Selection Neglect and the Cross-Section of Wine Returns*

Robbe Van Tillo
KU Leuven

Gertjan Verdickt
University of Auckland

First draft: November 1, 2023

This draft: July 19, 2024

Abstract

This paper investigates the asset-pricing implications of a violation of "the shared information assumption." We turn to the concept of selection neglect, which is the behavioral tendency characterized by decision-making on censored data. Using a dataset of more than 3 million wine transactions, we document negligent investors bear expected utility losses, and more exposed wines exhibit higher future returns. The evidence supports the information uncertainty hypothesis. Investor attention, time-varying preferences, or limits to arbitrage cannot explain these findings. We argue that selection neglect increases the price inefficiencies of infrequently traded assets, such as fine wine.

KEYWORDS: alternative investments, risk premia, wine, selection neglect, auctions
JEL CLASSIFICATION: G11, G12, G13, G14, G15, L66

---

* We are grateful to Henk Berkman, Frank De Jong, Alex Dickerson, Rik Frehen, Neal Galpin, Marina Gertsberg, David Hirshleifer, Dong Huang, Eric Le Fur, Jonathan Lewellen, Marta Khomyn, Kim Oosterlinck, Lyndon Moore, Helen Popper (discussant), Tarik Roukny, Joshua Shemesh, Christophe Spaenjers, Francesco Stradi, Jens Van Mele, Michael Weber, Rudiger Weber, Johannes Weytjens (discussant), Yongxin Xu, Adam Zaremba, and participants at Adelaide Business School, Annual Global Finance Conference (Cagliari, '24), Auckland Business School, the Belgian Financial Research Forum (Brussels, '24), KU Leuven, Open Universiteit (Heerlen), Macquarie University, Monash Business School, University of Antwerp, University of Queensland, UNSW Business School, and the World Finance Conference (Cyprus, '24).

We thank the assistance of Zofia Dalkowska (Sotheby's), Jochen De Vylder (Ampersand), Manuela Duroo, Wallace Kwong (British Library), Gloria Heesen, Calvin Hensgens, Victoria Lagae, Giovanna Magri, Hannah Merki, Chloë Paesen, Rakesh Sathiya, Albert Yan, Chengming Yang, Kevin Xu, Jingqi Xu, and James Weber.

Corresponding author: gertjan.verdickt@auckland.ac.nz. All remaining errors are our own.

1. Introduction

A central tenet of finance theory is the notion of 'rational expectations,' where economic agents are assumed to optimally process all information when forming beliefs and making (financial) decisions. However, a large body of empirical evidence has documented a departure from full rationality, including limited attention (Bordalo et al., 2013) and extrapolative beliefs (Barberis et al., 2015). While this past work has explored heterogeneous beliefs between investors, it has maintained the premise of a shared information set. Recently, Eyster et al. (2019) introduced a theoretical model in which investors neglect information. They document that deviations from this shared information assumption could lead to excessive trading. Motivated by this, we test the implications of such deviations on return expectations and portfolio selection.

One way differences in information sets manifest themselves is through selection neglect. This refers to a tendency of economic agents to base actions on biased observations. Agents observe a censored sample and infer properties of the population from it (Tversky & Kahneman, 1971). They do not correct for selection biases, as they would as econometricians (Barron et al., 2019; Heckman, 1979). Finance theory argues that this can lead to suboptimal (investment) strategies (e.g., Hirshleifer & Plotkin, 2021; Jehiel, 2018). Hence, this raises two empirical questions: Does selection neglect impact an investor's performance on the portfolio level, and is it priced on an asset level?

Studying selection neglect presents a challenge outside the laboratory (e.g., Koehler & Mercer, 2009; López-pérez et al., 2022). People could see specific information but choose not to act, thus mimicking neglect. As such, we need a more objective metric of differences in information sets. We focus on two return methods to approach the issue. First, we calculate the changes between subsequent observed prices. This only considers successful trades (i.e., a censored sample), so we define it as biased returns. Second, we correct these returns through a Markov Chain Monte Carlo (MCMC) model that considers the probability of trading and estimates latent prices even when no trade occurred (Korteweg & Sorensen, 2010). We refer to this as the selection-adjusted returns, defined as the change between these latent prices.

We use the wine market as a laboratory for our analysis. Wine is traded in sequential auctions. This implies that the supply side is observable and fixed, as sellers must submit their lots (and reserve prices) prior to trading (Backus & Lewis, 2024). This leads to two (potential) issues for investors. First, only the successful lots have observed prices. Investors with pessimistic views would not be able to trade, leading to sidelined investors (Goetzmann & Massa, 2005). This, in turn, would lead to upward pricing biases (Korteweg et al., 2016). Second, there is endogeneity in trading (Lovo & Spaenjers, 2018). Also, not every wine is offered at every auction event. The lack of trading also contains information (Giglio & Shue, 2014). If investors ignore these issues, they have to form expectations from asymmetric information (Eyster et al., 2019; Hirshleifer & Plotkin, 2021; Hong & Stein, 1999).

Over recent decades, the wine market has grown substantially as an asset class (Dimson et al., 2015). It has features similar to traditional financial assets, such as corporate bonds.[1] However, there are important differences relative to standard asset markets. First, wine produces no cash flows.[2] Similar to housing, dividends from owning wines come from consumption.[3] Therefore, investor expectations can be formed only from (past) prices.[4] This could enhance the impact of selection neglect. Second, wine lacks the institutional framework of traditional assets (Masset & Weisskopf, 2015).[5] For instance, buying large amounts of bonds is more efficient than buying large amounts of wine. Subsequently, selling short stocks is far easier than selling short wines. Hence, a lack of arbitrage opportunities ensures the existence of (potential) behavioral biases. Finally, there is limited data on production processes, wine supply, and quality. This enhances information frictions (Akerloff, 1970). Overall, the wine market provides an interesting lens to study the asset-pricing implications of selection neglect in an opaque information environment with slow information diffusion, private information, and short-selling constraints.

We introduce a new dataset of wine auction transactions from 41 auction houses. This database encompasses 3.3 million transactions between January 2003 and December 2022. We excluded

---

[1] Winemakers (firms) produce a fixed amount of bottles from a specific vintage (fixed dollar amount for a specific bond issue), they have a quality assessment (rating), a drinking window (maturity), are traded through auctions (secondary market), and have buy-and-consume agents (buy-and-hold investors) that take liquidity out of the market.
[2] Investors can only disagree on the discount rate channel, as no (future) cash flows stem from this asset. If investors miss past transactions, they have limited information to form expectations.
[3] In the case of wine, consumption is also a 'nonspeculative trade motive.' However, even in this setting, it is important to understand pricing dynamics, as consumers arguably do not want to overpay for wine.
[4] Since we require a small number of transactions to accurately calculate 'the biased returns' on an asset level, other real assets (e.g., art or housing) may not be perfectly suited. Nevertheless, our concepts are not wine-specific and can be generalized to other (real) assets.
[5] Fractional equity can make real assets more appealing to a wider audience (Whitaker & Kräussl, 2020).

wine with a price below $20 and above $50,000 to ensure that market microstructure effects do not drive the results. This leaves us with 142,439 individual bottles. We merge this with rating and drinking window information from Wine Advocate (by Robert Parker). This allows us to control for quality, maturity, and other characteristics that help to explain the cross-section of returns (Dickerson et al., 2023; Hadj Ali et al., 2008; McManus et al., 2013).

To test whether informational differences harm investment performance on the portfolio level, we introduce two investors: a negligent and non-negligent trader. Both apply a mean-variance optimization to obtain the tangency portfolio, which offers the highest Sharpe ratio among all combinations. The difference between them is the information they use to obtain this portfolio. Respectively, they use biased returns and selection-adjusted returns to calculate these optimal portfolio weights. In other words, in our framework, investors are exposed to selection neglect in overall *portfolios*. On average, we highlight that negligent investors give up performance by suffering from selection neglect, as measured by certainty-equivalent returns. Expected utility losses range from 0.4% to 0.8% per month. In other words, the differences in information could result in suboptimal investment strategies. This supports Jehiel's (2018) predictions.

Following Huang and Goetzmann (2023), we measure selection neglect on a wine-level as the previous 12-month difference between the biased and selection-adjusted return. We define this as *Selection Bias* (SB). In this framework, the investors are exposed to selection neglect for *future* wine-level gains and losses by looking at its *past* performance. Next, we sort wine into quintiles based on its SB value and calculate next-month returns to test two competing hypotheses. The sideline investor hypothesis argues that a high belief dispersion leads to current overvaluation, followed by low future returns (e.g., Cen et al., 2017; Diether et al., 2002; Goetzmann & Massa, 2005; Hillert et al., 2018). In turn, the information uncertainty theory implies that informational differences positively affect the subsequent returns (Hirshleifer & Plotkin, 2021; Hong & Stein, 1999; Jiang & Sun, 2014).

Wine is an appealing test asset for the two hypotheses due to the unique characteristics of its market structure and inherent properties. Auction markets (naturally) lead to a scenario where some investors are "sidelined" - those who participated but are unsuccessful in their bids. This aligns with the sideline investor hypothesis, as these investors may form opinions (about wine) based on observed auction results, potentially leading to belief dispersions. Furthermore, wine exhibits information uncertainties, such as future demand, provenance, production processes, quality, and quantity. This leads to an environment where investors could have different levels of knowledge or access to information. We can observe how these competing theories play out in financial markets using wine as a test asset.

Our empirical results support the information uncertainty hypothesis. We highlight that high-SB wines generate positive returns and low-SB wines earn negative returns, both significant at the 1% level. The return spread on the zero-cost strategy of buying high-SB and selling low-SB wine is economically large: 7.3% per month (significant at the 1% level). The return differences cannot be explained by traditional stock and bond and novel wine factors (based on illiquidity, momentum, and the excess wine market). The intuition is clear: Wines with positive SB values are appealing to investors, as they are overly optimistic about such wines due to the neglect of negative information. That is, negligent investors experience more positive returns (i.e., biased returns) than their non-negligent counterparts (i.e., selection-corrected returns). This arguably translates to overconfidence and, as a result, higher future returns.

To ensure that the selection neglect effect is not a repackaging of other anomalies, we construct double-sorted portfolios and wine-level regressions. Overall, Selection Bias retains significant explanatory power for future wine returns, even after we control for other wine characteristics. Furthermore, we find strong evidence of a long-term relationship between Selection Bias and future returns up to 12 months ahead. In other words, the evidence re-confirms the predictions of the information uncertainty hypothesis.

Other variables could explain the effect we highlight. First, we test if selection neglect interacts with momentum, as shown by Huang and Goetzmann (2023). To proxy for extrapolation, we calculate the wine's momentum value (defined as the cumulative returns over the previous 11 months ending two months before the current month). We document there are strong selection neglect effects among wines with weak and strong momentum. It is, however, more prominent for lower quintiles. As such, we establish the negative relationships between extrapolation and future returns in Da et al. (2023) and McManus et al. (2013).

Second, investor attention cannot explain the relationship between selection biases and future returns. We show that the effect is weaker for wines exhibiting more salient features and those with high abnormal returns. Interestingly, these are two metrics for which one needs to know the market return. In other words, one must expand its censored sample to experience salience and abnormal return. Similarly, wines featured in the Top 100 of Wine Spectator exhibit a more significant effect of selection bias on future returns. As such wines arguably attract attention – thus potentially restrict the sample one is looking at – they enhance the selection neglect effect.

Finally, we document a substantial relation between selection bias and future returns for wines with greater limits to arbitrage. The effect on future returns is more substantial for wines with high idiosyncratic risks. This aligns with the results of traditional assets, as Shleifer and Vishny (1997) posited that assets with higher idiosyncratic risks exhibit higher mispricings. Moreover, wines with a longer expiration time and more illiquidity have higher future returns. Similarly, this aligns with the evidence from traditional assets (Weber, 2018). Overall, trading constraints are an issue in the real and private-valued asset market. Indeed, investors cannot (easily) short-sell wine. Limits to arbitrage would, therefore, increase price inefficiencies (David et al., 2013). As such, we conclude that selection neglect further decreases the informational component of prices.

Our analysis reveals that selection neglect becomes more pronounced during higher economic uncertainty and elevated tail risks. This result indicates that selection neglect serves as a hedge against uncertainty. Furthermore, we document that the effect is amplified in the post-COVID-19 period. We argue that a shift to online events exacerbates information asymmetry (Brogaard et al., 2023; De Felice et al., 2022) and amplifies overconfidence (Barber & Odean, 2002). In sum, these combined results raise questions about price inefficiency, especially in light of behavioral biases. Indeed, we conclude that selection neglect increases inefficiencies, and is distinct from other behavioral biases, such as limited attention or extrapolative beliefs.

We contribute to the literature in several ways. First, the existing literature on selection neglect is either theoretical (Brundage et al., 2022; Hirshleifer & Plotkin, 2021; Jehiel, 2018; Ngangoué & Weiszacker, 2021) or experimental (Barron et al., 2019; Koehler & Mercer, 2009; López-pérez et al., 2022). One exception is Huang and Goetzmann (2023). They interact extrapolative beliefs with selection neglect in the market of non-fungible tokens to explain its role in boom and bust cycles. We differentiate ourselves from these papers by showing how selection neglect impacts both portfolio selection and return expectations using a comprehensive dataset of wine prices.

Also, our paper adds to the growing literature on the effects of selection biases on asset returns. Recent papers demonstrate that these returns are upward-biased when one ignores such issues (Breeden, 2022; Cameron et al., 2019; Korteweg et al., 2016). To the best of our knowledge, we are the first to provide evidence of the empirical asset-pricing implications in a cross-sectional framework. More importantly, we show that selection biases could affect future returns in two directions: upward and downward. This aligns with the information uncertainty theory.

Second, we contribute to the literature deviating from rational expectations, such as behavioral choice theories (Barberis et al., 2016, 2021; Cakici & Zaremba, 2022; Cosemans & Frehen, 2021). More specifically, we empirically examine the information neglect theory of Eyster et al. (2019). Where they focused on excessive trading (see Daniel & Hirshleifer, 2015; Huang & Goetzmann, 2023), we turn to the impact of selection neglect on return expectations and portfolio selection. Our paper highlights that information neglect can lead to (over)confident investors and higher asset prices.

Finally, we advance the literature on real assets (see Goetzmann et al., 2021). This includes the valuation of collectibles (Dimson et al., 2023; Dimson & Spaenjers, 2011; Korteweg et al., 2016; Masset & Henderson, 2010; Renneboog & Spaenjers, 2013), real estate (Giglio et al., 2015; Li et al., 2023), and general auction markets (Aubry et al., 2023; Lovo & Spaenjers, 2018). Le Fur and Outreville (2019) offer a good overview of the wine economics literature. We contribute to this literature by showing how selection neglect affects the price-setting of real assets, in particular fine wine.

2. Portfolio-level analysis

Selection neglect represents the manifestation of informational asymmetries between different investors. In fact, this bias refers to the propensity of investors to make decisions on a censored sample. Jehiel (2018) shows theoretically that being exposed to selection neglect could translate into suboptimal investment strategies. As such, this section introduces data and methodology to test this hypothesis empirically.

2.1. Data

The primary data for our study are wine auction transactions. In line with previous studies on art markets, we use public auction data as this is more reliable, easily obtainable, and generally accepted as the benchmark for private auctions (Korteweg et al., 2016; Renneboog & Spaenjers, 2013). For the period from January 2003 to December 2022, we collect past auction results from 41 auction houses (located in 11 countries) from their websites.

We argue that focusing on multiple countries and auction houses is essential. First, it has been shown that there is a home bias in auction prices (offerings) (Shi et al., 2017). Second, including multiple auctions reduces potential issues concerning price dispersions (Cardebat et al., 2017) and the impact of biased auctioneers (Aubry et al., 2023). Finally, buyers could search the price history of wine across auctions to form price expectations. At the same time, sellers can quickly sell their wines at other auction houses. As such, having a comprehensive dataset provides a complete picture of trading activity, particularly related to behavioral biases.

[FIGURE 1]

Figure 1 plots the sample coverage. Six auction houses exhibit data for the entire period: Acker Merrall & Condit, Bonhams, Brentwood, Christie's, Sotheby's, and Zachy's. In terms of dollar value of transactions, the most popular auction houses are Acker Merrall & Condit (20.354%), Sotheby's (17.866%), Zachy's (17.758%), Christie's (14.056%), and Hart Davis Hart (11.659%). We refer to these auction houses as the 'Top 5'. As mentioned above, these houses have events in multiple countries, which is the international heterogeneity we later exploit.

All auction houses in our sample follow the English model, referred to as an "ascending price" auction. This is a model in which prices start at low values and are raised until only one bidder remains, who then pays the "hammer price." This is a widespread auction type in both art and wine auctions (Ashenfelter, 1989). In our sample, there are two auction formats: in-person and online. This is essential in our setting, as one can imagine that economic agents in online events have more information at their disposal.

We link wine transaction data to information on producers, regions, styles, and grape varieties from Cellartracker. This is one of the most comprehensive databases with information for over 4.4 million wines. Matching this information to the transaction data enables us to study wine-specific details, such as the impact of color and the production location.

We adopt three filters. First, we exclude mixed lots. Breeden (2023) documents a sizeable price discount for lots with bottles of wine from different vintages, producers, or regions. Therefore, we only keep homogeneous lots, including multiple bottles of identical wine. This remarkably facilitates the calculation of the average price per bottle. Second, we require that a wine trades at least twice in the previous 12 months. This filter aligns with the practices in the literature on financial assets (Cosemans & Frehen, 2021; Dickerson et al., 2023). Finally, we exclude all wines below $20, as this wine is more likely for consumption, and above $50,000. Overall, this leaves us with 1.49 million prices for 142,439 individual bottles.[6] To mitigate currency-related issues, we convert all prices to U.S. dollars.

While forming our sample, we are conscious of one important (potential) bias. We assume that identical wines were sold because we do not have consistent information on bottle quality. In reality, differences could occur because of damages to the labels, an increase in ullage, etc. To minimize these effects, we run a battery of robustness tests in which we calculate returns using minimum, maximum, or average prices. Our conclusions hold in these different specifications.

[FIGURE 2]

Figure 2 displays the total number of transactions. The first observation is that the wine auction market has been growing steadily. The average growth rate is 14.29% per year. Although this is somewhat mechanical as the number of auction houses in our sample increases, the increase in the number of trading platforms highlights the growing interest in wine as an investment.

In Figure 2, we also split the sample into wines that trade at least once over the next 12 months. On average, 81% of wines get at least one (successful) trade in the next year. This suggests that wine is more liquid relative to other real assets. For instance, Korteweg et al. (2016) show that the time between art sales is, on average, 9.42 years. In our sample, the time between two sales is, on average, 9.91 months. A difference with art is that there is only one painting, while there are more bottles of identical wines. We consider this an advantage in our setting, as an identical bottle can be sold on multiple continents.

[FIGURE 3]

---

[6] We argue that bottles with prices below $20 are predominately stemmed for consumption. In turn, the bottles with prices above $50,000 could be only attainable by a (very) small group of investors with deep pockets. However, including these wines does not alter the qualitative conclusions.

Figure 3 plots all wine-producing countries with at least one wine trading in our sample. There are 47 wine-producing countries spread across the globe. There are two main takeaways. First, France is the most important producer. Indeed, 87.34% of all transactions are for French wines. Second, excluding France, the U.S. (58.95%), Italy (22.43%), Portugal (8.07%), and Spain (5.8%) also have large trading volumes. These countries represent over 99% of all wines traded in our sample.

2.2. Methodology

We calculate the return for individual wines as follows:

$$r_{i,t} = \frac{P_{i,t}}{P_{i,t-1}} - 1 \tag{1}$$

where $r_{i,t}$ is the return in month t for wine i, and $P_{i,t}$ is the price in month t for wine i. If a wine gets traded more than once a month, we focus on the last price.

Unlike traditional assets, real assets are generally sold with reserve prices – the minimum price the seller wants (Huang & Goetzmann, 2023; Korteweg et al., 2016). This is an essential feature when calculating price differences. As only the successful assets are sold, returns are generally upward biased. This, thus, creates a selection bias (e.g., Chen et al., 2022; Hirshleifer & Plotkin, 2021; Korteweg et al., 2016). Eq. 1 only considers realized trades, which are subject to selection biases. As such, we label this 'the biased return.'

The probability of a trade depends on investors' potential returns, which are not realized until that trade occurs. Moreover, the timing of a sale is endogenous. Hence, we follow the two-step approach by Korteweg and Sorensen (2010) to address this problem. The approach models the entire path of unobserved (latent) prices based on observed transactions and the likelihood of observing realized transactions at each point in time. This is recursively updated and repeated until the parameters are converged. In other words, it is a Bayesian estimation technique, also known as the Markov Chain Monte Carlo (MCMC) model.[7] In fact, we apply the Gibbs sampler to simulate the augmented posterior distribution to get the model parameters. That is,

[7] The method is frequently used in the area of infrequently traded assets, such as private equity, art, and NFTs (e.g., Huang & Goetzmann, 2023; Korteweg et al., 2016; Korteweg & Sorensen, 2010).

$$w_{i,t}^{L} = \sum_{k=1}^{K} g_k(p_{i,t}; Z_{i,t}) \cdot \beta_k \tag{2}$$

$$w_{i,t}^{L} = \beta_1 + (r_{i,t}^{p} \cdot \beta_2) + (h_{i,t} \cdot \beta_3) + (h_{i,t}^{2} \cdot \beta_4) + (k_t \cdot \beta_5) + \eta_{t,i} \tag{3}$$

where $p_{i,t}$ is the observed price for wine *i* in month *t*, β_k is the probability of sale given a set of observed covariates Z_t (e.g., the state of the economy, k_t), $r_{i,t}^{p}$ is the non-annualized log return on wine *i* since its last sale (referred to as 'the potential holding return'), $w_{i,t}^{L}$ is a latent selection variable, and $\eta_{t,i}$ is the error term with mean zero and variance equal to one. As in Huang and Goetzmann (2023) and Korteweg et al. (2016), we include $h_{i,t}$, defined as the number of months since its last transaction. To proxy for the state of the economy, we include excess stock returns for developed markets from Kenneth French. Figure A.3. provides additional insights into the convergence of MCMC parameters.

We use the latent selection variable, $w_{i,t}^{L}$, in a probit regression,

$$p_{i,t}^{L} = \begin{cases} p_{i,t} & w_{i,t}^{L} \geq 0 \\ unobserved & w_{i,t}^{L} < 0 \end{cases} \tag{4}$$

where $p_{i,t}$ is the observed price for wine *i* in month *t*, $p_{i,t}^{L}$ is the latent price for wine *i* in month *t* if $w_{i,t}^{L}$ is non-negative. Figure A.2. plots the time series of observed and unobserved prices for six randomly-chosen wines, including.

[FIGURE 4]

Figure 4 displays the probability of trading conditional on potential holding returns. Perhaps unsurprisingly, the probability is lower (higher) for negative (positive) potential returns. They align with other assets (Huang & Goetzmann, 2023; Korteweg et al., 2016).

Finally, this gives a selection model in which latent price $p_{i,t}^{L}$ is observed if $w_{i,t}^{L}$ is a non-negative number, as in Equation 4. The difference between latent prices is then referred to as 'selection-adjusted return.' Similar to the literature, we apply 6,000 simulations and discard the first 1,000 for burn-in (see Chen et al., 2022). We refer to Korteweg and Sorensen (2010) for the (technical) details.

[FIGURE 5]

Figure 5 shows the distribution of observed and latent prices. The average latent price ($247.68) for the entire sample period is $293.68 lower than the average observed price ($541.12). This is an economically large difference.

[TABLE 1]

Table 1 provides information about the price increases between consecutive sales (i.e., a biased return) and latent valuations (i.e., selection-adjusted returns). We show that the biased returns, on average, equal 2.2% per month, with a standard deviation of 25.9%. Negative returns occur relatively frequently. Indeed, the wine return distribution seems to be right skewed, as the 1st percentile equals -44.0% per month relative to 85.7% per month for the 99th percentile. Relative to art, these returns are sizeably smaller (Korteweg et al., 2016).

When we use latent valuations instead of only focusing on successful transactions, the average return decreases to 1.8% per month. Interestingly, the return's standard deviation increases to 28.5% per month. It also lowered the median return from 0% (i.e., biased return) to -1.4% (i.e., selection-adjusted returns). A takeaway from this table is that negligent investors, on average, experience more positive returns (less risk) than non-negligent investors. This could translate, for example, into different weights when forming the optimal portfolio. We examine this issue in the next section.

2.3. Portfolio optimization

On the portfolio level, the question pertains to whether investors incur a performance penalty when exposed to selection neglect. We investigate this empirically through the lens of a mean-variance (MV) paradigm. Two representative agents are posited: one prone to selection neglect (negligent investors), the other immune (non-negligent investors). Both maintain the tangency portfolio, but the weights diverge across investors. The negligent investors use biased returns, whereas non-negligent investors adjust for selection biases before forming portfolios. To make this procedure less computationally expensive, we separately focus on the most traded (wine-producing) countries (i.e., France, the USA, Italy, Spain, and Portugal).

Each month, investors estimate the optimal weights with the returns of the previous 36 months and use the weights to estimate the certainty-equivalent (CEQ) return, defined as the risk-free rate an investor is willing to accept rather than choosing the risky portfolio strategy,

$$\widehat{CEQ} = \hat{\mu}_k - \frac{\gamma}{2}\hat{\sigma}_k^2, \tag{5}$$

where $\hat{\mu}_k$ is the sample mean of out-of-sample excess returns over the forecasting period, $\hat{\sigma}_k^2$ is the out-of-sample variance, and γ is a risk aversion parameter (e.g., DeMiguel et al., 2009).

We use selection-adjusted returns as proxy for future returns as these are more representative of an investor experience (Korteweg et al., 2016). If we use biased returns, we implicitly assume that investors exhibit perfect asset-picking skills (as they could pick winners, those more likely to sell) and that there is no selection bias (Chen et al., 2022). Indeed, selection-adjusted returns measure the rise in the value of a representative portfolio of wines, both that are sold and those that are not.

Table 2 presents the results for CEQ returns for a one- and twelve-month horizon for investors, with risk aversion parameters of one and five. When extending the horizon to longer periods, we assume that traders rebalance at the same frequency as the forecasting horizon (similar to Demiguel et al., 2007). We define 'the expected utility gains' as the difference between the CEQ returns of non-negligent and negligent investors. This can also be interpreted as the economic value of not suffering from selection neglect.

[TABLE 2]

On average, the CEQ returns are always higher for the non-negligent investors. For instance, the average expected utility loss for negligent investors with a logarithmic utility function (i.e., $\gamma = 1$) equals 0.4% per month after one month. This corresponds to a relative decrease of 25%. In other words, the non-negligent investor has an expected utility gain of 0.4% per month. For a longer horizon, average expected utility gains increase to 0.6% per month. This difference is economically meaningful, given that the average return on wine is 1.8% per month (Table 1). This highlights that suffering from selection neglect leads to suboptimal investment decisions, as predicted by Jehiel (2018).

3. Cross-sectional-level analysis

The previous section showed that negligent investors could lose investment performance on a portfolio level. This section, in turn, examines how investors adjust their (return) expectations on the asset level. In other words, investors are exposed to this bias for future wine-level gains and losses by looking at its past performance. A cross-sectional analysis allows us to study two hypotheses. On the one hand, the information uncertainty theory documents that information differences positively impact future returns (Hirshleifer & Plotkin, 2021; Hong & Stein, 1999; Jiang & Sun, 2014). On the other hand, the sidelined-investor theory posits that the pessimistic investors will not trade when they have different information. Optimistic investors, therefore, will drive up current prices, leading to low future returns (Cen et al., 2017; Diether et al., 2002; Goetzmann & Massa, 2005; Hillert et al., 2018).

3.1. Variable definition

Our variable of interest at the asset-level is selection neglect. However, this is a very subjective concept. For instance, it is impossible to deduce which information investors did (not) observe outside the laboratory. In our framework, an investor can only be interested in an Italian wine. For instance, while still seeing French wines, he can ignore this information when forming his expectations for said Italian wine. Hence, we need a more objective metric of selection neglect.

As Huang and Goetzmann (2023), we define *Selection Bias* as the 12-month difference between biased and selection-adjusted returns for individual wines. This measure does not suffer from the informational assumption that the above example showcased. It compares the situation in which investors ignored the endogeneity of trading and only took into account the successful trades (biased returns) relative to the situation in which they consider all available information (selection-adjusted returns). In other words, the investor mentioned above only considers the successful transactions for a specific Italian wine, while ignoring the likelihood of trading and the failed trades for that exact Italian wine.

To ensure that consumption is not driving our results, we require that wines have at least two trades in the last 12 months. If consumers exhibit a buy-and-drink rather than a buy-and-hold strategy, we must ensure that these wines are not in the rest of the sample. In other words, this ensures that permanent illiquidity is not a potential driver for the results. We, however, control for temporary illiquidity by adding a measure capturing this, which is specified below.

3.1.1. Control variables

We measure momentum as the wine's cumulative return over an 11-month period ending two months prior to the current month. Reversal is measured as a wine's prior month's return. As in Lin et al. (2011), illiquidity is measured as the covariance of price changes between month *t* and month *t-12* multiplied by -1. Beta is a slope coefficient from a regression of a wine's excess return on the price-weighted excess market return (CAPM) using a 36-month rolling window. Finally, we create two metrics to proxy for lottery demand: MAX (MIN) is an individual wine's maximum (minimum) monthly return over the last 12 months.

We merge our transaction data with ratings and drinking window by Wine Advocate. Ratings are assigned a number from 60 (low quality) to 100 (high quality) compared to the peer group (e.g., style, region, and grape variety). Faulty wines, such as cork taint and light strike, are not rated. There are two small changes we make to the raw data: (1) we add 0.5 to the rating when it includes a plus (e.g., 99+ becomes 99.5)[8], and (2) we use the averages if they provide a range (e.g., 98-100 becomes 99). Wine Advocate also provides the 'drinking window.' These are two years that give the optimal utility for a consumer. We interpret the last year of this window as the expiration of that wine.

3.1.2. Summary statistics

[TABLE 1]

The average price of wine in our sample is $608.51. This includes both standard bottles (750ml) and other sizes. This partially explains the large standard deviation ($1,585.14).[9] Indeed, prices range from $28.51 in the 1st percentile to $7,499.58 in the 99th percentile. Unsurprisingly, ratings are high. For instance, a rating of 93.93 corresponds to "outstanding and extraordinary wines." The average time to expiration is 12.27 years, with a sizeable variation. This offers insights into wines sold at the auctions in our sample, which are not standard consumption wines but could have more investment purposes.

3.2. Univariate regressions

At the end of each month, we sort wine into quintile portfolios based on SB value and calculate price-weighted returns over the next month. This portfolio is rebalanced each month. As Chen et al. (2022), we calculate returns with latent valuations and apply the observed price as weight

[8] A plus sign implies that the reviewer believes it has the potential to improve in the bottle.
[9] Outreville (2013) documents that larger bottles trade at a discount relative to a combination of smaller bottles. As such, we run a robustness test focusing only on standard bottles (cfr. Table A.3.).

if available. Table 3 reports the time-series average for *Selection Bias*, the next-month selection-adjusted returns, and alphas for each quintile. Figure A.7. plots the time series for the *Selection Bias* metric per quintile. The last row reports the return spreads between Quintile 5 (High) and 1 (Low), which is the return on a zero-cost strategy of buying high-SB and selling low-SB wine.

Alphas are estimated through a five-factor model for stocks or bonds.[10] Previous literature has documented that wine returns positively correlate with stocks, particularly during recessions (e.g., Masset & Henderson, 2010). As mentioned above, wines have similar features as bonds. As such, we use existing traditional asset factors to correct for potential risk sources. Moreover, we calculate risk-adjusted returns with new wine factors. We include factors for illiquidity and momentum, as well as the excess wine market return.[11] The factors are constructed in line with the existing literature (e.g., Dickerson et al., 2023).

[TABLE 3]

The evidence in Table 3 strongly supports the hypothesis of the positive relationship between selection biases and future returns. The portfolio returns increased (monotonically) from -5.5% per month for lower-SB wine to 1.8% per month for higher-SB wine. The High-Low difference portfolio totals 7.3% monthly. Even if we control for stock, bond, and wine factors, the spreads do not decrease. This implies that the cross-sectional variation in SB explains changes in future returns, and traditional (risk) factors cannot explain these differences.[12]

There are two essential points to make with this return spread. First, as you cannot short wine, the High-Low difference portfolio is impossible to attain. However, the results hold if we look at the two components in isolation. Indeed, the High portfolio generates a 1.6% monthly risk-adjusted return (significant at the 1% level), whereas the Low quintile generates a risk-adjusted return of -5.7% (significant at the 1% level). This shows that the effect is significant in the long

[10] The five stock factors include the excess stock market return and factors on size (SMB), value (HML), operating profitability (RMW), and investment (CMA) for the developed markets from Kenneth French's website. Five bond factors include the excess corporate bond return and factors on downside risk, credit risk, duration, and long-term reversal from Dickerson et al. (2023).

[11] We include these two measures as they are shown to be important in the empirical wine literature.

[12] Table A.4. shows that the results are robust when we apply this methodology to individual countries. Indeed, suppose we limit the sample to only trades that occurred in a specific country. In that case, the risk-adjusted returns of the High-Low difference portfolio remain statistically significant at the 1% level.

leg (even when we control for other factors). Second, these risk-adjusted returns do not include storage, insurance, or other costs. This, however, does not take away from the main finding – that high SB wines outperform low SB wines – as such costs will be similar for both quintiles.

Koehler and Mercer (2009) show that investors are partially sensitive to selection neglect. That is, only if they are fully aware of the selection process, will they adjust their return expectations relative to negligent investors. This implies that negligent investors drive up prices if they are experiencing more positive returns than the non-negligent investors. In fact, positive selection neglect is capturing this, as biased returns are higher than selection-adjusted returns. Eyster et al. (2019) argue that information neglect makes negligent investors (over)confident. This could also explain our findings. That is, negligent investors become overly confident and, as a result, exhibit more optimistic expectations, leading to positive future returns.

One potential concern regarding *Selection Bias* is that it depends on a 12-month horizon, similar to Huang and Goetzmann (2023). We provide two approaches to show the robustness of these results. First, Table A.3. confirms our main conclusion for different time horizons or minimum number of trades. Second, Table A.5. employs a different measure. We estimate the ratio of the differences between observed and latent prices to latent prices. This has the key advantage of not depending on a time horizon.[13] The qualitative conclusions are robust to these changes.

[TABLE 4]

To better understand the composition of the quintile portfolios, we compute the cross-sectional average of various characteristics. Table 4 provides the time-series averages for price-weighted portfolios. Wines in the High portfolio have lower prices compared to the other quintiles. Their average price is 32.95% lower than the average wines in our sample. Interestingly, they exhibit a lower rating and shorter time to expiration. If we compare their past performances, the High portfolio has a significantly higher momentum and a lower past one-month return. Moreover, the most negative (positive) return over the last 12 months has been, on average, 6.90% (12.4%)

[13] The advantage of this metric is that it provides an intuitive interpretation, is normalized across wines, is time-independent, and is a direct measure of mispricing. The drawback is that nero-zero latent prices will lead to extensive values. As such, we only use this as a robustness test.

lower than the other extreme quintile. This suggests a higher volatility for these wines. Indeed, the High portfolio's idiosyncratic volatility has been higher than the Low quintile.

In conclusion, the univariate regressions offer preliminary evidence in favor of the information uncertainty hypothesis. We show a strong positive effect of selection bias on one-month-ahead returns. Interestingly, the significant spread between the two quintile portfolios is both due to the outperformance of High SB-wines and the underperformance of Low SB-wines. This result cannot be explained by traditional risk and novel wine factors. Table A.3. highlights that these findings are robust to changes in our empirical design, such as changes to the valuation model, sample splits, and auction house heterogeneity.

3.3. Bivariate regressions

The previous section finds that wines across quintile portfolios exhibit different characteristics. This section examines whether the heterogeneity in the features can explain the return spread across SB-sorted portfolios. To study this, we create double-sorted portfolios to control for the characteristics. Each month, we sort wines on a second variable into quintiles. In each quintile, we sort wine on its SB values. This provides subquintiles, where we test the differences in risk-adjusted returns conditional on the wine characteristic. In total, we create 25 portfolios. Similar to above, Low (High) represents either the quintile portfolios with low (high) SB values or the control variable.[14] The bottom row across all panels reports the 10-factor stock and bond alphas for the return spreads across these quintiles.

[TABLE 5]

Table 5 provides the results of the bivariate sorts. We highlight that the selection neglect effect remains statistically significant and economically large after controlling for the characteristics. The alphas range from 0.7% to 8.2% per month (significant at the 5% level). This indicates that other factors do not explain the positive relationship between Selection Bias and future returns.

There are several noteworthy findings. Panel G of Table 5 documents that the ten-factor alpha increases monotonically across all portfolios. For example, conditional on low prices, the alpha

[14] We label a wine as *investment-grade* (IG) if it received a rating equal to and above 95 and *non-investment grade* (NIG) with a rating below 90. We choose these groups as they are more distinct in their rating. The results, however, hold when we include those wines in the 90-95 rating bucket.

ranges from -2.4% for low values of selection bias to 5.8% in the highest quintile (all significant at the 1% level). Notably, the high-low portfolio alphas are significant. Interestingly, the return differences between the price groups are large. This is in line with the stock market results that high-priced stocks have lower future returns (Disli et al., 2019). There are two explanations for this. First, the dollar difference for higher-priced wine is larger than for low-priced wine, given the same percentage changes (Shue & Townsend, 2019). This can increase the behavioral bias of focusing on biased observations by setting faulty return expectations. Second, lower-priced wine is arguably more demand-elastic, which can explain the return difference between these groups.

We highlight that return reversal has strong effects on the wine market. In Panel H, wines with higher past one-month returns exhibit smaller future returns, even if we condition on selection bias. Even if we increase the horizon to past momentum returns, the higher-performing wines have lower future returns than lower-performing wines. Furthermore, this conclusion remains valid, considering the past minimum or maximum returns. However, even if we consider such effects, the High-Low difference portfolio produces significant alphas. This confirms the main conclusion of McManus et al. (2013) that there is a reversal in wine auction prices. Second, and more importantly, these findings indicate that the well-known cross-sectional findings in wine (and other assets) cannot explain the significant selection bias premium.

3.4. Wine-level regression

One of the advantages of using portfolios is the reduction of residual variance. The drawback, however, is that we throw out important asset-specific information. Moreover, using portfolio-level analysis, we can only test for information in the aggregate variation by a limited number of variables. This section runs panel regressions with the one-month ahead (Panel A) and long-term excess return (Panel B) as dependent variables, and selection bias and wine characteristics as independent variables. More specifically, we run:

$$R_{i,t+h} = \beta_0 + \beta_1 SB_{i,t} + \beta_k \gamma_{i,t} + \tau_i + \epsilon_{i,t+h}, \tag{6}$$

where $R_{i,t+h}$ is the return on wine *i* in month *t* to month *t+h*, $SB_{i,t}$ is the Selection Bias measure for wine *i* in month t, $\gamma_{i,t}$ is a vector of characteristics for wine *i* in month *t*, defined in Table 1, and τ_i are wine fixed effects. The *t*-statistics are calculated with wine-level clustering (Petersen, 2009).

[TABLE 6]

Table 6 reports that, in the univariate regression, the coefficient on Selection Bias is statistically significant at the 1% level. It is also economically large: an increase in SB coincides with a surge of 3.2% in future returns per month. The qualitative conclusion holds even if we include more control variables. Indeed, Columns 2 to 11 highlight that SB remains statistically significant (at the 1% level) in predicting future returns.

The evidence for the control variables is in line with the results mentioned above. For instance, Table 6 reports a negative relationship between prices and future returns. There is evidence of a reversal in wine prices, as indicated by the negative momentum and reversal coefficients on future returns. Wines with higher ratings earn high returns, which aligns with the conclusions of Hadj Ali et al. (2008). More importantly, adding the control variables does not alter the main results: there is a positive relationship between selection bias and future returns. This supports the prediction from the information uncertainty theory.

Panel B of Table 6 examines the long-term predictive power of selection bias. We find that the positive relationship with future returns remains significant up to 12 months ahead. In these specifications, we control for all wine characteristics (cfr. Column 11 of Panel A). The adjusted R-squared increases up to 15.5%, implying that there is still a sizeable portion of future returns that cannot be explained.

One potential concern regarding our methodology is that investors cannot attain the selection-adjusted returns, as they captured differences in latent valuations (Huang & Goetzmann, 2023; Korteweg et al., 2016). Nevertheless, it more accurately represents an investor's experience. In turn, biased returns assume that an investor has perfect asset-picking and market-timing skills, which is an unrealistic assumption. However, our findings are robust when we employ biased returns as the dependent variable, as demonstrated in Table A.7. The table highlights that the magnitude of SB diminishes largely to 1.3% per month (significant at the 1% level).

4. Channels of selection neglect

4.1. Investor attention

The definition of selection neglect is that 'one observes a censored sample and infers properties for the population.' As such, the evidence we documented can be driven by investor attention. To examine this, we use a panel regression with five measures capturing attention: (i) *Abnormal returns* is the 12-month return difference between an individual wine and the wine market, (ii) *Abnormal trading volume* is the ratio of the number of transactions between months *t* to *t-11* and months *t-12* to *t-23* for a specific wine, (iii) *Google* captures the global Google search trends for wine, and (iv) *Top 100* is an indicator variable equal to one from November, when it is included on the Top 100 list of Wine Spectator, or zero otherwise, and (v) *Salience* is defined in line with Cosemans and Frehen (2021):

$$Salience = \sigma(x_{is}, \bar{x}_s) = \frac{|x_{it} - \bar{x}_t|}{|x_{it}| + |\bar{x}_t| + \theta} \tag{7}$$

where x_{is} is the return of wine *i* in month *t*, $\bar{x}_t$ is the average return of all wine in month *t*, and θ is a constant equal to 0.1. We computed *Salience* over the last 12 months to match the horizon of *Selection Bias*. Using these variables, we run the following regression:

$$R_{i,t+1} = \beta_0 + \beta_1 SB_{i,t} + \beta_l SB_{i,t} \cdot Z_{i,t} + \beta_k Z_{i,t} + \beta_l \gamma_{i,t} + \tau_i + \epsilon_{i,t+1}, \tag{8}$$

where $R_{i,t+1}$ is the selection-adjusted return on wine *i* in month *t+1*, $SB_{i,t}$ is the selection bias value for wine *i* in month t, $Z_{i,t}$ is the vector of proxies for investor attention for wine *i* in month *t*, defined above, $\gamma_{i,t}$ is the vector of control variables for wine *i* in month *t*, defined in Table 1, and τ_i are wine fixed effects. The *t*-statistics are calculated with wine-level clustering (Petersen, 2009).

We expect that the first four measures positively impact future returns, conditional on SB. One must know the market performance to understand which wine has a different return than the market (e.g., abnormal returns and salience). In other words, investors would spend more time collecting data to spot outperformers. Therefore, they will be less likely to neglect information. In contrast, if a wine is included in the Top 100 list, investors are arguably more likely to miss other information if they are fixated on this acknowledgment. Hence, we hypothesize there is a positive relationship between Top 100 and future returns.

[TABLE 7]

The positive relationship between selection bias and future returns remains, even after adding investor attention measures (significant at the 1% level). Table 7 finds that the effect of selection bias on future returns is weaker for wines with more salient features or high abnormal returns. This confirms the hypothesis mentioned above. In turn, the effect is stronger for Top 100 wines. This can imply that investors *only* look at wines that have received attention and neglect other data about them. Overall, the evidence validates the predictions of the information uncertainty theory, and it indicates that the well-known cross-sectional anomalies in wine (or other assets) cannot explain the significant selection bias premium.

4.2. Limits to arbitrage

If selection neglect is the manifestation of irrational mispricing, it should be stronger for assets with more pronounced trading frictions. For traditional assets, multiple papers document that cross-sectional return variation can be explained by limits to arbitrage (e.g., Shleifer & Vishny, 1997). Consistent with this, we test whether assets with idiosyncratic risks and illiquidity drive our results. We define the idiosyncratic volatility (*IVOL*) as the standard deviation of the error term of CAPM. In the case of wine, we argue that two more frictions could be important: prices and time to expiration. These can be seen as a limit to arbitrage as the (potential) investor base shrinks if wine becomes too expensive or is outside the drinking window.

Without limits to arbitrage, a rational investor can correct the mispricing induced by selection neglect. To test for this, we include an interaction term with the above-mentioned variables in a panel regression. That is,

$$R_{i,t+1} = \beta_0 + \beta_1 SB_{i,t} + \beta_l SB_{i,t} \cdot Z_{i,t} + \beta_k Z_{i,t} + \beta_p \gamma_{i,t} + \tau_i + \epsilon_{i,t+1}, \quad (9)$$

where $R_{i,t+1}$ is the selection-adjusted return on wine *i* in month *t+1*, $SB_{i,t}$ is the selection bias value for wine *i* in month t, $Z_{i,t}$ is the vector of proxies for limits to arbitrage for wine *i* in month *t*, defined above, $\gamma_{i,t}$ is the vector of control variables for wine *i* in month *t*, defined in Table 1, and τ_i are wine fixed effects. The *t*-statistics are calculated with wine-level clustering (Petersen, 2009).

[TABLE 8]

Table 8 reports the regression coefficients. First, we confirm the positive relationship between our measure and future returns. Table 8 also shows that wine with more illiquidity and higher prices are more exposed to the selection neglect effect. In turn, idiosyncratic volatility and time to expiration do not have an effect on future returns. Second, the table highlights that limits to arbitrage have little effect on the predictive ability of *Selection Bias*. One potential reason is that these markets are characterized by trading constraints. As mentioned, investors cannot (easily) short-sell wine. Therefore, limits to arbitrage will increase the inefficiency in real asset markets (David et al., 2013).

4.3. Preferences

The interest in wine as an investment has been growing over time (Dimson et al., 2015; Masset & Weisskopf, 2015). However, it remains primarily a consumption good. This has two sizeable implications. First, consumption could be a non-speculative trade motive. This is not a crucial feature in our framework. If wine is consumed, it will leave our sample due to a lack of (future) liquidity. Since Selection Bias imposes (at least) two trades within the last twelve months, those wines intended for consumption do not have a lasting impact on future returns. Second, there can be (time-varying) preferences. Consumers enter this market, purchase their favorite wines, and leave. To ensure that such features do not drive our results, we focus on the assets lacking consumption applications: the wines outside their drinking window. Consumers will arguably leave this market for such assets, and only investors remain. This section revisits the approach in Section 3.2. and run a univariate regression solely for wines outside their drinking window.

[TABLE 9]

Table 9 confirms the evidence in the primary analysis: There is a positive relationship between selection bias and future returns. Even if we control for stock, bond, or wine factors, the High-Low difference portfolio alphas are statistically significant at the 5% level and are economically meaningful. Furthermore, this re-confirms the predictions from the information uncertainty theory and shows that consumer preferences cannot explain the observed results.

5. Economic mechanisms

5.1. Time variation in selection neglect

Having established the unconditional effect of selection neglect, we now turn our attention to its time variation. Following Cakici and Zaremba (2022), we study how it evolves over time in relation to economic uncertainty and market sentiment. This will allow us to understand better the effect of economic shocks (e.g., economic policy uncertainty), news (e.g., geopolitical risks), and market sentiment (e.g., disagreement) on our bias. The variables are defined in Table A.7. We run the following regression,

$$R_{i,t} = \beta_0 + \beta_l Z_{i,t-1} + \beta_k X_{t-1} + \epsilon_{i,t}, \tag{11}$$

where $R_{i,t}$ is the monthly return on the High-Low selection-bias-sorted difference portfolio in month t, $Z_{i,t-1}$ is the indicator in month *t-1,* defined in Table A.7., and X_{t-1} is the vector of five stock or bond factors, and three wine factors, as defined above. *T*-statistics are calculated with Newey-West (1987).

[TABLE 10]

Table 10 reports the regression coefficients. The findings generally highlight that the economic conditions play a role in explaining the variation of the selection neglect effect. For instance, it appeares more pronounced in periods of high economic policy uncertainty, high climate risks, and increased tail risks, as indicated by the positive coefficients for the variables. Interestingly, Table A.8. highlights the robustness of this conclusion for equally-weighted portfolios.[15]

These findings paint a clear picture: the more difficult it is to trade, the more pronounced the selection neglect effect. Indeed, Panel B of Table 10 finds a negative (positive) impact for credit spreads (dividend yield). While not all coefficients are statistically significant, the main pattern suggests that the High-Low portfolio serves as a hedge against economic uncertainty. Overall, this confirm that selection neglect strengthens when price inefficiencies are more pronounced (i.e., in weaker economic conditions), consistent with our earlier findings on limits to arbitrage. This also aligns with the stock market evidence on salience (Cakici & Zaremba, 2022).

Interestingly, there is a negative relationship between the short-interest index and the selection neglect effect. There are two potential explanations for this result. First, short selling could lead to temporary price pressure (and price inefficiencies) (Asquith et al., 2005). Second, and more

[15] One exception is *climate risk*, which is not significant in the equally-weighted portfolios.

importantly, short selling can reflect more informed trading. Indeed, it can imply less negative information or, in other words, with more optimistic expectations. Eyster et al. (2019) highlight the role of (over)confidence in light of information neglect. Therefore, the negative relationship between short interest, while measured forr the stock market, aligns with our main conclusion, and provide additional context for understanding the dynamics of selection neglect in the wine market. That is, selection neglect is correlated with strong price inefficiencies.

5.2. Conditional regressions

Following on the conditional regression presented above, we classify each month as a low and high sentiment month, where a low (high) state is defined as each month in which the previous month's VIX is above (below) its sample median. This is consistent with (asset-pricing) models using risk aversion as a non-linear function of volatility. Similar to Stambaugh et al. (2012), we run a regression with two indicator variables,

$$R_{i,t} = \beta_0^H \cdot d_t^H + \beta_1^L \cdot d_t^L + \beta_l X_t + \epsilon_{i,t}, \tag{10}$$

where $R_{i,t}$ is the monthly excess return on quintile 5 (long), quintile 1 (short), and the difference (long-short), d_t^H (d_t^L) is a dummy variable that equals one if VIX in month *t-1* is above (below) its sample median, and zero otherwise, and X_t is a vector of five stock or bond factors, or wine market factors. *T*-statistics are calculated with Newey-West (1987).

[TABLE 11]

There are three key takeaways from Table 11. First, we document that the risk-adjusted returns increase monotonically across high and low-market stress episodes. Second, the return spreads on the High-Low difference portfolio are significant for both traditional asset and wine factors. This confirms the evidence from the previous section. Finally, the differences across alphas in high and low-stress states are not statistically significant. This implies that the selection neglect effect is not non-linear relative to other risk-based measures, which suggests that time-varying risk does not explain the selection bias premium.

5.3. Neglect amplification

To further test the economic mechanisms of selection neglect, this section focuses on the impact of the COVID-19 pandemic. Not only did this lead to a switch from in-person to online events, it also created many opportunities for (new) investors. Indeed, online auctions could attract a broader audience, have a long bidding duration, and provide more information on the wines. If investors apply the additional features, selection neglect will *decrease*. However, recent stock market evidence indicates that human interaction *increases* information (Brogaard et al., 2023). In other words, the COVID-19 pandemic can lead to a scenario in which selection neglect will intensify due to lower informational transmission. To examine these two hypotheses, we apply the following regression:

$$R_{i,t} = \beta_0 + \beta_1 I_t + \beta_l X_t + \beta_k (X_t \cdot I_t) + \epsilon_{i,t}, \tag{10}$$

where $R_{i,t}$ is the monthly excess return on the quintile portfolios and the long-short difference, I_t is a dummy variable that yields one from March 2020 onwards, and zero otherwise, and X_t is a vector of five stock and bond factors or wine market factors. T-statistics are calculated with Newey-West (1987).

[TABLE 12]

Table 12 reports the regression coefficients. The results imply that the effect of selection neglect strengthened in the post-COVID-19 period, as evidenced by an increased spread between high and low SB wines. Indeed, the regression coefficient increases from -1.0% per month for lower-SB wine to 0.6% per month for high-SB wine when we control for wine factors. The High-Low regression coefficient equals 1.5% monthly (significant at the 1% level). This indicates that the selection neglect effect is substantially stronger in the post-COVID-19 period. The results align with the results of Brogaard et al. (2023), who note that price discovery was hampered due to a lack of in-person trading.

We believe this amplification has two channels. First, a shift to online events likely exacerbated information asymmetry. Following De Felice et al. (2022), we argue that absences of in-person events cause participants to rely more heavily on observable (past) data, which intensifies the bias. Second, the online format potentially heightens investor confidence. Reduced social cues, less feedback, or easier access for less experienced participants then reinforce selection neglect. This aligns with earlier (stock market) evidence (Barber & Odean, 2002).

Overall, the pandemic appeared to have created an environment where this bias becomes more pronounced. This result confirms our previous conclusions and highlights the dynamic nature of selection neglect, particularly in periods of market stress. It raises important questions about price inefficiency in the market of real assets and underscores the need for increased awareness of these biases among investors.

6. Conclusions

We study the asset-pricing implications of deviations from the shared information assumption that underlies recent finance theory (Eyster et al., 2019; Hirshleifer & Plotkin, 2021; Jehiel, 2018; Ngangoué & Weiszacker, 2021). Specifically, we turn to selection neglect, which is a tendency to make decisions based on biased observations. This paper uses assets that trade in sequential auctions with severe short-sale constraints and informational frictions: wine.

Using wine allows us to test two competing theories on the impact of informational differences on investor performances and return expectations: the sidelined investor theory (i.e., expecting negative future returns) and the information uncertainty theory (i.e., expecting positive future returns). Indeed, the auction market creates a scenario where some investors are "sidelined" – those who are less optimistic about future payoffs – and exhibit more information uncertainty.

First, we highlight that investors suffering from this bias forgo investment performance at the *portfolio-level*. Measured by certainty equivalent returns, we find that the negligent investors experience large expected utility losses. The losses increase with the investment horizon. This highlights the importance of this behavioral bias and confirms Jehiel's (2018) predictions.

Second, we find strong support for the information uncertainty hypothesis on an *asset level*. A zero-cost strategy of buying high-SB wine and selling low-SB wine documents positive future returns. Standard asset-pricing factors for stocks, bonds, and wine cannot explain this return spread. The effect remains significant if we control for other potential channels, such as limits to arbitrage, investor attention, preferences, or extrapolative beliefs. We argue that it is because negligent investors become overconfident and drive up prices. We show that periods of higher economic uncertainty and the COVID-19 pandemic intensified this effect.

We choose wine because of its growing importance in the investment industry (Dimson et al., 2015; Masset & Weisskopf, 2015) and its relatively liquid nature across other real alternatives.

Indeed, recent work highlights a switch of institutional investors towards real assets (Begenau et al., 2023; Kräussl et al., 2013). As such, based on the evidence from this study, policymakers should consider developing a (regulatory) framework for real asset markets, particularly those using auction-based trading systems. Indeed, they must contemplate potential spillover effects of selection neglect from real assets to financial markets (Goetzmann et al., 2021; Li et al., 2023; Lustig & Nieuwerburgh, 2005).

Although our paper uses wine as a test asset, it can easily be extended to other real and private-value assets (Dimson et al., 2023; Goetzmann et al., 2021). Other real assets suffer from a similar problem: their illiquid nature and market design lead to neglect of selection neglect. Moreover, these assets are characterized by severe market inefficiency (David et al., 2013). We argue that selection neglect adds to the inefficiencies. Nevertheless, an extension to other real and private-valued assets could be interesting for future research.

References

Akerloff, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanisms. *Quarterly Journal of Economics*, *84*, 488–500.

Ashenfelter, O. (1989). How Auctions Work for Wine and Art. *Journal of Economic Perspectives*, *3*(3), 23–36.

Asquith, P., Pathak, P. A., & Ritter, J. R. (2005). Short interest, institutional ownership, and stock returns. *Journal of Financial Economics*, *78*(2), 243–276.

Aubry, M., Kräussl, R., Manso, G., & Spaenjers, C. (2023). Biased Auctioneers. *The Journal of Finance*, *78*(2), 795–833.

Backus, M., & Lewis, G. (2024). Dynamic Demand Estimation in Auction Markets. *Review of Economic Studies*, 1–36.

Barber, B. M., & Odean, T. (2002). Does Online Trading Change Investor Behavior? *European Business Organization Law Revies*, *3*, 83–128.

Barberis, N. C., Greenwood, R., Jin, L., & Shleifer, A. (2015). X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics*, *115*(1), 1–24.

Barberis, N. C., Jin, L., & Wang, B. (2021). Prospect Theory and Stock Market Anomalies. *The Journal of Finance*, *76*(5), 2639–2687.

Barberis, N. C., Mukherjee, A., & Wang, B. (2016). Prospect Theory and Stock Returns : An

Empirical Test. *Review of Financial Studies*, *29*(11), 3068–3107.

Barron, K., Huck, S., & Jehiel, P. (2019). *Everyday econometricians: Selection neglect and overoptimism when learning from others* (SP II 2019-301; WZB Discussion Paper).

Begenau, J., Liang, P., & Siriwardane, E. (2023). *The Rise of Alternatives*. *July*.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and Asset Prices. *American Economic Review*, *103*(3), 623–628.

Breeden, J. L. (2022). How sample bias affects the assessment of wine investment returns. *Journal of Wine Economics*, 127–140.

Breeden, J. L. (2023). How heterogeneous wine lots result in a price discount relative to homogeneous lots. *International Journal of Wine Business Research*, *36*(1).

Brogaard, J., Ringgenberg, M. C., & Roesch, D. (2023). Does Floor Trading Matter? *The Journal of Finance*.

Brundage, M., Little, A. T., & You, S. S. (2022). *Selection Neglect and Political Beliefs*. *July*, 1–58.

Cakici, N., & Zaremba, A. (2022). Salience theory and the cross-section of stock returns: International and further evidence. *Journal of Financial Economics*, *146*, 689–725.

Cameron, L., Goetzmann, W. N., & Nozari, M. (2019). Art and gender : market bias or selection bias ? *Journal of Cultural Economics*, *43*(2), 279–307.

Cardebat, J.-M., Faye, B., Le Fur, E., & Storchmann, K. (2017). The Law of One Price ? Price Dispersion on the Auction Market for Fine Wine. *Journal of Wine Economics*, *12*(3), 302–331.

Cen, L., Wei, K. C. J., & Yang, L. (2017). Disagreement, Underreaction, and Stock Returns. *Management Science*, *63*(4), 1214–1231.

Chen, Y., Wu, C., & Yeh, C.-Y. (2022). Asset Pricing Tests of Infrequently Traded Securities : The Case of Municipal Bonds. *Review of Asset Pricing Studies*, *12*(January), 754–807.

Cosemans, M., & Frehen, R. (2021). Salience theory and stock prices: Empirical evidence. *Journal of Financial Economics*, *140*(2), 460–483.

Daniel, K., & Hirshleifer, D. (2015). Overconfident Investors, Predictable Returns, and Excessive Trading. *Journal of Economic Perspectives*, *29*(4—Fall), 61–88.

David, G., Oosterlinck, K., & Szafarz, A. (2013). Art market inefficiency. *Economics Letters*, *121*(1), 23–25.

De Felice, S., De C. Hamilton, A., Ponari, M., & Vigliocco, G. (2022). Learning from others is good , with others is better: the role of social interaction in human acquisition of new knowledge. *Philosophical Transactions B*, *378*(20210357).

Demiguel, V., Garlappi, L., & Uppal, R. (2007). Optimal Versus Naive Diversification: How Inefficient is the 1 / N Portfolio Strategy? *Review of Financial Studies*, *22*(5), 1915–1953.

Dickerson, A., Mueller, P., & Robotti, C. (2023). Priced risk in corporate bonds. *Journal of Financial Economics*, *150*(2), 103707.

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance*, *57*(5), 2113–2141.

Dimson, E., Pukthuanthong, K., & Vorsatz, B. (2023). *Convenience Yields of Collectibles*.

Dimson, E., Rousseau, P. L., & Spaenjers, C. (2015). The price of wine. *Journal of Financial Economics*, *118*(2), 431–449.

Dimson, E., & Spaenjers, C. (2011). Ex-post: The investment performance of collectible stamps. *Journal of Financial Economics*, *100*(2), 443–458.

Disli, M., Inghelbrecht, K., Schoors, K., & Stieperaere, H. (2019). *Stock Price Anchoring* (Issue March).

Eyster, E., Rabin, M., & Vayanos, D. (2019). Financial Markets Where Traders Neglect the Informational Content of Prices. *The Journal of Finance*, *LXXIV*(1), 371–399.

Giglio, S., Maggiori, M., & Stroebel, J. (2015). Very Long-Run Discount Rates. *Quarterly Journal of Economics*, *130*(1), 1–53.

Giglio, S., & Shue, K. (2014). No News is News: Do Markets Underreact to Nothing? *Review of Financial Studies*, *27*(12), 3389–3440.

Goetzmann, W. N., & Massa, M. (2005). Dispersion of opinion and stock returns. *Journal of Financial Markets*, *8*, 325–350.

Goetzmann, W. N., Spaenjers, C., & Van Nieuwerburgh, S. (2021). Real and Private-Value Assets. *Review of Financial Studies*, *34*(8), 3497–3526.

Hadj Ali, H., Lecocq, S., & Visser, M. (2008). The Impact of Gurus: Parker Grades and En Primeur Wine Prices. *Economic Journal*, *118*(June), 158–173.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, *47*(1), 153–161.

Hillert, A., Jacobs, H., & Müller, S. (2018). Journalist disagreement. *Journal of Financial Markets*, *41*, 57–76.

Hirshleifer, D., & Plotkin, J. B. (2021). Moonshots, investment booms, and selection bias in the transmission of cultural traits. *Proceedings of the National Academy of Sciences*, *118*(26), 1–9.

Hong, H., & Stein, J. C. (1999). A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets. *The Journal of Finance*, *LIV*(6).

Huang, D., & Goetzmann, W. N. (2023). *Selection-Neglect in the NFT Bubble* (Working Paper 31498; NBER Working Paper Series).

Jehiel, P. (2018). Investment Strategy and Selection Bias: An Equilibrium Perspective on Overoptimism. *American Economic Review*, *108*(6), 1582–1597.

Jiang, H., & Sun, Z. (2014). Dispersion in beliefs among active mutual funds and the cross-section of stock returns. *Journal of Financial Economics*, *114*(2), 341–365.

Koehler, J. J., & Mercer, M. (2009). Selection Neglect in Mutual Fund Advertisements. *Management Science*, *55*(7), 1107–1121.

Korteweg, A., Kräussl, R., & Verwijmeren, P. (2016). Does it Pay to Invest in Art? A Selection-corrected Returns Perspective. *Review of Financial Studies*, *29*(4), 1007–1038.

Korteweg, A., & Sorensen, M. (2010). Risk and Return Characteristics of Venture Capital-Backed Entrepreneurial Companies. *Review of Financial Studies*, *23*(10), 3738–3772.

Kräussl, R., Lehnert, T., & Rinne, K. (2013). The search for yield: Implications to alternative investments. *Journal of Empirical Finance*.

Le Fur, E., & Outreville, J.-F. (2019). Fine wine returns: a review of the literature. *Journal of Asset Management*, *20*(3), 196–214.

Li, Z., Van Nieuwerburgh, S., & Renxuan, W. (2023). *Understanding Rationality and Disagreement in House Price Expectations* (Working Paper 31516; NBER Working Paper Series).

Lin, H., Wang, J., & Wu, C. (2011). Liquidity risk and expected corporate bond returns. *Journal of Financial Economics*, *99*(3), 628–650.

López-pérez, R., Pintér, Á., & Sánchez-mangas, R. (2022). Some conditions (not) affecting selection neglect: Evidence from the lab. *Journal of Economic Behavior and Organization*, *195*, 140–157.

Lovo, B. S., & Spaenjers, C. (2018). A Model of Trading in the Art Market †. *American Economic Review*, *108*(3), 744–774.

Lustig, H. N., & Nieuwerburgh, S. G. Van. (2005). Housing Collateral, Consumption Insurance, and Risk Premia: An Empirical Perspective. *The Journal of Finance*, *60*(3), 1167–1219.

Masset, P., & Henderson, C. (2010). Wine as an Alternative Asset Class. *Journal of Wine Economics*, *5*(1), 87–118.

Masset, P., & Weisskopf, J.-P. (2015). Wine Funds: An Alternative Turning Sour? *Journal of Alternative Investments*, *17*(4), 6.

McManus, G., Sharma, R., & Tezel, A. (2013). Reversals in Wine Auction Prices. *Journal of Wine Economics*, *8*(2), 189–197.

Ngangoué, K. M., & Weiszacker, G. (2021). Learning from Unrealized versus Realized Prices. *American Economic Journal: Microeconomics*, *13*(2), 174–201.

Outreville, J.-F. (2013). The Price of Wine: Does the Bottle Size Matter? In M.-C. Pichery & E. Giraud-Héraud (Eds.), *Wine Economics: Quantitative Studies and Empirical Applications* (Applied Ec, pp. 88–101). Palgrave Macmillan UK.

Petersen, M. A. (2009). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *Review of Financial Studies*, *22*(1), 435–480.

Renneboog, L., & Spaenjers, C. (2013). Buying Beauty: On Prices and Returns in the Art Market. *Management Science*, *59*(1), 36–53.

Shi, Y., Xu, H., Wang, M., & Conroy, P. (2017). Home bias in domestic art markets : Evidence from China. *Economics Letters*, *159*, 201–203.

Shleifer, A., & Vishny, R. W. (1997). The Limits of Arbitrage. *The Journal of Finance*, *LII*(1), 35–55.

Shue, K., & Townsend, R. R. (2019). Can the Market Multiply and Divide? Non-Proportional Thinking in Financial Markets. *The Journal of Finance*, *76*(5), 2307–2357.

Tversky, A., & Kahneman, D. (1971). Beliefs in the Law of Small Numbers. *Psychological Bulletin*, *76*(2), 105–110.

Weber, M. (2018). Cash flow duration and the term structure of equity returns. *Journal of Financial Economics*, *128*(3), 486–503.

Whitaker, A., & Kräussl, R. (2020). Fractional Equity, Blockchain, and the Future of Creative Work. *Management Science*, *66*(10), 4594–4611.

Figures

Figure 1: Distribution of the number of auction houses

This figure presents the monthly distribution of the number of auction houses in the sample and the percentage of traded value per auction house over the entire sample period.

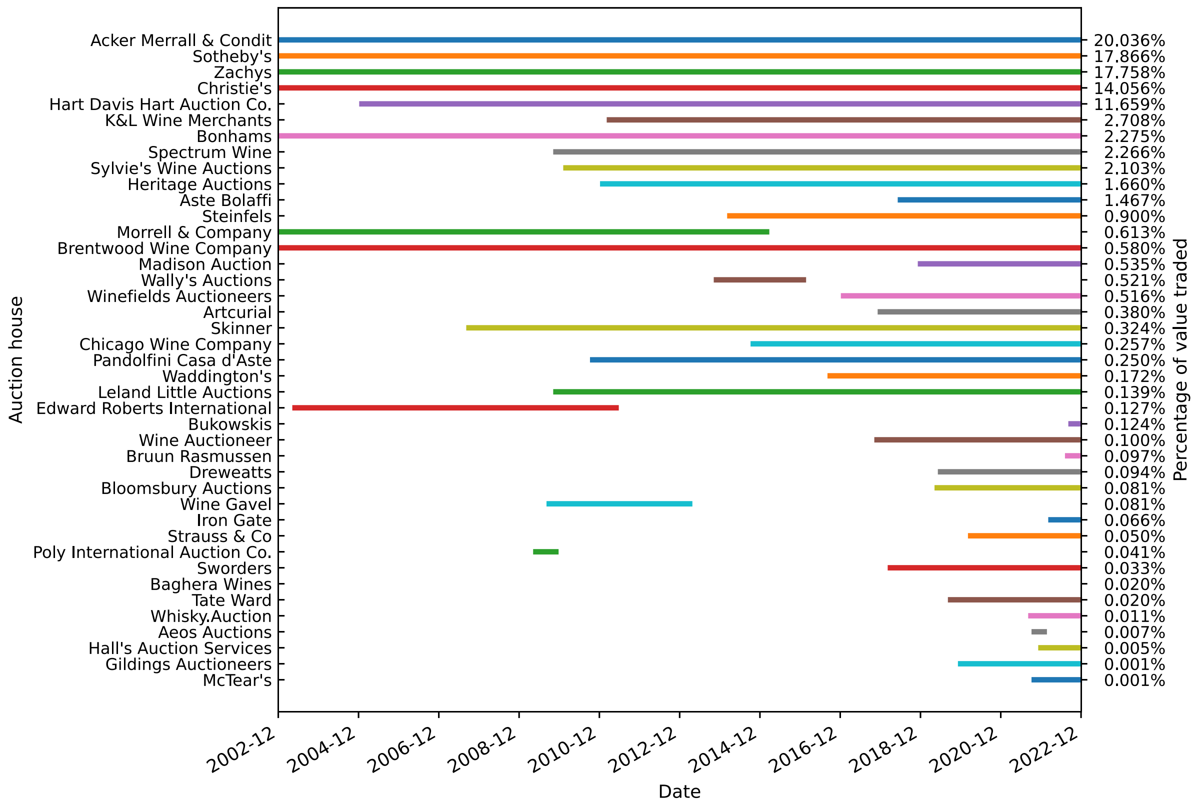


Figure 2: Research sample

This figure presents the monthly number of wines being auctioned for those with at least one trade in the next 12 months (orange) and those that do not have at least one trade in the next 12 months (blue).

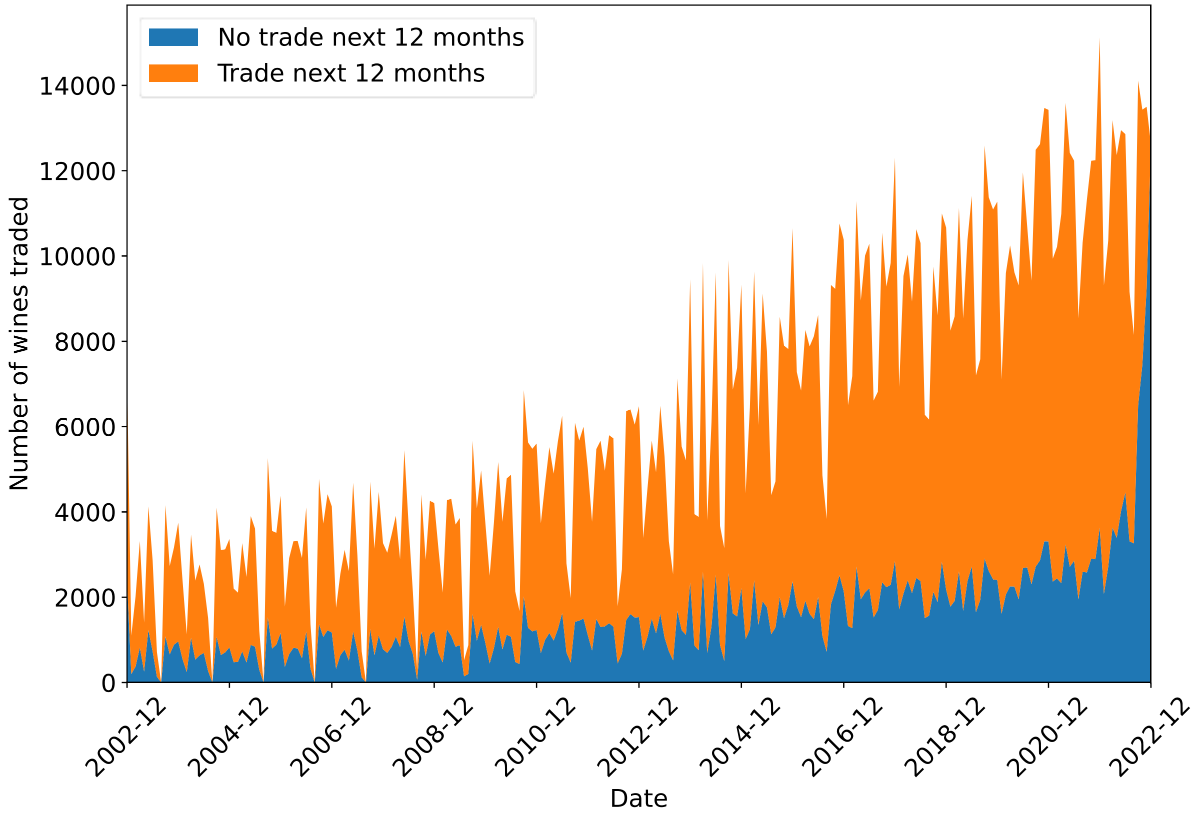


Figure 3: Geographical presence of wine

This figure plots all countries (blue) that have at least one wine traded from its respective countries in our sample.

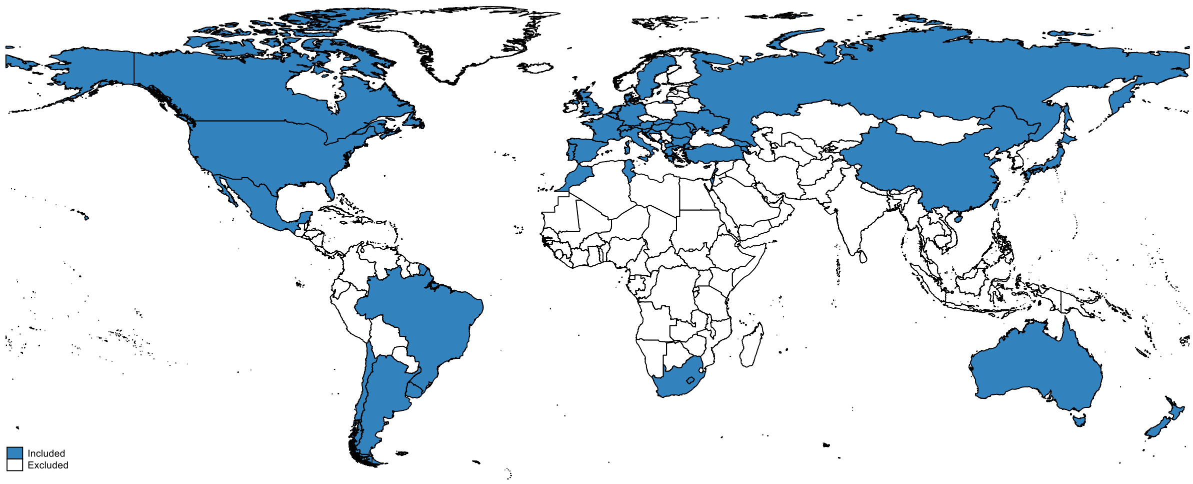


Figure 4: Probability of trading

This figure plots the probability of sales with respect to potential holding returns at various horizons (ranging from 1 to 12 months), estimated by the MCMC methodology by Korteweg and Sorensen (2010).

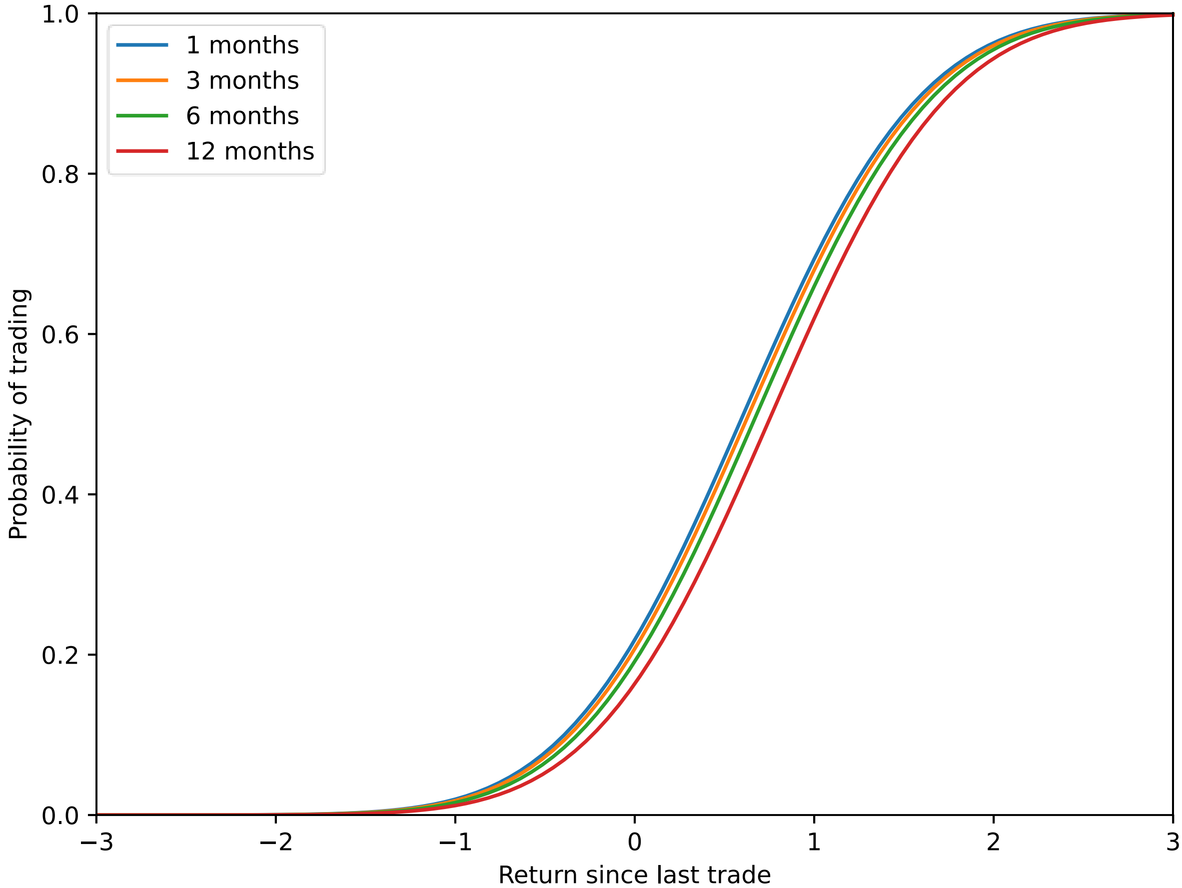
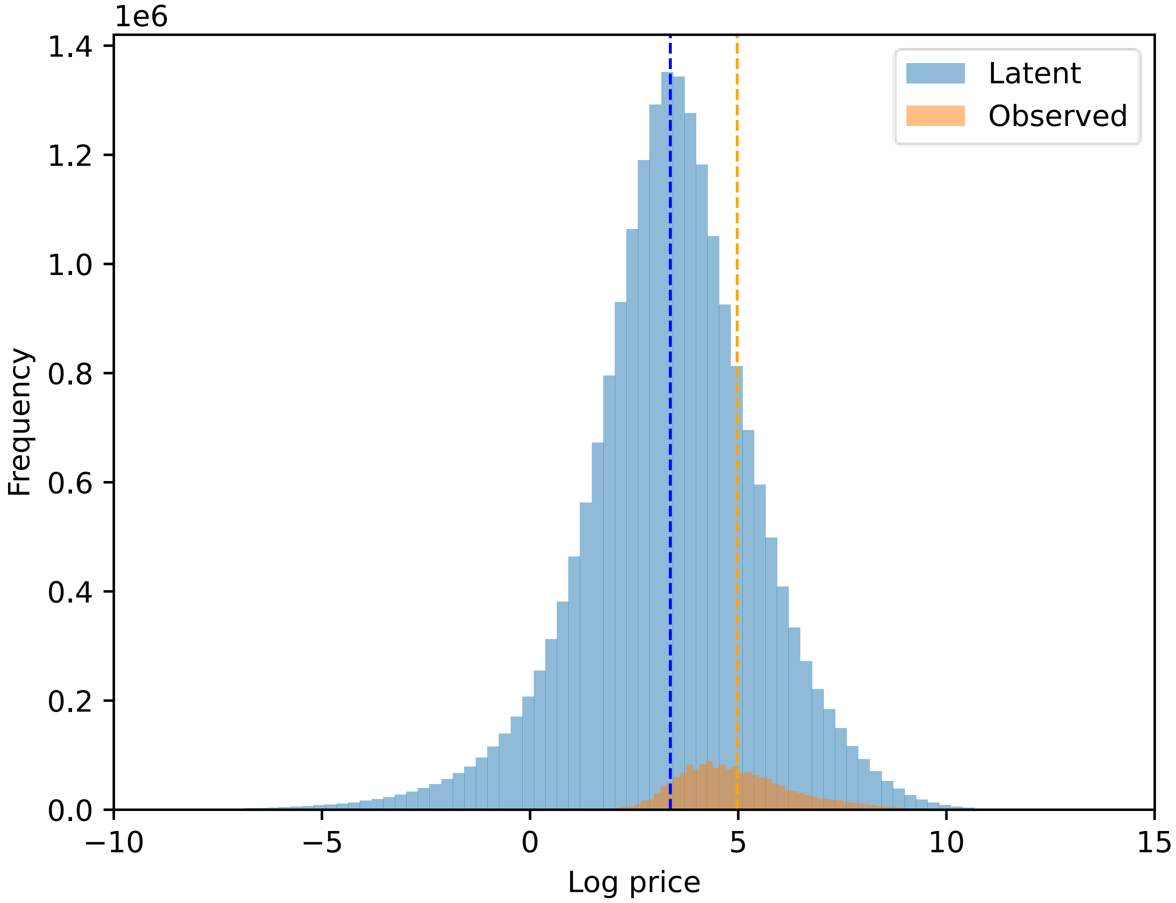


Figure 5: Distribution of latent and observed prices

This figure shows the distribution of observed transaction prices and the latent prices of untraded wines estimated by the MCMC selection adjustment procedure of Korteweg and Sorensen (2010).



Tables

Table 1: Summary statistics

This table reports the summary statistics. We define *Selection Bias* as the 12-month difference between the observed (*Biased return*) and the Markov Chain Monte Carlo's return (*Selection-adjusted return*). We require that the wine has at least two trades in the last 12 months and exclude wines below $20 and above $50,000. *Beta* is the coefficient of a regression model with excess price-weighted market returns as the independent variable (CAPM) using a 36-month rolling regression. *Illiquidity* is the autocovariance of returns over the last 12 months multiplied by -1. *MAX* is the maximum monthly return over the last 12 months. *MIN* is the minimum monthly return over the last 12 months. *Momentum* is the cumulative return over an 11-month period ending two months before the current month. *Price* is the dollar price. *Rating* is a number provided by Wine Advocate, where 60 is the lowest and 100 is the highest rating. *Reversal* is the last month's selection-adjusted return. The time to expiration (*TTE*) is the number of years to the end of the drinking window provided by Wine Advocate and zero otherwise. Period: January 2003 to December 2022.

			Percentiles		
	Mean	Std. Dev.	1st	50th	99th
Selection bias	-0.064	0.858	-2.793	0.000	1.371
Biased return	0.022	0.259	-0.440	0.000	0.857
Selection-adjusted return	0.018	0.285	-0.469	-0.014	0.823
Beta	0.536	3.353	-5.556	0.291	8.477
Illiquidity	0.011	0.027	-0.041	0.009	0.086
MAX	0.466	0.793	0.000	0.330	2.642
MIN	-0.225	0.171	-0.696	-0.212	0.000
Momentum	0.160	0.729	-0.585	0.060	2.038
Price	608.506	1,585.135	28.512	177.244	7,499.582
Rating	93.928	3.579	85	94	100
Reversal	0.022	0.259	-0.440	0.000	0.857
TTE	12.269	11.264	0	10	48

Table 2: Portfolio optimization

This table reports the performance statistics. Every month, we create portfolios in an MV optimization using either biased (Negligent) or selection-adjusted returns (Non-Negl.). Weights from the optimization are obtained for the previous 36 months and used to calculate one, six, and 12-month ahead (t) certainty-equivalent returns for investors with a risk aversion parameter equal to one (Panel A) and five (Panel B). We split the total sample into five countries: France, the USA, Italy, Spain, and Portugal.

Panel A: $\gamma = 1$

	t = 1		t = 6		t = 12	
	Non-Negl.	Negligent	Non-Negl.	Negligent	Non-Negl.	Negligent
France	0.026	0.026	0.034	0.030	0.035	0.032
USA	0.018	0.013	0.037	0.022	0.039	0.026
Italy	0.022	0.014	0.033	0.028	0.035	0.031
Spain	0.017	0.013	0.032	0.022	0.035	0.027
Portugal	0.017	0.016	0.029	0.025	0.031	0.028
Average	0.020	0.016	0.033	0.025	0.035	0.029

Panel B: $\gamma = 5$

	t = 1		t = 6		t = 12	
	Non-Negl.	Negligent	Non-Negl.	Negligent	Non-Negl.	Negligent
France	-0.018	-0.019	0.012	0.007	0.022	0.016
USA	-0.011	-0.017	0.019	0.004	0.027	0.012
Italy	-0.008	-0.023	0.010	-0.002	0.018	0.009
Spain	-0.035	-0.053	-0.004	-0.018	0.010	-0.002
Portugal	-0.029	-0.053	-0.003	-0.012	0.004	-0.003
Average	-0.020	-0.033	0.007	-0.004	0.016	0.006

Table 3: Univariate regressions

This table reports portfolio returns and alphas. At the end of every month, quintile portfolios are formed by sorting wines on selection bias (SB). Quintiles 1 (Low) and 5 (High) are portfolios with the lowest SB and highest SB values. Portfolio returns in month *t+1* are price-weighted using the wine's observed prices as weights if available and latent prices otherwise. We calculate alphas using stock, bond, or wine data. The stock factor model (*Stocks*) includes the excess stock return and factors on size, value, investments, and operating profitability for the developed markets. The bond factor model (*Bonds*) includes the excess corporate bond return and factors on downside risk, duration, credit risk, and long-term reversal. The wine factor model (*Wine*) includes the excess wine return and factors on momentum and illiquidity. Newey-West t-statistics are given in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

	SB_t	Ret_{t+1}	α_{t+1}^{Stocks}	α_{t+1}^{Bonds}	α_{t+1}^{Wine}
1 (Low)	-1.042	-0.055	-0.057***	-0.057***	-0.060***
			(-18.854)	(-20.047)	(-8.901))
2	-0.168	-0.012	-0.014***	-0.013***	--0.018***
			(-5.167)	(-5.106)	(-4.886)
3 (Mid)	0.010	0.002	0.001	0.001	0.014**
			(0.394)	(0.550)	(2.296)
4	0.200	0.025	0.023***	0.024***	0.024***
			(10.484)	(11.911)	(5.726)
5 (High)	0.692	0.018	0.016***	0.016***	0.040***
			(6.665)	(6.731)	(11.201)
High – Low	1.734	0.073	0.073***	0.073***	0.100***
			(18.852)	(20.028)	(13.865)

Table 4: Characteristics of SB-sorted portfolios

This table reports wine characteristics for portfolios formed on the basis of Selection Bias. At the end of every month, quintile portfolios are formed by sorting wine on selection bias, as defined in Table 1. Quintiles 1 (Low) and 5 (High) are portfolios with the lowest and highest SB values. Portfolios are price-weighted using the wine's observed prices as weights if available and latent prices otherwise. All other variables are defined in Table 1. . ∗, ∗∗, and ∗∗∗ indicate significance at the 10%, 5%, and 1% levels, respectively.

	1 (Low)	2	3 (Mid)	4	5 (High)	High - Low
Beta	0.415	0.523	0.758	0.478	0.597	0.182
IVOL	0.164	0.180	0.212	0.172	0.191	0.027
Illiquidity	0.009	0.012	0.018	0.012	0.008	-0.002
MIN	-0.310	-0.335	-0.342	-0.352	-0.379	-0.069
MAX	0.671	0.571	0.572	0.527	0.547	-0.124
Momentum	0.157	0.024	0.064	0.070	0.476	0.319
Price	612.188	525.605	551.169	428.084	407.971	-204.217
Rating	93.579	93.865	94.242	93.756	93.351	-0.228
Reversal	0.119	0.064	0.040	-0.038	-0.089	-0.034
TTE	12.680	13.554	15.330	13.148	11.807	-0.873

Table 5: Bivariate regressions

This table reports portfolio alphas. At the end of every month, quintile portfolios are formed by sorting wines on a control variable, as defined in Table 1. Then, within every quintile, wine is further sorted into quintiles on selection bias (SB). We calculate portfolio alphas using a 10-factor (stock and bond) model. In Panel H, we define investment-grade (IG) as wine with a rating above 95, and non-investment-grade (NIG) as wine with a rating below 90. In Panel K, we define *Short* and *Long* as the 20% number of wines with either the shortest or longest time to expiration (TTE). Newey-West *t*-statistics are given in parentheses. ∗, ∗∗, and ∗∗∗ indicate significance at the 10%, 5%, and 1% level, respectively.

	Panel A: Beta		**Panel B: Illiquidity**		**Panel D: MAX**	
	Low	High	Low	High	Low	High
Low	-0.029***	-0.036***	-0.014***	-0.014***	0.022***	-0.034***
	(-10.311)	(-16.446)	(-6.639)	(-6.589)	(9.683)	(-14.736)
High	0.012***	-0.012***	0.013***	0.014***	0.028***	-0.007***
	(3.903)	(-6.649)	(6.121)	(5.340)	(15.244)	(-2.723)
High – Low	0.041***	0.024***	0.027***	0.028***	0.007***	0.026***
	(12.786)	(9.651)	(11.234)	(13.964)	(2.604)	(10.844)

	Panel E: MIN		**Panel E: Momentum**		**Panel F: Price**	
	Low	High	Low	High	Low	High
Low	0.016***	-0.041***	0.010*	-0.071***	-0.024***	-0.060***
	(5.826)	(-18.553)	(1.948)	(-28.100)	(-19.655)	(-16.085)
High	0.021***	-0.010***	0.028***	-0.055***	0.058***	0.010***
	(10.098)	(-4.423)	(5.363)	(-28.896)	(21.182)	(3.665)
High – Low	0.005**	0.031***	0.018***	0.016***	0.082***	0.069***
	(2.569)	(15.092)	(3.765)	(5.613)	(35.621)	(14.740)

	Panel H: Rating		**Panel H: Reversal**		**Panel I: TTE**	
	NIG	IG	Low	High	Short	Long
Low	-0.003	-0.018***	0.048***	-0.059***	-0.001	-0.013***
	(-1.260)	(-4.313)	(18.869)	(-21.097)	(-0.311)	(-2.914)
High	0.009*	0.006**	0.010***	0.000	0.012**	0.002
	(1.958)	(1.979)	(5.225)	(0.017)	(2.089)	(0.350)
High – Low	0.013**	0.024***	-0.038***	0.059***	0.013*	0.014**
	(2.342)	(5.443)	(-13.669)	(15.557)	(1.952)	(2.523)

Table 6: Wine-level panel regressions

This table reports coefficients for panel regressions. Monthly cross-sectional regressions run on excess wine returns in month *t+1* (Panel A) and in month *t+h* (Panel B) on selection bias and a vector of characteristics measured at the end of month *t*, as defined in Table 1. We add wine fixed effects. The *t*-statistics in parentheses are based on standard errors clustered by individual wine. Adjusted R-squared (Adj. R2) is provided in the last row. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: One-month ahead

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
Selection Bias	0.032***	0.026***	0.031***	0.032***	0.039***	0.036***	0.034***	0.029***	0.030***	0.051***	0.055***
	(12.568)	(13.381)	(11.401)	(12.483)	(9.721)	(15.743)	(12.577)	(12.683)	(2.580)	(4.019)	(2.803)
Log(Price)		-0.008***									-0.006***
		(-38.686)									(-12.934)
Beta			-0.001***								-0.001*
			(-5.220)								(-1.900)
Illiquidity				-0.140***							-0.268***
				(-7.524)							(-10.379)
Momentum					-0.063***						-0.116***
					(-9.764)						(-35.184)
MAX						-0.026***					-0.025***
						(-19.963)					(-8.571)
MIN							-0.110***				0.029***
							(-7.299)				(3.968)
Reversal								-0.138***			-0.165***
								(-27.982)			(-28.840)
Rating									0.060***		0.101***
									(2.727)		(2.764)
TTE										-0.000	0.000
										(-0.557)	(0.146)
Constant	0.014	0.053	0.017	0.016	0.026	0.027	-0.010	0.017	-0.250	0.022	-0.418
	(84.177)	(56.245)	(79.587)	(59.083)	(25.455)	(38.646)	(-36.922)	(72.246)	(-2.522)	(15.595)	(-2.526)
Adj. R2	0.006	0.008	0.007	0.008	0.027	0.012	0.011	0.024	0.005	0.011	0.060

Panel B: Long-run returns

	t = 2	t = 3	t = 4	t = 5	t = 6	t = 7	t = 8	t = 9	t = 10	t = 11	t = 12
Selection Bias	0.075***	0.091***	0.102***	0.111***	0.116***	0.121***	0.124***	0.127***	0.129***	0.131***	0.133***
	(3.359)	(5.889)	(7.606)	(8.576)	(8.873)	(9.086)	(0.204)	(9.161)	(8.996)	(8.617)	(8.162)
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R2	0.159	0.158	0.158	0.157	0.157	0.156	0.156	0.156	0.156	0.156	0.155

Table 7: Investor attention

This table reports coefficients for panel regressions. Monthly cross-sectional regressions run on excess wine returns in month *t+1* on selection bias, a vector of characteristics, and five measures of investor attention. We use abnormal returns and volume, salience, Google Trends, and Top 100. *Abnormal returns* is the difference between an individual wine's return and the wine market return over the last 12 months. *Abnormal volume* is the number of trades over the last 12 months over the number of trades made one year before. We use a logarithmic transformation of abnormal volumes. *Google Trends* is the monthly year-on-year difference in worldwide Google searches for wines, available from 2015. *Top 100* is an indicator variable equal to one if the wine is included in the Wine Spectator's Top 100 from November in its respective year onwards and zero otherwise. We include wine control variables as defined in Table 1 and wine fixed effects. The *t*-statistics in parentheses are based on standard errors clustered by individual wine. Adjusted R-squared (Adj. R2) is provided in the last row. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	1.	2.	3..	4.	5.
Selection Bias	0.054***	0.053***	0.078***	0.053***	0.053***
	(4.702)	(4.468)	(5.870)	(4.487)	(4.379)
Abnormal Returns	-0.052***				
	(-7.284)				
Abnormal Volume		-0.000			
		(-1.602)			
SB x Salience			-0.106***		
			(-3.287)		
SB x Google				0.003***	
				(3.200)	
SB x Top 100					0.015**
					(2.041)
Controls	Yes	Yes	Yes	Yes	Yes
Adj. R2	0.057	0.056	0.060	0.056	0.056

Table 8: Limits to arbitrage

This table reports coefficients for panel regressions. Cross-sectional regressions are run of excess returns in month *t+1* on selection bias (SB), a vector of wine-level characteristics, and interaction terms between SB and proxies for limits to arbitrage constructed measured at the end of the previous month *t*. We include salience, price, illiquidity, beta, time to expiration (TTE), idiosyncratic volatility (IVOL), maximum monthly return (MAX), minimum monthly return (MIN), momentum, and reversal, as defined in Table 1. We add wine fixed effects. *T*-statistics in parentheses are based on standard errors clustered by individual wine. The adjusted R-squared (Adj. R2) is provided in the last row. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

	1.	2.	3.	4.
Selection Bias	0.053***	0.051***	0.056***	0.048***
	(5.226)	(8.845)	(4.738)	(4.124)
SB x Illiquidity	0.142**			
	(2.237)			
SB x TTE		0.000		
		(1.208)		
SB x IVOL			-0.007	
			(-0.806)	
SB x Log(Price)				0.004***
				(4.932)
Controls	Yes	Yes	Yes	Yes
Adj. R2	0.056	0.055	0.056	0.056

Table 9: Preferences

This table reports portfolio returns and alphas. At the end of every month, quintile portfolios are formed by sorting wines outside drinking window on selection bias (SB). Quintiles 1 (Low) and 5 (High) are portfolios with the lowest SB and highest SB values. Portfolio returns in month *t+1* are price-weighted using the observed prices as weights if available and latent prices otherwise. We calculate alphas using stock, bond, or wine market data. The stock factor model (*Stocks*) includes the excess stock return and factors on size, value, investments, and operating profitability for the developed markets. The bond factor model (*Bonds*) includes the excess corporate bond return and factors on downside risk, duration, credit risk, and long-term reversal. The wine factor model (*Wine*) includes the excess wine return and factors on momentum and illiquidity. Newey-West t-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	SB_t	Ret_{t+1}	α_{t+1}^{Stocks}	α_{t+1}^{Bonds}	α_{t+1}^{Wine}
1 (Low)	-1.071	-0.056	-0.058***	-0.059***	-0.039
			(-5.792)	(-6.533)	(-1.563)
2	-0.171	-0.012	-0.007	-0.006	-0.001
			(-0.653)	(-0.534)	(-0.307)
3 (Mid)	0.013	0.002	0.011	0.013	0.015*
			(1.271)	(1.601)	(1.712)
4	0.212	0.027	0.042***	0.035***	0.066***
			(4.556)	(4.775)	(2.760)
5 (High)	0.720	0.017	0.024**	0.027***	0.042*
			(2.527)	(3.536)	(1.863)
High – Low	1.791	0.073	0.082***	0.086***	0.081**
			(6.304)	(7.792)	(2.461)

Table 10: Time variation in selection neglect

This table reports the coefficients of High-Low *Selection-Bias*-sorted portfolio returns on economic metrics. Quintiles 1 (Low) and 5 (High) are portfolios with the lowest and highest SB values. Portfolio returns in month *t+1* are price-weighted using observed prices as weights if available, and latent prices otherwise. We control for five stock market factors, five bond market factors, and three wine market factors: The stock factors includes the excess stock return and factors on size, value, investments, and operating profitability for developed markets, whereas the bond factor model includes the excess corporate bond market return and factors on downside risk, duration, credit risk, and long-term reversal (*Traditional factors*). The wine factor model (*Wine*) includes the excess wine market return and factors on momentum and illiquidity. The economic measures are defined in Table A.7. Newey-West t-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	1.	2.	3.
Panel A: Significant results			
Economic policy uncertainty	0.019***	0.022***	0.018***
	(4.191)	(4.546)	(3.632)
Short interest	-0.008***	-0.007***	-0.007***
	(-3.478)	(-3.557)	(-3.679)
Climate risk index	0.011***	0.012***	0.008**
	(2.908)	(3.306)	(2.134)
Tail index	0.193**	0.209**	0.151*
	(2.244)	(2.399)	(1.672)
Panel B: Mixed results			
Short-term stock market returns	-0.147*	-0.191**	-0.094
	(-1.848)	(-2.287)	(-0.966)
Short-term wine market returns	-0.324*	-0.411*	-0.091
	(-1.681)	(-1.927)	(-0.394)
Long-term wine market returns	-0.254	-0.272*	0.063
	(-1.638)	(-1.786)	(0.355)
Dividend yield	0.011	0.022**	-0.052
	(1.155)	(2.108)	(-0.656)
Default spread	-0.039	-0.046	-0.064**
	(-0.567)	(-1.233)	(-2.054)
Long-term stock market returns	0.049	0.011	0.223*
	(0.487)	(0.093)	(1.759)
Panel C: Insignificant results			
Term spread	0.078	0.078	0.065
	(1.219)	(1.219)	(1.109)
Inflation	-0.662	-0.203	-0.239
	(-1.114)	(-0.339)	(-0.403)
Risk aversion	0.003	0.005	-0.000
	(1.081)	(1.626)	(-0.012)
Disagreement	0.004	-0.055	-0.007
	(0.939)	(-0.631)	(-0.713)
Implied variance	-0.463	-0.203	-0.673
	(-0.895)	(-0.339)	(-1.063)
Flight-to-safety	0.006	0.005	0.011
	(0.316)	(0.263)	(0.571)
Geopolitical risk	-0.001	-0.004	0.002
	(-0.089)	(-0.360)	(0.232)
Traditional factors	No	Yes	No
Wine factors	No	No	Yes

Table 11: Conditional asset-pricing models

This table reports the portfolio alphas following periods of high and low sentiment. High sentiment (low sentiment) is defined as those months in which the VIX in the previous month is above (below) the median value for the sample period. At the end of each month, wine is sorted into the quintile portfolios based on their SB values. Portfolio Low (High) contains wines in the lowest (highest) quintile for SB values. Returns in month *t+1* are price-weighted using the wine's observed prices as weights if available and latent prices otherwise. We calculate alphas using stock and bond or wine factors. The last row reports differences in the alpha between Quintile 5 (High) and 1 (Low). Newey-West t-statistics are given in parentheses. ∗, ∗∗, and ∗∗∗ indicate significance at the 10%, 5%, and 1% levels, respectively.

	Market Stress					
	Traditional factors			Wine factors		
	Low	High	High-Low	Low	High	High-Low
Low	-0.028***	-0.029***	0.002	-0.036***	-0.024***	0.012
	(-6.581)	(-6.172)	(0.186)	(-3.403)	(-2.202)	(0.569)
2	-0.004**	-0.009***	0.005*	-0.014***	-0.004	0.010
	(-2.365)	(-5.021)	(1.850)	(-3.888)	(-0.934)	(1.532)
3 (Mid)	0.003	-0.002	0.004*	0.010	0.004	-0.006
	(1.308)	(-1.164)	(1.658)	(1.280)	(0.633)	(-0.456)
4	0.014***	0.009***	0.005	0.013**	0.011*	-0.002
	(6.000)	(3.878)	(1.251)	(2.271)	(1.840)	(-0.209)
High	0.010***	0.007***	0.003	0.019***	0.020***	-0.001
	(4.325)	(3.348)	(0.866)	(3.518)	(3.520)	(-0.069)
High – Low	0.038***	0.036***	0.002	0.056***	0.044***	-0.012
	(6.247)	(5.926)	(0.134)	(4.610)	(3.128)	(-0.492)

Table 12: Neglect amplification

This table reports portfolio alphas. At the end of every month, wine is sorted into the quintile portfolios based on their SB values. Portfolio Low (High) contains wines in the lowest (highest) quintile for SB values. Returns in month *t+1* are price-weighted using wines' observed prices as weights if available and latent prices otherwise. We calculate alphas using stock and bond (Panel A) or wine factors (Panel B). We include an indicator variable (Post-COVID-19) that yields one from March 2020 onwards, and zero otherwise. Newey-West t-statistics are given in parentheses. ∗, ∗∗, and ∗∗∗ indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Traditional factors						
	Low	2.	Mid	3.	High	High-Low
Post-COVID-19	-0.001	-0.003	0.004	0.013*	0.013***	0.013**
	(-0.076)	(-0.405)	(0.575)	(1.761)	(2.873)	(2.149)
Intercept	-0.058***	-0.014***	-0.000	0.022***	0.014***	0.072***
	(-16.195)	(-4.813)	(-0.147)	(11.178)	(5.599)	(16.565)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Wine factors						
	Low	2.	Mid	3.	High	High-Low
Post-COVID-19	-0.010**	-0.014***	-0.010***	0.002	0.006	0.015***
	(-2.500)	(-4.439)	(-3.486)	(0.644)	(1.643)	(2.641)
Intercept	-0.060***	-0.017***	0.015**	0.024***	0.040***	0.100***
	(-20.254)	(-3.994)	(2.146)	(16.336)	(6.346)	(17.553)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Appendix

Robustness checks

To highlight that the qualitative results are robust for all choices we made, we ran a battery of robustness tests. We categorize the tests into six groups:

(1) **Return calculation**: extending the holding period from three, six, and 12 months, using only observed prices as weights, and calculating alphas with behavioral factors and the Q5 factor model.

(2) **Sample splits**: splitting the sample on bottle sizes (750ml. and other sizes), colour (red, white, sparkling, and fortified wine), location of production (New and Old world), and different subperiods (2004 to 2012 and 2013 to 2022).

(3) **Auction house heterogeneity**: the type of auction house (i.e., internet or live exchange), its continent (i.e., Europe, Asia, and other continents), and relative importance (i.e., top 5 or bottom 36). We use an MCMC regression in each specification to calculate selection bias and wine returns.

(4) **Portfolio sort changes**: We calculated selection bias for one or four trades in the last 12 months, two trades every six months, and one or two transactions in the last 24 months, or we sorted wines into deciles relative to quintiles.

(5) **Restricting the minimum amount of trades**: in our main specification, we require wine to be traded at least twice over the entire sample period. We increase this number to at least five, ten, 20, and 30 trades, as highlighted in Figure 5.

(6) **Valuation model**: smaller changes to the valuation model by using average, minimum, or maximum prices, using the 10-factor model to capture the state of the economy, and allowing for wider priors in the MCMC model, as in Huang and Goetzmann (2023).

Tables

Table A.1. Summary statistics

This table reports the sample means based on a subsample of a minimum amount of trades. Variables are as defined in Table 1.

	Minimum of trades					
	2	5	10	20	30	2- 30
Selection Bias	-0.064	-0.071	-0.070	-0.060	-0.050	-0.014***
						(-20.533)
Beta	0.536	0.540	0.564	0.619	0.683	-0.147***
						(-42.479)
Illiquidity	0.011	0.012	0.012	0.014	0.015	-0.004***
						(-136.846)
MIN	-0.225	-0.228	-0.235	-0.248	-0.259	0.034***
						(217.672)
MAX	0.466	0.468	0.472	0.486	0.502	-0.036***
						(-50.314)
Momentum	0.160	0.156	0.148	0.138	0.133	0.027***
						(40.651)
Price	608.506	613.026	626.084	644.715	658.537	-50.031***
						(-20.184)
Rating	93.928	93.959	94.046	94.178	94.209	-0.281***
						(-35.587)
Reversal	0.022	0.022	0.023	0.025	0.026	-0.004***
						(-16.153)
Salience	0.231	0.234	0.242	0.259	0.276	-0.045***
						(-172.412)
TTE	12.269	12.355	12.601	12.939	13.115	-0.846***
						(-32.384)

Table A.2. Correlation matrix

This table reports the correlation between wine characteristics, as defined in Table 1.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Beta	1									
2. IVOL	0.071	1								
3. Illiquidity	-0.003	-0.012	1							
4. MIN	-0.013	-0.276	-0.024	1						
5. MAX	0.05	0.423	-0.046	-0.153	1					
6. Momentum	0.038	0.193	-0.049	0.508	0.78	1				
7. Price	0.039	0.222	-0.019	0.24	0.404	0.503	1			
8. Rating	0.013	0.03	0.012	-0.016	0.026	0.017	0.086	1		
9. Reversal	-0.004	-0.052	-0.006	0.127	0.225	0.286	0.202	0.004	1	
10. Salience	0.033	-0.03	0.003	-0.167	0.205	0.081	0.058	0.014	0.232	1
11. TTE	-0.035	-0.084	-0.003	0.01	-0.056	-0.053	-0.171	0.102	-0.012	-0.035

Table A.3. Robustness evidence

This table reports the 10-factor alphas on the high-low selection-neglect storted portfolios. The presented robustness tests assume deviations from our baseline method. Changes are labeled as (1) return calculations, (2) sample splits, (3) wine auction heterogeneity, (4) portfolio sorts, and (5) minimum required trades. Portfolio Low (High) contains all wines with the lowest (highest) SB values. Portfolios are price- or equally-weighted. Newey-West t-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample period is January 2003 to December 2022.

	Price-weighted returns			**Equally-weighted returns**		
	Low	High	High-Low	Low	High	High-Low
			Panel A: Return calculations			
1. 3-month holding period	-0.139***	0.060***	0.199***	-0.107***	0.128***	0.236***
	(-26.948)	(9.185)	(30.902)	(-50.025)	(39.900)	(77.231)
2. 6-month holding period	-0.192***	0.112***	0.304***	-0.155***	0.217***	0.372***
	(-19.330)	(10.066)	(36.618)	(-40.379)	(37.904)	(83.315)
3. 12-month holding period	-0.198***	0.172***	0.369***	-0.185***	0.320***	0.504***
	(-9.592)	(5.947)	(23.978)	(-23.462)	(21.559)	(53.287)
4. Observed weighted	-0.045***	0.025***	0.070***	-0.035***	0.034***	0.069***
	(-6.417)	(3.549)	(6.995)	(-11.196)	(7.688)	(15.435)
5. Q5 risk factors	-0.057***	0.016***	0.073***	-0.044***	0.045***	0.088***
	(-18.398)	(6.615)	(19.563)	(-31.409)	(34.658)	(54.330)
6. Behavioral factors	-0.050***	0.021***	0.071***	-0.033***	0.064***	0.097***
	(-15.657)	(6.377)	(16.759)	(-16.269)	(31.577)	(38.579)
			Panel B: Sample splits			
7. Period 1: 2004 – 2013	-0.052***	0.011***	0.063***	-0.040***	0.045***	0.085***
	(-9.842)	(3.526)	(10.952)	(-31.917)	(29.818)	(52.706)
8. Period 2: 2014 – 2022	-0.063***	0.022***	0.084***	-0.047***	0.045***	0.092***
	(-18.396)	(9.231)	(27.371)	(-38.984)	(21.491)	(46.363)
9. 750ml bottle	-0.052***	0.017***	0.068***	-0.042***	0.048***	0.089***
	(-19.747)	(5.602)	(16.317)	(-36.844)	(39.024)	(57.220)
10. Other bottle sizes	-0.065***	0.016***	0.081***	-0.052***	0.036***	0.087***
	(-16.174)	(6.563)	(18.081)	(-24.743)	(26.683)	(46.544)
11. Red wine	-0.052***	0.019***	0.071***	-0.042***	0.048***	0.090***
	(-16.300)	(6.885)	(15.723)	(-33.615)	(42.598)	(57.714)
12. White wine	-0.058***	0.017***	0.075***	-0.048***	0.040***	0.088***
	(-17.278)	(3.800)	(14.225)	(-27.481)	(22.471)	(35.888)
13. Sparkling wine	-0.046***	0.016***	0.062***	-0.035***	0.041***	0.076***
	(-10.431)	(3.011)	(7.778)	(-17.345)	(11.345)	(18.479)
14. Fortified wine	-0.043***	-0.003	0.040***	-0.036***	0.022***	0.058***
	(-6.675)	(-0.424)	(4.423)	(-12.079)	(6.396)	(11.592)
15. Old World	-0.052***	0.016***	0.068***	-0.040***	0.045***	0.085***
	(-17.030)	(5.632)	(15.551)	(-29.268)	(38.018)	(49.920)
16. New World	-0.069***	0.041***	0.110***	-0.050***	0.051***	0.101***
	(-41.009)	(15.687)	(34.972)	(-40.319)	(25.862)	(51.200)

(Table A.3. continued)

Panel C: Auction house heterogeneity						
17. Top 5 auctions	-0.040***	0.004	0.043***	-0.034***	0.030***	0.064***
	(-10.769)	(1.351)	(8.607)	(-16.617)	(24.448)	(26.111)
18. Bottom 36 auctions	-0.058***	0.013***	0.071***	-0.041***	0.030***	0.071***
	(-34.185)	(4.652)	(22.382)	(-21.669)	(18.681)	(40.579)
19. Internet auctions	-0.057***	0.015***	0.072***	-0.041***	0.027***	0.069***
	(-31.960)	(8.494)	(30.972)	(-20.606)	(14.420)	(34.187)
20. Live auctions	-0.051***	0.009***	0.060***	-0.040***	0.038***	0.078***
	(-17.066)	(3.629)	(15.990)	(-25.106)	(28.390)	(48.137)
21. European auctions	-0.031***	-0.005**	0.026***	-0.026***	0.018***	0.044***
	(-10.931)	(-2.069)	(7.187)	(-13.791)	(12.180)	(18.025)
22. Asian auctions	-0.025***	-0.002	0.023***	-0.021***	0.016***	0.037***
	(-9.688)	(-0.461)	(4.467)	(-9.186)	(7.376)	(14.185)
23. U.S. auctions	-0.048***	0.013***	0.061***	-0.037***	0.029***	0.066***
	(-15.479)	(5.198)	(12.891)	(-27.295)	(28.592)	(45.849)
Panel D: Portfolio sorts						
24. Decile portfolios	-0.070***	0.004	0.074***	-0.058***	0.041***	0.098***
	(-21.271)	(1.400)	(15.640)	(-50.662)	(25.792)	(51.367)
25. All wines	-0.056***	0.013***	0.069***	-0.045***	0.040***	0.085***
	(-16.119)	(4.237)	(13.596)	(-39.658)	(34.370)	(59.743)
26. 2 trades every six months	-0.057***	-0.001	0.056***	-0.044***	0.027***	0.070***
	(-17.956)	(-0.412)	(10.624)	(-36.197)	(24.287)	(42.070)
27. 1 trade every 12 months	-0.062***	0.003	0.065***	-0.053***	0.030***	0.083***
	(-17.495)	(0.953)	(16.638)	(-40.202)	(25.886)	(65.164)
28. 4 trades every 12 months	-0.044***	0.029***	0.073***	-0.028***	0.058***	0.086***
	(-12.274)	(9.464)	(14.641)	(-19.132)	(46.426)	(58.151)
29. 1 trade every 24 months	-0.051***	0.017***	0.068***	-0.038***	0.045***	0.083***
	(-13.098)	(8.512)	(15.242)	(-28.198)	(36.040)	(54.185)
30. 2 trades every 24 months	-0.050***	0.020***	0.071***	-0.036***	0.053***	0.088***
	(-13.251)	(8.426)	(16.923)	(-32.924)	(48.484)	(63.507)
Panel E: Minimum amount of trades in the sample						
31. At least five trades	-0.057***	0.019***	0.075***	-0.042***	0.050***	0.092***
	(-17.847)	(7.596)	(18.111)	(-37.833)	(39.048)	(54.472)
32. At least ten trades	-0.054***	0.025***	0.078***	-0.038***	0.058***	0.096***
	(-16.497)	(8.754)	(17.659)	(-36.753)	(34.413)	(51.056)
33. At least 20 trades	-0.049***	0.031***	0.080***	-0.032***	0.069***	0.101***
	(-15.630)	(10.901)	(18.778)	(-29.314)	(37.725)	(53.038)
34. At least 30 trades	-0.045***	0.035***	0.080***	-0.027***	0.075***	0.102***
	(-13.766)	(13.389)	(20.628)	(-23.555)	(43.865)	(60.912)
Panel F: Valuation model						
35. Using the average price	-0.048***	0.010***	0.058***	-0.037***	0.036***	0.073***
	(-14.532)	(4.448)	(13.594)	(-28.576)	(35.208)	(53.402)
36. Using the minimum price	-0.057***	0.014***	0.072***	-0.042***	0.043***	0.085***
	(-21.671)	(6.809)	(21.959)	(-33.450)	(42.174)	(57.542)
37. Using the maximum price	-0.055***	0.014***	0.070***	-0.042***	0.046***	0.088***
	(-17.693)	(5.934)	(19.059)	(-31.148)	(39.627)	(62.305)
38. 10-Factor model	-0.055***	0.014***	0.069***	-0.042***	0.044***	0.085***
	(-16.076)	(5.266)	(15.441)	(-26.651)	(35.509)	(49.365)
40. Changing priors	-0.054***	0.014***	0.068***	-0.042***	0.044***	0.087***
	(-15.968)	(5.262)	(16.155)	(-28.794)	(40.875)	(47.172)

Table A.4. Country-level univariate regressions

This table reports country-portfolio alphas. Quintile portfolios are formed monthly from January 2003 to December 2022 by sorting wines on selection bias (SB) for each country. We define SB as the 12-month difference between the observed and Markov Chain Monte Carlo return. Quintiles 1 (Low) and 5 (High) are portfolios with the lowest and highest SB values. Portfolio returns in month *t+1* are price-weighted using the wine's observed prices as weights if available and latent prices otherwise. We calculate alphas using stock and bond market data. The stock factor market model (*Stocks*) includes the excess stock market return and factors on size, value, investments, and operating profitability for the developed markets. The bond factor model (*Bonds*) includes the excess corporate bond market return and factors on downside risk, duration, credit risk, and long-term reversal. Newey-West t-statistics are given in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

	Belgium	France	Hong Kong	Singapore	Switzerland	Netherlands	UK	US
Stocks	0.029***	0.002	0.022***	0.004	0.013***	0.012**	0.016***	0.038***
	(5.854)	(0.293)	(6.484)	(0.501)	(3.269)	(2.608)	(4.611)	(12.111)
Bonds	0.029***	0.003	0.020***	0.006	0.015***	0.011***	0.016***	0.037***
	(5.736)	(0.407)	(6.301)	(0.708)	(4.424)	(2.869)	(4.824)	(12.719)

Table A.5. Alternative measure

This table reports portfolio alphas. At the end of every month, quintile portfolios are formed by sorting wines on selection bias (SB). Selection bias is measured as the difference between the observed and latent valuations relative to the latent valuation. Quintiles 1 (Low) and 5 (High) are portfolios with the lowest and highest SB values. Portfolio returns in month *t+1* are price-weighted using the wine's observed prices as weights if available and latent prices otherwise. We calculate alphas using stock, bond, or wine market data. Stock factor market model (*Stocks*) includes the excess stock market return and factors on size, value, investments, and operating profitability for the developed markets. Bond factor model (*Bonds*) includes the excess corporate bond market return and factors on downside risk, duration, credit risk, and long-term reversal. Wine factor model (*Wine*) includes the excess wine market return and factors on momentum and illiquidity. Newey-West t-statistics are given in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

	α_{t+1}^{Stocks}	α_{t+1}^{Bonds}	α_{t+1}^{Wine}
1 (Low)	-0.048***	-0.047***	-0.055***
	(-11.835)	(-12.114)	(-4.813)
2	-0.026***	-0.025***	0.019
	(-10.540)	(-10.900)	(1.282)
3 (Mid)	-0.013***	-0.013***	0.036***
	(-6.638)	(-7.539)	(4.475)
4	0.027***	0.028***	0.038***
	(9.080)	(10.115)	(3.525)
5 (High)	0.053***	0.053***	0.051***
	(24.328)	(23.381)	(5.983)
High – Low	0.101***	0.101***	0.106***
	(20.248)	(20.526)	(9.120)

Table A.6. Wine-level panel regressions: biased returns

This table reports coefficients for panel regressions. Monthly cross-sectional regressions run on biased wine returns in month *t+1* (Panel A) and in month *t+h* (Panel B) on selection bias and a vector of characteristics measured at the end of month *t*, as defined in Table 1. We add wine fixed effects. The *t*-statistics in parentheses are based on standard errors clustered by individual wine. Adjusted R-squared (Adj. R2) is provided in the last row. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
Selection Bias	0.034***	0.021***	0.034***	0.034***	0.056***	0.037***	0.030***	0.015***	0.026***	0.026***	0.013***
	(12.052)	(9.127)	(10.266)	(11.966)	(7.836)	(12.903)	(10.907)	(5.142)	(5.706)	(7.635)	(3.116)
Log(Price)		-0.613***									-0.645***
		(-62.484)									(-17.911)
Beta			-0.001								0.001*
			(-1.365)								(1.776)
Illiquidity				-0.093*							-0.333***
				(-1.766)							(-4.205)
Momentum					-0.266***						-0.039***
					(-16.092)						(-4.132)
MAX						-0.053***					0.002
						(-8.309)					(0.271)
MIN							-0.394***				-0.154***
							(-47.407)				(-7.776)
Reversal								-0.315			-0.105***
								(-14.623)			(-8.029)
Rating									-0.020		0.783***
									(-0.365)		(0.001)
TTE										0.002	0.006
										(0.625)	(0.937)
Constant	0.048	3.433	0.050	0.051	0.086	0.078	-0.066	0.066	0.132	0.036	0.091
	(208.495)	(63.366)	(75.111)	(43.793)	(38.494)	(21.892)	(-27.283)	(54.127)	(0.533)	(5.782)	(0.083)
Adj. R2	0.023	0.217	0.025	0.024	0.116	0.030	0.039	0.129	0.015	0.011	0.289

Table A.7. Variable definition

This table describes the variables used in the analysis of the time-variation of the selection neglect effect

Variable	Description
Climate risk index	New York Times Climate Risk Index by Giglio et al. (2024)
Default spread	Difference between BAA and AAA-rated corporate bond yields.
Disagreement	Individual stock analyst forecast dispersions
Dividend yield	Difference between the log of dividends and the log of lagged prices.
Economic policy uncertainty	Economic policy uncertainty index by Baker, Bloom, and Davis (2016)
Flight-to-safety	Percentage of the number of flight-to-safety episodes in a given month, from Baele et al. (2020)
Geopolitical risk	Geopolitical Risk index of Caldara and Iacoviollo (2018)
Implied variance	Implied stock market variance
Inflation	Consumer Price Index (All Urban Consumers) from the Bureau of Labor Statistics.
Long-term stock market returns	Average 12-month return on the value-weighted excess stock market return
Long-term wine market returns	Average 12-month return on the price-weighted excess wine market return
Risk aversion	Financial proxy to risk aversion by Bekaert, Engstrom, and Xu (2022)
Short interest	Short interest
Short-term stock market returns	Average three-month return on the value-weighted excess stock market return
Short-term wine market returns	Average three-month return on the price-weighted excess wine market return
Tail index	Time-varying tail risk of cross-sectional returns from Kelly and Jiang (2016)
Term spread	Difference between the long term yield on government bonds and the Treasury-bill.

Table A.8. Time variation in selection neglect: Equally-weighted returns

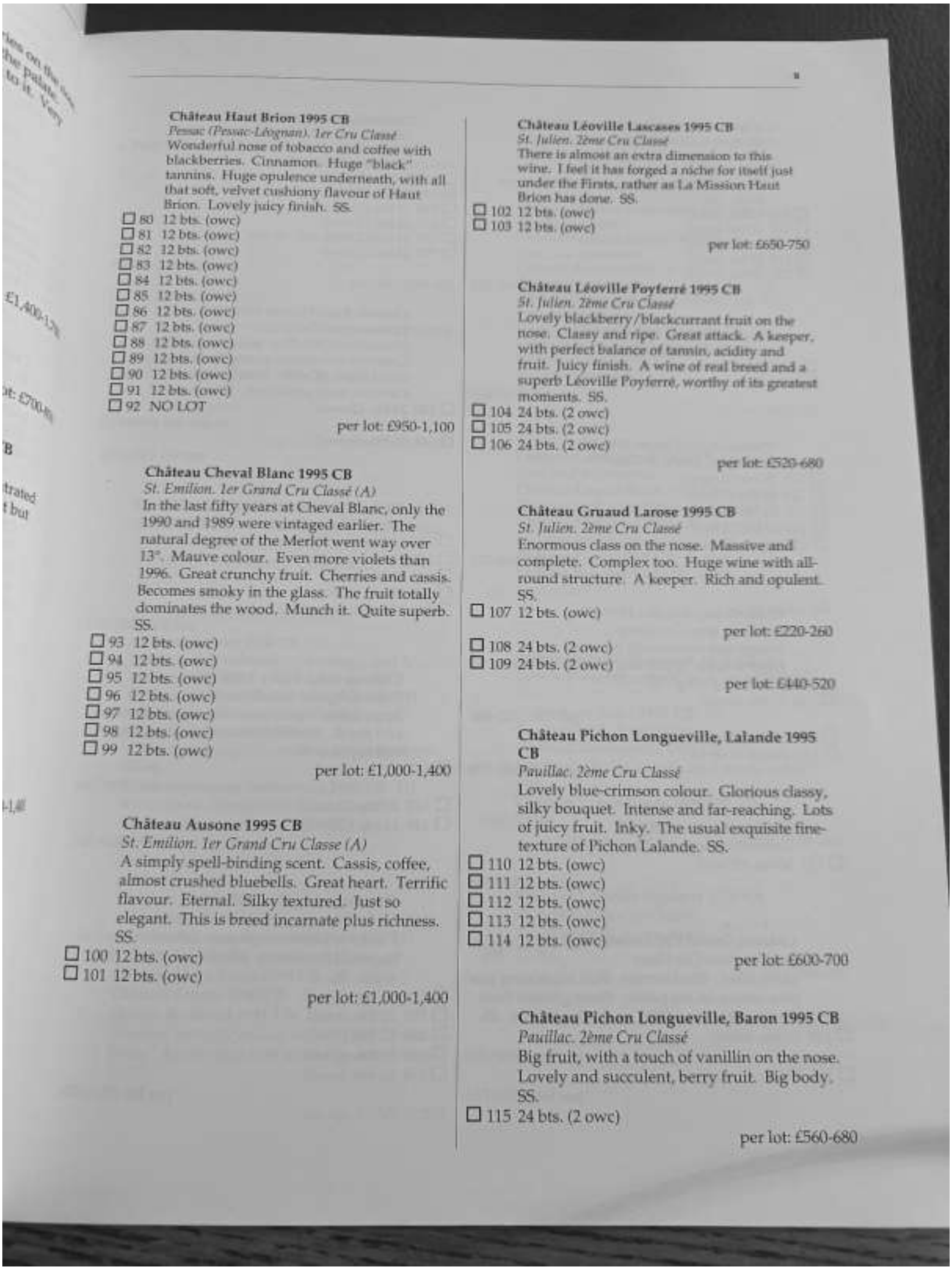
This table reports the coefficients of time-series regressions of High-Low Selection-Bias-sorted portfolios on various economic measures. Quintiles 1 (Low) and 5 (High) are portfolios with the lowest and highest SB values. Portfolio returns in month *t+1* are equally-weighted. The stock factor model includes the excess stock return and factors on size, value, investments, and operating profitability for developed markets. In contrast, the bond factor model includes the excess corporate bond market return and factors on downside risk, duration, credit risk, and long-term reversal (*Traditional factors*). The wine factor model (*Wine*) includes the excess wine market return and factors on momentum and illiquidity. The economic measures are defined in Table A.7. Newey-West t-statistics are given in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	1.	2.	3.
Tail index	0.113***	0.118***	0.093**
	(3.087)	(3.321)	(2.530)
Economic policy uncertainty	0.006**	0.006***	0.004*
	(2.658)	(2.854)	(1.714)
Climate risk index	0.001	-0.000	-0.001
	(0.056)	(-0.216)	(-0.426)
Short interest	-0.001	-0.002*	-0.008
	(-1.458)	(-1.683)	(-0.900)
Short-term stock market returns	-0.097***	-0.097**	-0.086**
	(-2.758)	(-2.447)	(-2.392)
Dividend yield	0.020***	0.022***	0.017**
	(2.953)	(3.145)	(2.414)
Short-term wine market returns	-0.115	-0.128	-0.100
	(-1.064)	(-1.134)	(-0.911)
Long-term wine market returns	-0.103	-0.103	-0.015
	(-1.544)	(-1.557)	(-0.193)
Default spread	-0.039	0.007	0.009
	(-1.185)	(0.418)	(0.558)
Long-term stock market returns	-0.075	-0.068	-0.049
	(-1.162)	(-1.019)	(-0.747)
Term spread	0.073***	0.072**	0.073***
	(2.692)	(2.567)	(2.814)
Inflation	-0.689***	-0.719***	-0.604**
	(-2.859)	(-2.742)	(-2.658)
Disagreement	0.003	-0.048	-0.084
	(0.191)	(-1.309)	(-0.357)
Flight-to-safety	0.013	0.012	0.011
	(1.354)	(1.188)	(1.138)
Geopolitical risk	-0.001	-0.000	0.000
	(-0.289)	(-0.067)	(0.041)
Implied variance	-0.050	-0.157	-0.084
	(-0.175)	(-0.538)	(-0.357)
Risk aversion	0.001	0.001	0.000
	(1.171)	(0.785)	(0.458)
Traditional factors	No	Yes	No
Wine factors	No	No	Yes

Figures

Figure A.1. Excerpt of the data

This figure displays the information from the Sotheby's archive.



Château Haut Brion 1995 CB
Pessac (Pessac-Léognan). 1er Cru Classé
Wonderful nose of tobacco and coffee with blackberries. Cinnamon. Huge "black" tannins. Huge opulence underneath, with all that soft, velvet cushiony flavour of Haut Brion. Lovely juicy finish. SS.

☐ 80 12 bts. (owc)
☐ 81 12 bts. (owc)
☐ 82 12 bts. (owc)
☐ 83 12 bts. (owc)
☐ 84 12 bts. (owc)
☐ 85 12 bts. (owc)
☐ 86 12 bts. (owc)
☐ 87 12 bts. (owc)
☐ 88 12 bts. (owc)
☐ 89 12 bts. (owc)
☐ 90 12 bts. (owc)
☐ 91 12 bts. (owc)
☐ 92 NO LOT

per lot: £950-1,100

Château Cheval Blanc 1995 CB
St. Emilion. 1er Grand Cru Classé (A)
In the last fifty years at Cheval Blanc, only the 1990 and 1989 were vintaged earlier. The natural degree of the Merlot went way over 13°. Mauve colour. Even more violets than 1996. Great crunchy fruit. Cherries and cassis. Becomes smoky in the glass. The fruit totally dominates the wood. Munch it. Quite superb. SS.

☐ 93 12 bts. (owc)
☐ 94 12 bts. (owc)
☐ 95 12 bts. (owc)
☐ 96 12 bts. (owc)
☐ 97 12 bts. (owc)
☐ 98 12 bts. (owc)
☐ 99 12 bts. (owc)

per lot: £1,000-1,400

Château Ausone 1995 CB
St. Emilion. 1er Grand Cru Classe (A)
A simply spell-binding scent. Cassis, coffee, almost crushed bluebells. Great heart. Terrific flavour. Eternal. Silky textured. Just so elegant. This is breed incarnate plus richness. SS.

☐ 100 12 bts. (owc)
☐ 101 12 bts. (owc)

per lot: £1,000-1,400

Château Léoville Lascases 1995 CB
St. Julien. 2ème Cru Classé
There is almost an extra dimension to this wine. I feel it has forged a niche for itself just under the Firsts, rather as La Mission Haut Brion has done. SS.

☐ 102 12 bts. (owc)
☐ 103 12 bts. (owc)

per lot: £650-750

Château Léoville Poyferré 1995 CB
St. Julien. 2ème Cru Classé
Lovely blackberry/blackcurrant fruit on the nose. Classy and ripe. Great attack. A keeper, with perfect balance of tannin, acidity and fruit. Juicy finish. A wine of real breed and a superb Léoville Poyferré, worthy of its greatest moments. SS.

☐ 104 24 bts. (2 owc)
☐ 105 24 bts. (2 owc)
☐ 106 24 bts. (2 owc)

per lot: £520-680

Château Gruaud Larose 1995 CB
St. Julien. 2ème Cru Classé
Enormous class on the nose. Massive and complete. Complex too. Huge wine with all-round structure. A keeper. Rich and opulent. SS.

☐ 107 12 bts. (owc)

per lot: £220-260

☐ 108 24 bts. (2 owc)
☐ 109 24 bts. (2 owc)

per lot: £440-520

Château Pichon Longueville, Lalande 1995 CB
Pauillac. 2ème Cru Classé
Lovely blue-crimson colour. Glorious classy, silky bouquet. Intense and far-reaching. Lots of juicy fruit. Inky. The usual exquisite fine-texture of Pichon Lalande. SS.

☐ 110 12 bts. (owc)
☐ 111 12 bts. (owc)
☐ 112 12 bts. (owc)
☐ 113 12 bts. (owc)
☐ 114 12 bts. (owc)

per lot: £600-700

Château Pichon Longueville, Baron 1995 CB
Pauillac. 2ème Cru Classé
Big fruit, with a touch of vanillin on the nose. Lovely and succulent, berry fruit. Big body. SS.

☐ 115 24 bts. (2 owc)

per lot: £560-680

Figure A.2. Latent prices

This figure plots the time series of latent (blue) and observed (red) prices for selected wines.

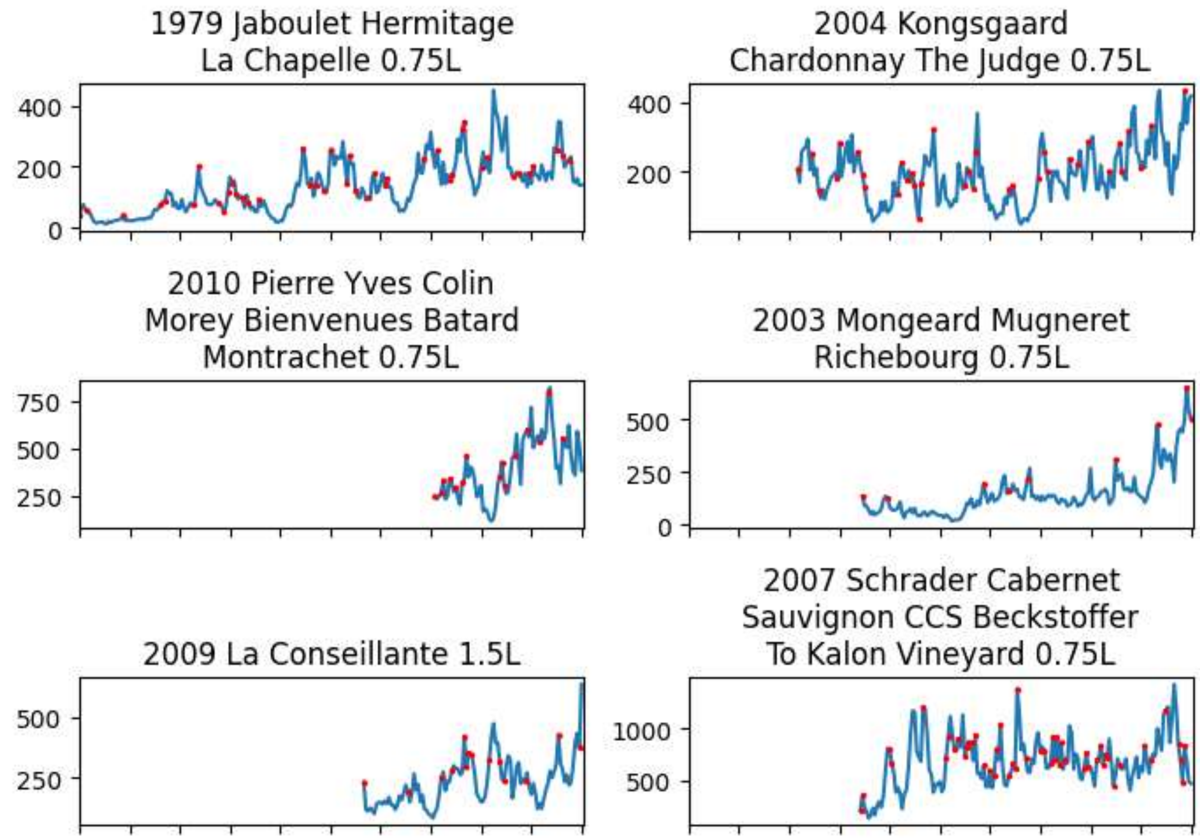
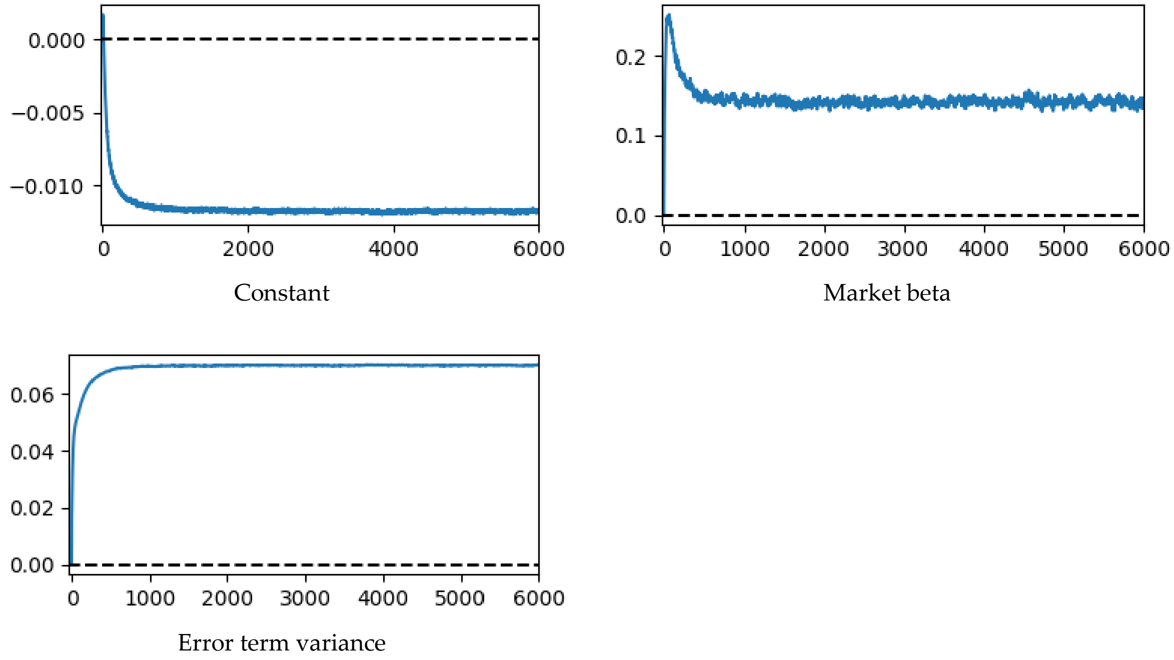


Figure A.3. Convergence

This figure shows the convergence of the parameters in the MCMC procedure.

Panel A: Valuation model



Panel B: Selection model

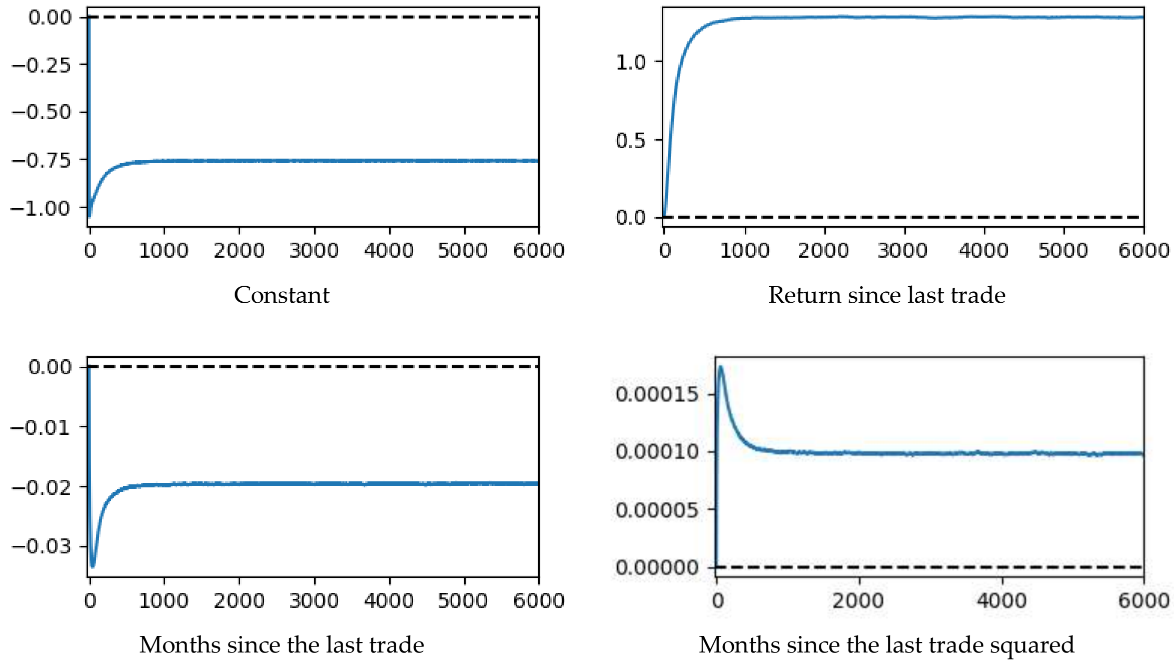
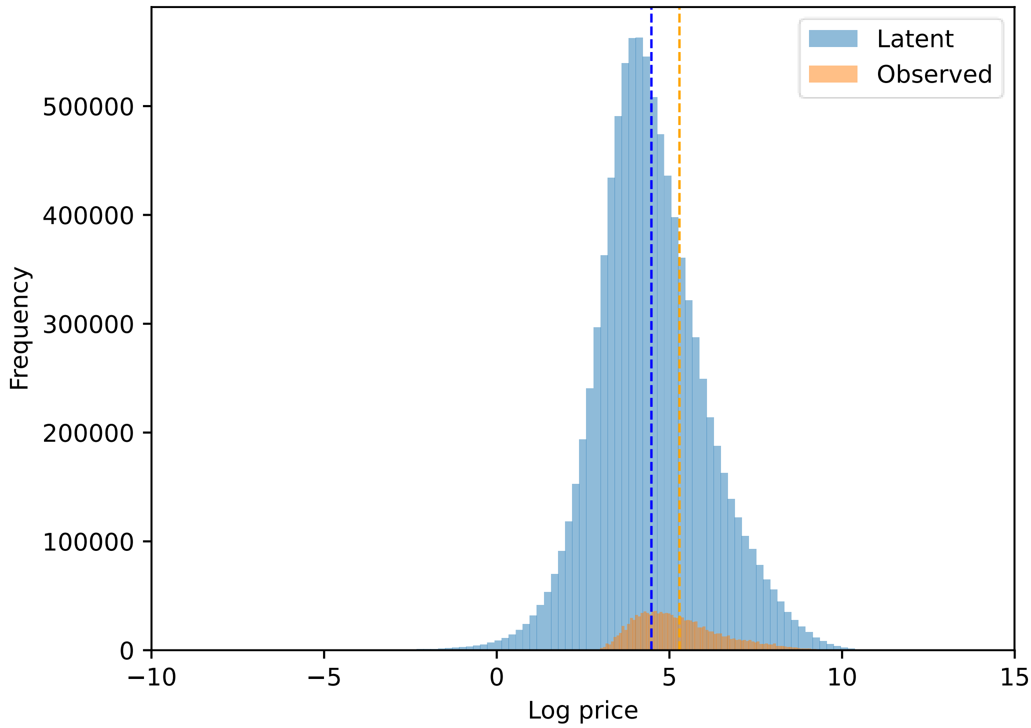
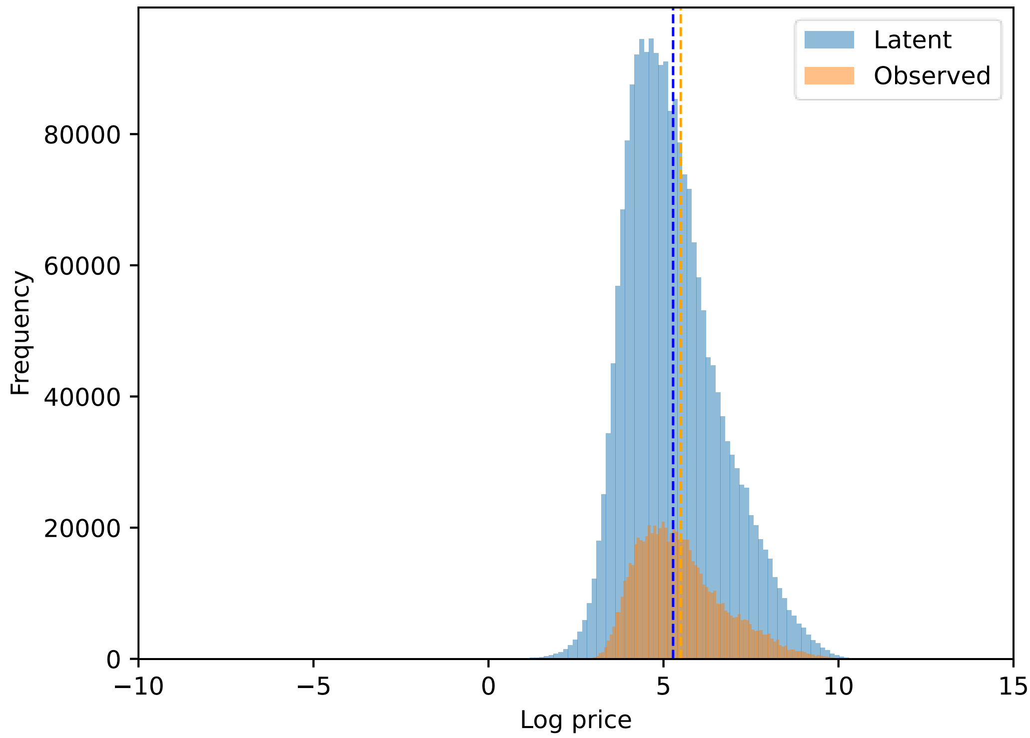


Figure A.4. Distribution of latent and observed prices

This figure shows the distribution of observed transaction prices and the latent prices of untraded wines estimated by the MCMC selection adjustment procedure of Korteweg et al. (2016) for at least five (Panel A) or 30 trades (Panel B) across the sample.



Panel A: At least five traded over the entire sample period



Panel B: At least 30 trades over the entire sample period

Figure A.5. Observed and Markov Chain Monte Carlo price index

This figure plots the monthly price index of wine estimated from observed returns and Markov Chain Monte Carlo (MCMC) models. In the different specifications, we require wine to have at least two, five, ten, 20, or 30 trades over the entire sample period. The MCMC series follows Korteweg and Sorensen's (2010) method.

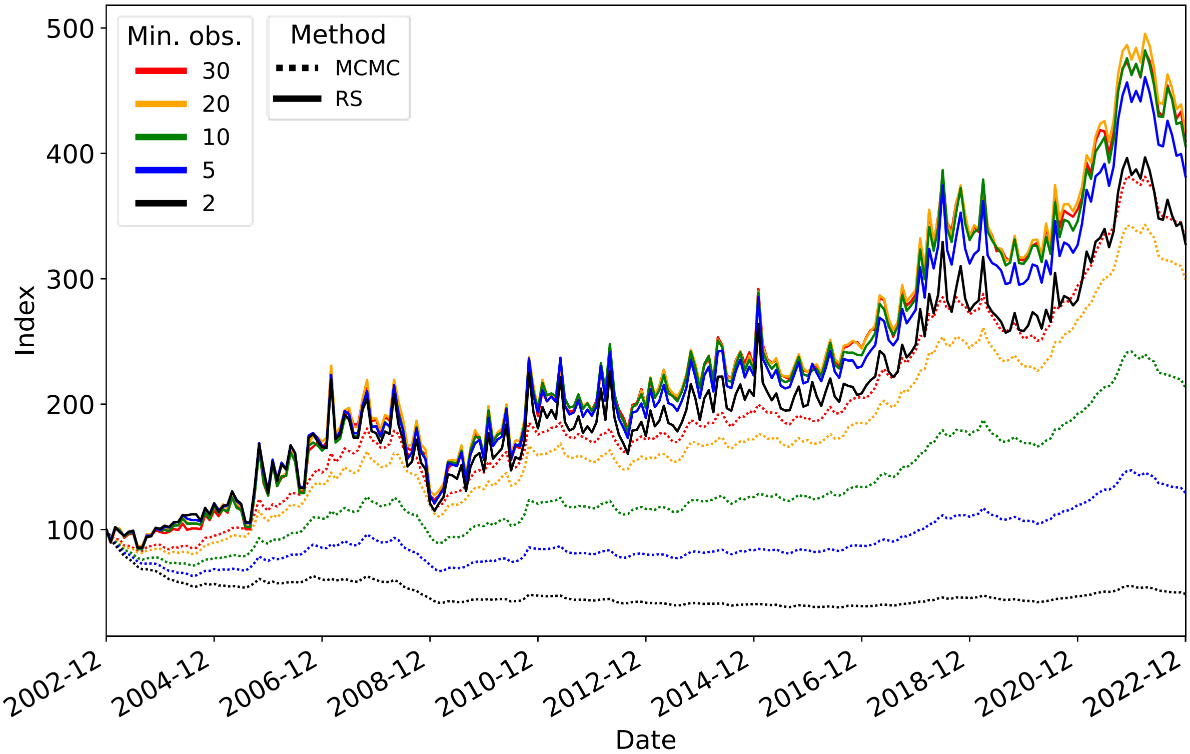


Figure A.6. Transition matrix

This figure plots the transition matrix. This equals the relative frequency at which a wine is sorted in selection-bias quintile portfolio *i* in month *t*, given that it was in selection-bias quintile portfolio *j* in month *t-12*.

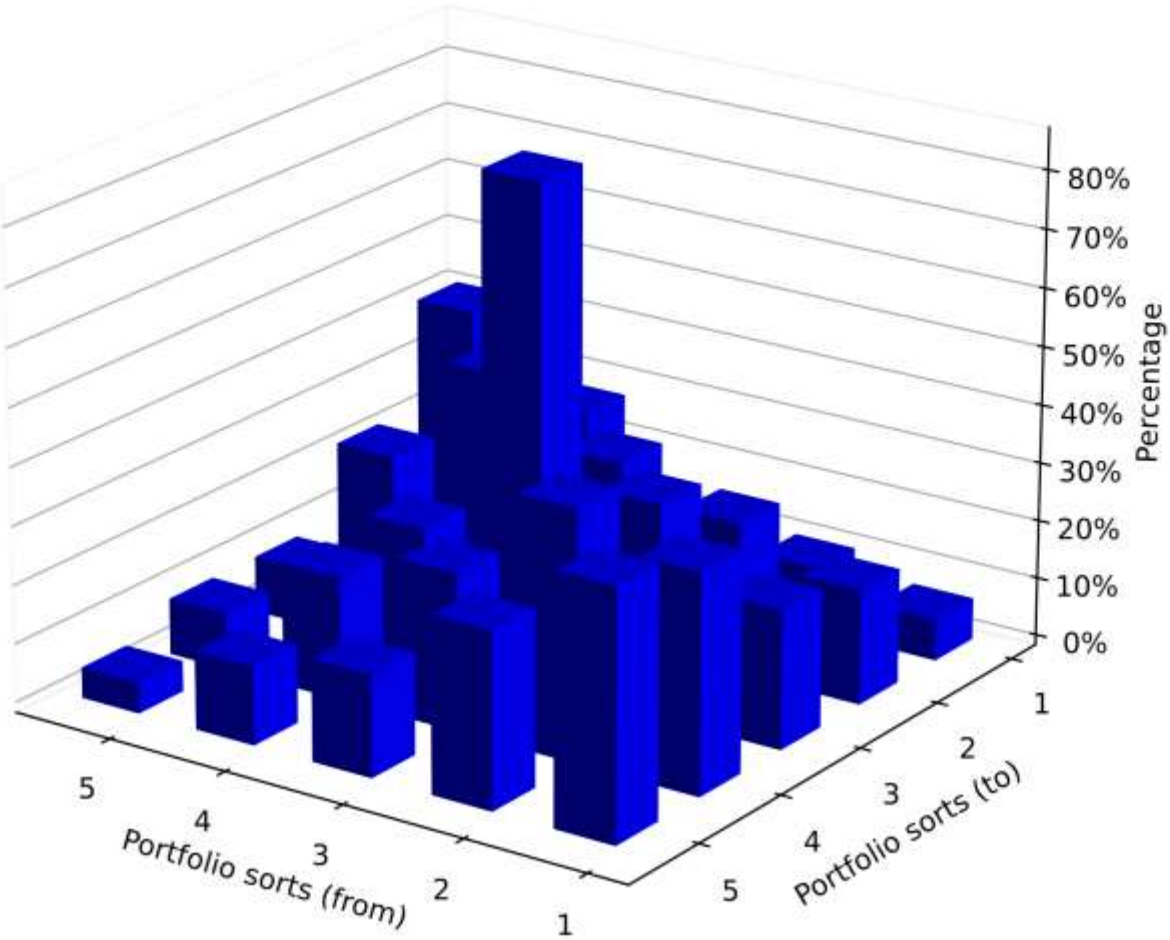


Figure A.7. Time-series variation in selection bias

This figure shows the time series of selection bias values across selection bias-sorted quintile portfolios.