maximum sublot forecast divided by the total lot forecast. As label dominance approaches 1.0, it means that all of the expected auction value for the heterogeneous lot is concentrated in one label. Popularity is an indicator variable for the popularity tier of the most popular wine. Popularity tiers are taken from the previous work by Breeden (2022). If one of the labels was a Top 25 label, the lot was assigned Tier 1. The least popular wines were in the 500+ segment, and a lot containing only those was assigned Tier 4. Lot structure refers to whether it is horizontal, vertical or diverse, according to the definitions in the introduction. The number of labels in a lot was considered.

4. Analysis methods

The modeling process starts by creating models of log(price.per.bottle) from homogeneous lots sold at auction. Those models are used to forecast log(price) for the bottles within a heterogeneous lot. Multivariate regression models were then created to explain the price



Notes: (a) Number of labels within Lot; (b) structure; (c) popularity; Horizontal = multiple wines of the same vintage, Vertical = multiple vintages of the same wine, and Diverse = an unstructured collection of labels. Popularity refers to the trading tier of the label: Tier 1 = Top 25, Tier 2 = 26 to 100, Tier 3 = 101 to 500, Tier 4 = 500+

Source: Figure by author

Figure 3. Structure refers to the design of the mixed lot

discount observed in actual auction prices of heterogeneous lots relative to the predicted price from the homogeneous lots models.

4.1 Homogeneous lots models

The homogeneous lots forecast models (HLM) are based upon an APC approach (Breeden and Liang, 2017; Breeden, 2010) that quantifies log(price.per.bottle) versus three nonparametric functions: the age of the bottle denoted as $F(a)$, the market index at the time of the auction denoted as H(t), a quality adjustment function $G(v)$ of the bottle's vintage v and the specific lot labeled as i:

$$\log(p(i, a, v, t)) = F(a) + G(v) + H(t) \tag{1}$$

APC decomposition is often performed with spline basis functions. The analysis in this paper uses Bayesian estimation of the functions represented nonparametrically with one parameter at each point in age, vintage and calendar date, as described in Schmid and Held (2007). Bayesian APC was preferred here to avoid premature smoothing from spline approximations. Note that most research on wine price appreciation lifecycles use linear or simple polynomial functions of age, which miss the detailed structure and can destabilize for large ages where less data is available.

Previous work has shown that not all labels experience the same price appreciation lifecycle. Instead, segmentation by trading popularity distinguishes lifecycle shapes for label price appreciation. The segmentation used here follows the work and segment definitions in Breeden (2022). That work performs a similar APC analysis of trading frequency to measure the relative popularity of labels.

APC analysis is usually viewed as a form of panel regression. The nuance with wine price analysis is that we are not observing repeated sales of the same bottle, but rather repeat sales of bottles that are assumed to be similar enough to be treated as repeat observations of a commodity. With differences in provenance, this may not be perfectly true, but is a reasonable approximation. APC models can be viewed as similar to hedonic regression (Rosen, 1974) and repeat sales regression (Bailey *et al.*, 1963).

The vintage function from APC has no value in predicting label prices, but is included in the analysis to eliminate possible frailty in the estimates of lifecycle and environment (Balan and Putter, 2020). "Frailty" is an effect where sample bias from an unobserved variable can distort the shape of a hazard function in survival analysis. The same principle arises here and is resolved during the lifecycle estimation by including a fixed effect (dummy variables) for vintage years.

A second-stage model is created that retains $F(a)+ G(v)$ as a fixed input and estimates coefficients for each of the N_{labels} and auction factors of bottle size, auction house, location and whether it was online:

$$\log\big(price\ \ per\ \ bottle(a, t, p_i, k) = F(a) + H(t) + c_1 bottle\ \ size_i + c_2 house_i + c_3 location_i \\ + c_4 online_i + b_k \tag{2}$$

In equation (2), p_i are the values of bottle size, auction house, location and online for homogeneous lot or sublot i. For each label, a unique intercept is estimated, b_k, to adjust for label-specific pricing power. Previous research on infrequently traded wines has shown that with as few as two or three auction results, the b_k in equation (2) can be estimated more reliably than with overall estimates by brand and vintage.

This final homogeneous lot model is most similar to the work of Dimson, Rousseau and Spaenjers (Dimson *et al.*, 2015), where the impact of aging on wine prices was studied using a value-weighted arithmetic repeat-sales regression over 1900–2012 for vintages of Haut-Brion, Lafite-Rothschild, Latour, Margaux and Mouton-Rothschild. Their model considered lifecycle with wine age, environmental impacts with calendar date, production and quality as explanatory variables to explain vintage variation.

4.2 Heterogeneous lot discount

As will be seen in the results, heterogeneous lots sell for a significant discount relative to the expected price from the homogeneous lot model. Given the available information on the composition of the lots, the analysis tested whether useful models could be created to predict the amount of discount.

The discount for a heterogeneous lot was measured in two ways. First, as a log-discount, equation (3), the discount is the log of the ratio of predicted to actual price:

$$log - discount(i) = log\left(\frac{estimated\ \ price(i)}{actual\ \ price(i)}\right) \tag{3}$$

In the economics literature, the dependent variable for modeling discounts is often defined as the percentage discount, equation (4):

$$percentage\ \ discount(i) = \left(\frac{estimated\ \ price(i)}{actual\ \ price(i)}\right) - 1 \tag{4}$$

The discount models are estimated using ordinary least squares regression, which assumes normal distributions. When the discounts are small, the difference between log discount and percentage discount is negligible, and the models will be roughly equivalent. In this case where the discounts were significant, the non-normality of a distribution of percentage errors becomes problematic. Further, percentage errors are asymmetric, meaning that if discounts of 10% and -10% are added, the result is not 0. For log discount, $0.1 + (-0.1) = 0$. To properly model an asymmetric dependent variable such as percentage discount, a nonlinear model would be required. Prior modeling experience has shown that using log-discount as the dependent variable solves both of these problems, resulting in better forecast accuracy within a simple linear construct. This premise will be tested in the current context.

With the chosen dependent variable for discount denoted, y, a regression equation will be created which considers lot structure: vertical, horizontal, diverse or homogeneous.

The models are compared using the median absolute percentage error (MAPE). R^2 and F-statistics are not comparable across different dependent variables, whereas MAPE is based upon the forecasts generated by the model, equation (5). MAPE is the median of the percentage forecast errors for lots i:

$$MAPE = Median_i\left[\left(\frac{predicted\ \ price(i)}{actual\ \ price(i)}\right) - 1\right] \tag{5}$$

The $predicted\ \ price(i) = HLM(i) + predicted\ \ discount(i)$, where predicted discount converts the predicted log discount or percentage discount to dollars.

5. Results

Summed sublot forecasts from the homogeneous lot models averaged 62% higher than actual heterogeneous lot prices. This indicates that realized prices were substantially less than would be expected based upon an in-sample forecast of the value of the labels within the lot when incorporating the auction house, wine age and market environment at the time of that auction.

5.1 Predicting log discount

To understand this seeming market inefficiency, several predictive factors were explored. When segmented by auction house, no difference was observed. However, factors measuring lot structure, label dominance, the number of unique labels, the total number of bottles and label popularity were all predictive. An initial feature design phase carefully studied each of the potential input factors in a series of univariate models. Continuous variables were finely binned to develop an understanding of the nonlinearities present, which led to the use of logarithmic transformations for number of labels, lot size in bottles and HLM-forecasted price. These univariate models are not shown, but they informed the design choices in the multivariate models.

Table 2 shows the details of two candidate models, each created to predict log discount. The key distinction between the models is the trade-off between label dominance in Discount Models 1 and log(Price Forecast) and popularity tier in Discount Model 2. Overall, the model with log(Price Forecast) and popularity was more accurate, but with more data, label dominance may yet have unique information.

The price predicted by the homogeneous lot model was the greatest predictor of discount. The realized price for the lot was negatively correlated to the log of the expected value. Expensive mixed lots sell for the greatest discount. This suggests that buyers are less willing to take chances on what should otherwise have been expensive bottles individually,

Discount model 1		Discount model 2	
(Intercept)	2.912 (44.60)***	(Intercept)	2.373 (35.78)***
log(Price.Forecast)	−0.479 (−66.33)***	log(Price.Forecast)	−0.447 (−60.26)***
Label.Dominance	−0.152 (−3.78)***	Popularity = Tier.1	*Reference level*
Structure.Horizontal	*Reference level*	Popularity = Tier.2	−0.048 (−1.72)
Structure.Vertical	0.034 (2.40)*	Popularity = Tier.3	−0.169 (−6.54)***
Structure.Diverse	−0.197 (−12.04) ***	Popularity = Tier.4	−0.225 (−9.76)***
		log(Num.Labels): Structure.Horizontal	−0.113 (−5.09)***
		log(Num.Labels): Structure.Vertical	−0.077 (−3.98)***
		log(Num.Labels): Structure.Diverse	−0.256 (−14.00)***
		log(Lot.Size)	0.275 (26.81)***
Observations	5,865	Observations	5,865
Adj R^2	0.446	Adj R^2	0.540
MAPE	28.2%	MAPE	26.3%

Notes: Log(Price.Forecast) is the logarithm of the lot price predicted from the homogeneous lot model (HLM). Label dominance is the ratio of the highest value sublot to the total lot, where a sublot is a group of bottles of the same label. Structure refers to the design of the mixed lot. Horizontal = multiple wines of the same vintage, Vertical = multiple vintages of the same wine and Diverse = an unstructured collection of labels. Popularity refers to the trading tier of the label: Tier 1 = Top 25; Tier 2 = 26–100; Tier 3 = 101–500 and Tier 4 = 500+; ***; ** and * denote statistical significance at the 0.1; 1 and 5% levels, respectively. Robust *t*-ratios are reported in parentheses
Source: Table by the author

Table 2. Summary of the final model to predict the discount for heterogeneous lots using log discount as the dependent variable

and for expensive lots, the realized price is substantially less than it should have been. Perhaps, such lots are created when the provenance of the bottles is less certain, or at least when buyers perceive such risks.

The structure of the heterogeneous lot was important to the result in Table 2. This variable worked best when interacted with the log of the number of labels in the lot, log (Num.Labels). The results show that having more unique labels in a lot increases the discount. However, the price is less sensitive to the number of labels for vertical and horizontal lots, presumably because the added structure makes the pricing more intuitive for bidders. Coefficients for vertically and horizontally structured lots were statistically equivalent. Note that in the regression equation, horizontally structured lots served as the reference level, meaning that the assigned coefficient is 0.

The popularity of the labels in the lot was based on the segmentation by trading frequency. Popularity is a discrete variable with four levels, each with a coefficient predicting the discount. Lots in Popularity Tier 1 had the least discount, whereas Popularity Tier 4 had the most. This suggests that the most familiar wines trade closest to their expected value when in a heterogeneous lot.

Even with the number of labels in the model, the lot size in bottles was also significantly predictive. Univariate analysis showed that the relationship between lot size and discount fit very well to a logarithmic function. Logarithmic functions appear empirically several times in this modeling, suggesting that this is a basic feature of how human traders process information and perceive risk.

Discount Models 1 and 2 in Table 2 had MAPEs of 28.2% and 26.3%, respectively. This compares favorably to a MAPE of 37.1% that is obtained when a log-discount model is created with only a constant term, meaning the original HLM forecast is used with only an added constant. Thus, Discount Model 2 is a 29.1% improvement over the adjusted HLM forecast and a 62.3% reduction compared with the unadjusted HLM forecast.

5.2 Predicting percentage discount

Using percentage discount as the dependent variable, the preceding analysis was rerun. The same predictive structure holds and the revised coefficients are given in Table 3. The coefficients changed because of the change in dependent variable definition, but the observed model design supported the same interpretation as above.

As seen in Table 3, MAPE for Discount Model 3 was 30.2%. This is an improvement over the adjusted HLM result, but a 14.8% increase in forecast error relative to the log discount model in Table 2. The increased forecast error is easily understood from the less ideal nature of predicting percentage discount.

Note that even though using MAPE means the performance measure is percentage error, directly predicting percentage error is not the best way to create a linear regression model. With a significantly more nonlinear model or machine learning method, this distinction usually disappears. Linear regression was used in creating the discount models, because the primary goal was to develop an intuitive understanding of the dynamics of heterogeneous lot trading.

6. Conclusions

Heterogeneous (or mixed) lots do not create a perception of greater value or "drag along" wines of lesser value to boost the realized price. Rather, they appear to have the opposite effect of creating uncertainty in the true value of the lot and therefore return lower overall value. Without detailed data or models on the expected value of the heterogeneous lot, potential buyers are uncertain, and uncertainty decreases perceived value. The effect

	Discount model 3
(Intercept)	0.864 (23.43)***
log(Price.Forecast)	−0.189 (−45.82)***
Popularity = Tier.1	*Reference level*
Popularity = Tier.2	−0.017 (−1.12)
Popularity = Tier.3	−0.078 (−5.46)***
Popularity = Tier.4	−0.066 (−5.13)***
log(Num.Labels): Structure.Horizontal	−0.076 (−6.17)***
log(Num.Labels): Structure.Vertical	−0.056 (−5.16)***
log(Num.Labels): Structure.Diverse	−0.131 (−12.89)***
log(Lot.Size)	0.119 (20.92)***
Observations	5,865
Adj R^2	0.390
MAPE	30.2%
	0.148

Notes: Log(Price.Forecast) is the logarithm of the lot price predicted by the homogeneous lot model (HLM). Structure refers to the design of the mixed lot: Horizontal = multiple wines of the same vintage, Vertical = multiple vintages of the same wine and Diverse = an unstructured collection of labels. Popularity refers to the trading tier of the label: Tier 1 = Top 25; Tier 2 = 26–100; Tier 3 = 101–500 and Tier 4 = 500+; ***; ** and * denote statistical significance at the 0.1; 1 and 5% levels, respectively. Robust *t*-ratios are reported in parentheses
Source: Table by the author

Table 3. Summary of the final model to predict the discount for heterogeneous lots using percentage discount as the dependent variable

becomes more severe with more complex lots and more expensive lots. The effect is less severe when the lot has some structure, such as a common vintage (horizontal lot) or a common wine with multiple vintages (vertical lot).

After obtaining these results, some wine auction houses were consulted for their interpretations. The market inefficiencies were not considered surprising, as those who will spend significant amounts on a lot prefer the clarity of homogeneous (or solid) lots. Nevertheless, mixed lots persist because of decisions around maintaining auction flow and creating lots that will not be so small as to be dominated by a minimum buyer's premium.

Further, much of the historic auction data lacks information on condition or is free text that would require specific analysis to incorporate. Without being able to confirm the condition of the bottles, one possible explanation is that bottles with condition problems are incorporated into mixed lots, which buyers know or suspect and consider in their bids.

As anecdotally reported, another contributing factor may be that restaurant buyers, which are a significant component of the market, are much less likely to bid on mixed lots, because these are likely to include labels they do not expect to be able to resell.

Although larger buyers such as premier restaurants may avoid mixed lots, this analysis shows that a discount may be present in heterogeneous lots for more adventurous collectors. The wine sellers are expected to want the highest possible return on the auction, but auction house requirements about auction flow and minimum lot value may force sellers into a discounted situation. Further, infrequently traded wines may be included in heterogeneous lots, because neither buyer nor seller knows the true value, and the inclusion is actually intended as a discount from the seller in hopes of raising interest.

Given the perspectives of sellers, auction houses, restaurant buyers and collectors, these discounts are likely to persist until all parties have access to better real-time predictive tools that can increase the available information.

References

Ashenfelter, O. (1989), "How auctions work for wine and art", *Journal of Economic Perspectives*, Vol. 3 No. 3, pp. 23-36.

Ashenfelter, O. (2010), "Predicting the quality and prices of Bordeaux wine", *Journal of Wine Economics*, Vol. 5 No. 1, pp. 40-52.

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*, Vol. 58 No. 304, pp. 933-942.

Balan, T.A. and Putter, H. (2020), "A tutorial on frailty models", *Statistical Methods in Medical Research*, Vol. 29 No. 11, pp. 3424-3454.

Bouri, E., Gupta, R., Wong, W.K. and Zhu, Z. (2018), "Is wine a good choice for investment?", *Pacific-Basin Finance Journal*, Vol. 51, pp. 171-183.

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

Breeden, J.L. (2022), "How sample bias affects the assessment of wine investment returns", *Journal of Wine Economics*, Vol. 17 No. 2.

Breeden, J.L. and Liang, S. (2017), "Auction-price dynamics for fine wines from age-period-cohort models", *Journal of Wine Economics*, Vol. 12 No. 2, pp. 173-202.

Brilliant, I. (1999), "Beethoven auction report (may 1998-May 1999)", *The Beethoven Journal*, Vol. 14 No. 1, p. 33.

Burton, B.J. and Jacobsen, J.P. (2001), "The rate of return on investment in wine", *Economic Inquiry*, Vol. 39 No. 3, pp. 337-350.

Chakraborty, I. (1999), "Bundling decisions for selling multiple objects", *Economic Theory*, Vol. 13 No. 3, pp. 723-733.

Chen, Y. and Riordan, M.H. (2013), "Profitability of product bundling", *International Economic Review*, Vol. 54 No. 1, pp. 35-57.

Chu, P.K.K. (2014), "Study on the diversification ability of fine wine investment", *The Journal of Investing*, Vol. 23 No. 1, pp. 123-139.

Dimson, E., Rousseau, P.L. and Spaenjers, C. (2015), "The price of wine", *Journal of Financial Economics*, Vol. 118 No. 2, pp. 431-449.

Fogarty, J.J. (2010), "Wine investment and portfolio diversification gains", *Journal of Wine Economics*, Vol. 5 No. 1, pp. 119-131.

Fogarty, J.J. and Sadler, R. (2014), "To save or savor: a review of approaches for measuring wine as an investment", *Journal of Wine Economics*, Vol. 9 No. 3, pp. 225-248.

Holford, T. (1983), "The estimation of age, period, and cohort effects for vital rates", *Biometrics*, Vol. 39 No. 2, pp. 311-324.

Holford, T. (2005), "Age-period-cohort analysis", *Encyclopedia of Statistics in Behavioral Science*, Wiley.

Jaeger, E. (1981), "To save or savor: the rate of return to storing wine", *Journal of Political Economy*, Vol. 89 No. 3, pp. 584-592.

Krasker, W.S. (1979), "The rate of return to storing wines", *Journal of Political Economy*, Vol. 87 No. 6, pp. 1363-1367.

McAfee, R.P., McMillan, J. and Whinston, M.D. (1989), "Multiproduct monopoly, commodity bundling, and correlation of values", *The Quarterly Journal of Economics*, Vol. 104 No. 2, pp. 371-383.

Mason, W.M. and Fienberg, S. (1985), *Cohort Analysis in Social Research: Beyond the Identification Problem*, Springer.

Masset, P. and Henderson, C. (2010), "Wine as an alternative asset class", *Journal of Wine Economics*, Vol. 5 No. 1, pp. 87-118.

Masset, P. and Weisskopf, J.P. (2018), "Raise your glass: wine investment and the financial crisis", in *World Scientific Reference on Handbook of the Economics of Wine: Volume 1: Prices, Finance, and Expert Opinion*, pp. 271-295.

Outreville, J.-F. and Le Fur, E. (2020), "Hedonic price functions and wine price determinants: a review of empirical research", *Journal of Agricultural and Food Industrial Organization*, Vol. 18 No. 2.

Palfrey, T. (1983), "Bundling decisions by a multiproduct monopolist with incomplete information", *Econometrica*, Vol. 51 No. 2, pp. 463-484.

Rosen, S. (1974), "Hedonic prices and implicit markets: product differentiation in pure competition", *Journal of Political Economy*, Vol. 82 No. 1, pp. 34-55.

Sanning, L.W., Shaffer, S. and Sharratt, J. (2008), "Bordeaux wine as a financial investment", *Journal of Wine Economics*, Vol. 3 No. 1, pp. 51-71.

Schmid, V.J. and Held, L. (2007), "Bayesian age-period-cohort modeling and prediction – BAMP", *Journal of Statistical Software*, Vol. 21 No. 8.

Stigler, G.J. (1968), "A note on block booking", *The Organization of Industry*, Vol. 165.

Wood, D. and Anderson, K. (2006), "What determines the future value of an icon wine? New evidence from Australia", *Journal of Wine Economics*, Vol. 1 No. 2, pp. 141-161.

Further reading

Storchmann, K. (2012), "Wine economics", *Journal of Wine Economics*, Vol. 7 No. 1, pp. 1-33.

Corresponding author

Joseph L. Breeden can be contacted at: breeden@deepfutureanalytics.com