A CONDITIONAL SHRINKAGE METHOD FOR PORTFOLIO SELECTION

Nathan Lassance, Alberto Martin-Utrera

REPRINT • 2026 / 08



UCLouvain

LFIN | LIDAM
Louvain Institute of Data Analysis and Modeling in economics and statistics

LFIN

Voie du Roman Pays 34, L1.03.01 B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/lfin/publications

A Conditional Shrinkage Method for Portfolio Selection*

Nathan Lassance Alberto Martín-Utrera

September 15, 2026

Abstract

We develop a conditional shrinkage method for portfolio selection. Our approach builds on a decomposition of the mean-variance portfolio into the 1/N portfolio and an alpha component that offers large efficiency gains but is exposed to substantial estimation risk. We show that a linear shrinkage covariance matrix determines the portfolio's relative exposure to these components, and our framework allows the shrinkage intensity to depend on a state variable that predicts their relative performance. We implement the method using investor sentiment and find that the estimated relation predominantly leads to greater shrinkage toward 1/N when sentiment is low, improving mean-variance efficiency while maintaining low turnover.

Keywords: 1/N portfolio, alpha component, estimation risk, mean-variance portfolio, sentiment.
JEL Classification: G11, G12.

*Lassance, LFIN/LIDAM, UCLouvain, e-mail: nathan.lassance@uclouvain.be; Martín-Utrera, Iowa State University, corresponding author, e-mail: amutrera@iastate.edu. The authors thank helpful comments from Ian Dew-Becker (the Editor), an anonymous Associate Editor, an anonymous referee, Alexandre Corhay, Mathijs Cosemans, Victor DeMiguel, Zhenzhen Fan, Nikola Gradojevic, Raymond Kan, Olivier Ledoit, Xiaowen Lei, Emiel Lemahieu, Jiacui Li, Fred Liu, Peng Liu, Charles Martineau, Yoshio Nozawa, Michael O'Doherty, André Portela-Santos, Majeed Simaan, Thanasis Stengos, Ilias Tsiakas, Raman Uppal, Rodolphe Vanderveken, Xiaolu Wang, Alex Wassensteiner, Michael Wolf, Guofu Zhou, conference participants at 3rd Frontiers of Factor Investing, INFORMS Annual Meeting, 17th International Conference on Computational and Financial Econometrics, Silicon Prairie Finance Conference, and seminar participants at Iowa State University, New Jersey Institute of Technology, Stevens Institute of Technology, UCLouvain, University of Guelph, and University of Toronto.

1. Introduction

The mean-variance portfolio framework of Markowitz (1952) is a cornerstone of asset pricing theory, as it provides a representation of the stochastic discount factor that prices all assets. However, estimating mean-variance portfolios is challenging in practice. A key difficulty lies in incorporating the full set of conditioning information available to investors when constructing optimal portfolios from a large universe of assets.[1] Another critical issue is the detrimental effect of estimation risk on the out-of-sample performance of optimal mean-variance portfolios (DeMiguel, Garlappi, and Uppal 2009).

To address these challenges, researchers often rely on stock-specific characteristics that are informative about the cross-sectional distribution of stock returns, construct managed portfolios, and solve the mean-variance problem for the managed portfolios unconditionally using shrinkage techniques (Ledoit and Wolf 2003; 2004; DeMiguel et al. 2009; Bryzgalova, Pelger, and Zhu 2025). While this is an effective approach for spanning the unconditional mean-variance efficient frontier of managed portfolios, it does not explicitly account for the conditional nature of the performance of characteristic-sorted portfolios, as highlighted in recent asset pricing literature (Stambaugh, Yu, and Yuan 2012; Moreira and Muir 2017; Barroso, Detzel, and Maio 2025; DeMiguel, Martín-Utrera, and Uppal 2024). This disconnect arises because shrinkage is typically implemented unconditionally, abstracting from time variation in expected returns. We bridge this gap by introducing a conditional shrinkage framework that dynamically adjusts portfolio weights in response to a state variable that captures this time variation.

[1] Cochrane (2009, Chapter 8) says that *"asset pricing theory really describes prices in terms of conditional moments"* derived from *"the investor's first-order conditions [...] using expectations conditional on the investor's time-t information."*

In particular, we introduce a shrinkage methodology for estimating the return covariance matrix that conditions on investor sentiment.[2] Our framework is not restricted to sentiment and can, in principle, incorporate any state variable that captures time variation in expected returns. However, we focus on sentiment because it is especially well-suited for this purpose. Prior work shows that sentiment predicts the expected returns of mispricing-based portfolios (Stambaugh, Yu, and Yuan 2012) as well as those of systematic factors (Huang et al. 2015; Shen, Yu, and Zhao 2017).[3] Because both mispricing-based and systematic sources of return variation contribute to the performance of mean-variance portfolios, conditioning shrinkage on sentiment aligns portfolio construction with variation in expected returns. Unlike existing shrinkage estimators that rely primarily on statistical considerations (Ledoit and Wolf 2003; 2004; 2017; 2020), our approach is explicitly guided by economic intuition.

Our methodology builds on three insights. First, we exploit a decomposition of the mean-variance portfolio into the 1/N portfolio (i.e., systematic component) and an alpha portfolio (i.e., a mispricing-based component orthogonal to 1/N). In Proposition 1, we show that the influence that each of these two components has on the mean-variance portfolio, measured by the return correlation between the mean-variance portfolio and each component, is determined by the component's in-sample Sharpe ratio. An interesting observation from this theoretical result is that the mean-variance portfolio can intrinsically tilt its performance toward 1/N or the alpha portfolio, and this tilt depends on each of these two components' performance.

Second, we find empirically that mean-variance portfolio returns are driven mainly by the alpha portfolio. This is because the alpha portfolio tends to largely outperform 1/N in sample, and

[2] Baker and Wurgler (2007) define sentiment as asset demand based on "*belief about future cash flows and investment risks that is not justified by the facts at hand*."

[3] In addition, we show in Section 4.2.1 that sentiment predicts the relative performance of the 1/N portfolio (our proxy for systematic risk) and an alpha portfolio better than other macroeconomic variables, providing further support for its use in our conditional shrinkage methodology.

thus, the return correlation between the mean-variance portfolio and the alpha portfolio is much larger than that with 1/N. The large exposure of the mean-variance portfolio to its alpha component is a problematic issue because this component suffers from substantial estimation risk.[4]

Third, we show in Proposition 2 that a linear shrinkage estimator of the covariance matrix enables us to bound the alpha portfolio's influence on the overall performance of the mean-variance portfolio. We build on these three insights to propose a sentiment-based shrinkage (SSH for short) methodology. Specifically, we derive the shrinkage intensity from the ex-ante probability that 1/N will outperform the alpha component out of sample. We estimate this probability using a logistic regression conditioned on the partial least squares sentiment index of Huang et al. (2015), which we reconstruct iteratively to avoid look-ahead bias. The data typically favor shrinking against sentiment; that is, increasing the 1/N portfolio's influence when sentiment is low. This pattern can be explained by the empirically observed negative (positive) relation between sentiment and the subsequent performance of the 1/N (alpha) portfolio.[5]

Our shrinkage methodology is economically motivated by a population no-arbitrage restriction in the sense of Ross (1976). Under a one-factor representation, Proposition 3 links this restriction to a bound on the alpha portfolio's population squared Sharpe ratio. Proposition 4 shows that, under its stated conditions, covariance shrinkage reduces this squared Sharpe ratio. SSH builds on this insight to regularize the corresponding sample quantity.

We evaluate the performance gains of our sentiment-based shrinkage methodology in simulated and empirical data. Importantly, our approach does not impose how sentiment should affect

[4]We analytically characterize the out-of-sample Sharpe ratio loss of the alpha portfolio in Section A of the Internet Appendix.

[5]Huang et al. (2015) document a negative relation between sentiment and future market returns. We find that their insights extend to 1/N returns, while alpha returns exhibit a strong positive relation with sentiment (Stambaugh, Yu, and Yuan 2012). We provide a theoretical foundation for this relationship in Section 4.1, using a model with rational and sentiment investors in which asset mispricing (alpha) is a function of sentiment.

the performance of either the 1/N or the alpha portfolio; instead, it lets the data determine, ex-ante, the optimal shrinkage intensity. In the simulation analysis, we consider a controlled environment where sentiment traders can induce either overpricing or underpricing. For instance, if most sentiment investors are optimistic, the overall market is overpriced, and therefore the 1/N portfolio delivers lower performance as sentiment increases. In this case, our methodology yields a lower shrinkage intensity when sentiment is high. On the other hand, if most sentiment investors are pessimistic, the overall market is underpriced, and therefore the 1/N portfolio delivers higher performance with investor sentiment. In this setting, both the 1/N and the alpha portfolio's Sharpe ratios increase with sentiment, and thus, the direction of shrinkage is, in principle, ambiguous. Nevertheless, in our simulation, the methodology yields a higher shrinkage intensity when sentiment is high; we provide a theoretical explanation for this result in Section 4.1. Overall, the simulations confirm that our shrinkage methodology performs consistently well across sentiment regimes independently of the directional effect of sentiment-driven mispricing.

Our analysis with empirical data using six datasets of characteristic-sorted portfolios illustrates that our sentiment-based shrinkage is a statistically and economically motivated portfolio selection approach. Across the six datasets, and after accounting for transaction costs of 30 basis points, our shrinkage portfolio delivers an average Sharpe ratio that is 113%, 65%, 42%, 27%, and 67% higher than those of the five benchmark mean-variance portfolios based on the estimators of Ledoit and Wolf (2003), Ledoit and Wolf (2004), Chen and Yuan (2016), Ao, Li, and Zheng (2019), and Ledoit and Wolf (2020), respectively. Its average Sharpe ratio is also 79% and 134% higher than those of the reward-to-risk timing strategy of Kirby and Ostdiek (2012) and the 1/N portfolio, respectively.[6]

[6]These relative improvements of SSH versus the benchmarks hold under the sample mean estimator. In Section 4.2.5, we show that SSH continues to dominate under a shrinkage estimator of the mean vector.

Moreover, the five benchmark mean-variance portfolios require a turnover that is on average, close to six times higher than that of SSH.[7]

Finally, we investigate the sources of the empirical performance gains of our approach by evaluating the out-of-sample Sharpe ratio of several portfolios across sentiment regimes. On average, the 1/N annualized Sharpe ratio is 0.24 and 0.87 in high and low-sentiment periods, respectively, and that of the alpha portfolio is 1.37 and 0.54 in high and low-sentiment months. In contrast, our sentiment-based shrinkage portfolio performs consistently well across sentiment regimes. Specifically, it delivers on average a Sharpe ratio of 1.46 and 0.88 in high and low-sentiment periods, which is 9.5% and 51% higher than those attained by the mean-variance portfolio exploiting the Ledoit and Wolf (2004) covariance matrix, respectively. Overall, the economic gains delivered by our sentiment-based shrinkage portfolio relative to the benchmark mean-variance portfolios stem from its substantial outperformance during low-sentiment regimes.

Although we focus on covariance-matrix shrinkage to establish the relative importance of the 1/N and alpha components conditional on sentiment, similar timing effects can also be implemented by shrinking the mean vector or the portfolio weights, as we document in Section 4.2.4. Empirically, however, covariance-matrix shrinkage yields lower turnover and generally higher net-of-cost Sharpe ratios than these alternative implementations.

Our work is closely related to the literature on shrinkage estimators of the covariance matrix in portfolio optimization (Ledoit and Wolf 2003; 2004; 2017; 2020). Instead of seeking to improve the estimation of the return covariance matrix, our approach assesses the ex-ante relation between a macroeconomic variable—investor sentiment—and the performance of the 1/N and alpha components to define the shrinkage intensity of the covariance matrix for the mean-variance portfolio,

[7] In Section B of the Internet Appendix, we show that SSH continues to dominate the benchmark strategies when imposing a monthly turnover cap.

which, as we show in Proposition 2, allows us to determine the contribution that each component has to the overall portfolio performance.

Our approach builds on recent literature that introduces a decomposition of mean-variance efficient portfolios to design robust investment strategies. The work most closely related to ours is Raponi, Uppal, and Zaffaroni (2023), who develop a general framework in which the mean-variance optimal portfolio can be decomposed into a benchmark (beta) portfolio and an orthogonal minimum-variance alpha portfolio (their Proposition 4.1). Their alpha portfolio is defined relative to a general benchmark, whereas our alpha component is specifically defined relative to 1/N: it is the optimal mean-variance portfolio orthogonal to 1/N, obtained by treating the excess return on 1/N as the single benchmark factor. Thus, our decomposition is a special case of their analysis. In this one-factor setting, Raponi, Uppal, and Zaffaroni (2023) show that when pricing errors are nonzero, the benchmark and the orthogonal alpha portfolio span the efficient frontier (two-fund separation; their Proposition 4.3), and they further document empirically that the alpha component typically accounts for a large share of the attainable Sharpe ratio. Their analysis of model misspecification imposes a no-arbitrage restriction that shrinks this benchmark-relative alpha portfolio. In contrast, we discipline our 1/N-relative alpha component by regularizing the covariance matrix, which also bounds its maximum squared Sharpe ratio (Proposition 4). Our contribution builds on these insights by using investor sentiment to construct a conditional strategy that dynamically regulates the relative exposure to 1/N and our alpha component, thereby providing an economically motivated, state-dependent implementation of shrinkage in mean-variance portfolios.

Our approach relates to Bayesian and robust-control portfolio choice (Pástor 2000; Maenhout 2004), where priors or ambiguity aversion attenuate the sensitivity of optimal portfolios to estimation error. Our approach regularizes the covariance matrix to bound the alpha component's

influence and tilt mean-variance portfolios toward 1/N, with an intensity that varies with investor sentiment. This state dependence distinguishes our approach from frameworks with fixed shrinkage intensity.

Our work is related to the literature that provides empirical and theoretical support for the relation between sentiment and asset prices; see Dumas, Kurshev, and Uppal (2009), Yu and Yuan (2011), Stambaugh, Yu, and Yuan (2012), Huang et al. (2015), and Shen, Yu, and Zhao (2017). These papers show that sentiment predicts market returns negatively and long–short mispricing returns positively. In contrast, our focus is on portfolio optimization, and we introduce a simple approach to accommodate sentiment in the construction of mean-variance portfolios.

Our manuscript is also linked to the literature documenting the economic gains in portfolio selection of modern statistical methods that exploit the predictability of stock returns (DeMiguel, Nogales, and Uppal 2014; Gu, Kelly, and Xiu 2020; Guijarro-Ordonez, Pelger, and Zanotti 2024). Indeed, serial autocorrelation and machine learning techniques can substantially improve performance, but often at the expense of higher portfolio turnover. In contrast to this literature, we focus on the sentiment-driven predictability of the two components of mean-variance portfolios. In addition, our methodology performs better than alternative shrinkage techniques with a substantially lower turnover.

Finally, our shrinkage portfolio can also be cast as a regression problem with norm constraints (Britten-Jones 1999; DeMiguel et al. 2009). Therefore, our method shares elements with Campbell and Thompson (2008), Pettenuzzo, Timmermann, and Valkanov (2014), and Bakalli, Guerrier, and Scaillet (2023), who impose economic constraints in forecasting regressions for the equity premium. Like them, we exploit the relation between an economic variable and asset returns to

impose a constraint on our vector of decision variables (i.e., the portfolio weights). In particular, we use sentiment to tilt the mean-variance portfolio toward 1/N or the alpha component.

2. Data

Our investor sentiment measure builds on the work of Huang et al. (2015). In particular, we obtain our sentiment index by applying partial least squares (PLS) to five observable investor sentiment proxies used by Baker and Wurgler (2006): (1) the closed-end fund discount, (2) the number of IPOs, (3) the first-day returns on IPOs, (4) the equity share in new issues, and (5) the dividend premium. Our sample spans from July 1965 to June 2022.[8] Following Baker and Wurgler (2006), we transform the raw sentiment proxies before constructing the sentiment index. Specifically, the dividend premium is lagged by twelve months, first-day IPO returns are replaced by the IPO-volume-weighted average over the prior twelve months and then lagged by twelve months, and IPO volume is replaced by the sum of IPOs over the prior twelve months. The closed-end fund discount and the equity share in new issues are lagged by one month, with the latter already defined over the prior twelve months.

We test the performance of our proposed methodology on six empirical datasets of monthly returns of characteristic-sorted portfolios. The first three datasets are double-sorted portfolios spanning July 1965 to June 2022: (i) 25 portfolios sorted on size and market beta (25SBETA), (ii) 25 portfolios sorted on size and book-to-market (25SBTM), (iii) 25 portfolios sorted on operating profitability and investment (25OPINV). The next two datasets are constructed from the 23 anomalies considered by Novy-Marx and Velikov (2016): (iv) the long and short legs of eight low-turnover anomalies (16ANOM) from July 1965 to December 2013 and (v) the long and short

[8] Our sample period does not include the 2022-2024 sentiment reversal.

legs of all the 23 anomalies (46ANOM) from July 1973 to December 2013. Finally, we consider a high-dimensional dataset: (vi) 100 portfolios formed on size and investment (100SINV) from July 1965 to June 2022.[9]

3. Shrinkage Aligned with Sentiment

In this section, we introduce our shrinkage portfolio methodology that exploits investor sentiment. First, we decompose the mean-variance portfolio into a 1/N and an alpha component. Second, we show how shrinking the covariance matrix of asset returns is a mechanism to determine the exposure of the mean-variance portfolio to its components. Third, we introduce a new methodology that exploits investor sentiment to determine the shrinkage intensity of the covariance matrix. Finally, we provide an economic rationale for shrinking the covariance matrix. A guide to the notation we use in the main body of the manuscript is available in Table 1.

3.1 The 1/N and Alpha Components of Efficient Portfolios

Let $\boldsymbol{r}_{t+1} \in \mathbb{R}^N$ be the vector of asset excess returns at time $t+1$. Given a risk-aversion coefficient $\gamma > 0$, we define the optimal mean-variance portfolio as

$$\boldsymbol{w}_{mv} = \arg\max_{\boldsymbol{w}} \ \boldsymbol{w}^\top \boldsymbol{\mu} - \frac{\gamma}{2} \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w} = \frac{1}{\gamma} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}, \tag{1}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the vector of means and covariance matrix of $\boldsymbol{r}_{t+1}$, respectively, and we assume $\boldsymbol{\Sigma}$ is positive definite. An important property of portfolio $\boldsymbol{w}_{mv}$ is that it delivers the maximum

[9] In Section C of the Internet Appendix, we test the robustness of our empirical results using 10 and 49 industry portfolios. This allows us to show the benefits of our methodology also in datasets where sentiment does not have a direct or well-established role in the performance of the underlying test assets.

Table 1: Guide to notation

This table describes the mathematical notation we use in the main body of the manuscript. The first column lists the symbol, and the second column defines it.

Notation	Definition
N	Number of assets
$\boldsymbol{r}_t$	Vector of asset excess returns at time t
$\boldsymbol{\mu}$	Vector of means
$\boldsymbol{\Sigma}$	Covariance matrix
γ	Risk-aversion coefficient
$\boldsymbol{w}_{mv}$	Optimal mean-variance portfolio in (1)
$\boldsymbol{w}_{ew}$	1/N portfolio in (3)
$\boldsymbol{w}_{\alpha}$	Optimal alpha portfolio in (4)
λ	Ratio between price of risk of 1/N and γ
$\text{SR}^2(\boldsymbol{w})$	Squared Sharpe ratio of portfolio $\boldsymbol{w}$
$\boldsymbol{\Sigma}_s$	Linear shrinkage of the covariance matrix toward the identity in (11)
δ	Shrinkage intensity
$\boldsymbol{w}_{smv}$	Shrinkage mean-variance portfolio in (12)
$\boldsymbol{w}_{s\alpha}$	Shrinkage alpha portfolio in (13)
λ_s	Shrinkage version of λ in (14)
$\boldsymbol{v}_i$	ith eigenvector of $\boldsymbol{\Sigma}$
$\sigma^2_{PC_i}$	ith eigenvalue of $\boldsymbol{\Sigma}$
$s_i(\delta)$	Shrinkage factor applied to lower-variance principal components in (17)
p_t	Probability that scaled 1/N outperforms alpha component out of sample; see Eq. (18)
Sentiment_t	Value of sentiment index at time t
$\hat{\boldsymbol{w}}_{ssh,t}$	Proposed SSH portfolio at time t
$\hat{\delta}_{ssh,t}$	SSH shrinkage intensity at time t
$\boldsymbol{\alpha}$	Vector of intercepts in the one-factor model (21)
$\boldsymbol{\beta}$	Vector of factor loadings
$\boldsymbol{\epsilon}_t$	Vector of innovations at time t
$\boldsymbol{d}_t$	Vector of random dividends at time t in the model of Assumption 1
ϕ	Mass of investors with biased beliefs
$\boldsymbol{b}$	Vector of biases
$\boldsymbol{w}_{mkt}$	Equilibrium market portfolio
$\boldsymbol{Q}$	Matrix governing the alpha portfolio Sharpe ratio in (28)
f	Proportion of assets held by optimistic investors

in-sample squared Sharpe ratio, given by the following expression:

$$\text{SR}^2(\boldsymbol{w}_{mv}) = \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}. \tag{2}$$

We now decompose $\boldsymbol{w}_{mv}$ into two components. First, the 1/N portfolio,

$$\boldsymbol{w}_{ew} = \mathbf{1}/N, \tag{3}$$

where $\mathbf{1}$ is an N-dimensional vector of ones. Second, the alpha component, defined as the optimal mean-variance portfolio orthogonal to the 1/N component:

$$\begin{aligned} \boldsymbol{w}_\alpha &= \arg\max_{\boldsymbol{w}} \ \boldsymbol{w}^\top \boldsymbol{\mu} - \frac{\gamma}{2} \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w} \quad \text{s.t.} \quad \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{ew} = 0 \\ &= \frac{1}{\gamma} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \lambda \boldsymbol{w}_{ew}, \end{aligned} \tag{4}$$

where $\lambda = \mu_{ew}/(\gamma \sigma_{ew}^2) = (\boldsymbol{w}_{ew}^\top \boldsymbol{\mu})/(\gamma \boldsymbol{w}_{ew}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{ew})$. A linear combination of the 1/N and alpha components delivers the mean-variance portfolio:

$$\boldsymbol{w}_{mv} = \lambda \boldsymbol{w}_{ew} + \boldsymbol{w}_\alpha. \tag{5}$$

Moreover, $\text{SR}^2(\boldsymbol{w}_{mv})$ in (2) is the sum of the squared Sharpe ratio of each component:[10]

$$\text{SR}^2(\boldsymbol{w}_{mv}) = \text{SR}^2(\boldsymbol{w}_{ew}) + \text{SR}^2(\boldsymbol{w}_\alpha). \tag{6}$$

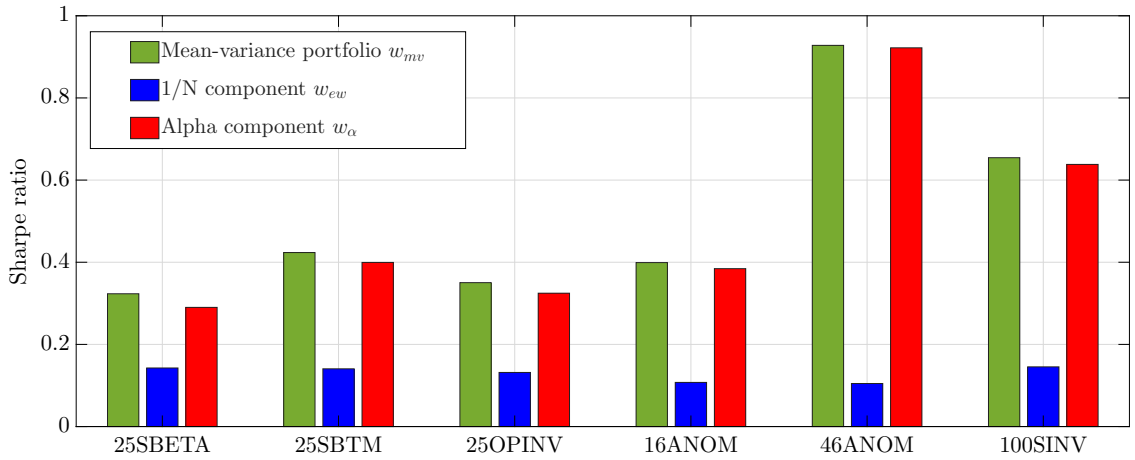
Equation (6) shows that each component has an independent contribution to the maximum in-sample squared Sharpe ratio attained by the mean-variance portfolio.

Example 1 *We illustrate, in a simple three-asset example, the portfolio decomposition* $\boldsymbol{w}_{mv} = \lambda \boldsymbol{w}_{ew} + \boldsymbol{w}_\alpha$ *and the corresponding Sharpe ratios. We set annualized mean returns* $(\mu_1, \mu_2, \mu_3) =$

[10] This is a well-known decomposition of the maximum squared Sharpe ratio in the asset pricing literature; see Gibbons, Ross, and Shanken (1989, Equation (23)).

Figure 1: In-sample Sharpe ratios of mean-variance portfolio components

This figure depicts the in-sample monthly Sharpe ratios of the mean-variance portfolio ($\boldsymbol{w}_{mv}$ in (1)), its 1/N component ($\boldsymbol{w}_{ew}$ in (3)), and its alpha component ($\boldsymbol{w}_{\alpha}$ in (4)), evaluated using the full sample of the six datasets described in Section 2. Across all six datasets, the alpha component delivers a substantially larger in-sample Sharpe ratio than that of 1/N, indicating that most of the attainable mean-variance efficiency originates from the alpha component.
Alt text: Grouped bars across six datasets show similar Sharpe ratios for the mean-variance and alpha portfolios, both exceeding the equal-weighted portfolio. Performance is highest for forty-six anomalies.



$(0.05, 0.10, 0.15)$*, annualized standard deviations* $(\sigma_1, \sigma_2, \sigma_3) = (0.10, 0.15, 0.20)$*, and a uniform pairwise correlation* $\rho = 0.7$*. In this case, the portfolio weights are*

$$\boldsymbol{w}_{mv} = (-1.97, 2.39, 3.18)^{\top}, \quad \lambda \boldsymbol{w}_{ew} = (1.83, 1.83, 1.83)^{\top}, \quad \boldsymbol{w}_{\alpha} = (-3.80, 0.56, 1.35)^{\top}, \tag{7}$$

and their Sharpe ratios are

$$\mathrm{SR}(\boldsymbol{w}_{mv}) = 0.79, \quad \mathrm{SR}(\boldsymbol{w}_{ew}) = 0.74, \quad \mathrm{SR}(\boldsymbol{w}_{\alpha}) = 0.26, \tag{8}$$

which satisfy $\mathrm{SR}^2(\boldsymbol{w}_{mv}) = \mathrm{SR}^2(\boldsymbol{w}_{ew}) + \mathrm{SR}^2(\boldsymbol{w}_{\alpha})$.

To have a better understanding of which component has the biggest contribution to maximum mean-variance efficiency, Figure 1 depicts the in-sample monthly Sharpe ratios of the mean-

variance portfolio and its 1/N and alpha components in our six datasets.[11] We observe that the Sharpe ratio of the alpha component substantially exceeds that of 1/N, indicating that most of the in-sample mean-variance efficiency originates from the alpha component. In the next proposition, we formally prove that the substantial outperformance of the alpha component relative to the 1/N component implies that the mean-variance portfolio-return variation is mostly driven by the alpha component.

Proposition 1 *The squared correlations between the mean-variance portfolio return and that of its 1/N and alpha components are given by*

$$\text{Corr}^2(\boldsymbol{w}_{mv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}) = \frac{\text{SR}^2(\boldsymbol{w}_{ew})}{\text{SR}^2(\boldsymbol{w}_{mv})}, \tag{9}$$

$$\text{Corr}^2(\boldsymbol{w}_{mv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{\alpha}^\top \boldsymbol{r}_{t+1}) = \frac{\text{SR}^2(\boldsymbol{w}_{\alpha})}{\text{SR}^2(\boldsymbol{w}_{mv})} = 1 - \text{Corr}^2(\boldsymbol{w}_{mv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}). \tag{10}$$

Proposition 1 shows that the correlation between the mean-variance portfolio and its components depends on their in-sample Sharpe ratios, which suggests that the mean-variance portfolio self-controls the exposure to its 1/N and alpha components. Figure 1 shows that the alpha component delivers a substantially larger Sharpe ratio than that of 1/N. Thus, the correlation between the mean-variance portfolio and the alpha component is substantially larger than that with 1/N.[12] However, this large exposure to the alpha component may be undesirable because, unlike the 1/N component, the alpha portfolio suffers important out-of-sample performance losses due to estimation risk. We theoretically characterize the out-of-sample performance losses of the alpha portfolio in Section A of the Internet Appendix.

[11] For visibility purposes, we depict the Sharpe ratios instead of the squared Sharpe ratios. However, this representation still allows us to see which of the two components has the largest contribution to the maximum squared Sharpe ratio delivered by the mean-variance portfolio.

[12] In our datasets, the correlation between mean-variance and alpha returns is above 90%.

3.2 Shrinking Toward the 1/N Component

We now derive the return correlation between the mean-variance portfolio and its two components when we apply shrinkage to the covariance matrix. In particular, we use a shrinkage covariance matrix of asset returns similar to that in Ledoit and Wolf (2004):

$$\boldsymbol{\Sigma}_s = (1-\delta)\boldsymbol{\Sigma} + \delta\bar{\sigma}^2\boldsymbol{I}, \tag{11}$$

where $\delta \in [0,1]$ is the shrinkage intensity, $\bar{\sigma}^2 = \frac{1}{N}\text{tr}(\boldsymbol{\Sigma})$ is the average of asset variances, and $\boldsymbol{I}$ is the identity matrix. The resulting shrinkage mean-variance portfolio is

$$\boldsymbol{w}_{smv} = \frac{1}{\gamma}\boldsymbol{\Sigma}_s^{-1}\boldsymbol{\mu}, \tag{12}$$

which can be decomposed into a 1/N component and a shrinkage alpha component as

$$\boldsymbol{w}_{smv} = \lambda_s \boldsymbol{w}_{ew} + \boldsymbol{w}_{s\alpha}, \tag{13}$$

where λ_s is defined such that the two components are orthogonal, i.e., $\boldsymbol{w}_{s\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{ew} = 0$, yielding

$$\lambda_s = \frac{\boldsymbol{w}_{ew}^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}_s^{-1}\boldsymbol{\mu}}{\gamma\sigma_{ew}^2}. \tag{14}$$

The following proposition generalizes the result in Proposition 1 obtained for a shrinkage intensity $\delta = 0$.

Proposition 2 *Let $\boldsymbol{v}_i$ and $\sigma_{PC_i}^2$ be the ith eigenvector and eigenvalue of the covariance matrix $\boldsymbol{\Sigma}$. Then, the squared correlations between the shrinkage mean-variance portfolio return and that of its 1/N and shrinkage alpha components are given by*

$$\text{Corr}^2\left(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}\right) = \frac{\text{SR}^2(\boldsymbol{w}_{ew}) \times \left(1+\varepsilon(\delta)\right)^2}{\text{SR}^2(\boldsymbol{v}_1) + \sum_{i=2}^N s_i^2(\delta)\text{SR}^2(\boldsymbol{v}_i)}, \tag{15}$$

$$\text{Corr}^2\left(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{s\alpha}^\top \boldsymbol{r}_{t+1}\right) = 1 - \text{Corr}^2\left(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}\right), \tag{16}$$

where $\varepsilon(\delta)$ *is defined in Equation* (A53) *of the Internet Appendix and* $s_i(\delta) \in [0,1]$ *is a shrinkage factor applied to lower-variance principal components:*

$$s_i(\delta) = \frac{\sigma_{PC_i}^2}{\sigma_{PC_i}^2 + \kappa(\delta)(\sigma_{PC_1}^2 - \sigma_{PC_i}^2)} \quad \textit{with} \quad \kappa(\delta) = \frac{\delta \bar{\sigma}^2}{\delta \bar{\sigma}^2 + (1-\delta)\sigma_{PC_1}^2}. \tag{17}$$

Furthermore, the shrinkage factor $s_i(\delta)$ *is monotonically decreasing in both* i *and* δ*. When the first eigenvector* $v_1 = \mathbf{1}/\sqrt{N}$*, we have* $\varepsilon(\delta) = 0$*; in this case, a higher* δ *increases the squared correlation with the 1/N component and decreases that with the alpha component.*

Proposition 2 implies that shrinking the covariance matrix allows us to *control* the mean-variance portfolio exposure to its 1/N and alpha components. In particular, under the eigenvector condition in Proposition 2, increasing the shrinkage intensity δ raises the squared correlation with 1/N by reducing the weights $s_i^2(\delta)$ on the squared Sharpe ratios of lower-variance PCs, $\text{SR}^2(\boldsymbol{v}_i)$ for $i \geq 2$.[13]

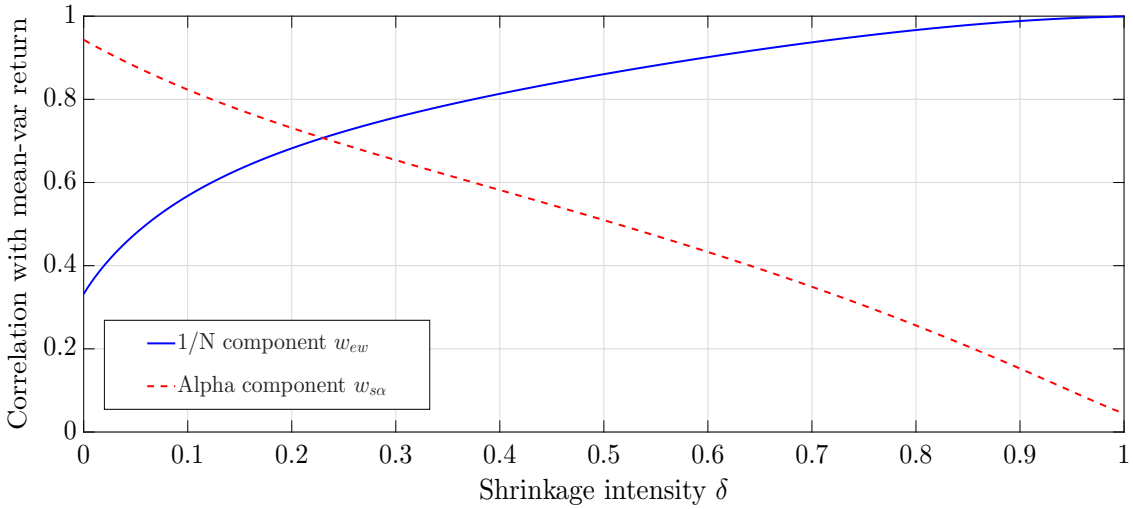
Figure 2 depicts, for the 25SBTM dataset, the correlation between the shrinkage mean-variance portfolio and its two components as a function of δ. In line with Proposition 2, we find that the correlation with the 1/N component increases with δ, from 33% when $\delta = 0$ to almost 100% when $\delta = 1$. In contrast, the correlation with the shrinkage alpha component is 94% when $\delta = 0$, but decreases to a mere 4.5% when $\delta = 1$.

The results in this section demonstrate that shrinkage is a statistical tool that can be used to tilt mean-variance portfolios toward their 1/N and alpha components. This is economically meaningful

[13] When the shrinkage intensity $\delta = 1$, the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ is proportional to the vector of mean returns $\boldsymbol{\mu}$. In Section D of the Internet Appendix, we study theoretically the relation between the mean-variance portfolio and its 1/N component when $\delta = 1$ and demonstrate an equivalence when there is no cross-sectional variation in average returns. We also show empirically that the correlation with the 1/N component when $\delta = 1$ is consistently high on average across datasets.

Figure 2: Impact of shrinkage intensity on correlations with 1/N and alpha components

This figure depicts the theoretical formulas in Proposition 2 for the correlations between the return of the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ in (12) and that of its 1/N component $\boldsymbol{w}_{ew}$ and shrinkage alpha component $\boldsymbol{w}_{s\alpha}$ in (13). The correlation formulas depend on δ, $\boldsymbol{\mu}$, and $\boldsymbol{\Sigma}$. We vary δ from zero to one and we set $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ equal to the sample mean and covariance matrix computed on a dataset of monthly excess returns on 25 portfolios sorted on size and book-to-market spanning the period from July 1965 to June 2022. As the shrinkage intensity increases, the correlation with the 1/N component rises from 33% to nearly 100%, while that with the alpha component falls from 94% to 4.5%, illustrating that shrinkage provides a mechanism to tilt mean-variance portfolios toward 1/N.
Alt text: As shrinkage increases, the solid curve rises toward perfect correlation with the equal-weighted portfolio, while the dashed curve falls toward zero correlation with the alpha portfolio.



because we can time the performance of 1/N and the alpha portfolio with economic variables and define the mean-variance portfolio's tilt toward its components accordingly. Motivated by this insight, we propose in the next section a framework to exploit investor sentiment in constructing the shrinkage covariance matrix.

3.3 Sentiment-based Shrinkage

We propose to shrink the covariance matrix and adjust the mean-variance portfolio's exposure to its 1/N and alpha components conditional on sentiment.[14] Prior research shows that sentiment plays a substantial role in the performance of systematic risk factors and alpha portfolios (Stambaugh, Yu, and Yuan 2012; Huang et al. 2015; Shen, Yu, and Zhao 2017), providing empirical support for its use in our conditional shrinkage approach.[15]

The proposed criterion for determining the shrinkage intensity applied to the return covariance matrix employs logistic regression (Hastie, Tibshirani, and Friedman 2009) to map sentiment onto a probability space. Specifically, given the level of sentiment at time t, we shrink the covariance matrix, and thus, tilt the mean-variance portfolio toward its 1/N component, with an intensity determined from the estimated probability that the 1/N portfolio return will exceed the alpha portfolio's out-of-sample return at time $t+1$. We name our shrinkage methodology as *sentiment-based shrinkage* (SSH).

Given a dataset of τ observations on N assets, we compute the SSH shrinkage intensity δ over time using an expanding estimation window composed of $M_t = T + t - 1$ observations at time t, where T is the initial sample size and $t = 1, \ldots, \tau - T$ are the portfolio rebalancing dates. Given the estimation window of size M_t at rebalancing date t, we use the first $\lceil M_t/2 \rceil$ observations to estimate the two components of the mean-variance portfolio in (5), $\lambda_t \boldsymbol{w}_{ew}$ and $\boldsymbol{w}_\alpha$.[16] Then, we evaluate the returns of these two portfolios on the second half of the estimation window of size $\lfloor M_t/2 \rfloor$, which we use to estimate the probability of the event $\lambda_t r_{ew,t+1} > r_{\alpha,t+1}$ conditional on

[14] Section 4.2.4 compares the empirical performance of analogous SSH implementations for the mean vector and portfolio weights, while Section E of the Internet Appendix provides their theoretical foundations.

[15] We provide further support for the use of sentiment as a conditioning variable in Section 4.2.1.

[16] If $N \geq \lceil M_t/2 \rceil$ so that the sample covariance matrix is singular, we use its pseudo-inverse to estimate $\boldsymbol{w}_\alpha$.

lagged sentiment using a logistic function:[17]

$$p_t = \mathrm{P}\big(\lambda_t r_{ew,t+1} > r_{\alpha,t+1} \,\big|\, \text{Sentiment}_t\big) = \frac{e^{\beta_0+\beta_1 \text{Sentiment}_t}}{1+e^{\beta_0+\beta_1 \text{Sentiment}_t}}, \tag{18}$$

where Sentiment_t is the Huang et al. (2015) sentiment index, which we reconstruct using only information up to rebalancing date t to avoid look-ahead biases.[18] The coefficients β_0 and β_1 in (18) are estimated via maximum likelihood; see, e.g., Hastie, Tibshirani, and Friedman (2009).[19]

Our method then exploits the estimated probability p_t to calibrate the shrinkage intensity δ applied to the covariance matrix (11).[20] We achieve this by using p_t to determine the desired correlation between the return of the shrinkage mean-variance portfolio and its 1/N component. That is, if p_t is high, our approach imposes a high correlation between 1/N and the shrinkage mean-variance portfolio. In particular, we use the predicted probability p_t to linearly interpolate the squared correlation between the return of 1/N and that of the shrinkage mean-variance portfolio:

$$\text{Corr}^2(p_t) = (1-p_t)\,\text{Corr}_0^2 + p_t\,\text{Corr}_1^2, \tag{19}$$

where Corr_0^2 and Corr_1^2 are the squared correlations in Equation (15) when the shrinkage intensity δ is zero and one, respectively. Then, using (15), we extract the value of the shrinkage intensity δ that

[17] Holding the level and alpha components fixed at their first-subsample estimates prevents the alpha component's inflated in-sample Sharpe ratio from mechanically contaminating the probability estimate, which is used only to calibrate the shrinkage intensity δ. The final estimated vector of means and covariance matrix in (20) use the full estimation window, and the procedure uses no look-ahead information. In Section F of the Internet Appendix, we confirm the robustness of our results to alternative probability model specifications, and Section G of the Internet Appendix assesses the sensitivity of our results to split ratios other than 50/50.

[18] We download the variables we need to construct the sentiment index from Jeffrey Wurgler's website.

[19] In Section H of the Internet Appendix, we estimate the logistic regression with rolling windows instead of expanding windows, which increases estimation error in the coefficients β_0 and β_1 but increases the responsiveness to changes in the relation between sentiment and portfolio performance. We find that although a rolling-window specification can improve the performance of our SSH portfolio before transaction costs, it consistently leads to higher portfolio turnover and lower net-of-cost out-of-sample Sharpe ratios.

[20] Although we only shrink the covariance matrix, our shrinkage methodology accounts for the estimation errors affecting the alpha portfolio, which stem from *both* the vector of means and the covariance matrix.

gives the desired squared correlation with 1/N, $\text{Corr}^2(p_t)$.[21] This calibration approach recognizes that the mean-variance portfolio exposure to its 1/N and alpha components, in terms of correlation, is not linear in δ.

Finally, the proposed SSH portfolio at time t is determined as

$$\hat{\boldsymbol{w}}_{ssh,t} = \frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}_{s,t}^{-1}\hat{\boldsymbol{\mu}}_t, \tag{20}$$

where $\widehat{\boldsymbol{\Sigma}}_{s,t} = (1-\hat{\delta}_{ssh,t})\widehat{\boldsymbol{\Sigma}}_t + \hat{\delta}_{ssh,t}\hat{\bar{\sigma}}_t^2\boldsymbol{I}$, $\hat{\bar{\sigma}}_t^2 = \frac{1}{N}\text{tr}(\widehat{\boldsymbol{\Sigma}}_t)$, $(\hat{\boldsymbol{\mu}}_t, \widehat{\boldsymbol{\Sigma}}_t)$ are the sample estimates of $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ from a sample of size M_t, and the SSH shrinkage intensity $\hat{\delta}_{ssh,t}$ is estimated as described above.

Figure 3 provides a step-by-step visualization of how the SSH portfolio is constructed at each rebalancing date. The diagram highlights how sentiment information enters the estimation process,[22] how the expanding window is partitioned to separate estimation from evaluation, and how the resulting probability mapping governs the calibration of the shrinkage intensity. At each rebalancing date t, the logistic regression is estimated using the second half of the observations available up to rebalancing date t within the expanding window. As the sample grows, the parameter estimates become progressively more stable, reducing the sensitivity of the shrinkage intensity to noisy short-lived return realizations.[23]

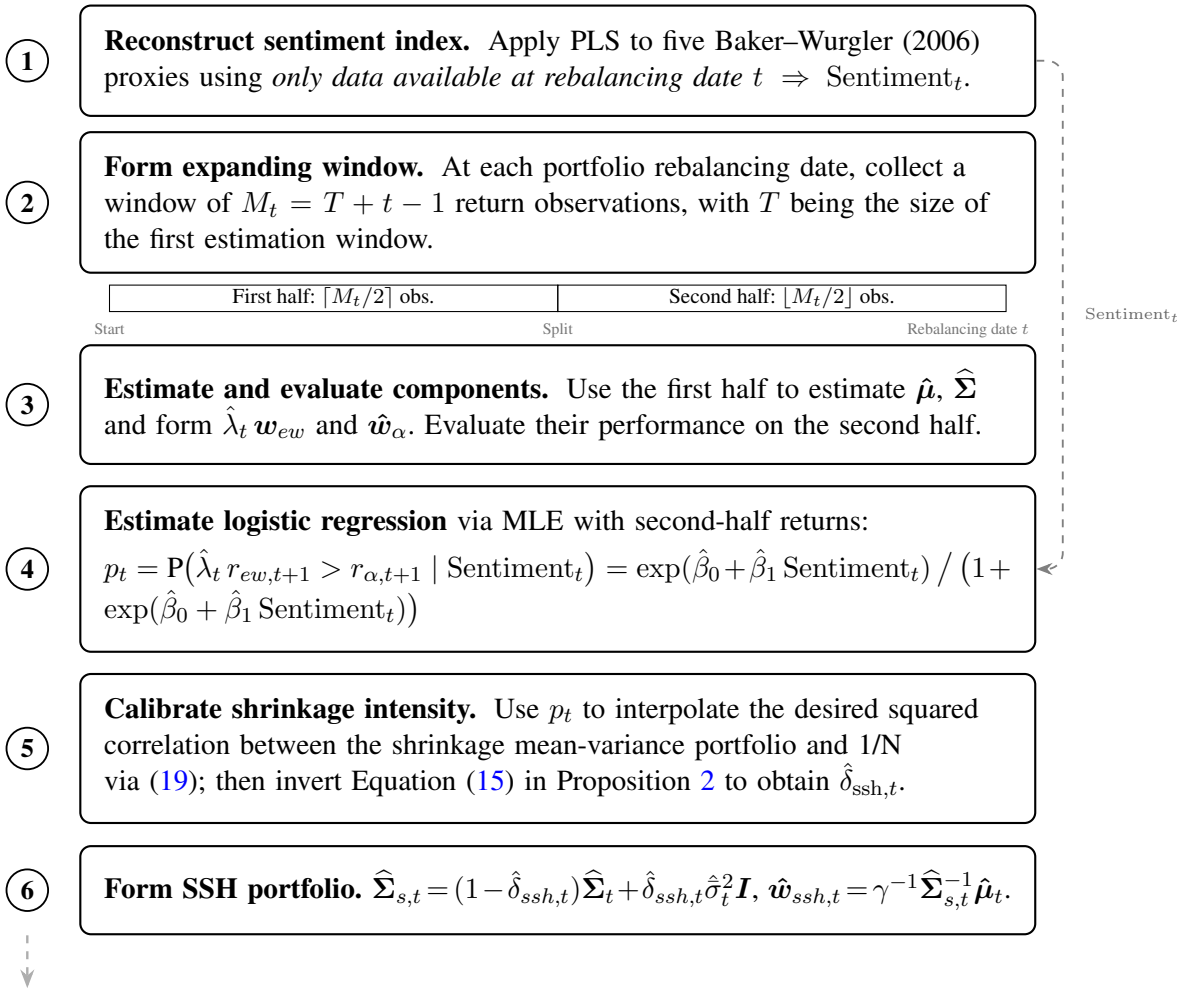
[21] We interpolate the squared correlation because the squared correlations between the shrinkage mean-variance portfolio return and that of its two components add up to one; see Proposition 2. Therefore, relation (19) also holds for the squared correlation with the shrinkage alpha portfolio return.

[22] The PLS sentiment index is estimated using an *expanding* window, rather than a rolling window. At each rebalancing date t, the PLS coefficients are estimated using all available observations from the beginning of the sample up to time t, ensuring that the sentiment measure is strictly out-of-sample and free of look-ahead bias. This mirrors the expanding-window approach adopted throughout the manuscript for portfolio optimization and avoids introducing additional estimation noise that would arise from short rolling windows. To obtain investor sentiment, our PLS sentiment index uses the return of the 1/N portfolio instead of the value-weighted market return as in Huang et al. (2015). Section I of the Internet Appendix shows that our results are robust to the construction of the sentiment index: we consider a PLS index that targets the value-weighted market excess return rather than the 1/N return, and the raw Baker and Wurgler (2006) index instead of our PLS approach.

[23] As illustrated in Figure 3, the logistic model is re-estimated at every rebalancing date t. In Section J of the Internet Appendix, we show that our results are robust to a less frequent re-estimation schedule in which the model is updated only once every five years.

Figure 3: Construction of SSH portfolio

This figure illustrates the six steps involved in constructing the proposed SSH portfolio at each portfolio rebalancing date. The procedure is repeated at each rebalancing date using an expanding window.
Alt text: Six steps reconstruct sentiment, expand the sample, evaluate portfolio components, estimate performance probabilities, calibrate shrinkage, and form the portfolio. The process repeats at each rebalancing date.



In the baseline implementation of our SSH portfolio, covariance-matrix shrinkage is the only regularization employed.[24] Specifically, the SSH covariance estimator takes the standard shrinkage form

$$\widehat{\boldsymbol{\Sigma}}_{s,t} = (1 - \hat{\delta}_{ssh,t})\,\widehat{\boldsymbol{\Sigma}}_t + \hat{\delta}_{ssh,t}\,\hat{\bar{\sigma}}_t^2\,\boldsymbol{I}.$$

Because the shrinkage target $\hat{\bar{\sigma}}_t^2\,\boldsymbol{I}$ is positive definite, any positive shrinkage intensity $\delta_t > 0$ guarantees that $\widehat{\boldsymbol{\Sigma}}_{s,t}$ is positive definite and therefore invertible, even when $N > T$ and the sample covariance matrix $\widehat{\boldsymbol{\Sigma}}_t$ itself is singular.

While shrinkage guarantees invertibility, the matrix could, in principle, still be ill-conditioned. To assess this possibility, we report in Figure 4 the condition number of the sample covariance matrix, the linear shrinkage covariance matrix of Ledoit and Wolf (2004), and our proposed SSH covariance matrix across expanding windows with an initial size of 120 months. Each subplot corresponds to one dataset. We observe that the SSH estimator delivers by far the lowest condition number and that it remains stable over time. These results indicate that the SSH covariance matrix is well-conditioned in practice, implying that its inversion is numerically stable.

Finally, Figure 5 illustrates the relation between the Huang et al. (2015) sentiment index and the SSH shrinkage intensity $\hat{\delta}_{ssh,t}$ obtained from the 25SBTM dataset. The correlation between the shrinkage intensity and investor sentiment is around -45%.

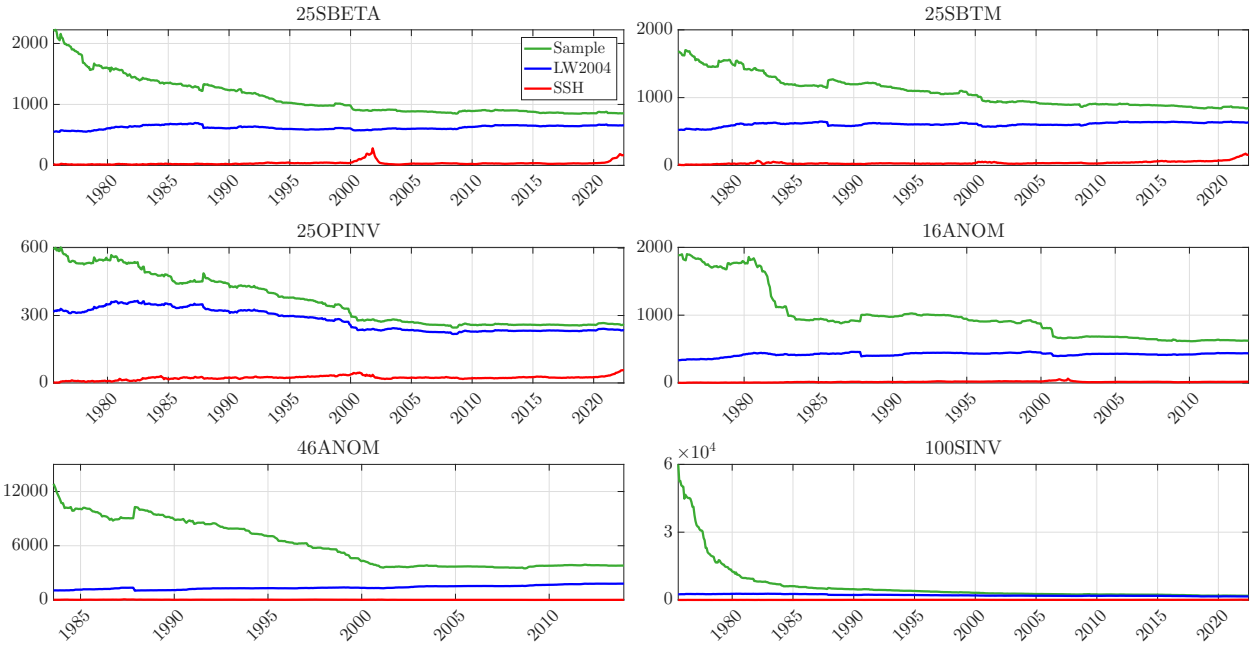
3.4 The Economic Rationale for Shrinkage

We now provide an economic rationale for covariance shrinkage based on a population no-arbitrage restriction in the sense of Ross (1976). This restriction bounds the population squared Sharpe ratio

[24] We also consider alternative shrinkage estimators. Section 4.2.4 compares sentiment-based shrinkage applied through the covariance matrix, mean vector, and portfolio weights, while Section 4.2.5 studies a statistical shrinkage estimator of the mean vector. Section E of the Internet Appendix provides the theoretical foundations for the alternative SSH implementations.

Figure 4: Condition number of different covariance matrix estimators

This figure compares the condition number of the sample covariance matrix, the Ledoit and Wolf (2004) linear shrinkage covariance matrix (LW2004), and our proposed SSH covariance matrix. Each subplot corresponds to a given dataset described in Section 2. The covariance matrices are estimated using expanding windows with an initial estimation period of $T = 120$ months. The figure shows that the SSH covariance matrix has the smallest condition number across all configurations.
Alt text: Six time-series panels compare covariance condition numbers. Sentiment-based shrinkage produces the lowest values, while the sample covariance generally has the highest, indicating poorer numerical conditioning.

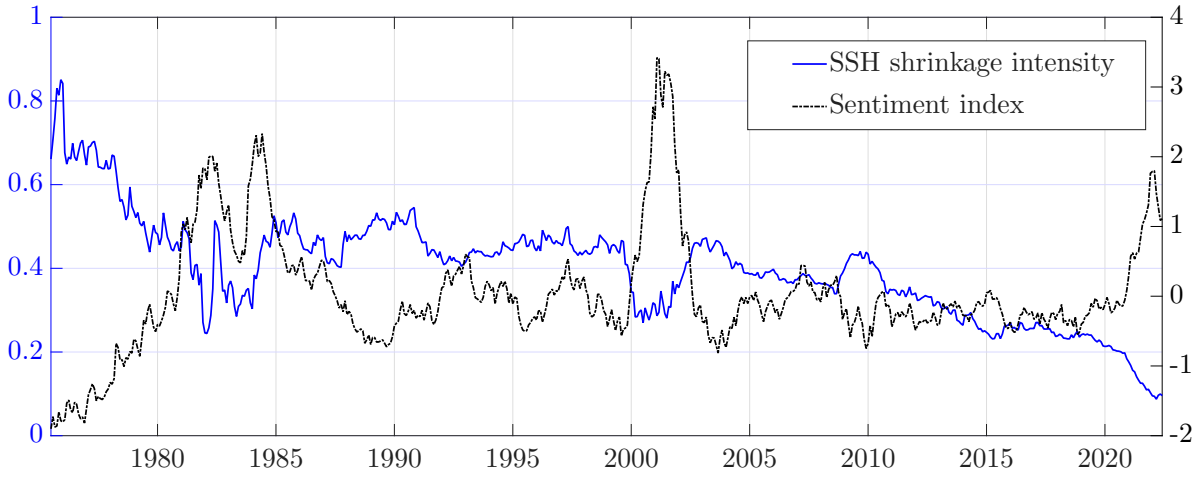


of the alpha component and motivates regularizing the corresponding sample estimate. Without loss of generality, let us assume that the vector of asset returns $\boldsymbol{r}_{t+1}$ follows the one-factor model

$$\boldsymbol{r}_{t+1} = \boldsymbol{\alpha} + \boldsymbol{\beta} r_{ew,t+1} + \boldsymbol{\epsilon}_{t+1}, \tag{21}$$

Figure 5: SSH shrinkage intensity versus sentiment index

This figure depicts the sentiment-based shrinkage intensity (left axis) obtained from the procedure described in Section 3.3, together with the sentiment index of Huang et al. (2015) (right axis). We consider the dataset of 25 portfolios sorted on size and book-to-market spanning the period from July 1965 to June 2022.
Alt text: Two time series generally move in opposite directions: shrinkage intensity tends to rise when investor sentiment falls. Shrinkage uses the left axis; sentiment uses the right.



where $r_{ew,t+1}$ is the excess return of the 1/N portfolio and $\boldsymbol{\epsilon}_{t+1}$ is a vector of innovations with zero mean. We define the no-arbitrage condition in this one-factor model as

$$\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} < \pi, \tag{22}$$

where $\pi > 0$ represents a bound on the amount of mispricing $\boldsymbol{\alpha}$. The following proposition links this no-arbitrage condition with the Sharpe ratio of the alpha component.

Proposition 3 *Under the one-factor model representation of asset returns in* (21)*, the no-arbitrage condition* (22) *is equivalent to*

$$\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} = \mathrm{SR}^2(\boldsymbol{w}_\alpha) < \pi. \tag{23}$$

Proposition 3 shows that the no-arbitrage condition in Equation (22) sets an upper bound on the population squared Sharpe ratio of the alpha component. Because population moments are unobserved, our shrinkage methodology does not impose this population restriction directly. Instead, it regularizes the corresponding sample estimate. We next show that shrinking the covariance matrix of asset returns as in (11) reduces, and thus can bound, the in-sample squared Sharpe ratio of the alpha component. Such regularization is well documented to improve out-of-sample portfolio performance (DeMiguel et al. 2009).

Proposition 4 *Assume that the first eigenvector of* $\boldsymbol{\Sigma}$ *is* $\boldsymbol{v}_1 = \mathbf{1}/\sqrt{N} \propto \boldsymbol{w}_{ew}$. *Then, the squared Sharpe ratio of the shrinkage alpha portfolio* $\boldsymbol{w}_{s\alpha}$ *in* (13) *is given by*

$$\mathrm{SR}^2(\boldsymbol{w}_{s\alpha}) = \frac{\left[\sum_{i=2}^{N} s_i(\delta)\mathrm{SR}^2(\boldsymbol{v}_i)\right]^2}{\sum_{i=2}^{N} s_i^2(\delta)\mathrm{SR}^2(\boldsymbol{v}_i)}, \tag{24}$$

where $s_i(\delta)$ *is defined in* (17). *Moreover,* $\mathrm{SR}^2(\boldsymbol{w}_{s\alpha})$ *in* (24) *decreases with the shrinkage intensity* δ *if and only if*

$$\sum_{i=2}^{N} -s_i'(\delta)\mathrm{SR}^2(\boldsymbol{v}_i)L_i \geq 0 \quad \textit{with} \quad L_i = \sum_{j=2}^{N}(s_j(\delta) - s_i(\delta))s_j(\delta)\mathrm{SR}^2(\boldsymbol{v}_j), \tag{25}$$

where $-s_i'(\delta) \geq 0$ *and* L_i *increases with the index* i.

Proposition 4 implies that the alpha Sharpe ratio typically decreases with δ, because L_i is typically positive for most indices i.[25] In particular, for each of the six datasets, we test the condition across all shrinkage intensities $\delta \in [0, 1]$ over time using (i) 120-month rolling windows and (ii) expanding windows with an initial estimation period of 120 months. In unreported results, we

[25] To understand why L_i is typically positive for most i, note that the summands of L_i in (25) are proportional to $(s_j(\delta) - s_i(\delta))s_j(\delta)$. Now, $s_i(\delta)$ decreases with i, as shown in Proposition 2. Therefore, as i increases, there are more indices j for which $s_j(\delta) - s_i(\delta) > 0$, say $j = 2$ up to some k, and these positive values are multiplied by a factor $s_j(\delta)$ that is large, whereas for the indices $j = k + 1$ up to N, we have that $s_j(\delta) - s_i(\delta) < 0$, but these negative values are multiplied by a small $s_j(\delta)$. Therefore, L_i is typically positive for most indices i except the first few, and the condition in (25) under which $\mathrm{SR}^2(\boldsymbol{w}_{s\alpha})$ decreases with δ is typically satisfied.

find that the condition is always satisfied. This result means that shrinking the covariance matrix imposes a tighter bound on the in-sample alpha Sharpe ratio, which is equivalent to imposing a no-arbitrage constraint as shown in Proposition 3. This shrinkage effect is useful because, when we do not apply shrinkage, the large in-sample Sharpe ratio of the alpha component experiences a dramatic decrease out of sample due to estimation risk; see Section 4.2.1 and Section A of the Internet Appendix.

4. Performance Analysis

In this section, we assess the performance of the shrinkage mean-variance portfolio that exploits the shrinkage covariance matrix proposed in Section 3.3. In Section 4.1, we evaluate the performance of our proposed methodology with simulated data based on a simple model of the economy with sentiment investors. In Section 4.2, we evaluate the performance of our shrinkage mean-variance portfolio across the six empirical datasets described in Section 2.[26]

4.1 Simulated Data

First, we study the performance of our proposed methodology on simulated data. To that end, we consider the following model of the economy with heterogeneous investors.

Assumption 1 *The economy satisfies the following characteristics:*

1. *Assets live for two periods,* t *and* $t+1$.

[26] In Section C of the Internet Appendix, we demonstrate the robustness of our empirical results using an alternative set of test assets based on industry portfolios.

2. *There are N risky assets that pay random dividends at time $t+1$,*

$$\boldsymbol{d}_{t+1} = \boldsymbol{\mu_d} + \boldsymbol{u}, \tag{26}$$

where $\boldsymbol{\mu_d} \in \mathbb{R}^N$ is the vector of average dividends and $\boldsymbol{u} \in \mathbb{R}^N$ follows a zero-mean multivariate distribution with covariance matrix $\boldsymbol{\Sigma}$.

3. *All investors have mean-variance preferences with risk-aversion coefficient $\gamma > 0$.*

4. *A mass $\phi \in [0,1]$ of investors hold biased beliefs about future cash flows, i.e., they believe that $\mathrm{E}[\boldsymbol{d}_{t+1}] = \boldsymbol{\mu_d} + \boldsymbol{b}$ with at least one $b_i \neq 0$. The remaining mass $1-\phi$ are investors with unbiased beliefs, i.e., they believe that $\mathrm{E}[\boldsymbol{d}_{t+1}] = \boldsymbol{\mu_d}$.*

Building on the economy described in Assumption 1, we characterize the Sharpe ratios of the equilibrium market portfolio and its orthogonal alpha portfolio. In particular, the following proposition shows how the Sharpe ratios of the market and alpha portfolios depend on the mass of sentiment investors, ϕ, and their bias vector, $\boldsymbol{b}$.

Proposition 5 *Let Assumption 1 hold and $\boldsymbol{w}_{mkt}$ be the equilibrium market portfolio. Then, the vector of equilibrium expected returns is*

$$\boldsymbol{\mu} = \gamma \boldsymbol{\Sigma} \boldsymbol{w}_{mkt} - \phi \boldsymbol{b}, \tag{27}$$

and thus, μ_i decreases with b_i and decreases (increases) with ϕ if $b_i \geq 0$ ($b_i < 0$). Moreover, the market and alpha Sharpe ratios are

$$\mathrm{SR}(\boldsymbol{w}_{mkt}) = \gamma \sigma_{mkt} - \frac{\phi \boldsymbol{b}^\top \boldsymbol{w}_{mkt}}{\sigma_{mkt}} \quad \textit{and} \quad \mathrm{SR}(\boldsymbol{w}_{\alpha}) = \phi \sqrt{\boldsymbol{b}^\top \boldsymbol{Q} \boldsymbol{b}}, \tag{28}$$

where the alpha portfolio is $\boldsymbol{w}_\alpha = \boldsymbol{w}_{mv} - (\mu_{mkt}/(\gamma\sigma^2_{mkt}))\boldsymbol{w}_{mkt}$, $\sigma^2_{mkt} = \boldsymbol{w}^\top_{mkt}\boldsymbol{\Sigma}\boldsymbol{w}_{mkt}$ *and* $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{w}_{mkt}\boldsymbol{w}^\top_{mkt}/\sigma^2_{mkt}$ *is positive semidefinite.*

Proposition 5 characterizes expected returns and Sharpe ratios for an arbitrary equilibrium market portfolio. The following corollary considers the specific case in which the equilibrium market portfolio is 1/N and provides a simple characterization of how the Sharpe ratios of the 1/N and alpha components depend on the mass of sentiment investors, ϕ, and their bias vector, $\boldsymbol{b}$.

Corollary 1 *Let Assumption 1 hold and the equilibrium market portfolio be the 1/N portfolio,* $\boldsymbol{w}_{mkt} = \boldsymbol{w}_{ew}$. *Then, in equilibrium, the 1/N and alpha Sharpe ratios are*

$$\text{SR}(\boldsymbol{w}_{ew}) = \gamma\sigma_{ew} - \frac{\phi\bar{b}}{\sigma_{ew}} \quad \textit{and} \quad \text{SR}(\boldsymbol{w}_\alpha) = \phi\sqrt{\boldsymbol{b}^\top\boldsymbol{Q}\boldsymbol{b}}, \tag{29}$$

where $\bar{b} = \boldsymbol{b}^\top\boldsymbol{w}_{ew}$ *is the average bias and* $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{w}_{ew}\boldsymbol{w}^\top_{ew}/\sigma^2_{ew}$. *For a fixed* ϕ, $\text{SR}(\boldsymbol{w}_{ew})$ *decreases with* $\bar{b}$ *and, assuming* $\boldsymbol{b} = \bar{b}\mathbf{1}$, $\text{SR}(\boldsymbol{w}_\alpha)$ *increases with* $|\bar{b}|$. *For a fixed* $\boldsymbol{b}$, $\text{SR}(\boldsymbol{w}_\alpha)$ *increases with* ϕ *and* $\text{SR}(\boldsymbol{w}_{ew})$ *decreases (increases) with* ϕ *if optimism (pessimism) dominates among sentiment investors, i.e.,* $\bar{b} > 0$ $(\bar{b} < 0)$.

Corollary 1 shows that when the market portfolio is the $1/N$ portfolio, the Sharpe ratio of $1/N$ decreases with the average bias among sentiment investors, $\bar{b}$, and with the mass of sentiment investors, ϕ, provided that sentiment is optimistic on average, i.e., $\bar{b} > 0$. These results reflect that a positive bias in asset cash flows ($\bar{b} > 0$) leads to overpricing, thereby lowering expected returns. In contrast, the alpha portfolio captures the mispricing generated by both optimistic and pessimistic investors, as its squared Sharpe ratio is a quadratic form in the bias vector $\boldsymbol{b}$ and it increases with the mass of sentiment investors, ϕ.

The case in which most sentiment investors are optimistic, i.e., $\bar{b} > 0$, is empirically more relevant. Indeed, when investors face short-sale impediments, stocks subject to high levels of dis-

agreement are primarily held by optimistic investors in equilibrium because pessimistic investors are sidelined due to their inability to sell short; see Miller (1977) and Hong and Sraer (2016).[27] This mechanism also implies that, when short-sale constraints bind, pessimistic beliefs are not fully reflected in equilibrium prices, so that the effective impact of sentiment becomes asymmetric and is primarily driven by optimistic investors.

We now simulate return data from the model of the economy studied above.[28] Specifically, we simulate asset returns $\boldsymbol{r}_t$, where $t = 1, \ldots, \tau$ with $\tau = 10{,}000$, as follows. First, we set the number of assets $N = 100$. Second, we set the dividend vector at time t as $\boldsymbol{d}_t = \boldsymbol{\mu_d} + \boldsymbol{u}_t$, where $\boldsymbol{\mu_d} = \mathbf{0}$ because, as we show in Proposition 5, this parameter is irrelevant for the performance of the 1/N and alpha components. In addition, we define the vector of dividend innovations as $\boldsymbol{u}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, and we set $\boldsymbol{\Sigma} = N\boldsymbol{I}$ and $\gamma = 1$ to scale the performance of the 1/N component such that $\text{SR}(\boldsymbol{w}_{ew}) = 1$ when $\phi = 0$; see Equation (29). Third, regarding the vector of sentiment investors' biases $\boldsymbol{b}$, we assume that a fraction $f \in [0, 1]$ of the assets are held by optimistic investors and their bias is equal to $c > 0$. That is, $\lfloor Nf \rfloor$ entries of the vector $\boldsymbol{b}$ are equal to c and $N - \lfloor Nf \rfloor$ entries to $-c$, and we set $c = 2$ and $f \in \{20\%, 50\%, 80\%\}$. Fourth, we use the value of ϕ at time t to represent the level of sentiment in the economy at time t. We first simulate a time series x_t from a gaussian autoregressive process with one lag and a first-order autocorrelation of 0.9837, which is the first-order autocorrelation of the Huang et al. (2015) sentiment index over the period July 1965 to June 2022. Then, we transform x_t in the $[0, 1]$ interval to obtain ϕ_t as $\phi_t = (x_t - \min(x))/(\max(x) - \min(x))$. Finally, as explained in the proof of Proposition 5, the asset returns

[27] Section K of the Internet Appendix extends the model to incorporate short-sale-constrained investors. We consider three groups of investors: optimistic and pessimistic short-sale-constrained sentiment investors, and unconstrained rational investors. When disagreement between optimistic and pessimistic sentiment investors is sufficiently large, the latter are sidelined in equilibrium due to short-sale constraints, so that only the optimistic component of sentiment is priced.

[28] In addition, Section L of the Internet Appendix studies the empirical link between the model-implied and observable sentiment measures.

at time t are $\boldsymbol{r}_t = \boldsymbol{d}_t - \boldsymbol{P}_t$, where the equilibrium prices are $\boldsymbol{P}_t = \boldsymbol{\mu_d} - \gamma\boldsymbol{\Sigma}\boldsymbol{w}_{ew} + \phi_t\boldsymbol{b}$, which gives $\boldsymbol{r}_t = \boldsymbol{u}_t + \gamma\boldsymbol{\Sigma}\boldsymbol{w}_{ew} - \phi_t\boldsymbol{b}$.[29]

Given a dataset of asset returns $\boldsymbol{r}_t$, where $t = 1, \ldots, \tau$ with $\tau = 10{,}000$, and a given portfolio strategy, we construct a series of out-of-sample portfolio returns using expanding windows. Specifically, at time t, we use an expanding window of $T + t - 1$ observations, where $T = 120$ is the initial sample size. Then, we evaluate the out-of-sample portfolio return at time $T + t$. We apply this process for all $t = 1, \ldots, \tau - T$ observations. As a robustness check, we also conduct simulations in a higher-dimensional setting with $N = 250$ and $\tau = 500$.

In Figure 6, we depict the out-of-sample Sharpe ratios of the SSH portfolio and the 1/N and alpha components. Panels A, B, and C consider the case where the proportion of optimistic investors is $f = 80\%$, 50%, and 20%, respectively. As explained in Section 4.1, the case with a larger fraction of optimistic investors is the more relevant one in practice. We depict the Sharpe ratios across sentiment deciles, i.e., deciles sorted on ϕ_t. In addition, we depict the SSH shrinkage intensity δ across sentiment deciles. To gain intuition about the results in Figure 6, we first provide the theoretical Sharpe ratios of the 1/N and alpha components in Corollary 1 under the specified parameters, which can be shown to be given by

$$\text{SR}(\boldsymbol{w}_{ew}) = 1 - 2\phi c(f - 1/2), \tag{30}$$

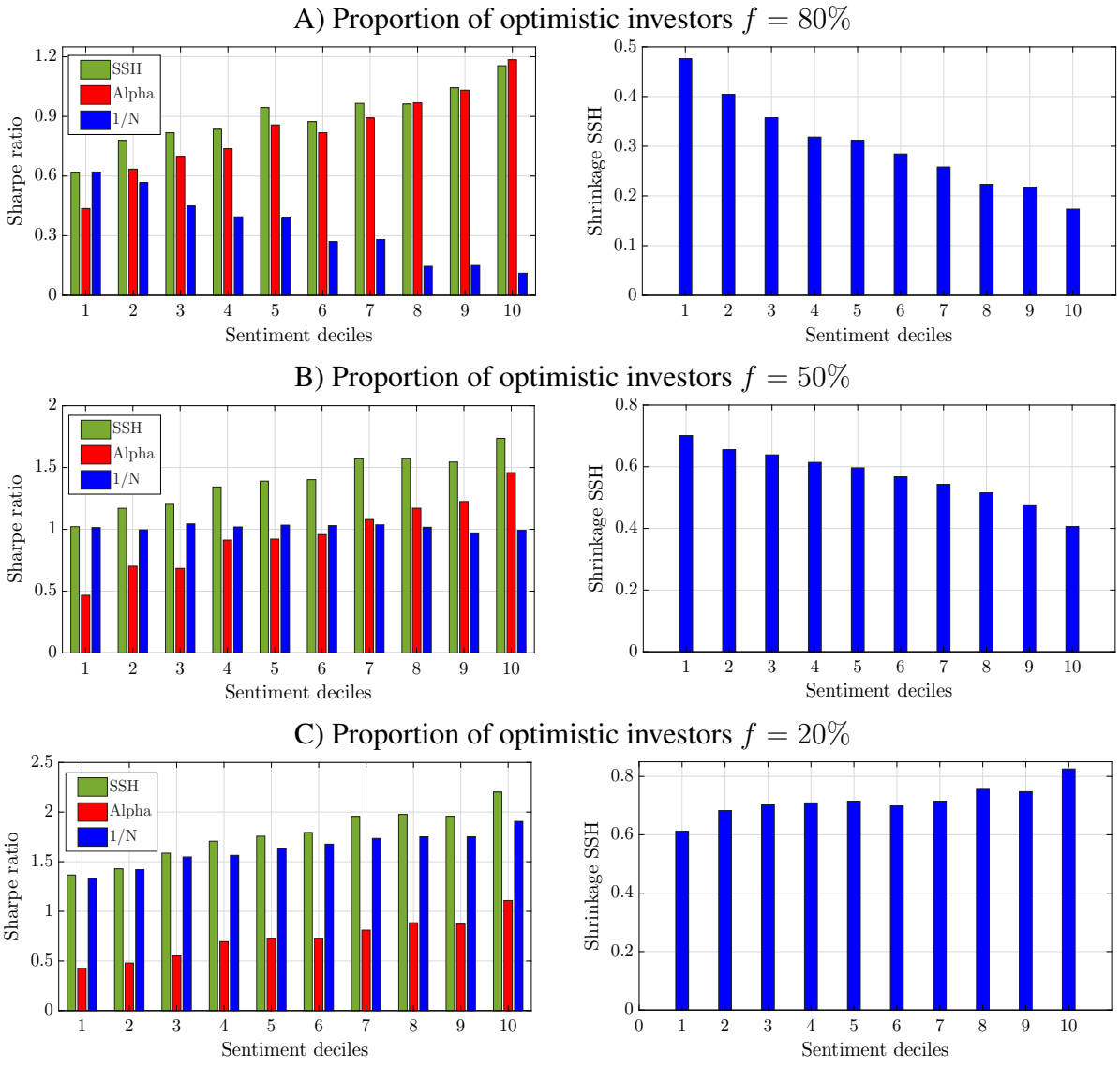
$$\text{SR}(\boldsymbol{w}_{\alpha}) = 2\phi c\sqrt{f(1-f)}. \tag{31}$$

We make the following observations from Figure 6. First, the 1/N Sharpe ratio decreases with sentiment when the proportion of optimistic investors $f = 80\%$, stays around one when $f = 50\%$, and increases with sentiment when $f = 20\%$, which is consistent with how $\text{SR}(\boldsymbol{w}_{ew})$ depends

[29] Under the considered parameters, the unconditional mean return of asset i is $\text{E}[r_{it}] = 1 - \text{E}[\phi_t]b_i = 1 - b_i/2$, which, for $c = 2$, is equal to zero (two) if the asset is held by optimistic (pessimistic) investors.

Figure 6: Sharpe ratios and SSH shrinkage intensity in simulated data

This figure depicts the out-of-sample Sharpe ratios of the proposed sentiment-based shrinkage (SSH) portfolio and the 1/N and alpha components, as well as the SSH shrinkage intensity, in the simulation experiment described in Section 4.1. Panels A, B, and C consider a proportion of optimistic investors $f = 80\%$, 50%, and 20% among sentiment investors, respectively. The Sharpe ratios and shrinkage intensities are depicted across sentiment deciles, i.e., deciles sorted on the mass of sentiment investors ϕ_t. In each panel, we simulate 10,000 returns on $N = 100$ assets. The out-of-sample Sharpe ratios are obtained using expanding windows with an initial size of $T = 120$. The SSH portfolio performs consistently well across sentiment deciles by strategically adjusting the shrinkage intensity: it shrinks more (less) when 1/N outperforms (underperforms) the alpha portfolio, independently of the directional effect of sentiment-driven mispricing.
Alt text: Left charts compare portfolio Sharpe ratios across sentiment deciles. Right charts show shrinkage generally decreasing in Panels A and B, but increasing in Panel C.



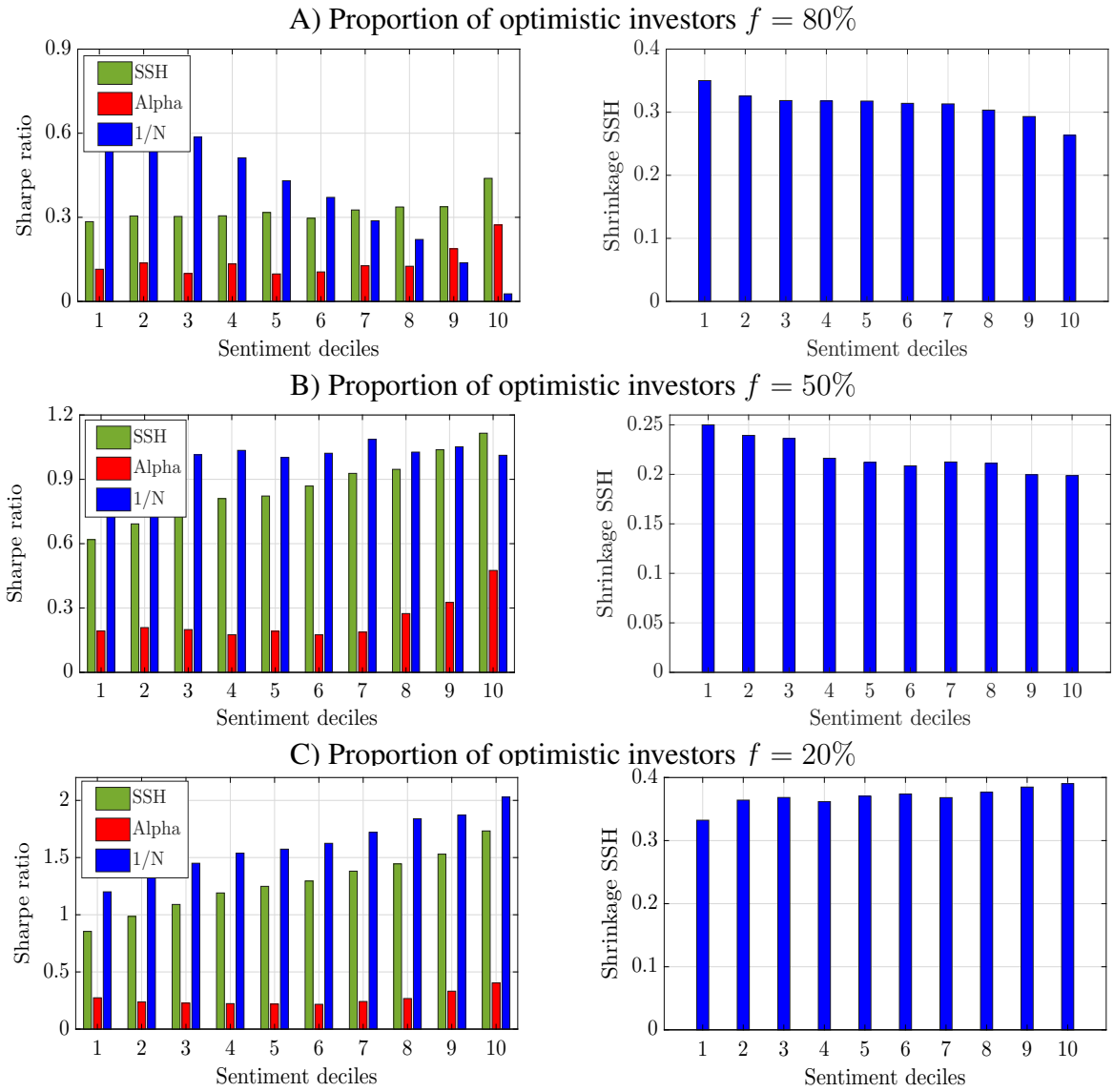
on f and ϕ in (30). In contrast, the Sharpe ratio of the alpha portfolio increases with sentiment for all considered values of f and is more sensitive to sentiment when $f = 50\%$, in line with how $\text{SR}(\boldsymbol{w}_\alpha)$ depends on f and ϕ in (31). Second, the SSH shrinkage intensity decreases with sentiment when $f = 80\%$ and 50% because, in these cases, alpha performance increases with sentiment while the 1/N performance decreases or stays flat with sentiment. In contrast, when $f = 20\%$, both the 1/N and alpha performance increase with sentiment, and the increase in 1/N performance is large enough for the SSH shrinkage intensity to increase with sentiment.[30] Overall, the SSH methodology achieves its objective by strategically tilting the mean-variance portfolio toward its two components conditional on sentiment.

We conclude the simulation analysis by considering a higher-dimensional setting where N is large relative to the available sample size. Specifically, we replicate Figure 6 by keeping the same baseline setting but increasing N to 250 and simulating only $\tau = 500$ returns instead of 10,000, so that this dimensionality remains substantial throughout the experiment. We repeat the experiment 100 times and report results averaged across the 100 repetitions. We depict the results in Figure 7. Overall, we find that the optimal shrinkage direction is similar to that in the baseline lower-dimensional setting of Figure 6.

[30] This is because the SSH shrinkage intensity is calibrated from the conditional probability $p_t = \Pr(\lambda r_{ew,t+1} > r_{\alpha,t+1} \mid \text{Sentiment}_t)$ rather than from a comparison of Sharpe ratios. In particular, what matters is the gap between the realized returns of the two components of the mean-variance portfolio. Under the calibration of Panel C of Figure 6 ($f = 20\%$, $c = 2$, $\gamma = 1$, $\boldsymbol{\Sigma} = N\boldsymbol{I}$), Corollary 1 delivers $\lambda(\phi) = \mu_{ew}(\phi) = 1 + 1.2\phi$, $\sigma^2_{ew} = 1$, and $\mu_\alpha(\phi) = 2.56\phi^2$, so that $\mathbb{E}[\lambda r_{ew,t+1} - r_{\alpha,t+1}] = 1 + 2.4\phi - 1.12\phi^2$, which is strictly increasing in ϕ on $[0, 1]$. This increasing expected return difference is consistent with the increasing shrinkage intensity observed in Panel C. When optimism dominates instead, as in Panel A ($f = 80\%$), the same algebra implies $\mathbb{E}[\lambda r_{ew,t+1} - r_{\alpha,t+1}] = 1 - 2.4\phi - 1.12\phi^2$, which is strictly decreasing in ϕ on $[0, 1]$, consistent with the decreasing SSH shrinkage intensity observed in Panel A. Either way, our simulation shows that, without imposing a priori any directional effect of investor sentiment on the relative performance of the $1/N$ and alpha components, our methodology correctly captures how the shrinkage intensity should respond to sentiment in order to deliver an SSH portfolio that performs consistently well across sentiment regimes.

Figure 7: Sharpe ratios and SSH shrinkage intensity in simulated data with a higher-dimensional setting

This figure depicts the out-of-sample Sharpe ratios of the proposed sentiment-based shrinkage (SSH) portfolio and the 1/N and alpha components, as well as the SSH shrinkage intensity, in the simulation experiment described in Section 4.1. Panels A, B, and C consider a proportion of optimistic investors $f = 80\%$, 50%, and 20% among sentiment investors, respectively. The Sharpe ratios and shrinkage intensities are depicted across sentiment deciles, i.e., deciles sorted on the mass of sentiment investors ϕ_t. In each panel, we simulate 500 returns on $N = 250$ assets and report averaged results across 100 repetitions of the experiment. The out-of-sample Sharpe ratios are obtained using expanding windows with an initial size of $T = 120$. The SSH portfolio strategically adjusts the shrinkage intensity: it shrinks more (less) when the performance of 1/N relative to the alpha portfolio improves (worsens), independently of the directional effect of sentiment-driven mispricing. Moreover, 1/N is hard to outperform in this high-dimensional setting.
Alt text: Across sentiment deciles, left charts compare average portfolio Sharpe ratios. Right charts show shrinkage declining in Panels A and B, and rising slightly in Panel C.



4.2 Empirical Data

We now turn to the empirical data described in Section 2. In Section 4.2.1, we study the performance of the 1/N and alpha components across sentiment regimes. In Section 4.2.2, we evaluate the out-of-sample performance of our proposed SSH shrinkage methodology relative to several benchmarks. In Section 4.2.3, we investigate the sources of the empirical gains of our shrinkage mean-variance portfolio. Section 4.2.4 compares implementations that apply sentiment-based shrinkage to the covariance matrix, mean vector, and portfolio weights, and Section 4.2.5 evaluates all portfolio strategies using a statistical shrinkage estimator of the mean vector.

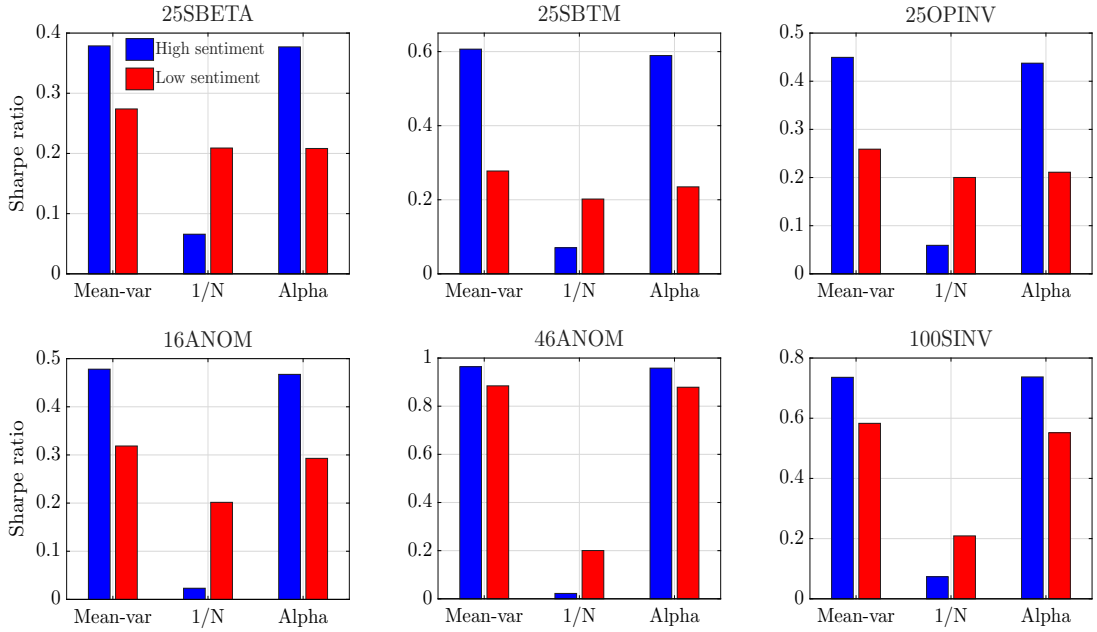
4.2.1 Role of Sentiment in 1/N and Alpha Components

Our main contribution is to propose a *conditional* shrinkage framework that can, in principle, accommodate any observable state variable predicting the relative future performance of the 1/N and alpha components. We use investor sentiment as our conditioning variable because of its documented relation with the returns of systematic and mispricing portfolios (Stambaugh, Yu, and Yuan 2012; Huang et al. 2015; Shen, Yu, and Zhao 2017). In this section, we provide empirical support for this choice.

First, we study the role of sentiment in the performance of the 1/N and alpha portfolios. Figure 8 depicts, across the six datasets, the in-sample monthly Sharpe ratios of the mean-variance portfolio and its 1/N and alpha components in high and low-sentiment regimes. Following Barroso and Detzel (2021), we define high-sentiment regimes as those years for which the Huang et al. (2015) sentiment index at the end of the prior year is above the historical median. We observe that the alpha component delivers a larger Sharpe ratio during high-sentiment regimes, whereas the 1/N performance is better during low-sentiment periods.

Figure 8: In-sample Sharpe ratios across sentiment regimes

This figure depicts the in-sample monthly Sharpe ratios of the mean-variance portfolio ($\boldsymbol{w}_{mv}$ in (1)), its 1/N component ($\boldsymbol{w}_{ew}$ in (3)), and its alpha component ($\boldsymbol{w}_{\alpha}$ in (4)), across the six datasets described in Section 2. The Sharpe ratios are computed separately in high and low-sentiment regimes. Like Barroso and Detzel (2021), we define high-sentiment regimes as those years for which the Huang et al. (2015) sentiment index at the end of the prior year is above the historical median. The alpha component delivers a larger in-sample Sharpe ratio during high-sentiment regimes, whereas the 1/N component performs better during low-sentiment periods, consistent with higher investor sentiment being associated with market overpricing.
Alt text: Six panels compare Sharpe ratios across sentiment regimes. Mean-variance and alpha portfolios perform better during high sentiment, while the equal-weighted portfolio performs better during low sentiment.



Second, investor sentiment predicts the relative out-of-sample performance of the 1/N and alpha components, as we show in a horse race against standard macroeconomic variables. At each rebalancing date t, we form the next-month return difference between the scaled 1/N and alpha components, $S_{t+1} = \lambda_t r_{ew,t+1} - r_{\alpha,t+1}$, whose sign defines the binary event underlying the probability in Equation (18). For each candidate conditioning variable, we estimate a linear predictive regression for S_{t+1} and the logistic specification in Equation (18), using the candidate variable

instead of sentiment. We compare the PLS sentiment index with five macroeconomic variables, each lagged by one month: realized market volatility, the Moody's Baa−Aaa credit spread, the aggregate dividend yield, industrial production growth, and initial claims growth.[31]

Table 2 reports the univariate and multivariate slope coefficients from pooled regressions, with S_{t+1} and the predictors standardized within each dataset in the linear regressions and the predictors standardized within each dataset in the logistic regressions. The reported t-statistics are based on standard errors clustered by calendar month to allow for contemporaneous dependence across the six dataset observations. Sentiment is negatively and statistically significantly related to S_{t+1} in the linear specification and to the probability that $S_{t+1} > 0$ in the logistic specification. It also delivers the highest univariate R^2 in the linear regressions and remains significant when all variables are included jointly. In the multivariate logistic regression, sentiment is the only variable significant at the 5% level. Some macroeconomic variables, such as initial claims growth, also contain predictive information, but sentiment remains significant when they are included jointly. Overall, the results indicate that sentiment is an economically motivated and empirically informative conditioning variable.[32]

4.2.2 Out-of-sample Performance of SSH Portfolio Methodology

We now evaluate the out-of-sample performance of our proposed sentiment-based shrinkage portfolio (SSH) across the six empirical datasets described in Section 2. We evaluate the out-of-sample portfolio performance using expanding windows with an initial size of $T = 120$ months as in the

[31] The market return is obtained from Kenneth French's Data Library, the dividend yield from Robert Shiller's website, and the remaining macroeconomic series from the FRED-MD database (series `BAAFFM−AAAFFM`, `INDPRO`, and `CLAIMSx`).

[32] In Section M of the Internet Appendix, we examine SSH performance when the logistic model is augmented with each macroeconomic variable.

Table 2: Predicting the relative performance of the scaled 1/N and alpha components

Panel A reports linear regressions of the one-month-ahead return difference between the scaled 1/N and alpha components, $S_{t+1} = \lambda_t r_{ew,t+1} - r_{\alpha,t+1}$, on each conditioning variable. Panel B reports logistic regressions for the binary event $\mathbf{1}_{S_{t+1}>0}$ underlying Equation (18). The conditioning variables are the PLS sentiment index and five macroeconomic variables, each lagged by one month: realized market volatility (the standard deviation of daily excess market returns within the month), the Moody's Baa-Aaa credit spread, the Shiller dividend yield D/P, industrial production growth (log difference), and initial claims growth (log difference). Observations are pooled across the six datasets described in Section 2, with S_{t+1} and all predictors standardized within each dataset. Column Univariate reports the slope from a regression containing that variable alone, whereas column Multivariate reports the slope from a regression including all six conditioning variables. t-statistics, based on standard errors clustered by calendar month to allow for contemporaneous dependence across datasets, are reported in square brackets. The final column reports the univariate R^2 in Panel A and McFadden's pseudo-R^2 in Panel B.

Conditioning variable	Univariate	Multivariate	R^2
Panel A: Linear regression for the 1/N-minus-alpha spread			
Sentiment	-0.124	-0.128	0.0154
	$[-3.34]$	$[-3.38]$	
Market volatility	0.062	0.048	0.0038
	[1.97]	[1.46]	
Credit spread	0.021	0.018	0.0004
	[0.64]	[0.46]	
Dividend yield	-0.011	-0.002	0.0001
	$[-0.34]$	$[-0.07]$	
Industrial production growth	-0.020	0.023	0.0004
	$[-0.87]$	[1.00]	
Initial claims growth	0.081	0.075	0.0065
	[4.67]	[3.32]	
Panel B: Logistic regression for the probability that 1/N outperforms alpha			
Sentiment	-0.140	-0.143	0.0033
	$[-2.59]$	$[-2.63]$	
Market volatility	0.083	0.053	0.0013
	[1.56]	[0.88]	
Credit spread	0.023	0.035	0.0001
	[0.49]	[0.55]	
Dividend yield	-0.042	-0.042	0.0003
	$[-0.84]$	$[-0.64]$	
Industrial production growth	-0.043	0.014	0.0003
	$[-1.15]$	[0.32]	
Initial claims growth	0.109	0.102	0.0019
	[2.45]	[1.90]	

simulation analysis. Every month t, we estimate the optimal mean-variance portfolio,

$$\boldsymbol{w}_{mv} = \frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}, \tag{32}$$

using information up to month t. We set a risk-aversion coefficient $\gamma = 1$.[33] We consider six strategies that estimate $\boldsymbol{w}_{mv}$ in different ways, as well as the low-turnover reward-to-risk and 1/N portfolios. The first five strategies estimate $\boldsymbol{\mu}$ using the sample mean estimator, and the covariance matrix $\boldsymbol{\Sigma}$ is estimated using the following shrinkage methods:[34]

1. The proposed covariance matrix in (11) that shrinks the sample covariance matrix toward the identity with the shrinkage intensity δ estimated from the sentiment-based shrinkage methodology described in Section 3.3, which gives the SSH portfolio.

2. The shrinkage covariance matrix proposed by Ledoit and Wolf (2004), which gives the MV-ID portfolio. This is the benchmark that is methodologically closest to ours because it also exploits a linear shrinkage covariance matrix, where the sample covariance matrix is shrunk toward the identity.

3. The shrinkage covariance matrix proposed by Ledoit and Wolf (2003) that shrinks the sample covariance matrix toward the CAPM-implied covariance matrix, which gives the MV-MKT portfolio.

4. The nonlinear shrinkage covariance matrix proposed by Ledoit and Wolf (2017; 2020), which gives the MV-NL portfolio.

[33] In Section N of the Internet Appendix, we show that our conclusions are robust to considering different portfolio scaling methods.

[34] Section O of the Internet Appendix considers a conditional EWMA covariance estimator designed to incorporate recent covariance information. The estimator is shrunk toward the identity matrix using either the SSH intensity or the Ledoit and Wolf (2004) intensity to alleviate estimation error.

5. The covariance matrix proposed by Chen and Yuan (2016) that mutes the contribution of a specific number of low-variance principal components, which gives the MV-PCA portfolio.

The sixth benchmark portfolio is the MAXSER portfolio by Ao, Li, and Zheng (2019) that estimates the optimal mean-variance portfolio in (32) via a least-squares regression representation and a LASSO constraint.[35] The last two benchmarks are simple low-turnover strategies:

7. The reward-to-risk (RTR) timing strategy of Kirby and Ostdiek (2012) that uses the sample estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ but setting all the sample covariances and negative mean returns equal to zero.

8. The 1/N portfolio $\boldsymbol{w}_{ew}$, which is a key benchmark because DeMiguel, Garlappi, and Uppal (2009) argue that the benefits of optimal diversification can be offset by estimation errors.

For all portfolio strategies, when the sum of the portfolio weights is larger than one in absolute value, we scale down the exposure to the risky assets so that the sum of the weights is capped at one, which avoids extreme leverage and portfolio weights and helps improve out-of-sample performance (Kirby and Ostdiek 2012). Moreover, in Section B of the Internet Appendix, we show that our results are robust to capping the monthly turnover of each strategy at 15%.

We report bootstrapped two-sided p-values of the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. To compute the p-values, we generate $1{,}000$ bootstrap samples using the stationary block bootstrap approach of Politis and Romano (1994) with an average block size of five and use the methodology of Ledoit and Wolf (2008, Remark 3.2.). 1/N serves as a natural benchmark to which we can compare all the other strategies because it plays a key role

[35]The original MAXSER by Ao, Li, and Zheng (2019) is designed to estimate the mean-variance portfolio that targets a given volatility σ, i.e., the portfolio $\boldsymbol{w} = \sigma(\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})^{-1/2}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. Cai et al. (2024) explain how to adapt MAXSER to estimate the portfolio $\boldsymbol{w} = \gamma^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ instead, and we follow their procedure here.

in our theory as the component of the mean-variance portfolio that does not suffer from estimation risk, and it delivers a notoriously strong balance between performance and turnover.[36]

Table 3 reports the annualized out-of-sample gross Sharpe ratios of the SSH portfolio and the seven benchmark strategies. The SSH portfolio achieves Sharpe ratios of 0.83, 0.91, 0.94, 1.10, 2.04, and 1.08 for the 25SBETA, 25SBTM, 25OPINV, 16ANOM, 46ANOM, and 100SINV datasets, respectively. Relative to the benchmark mean-variance portfolios, SSH delivers economically large improvements, with average outperformance across the six datasets of 38%, 61%, 38%, 32%, and 15% over MV-ID, MV-MKT, MV-NL, MV-PCA, and MAXSER, respectively. The only exceptions arise in the 46ANOM and 100SINV datasets. In 46ANOM, MV-NL and MAXSER achieve Sharpe ratios of 2.08 and 2.18, respectively, compared to 2.04 for SSH; in 100SINV, MAXSER attains a Sharpe ratio of 1.28 relative to 1.08 for SSH. Importantly, these differences largely disappear once transaction costs are incorporated, as the higher turnover of the benchmark portfolios erodes their apparent advantage (see Table 5).[37]

The performance gains of SSH are even more pronounced relative to simpler strategies: the average improvement in gross Sharpe ratio is 88% and 146% relative to RTR and 1/N, respectively. At the same time, the SSH portfolio weights are substantially more stable than the benchmark mean-variance portfolios. In particular, Table 4 shows that the benchmark mean-variance portfolios exhibit a turnover that is on average, close to six times higher than that of SSH.[38]

[36] In Section P of the Internet Appendix, we additionally report pairwise Sharpe-ratio tests between SSH and each mean-variance benchmark for each dataset, using the Ledoit and Wolf (2008) bootstrap methodology.

[37] Section Q of the Internet Appendix constructs volatility-managed versions of SSH and the mean-variance benchmarks. Every volatility-managed benchmark continues to have a lower net-of-cost Sharpe ratio than baseline SSH in all six datasets.

[38] Section R of the Internet Appendix provides a turnover decomposition to better understand the turnover differences between SSH and the mean-variance benchmark portfolios. In Section S of the Internet Appendix, we consider additional implementation aspects, including shorting costs and portfolio concentration.

Table 3: Out-of-sample Sharpe ratios

This table reports the annualized out-of-sample Sharpe ratio of the proposed SSH portfolio and seven benchmarks. The first four benchmarks are mean-variance portfolios constructed with the shrinkage covariance matrices of Ledoit and Wolf (2004) (MV-ID), Ledoit and Wolf (2003) (MV-MKT), Ledoit and Wolf (2020) (MV-NL), and Chen and Yuan (2016) (MV-PCA). The remaining ones are the MAXSER portfolio of Ao, Li, and Zheng (2019), the reward-to-risk timing strategy of Kirby and Ostdiek (2012) (RTR), and the 1/N portfolio. The table also reports in parentheses the two-sided p-values for the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The six datasets are described in Section 2. The numbers in bold font identify the portfolio strategies with the highest Sharpe ratio.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
SSH	**0.83**	**0.91**	**0.94**	**1.10**	2.04	1.08
	(0.08)	(0.03)	(0.01)	(0.00)	(0.00)	(0.00)
MV-ID	0.53	0.47	0.67	1.00	2.00	0.87
	(0.93)	(0.73)	(0.53)	(0.04)	(0.00)	(0.15)
MV-MKT	0.43	0.34	0.63	0.94	1.86	0.84
	(0.61)	(0.38)	(0.69)	(0.06)	(0.00)	(0.26)
MV-NL	0.51	0.47	0.65	0.93	2.08	0.99
	(0.86)	(0.75)	(0.60)	(0.08)	(0.00)	(0.05)
MV-PCA	0.62	0.73	0.80	0.88	1.26	0.84
	(0.76)	(0.43)	(0.12)	(0.11)	(0.00)	(0.13)
MAXSER	0.62	0.68	0.72	0.98	**2.18**	**1.28**
	(0.75)	(0.55)	(0.41)	(0.05)	(0.00)	(0.00)
RTR	0.64	0.61	0.62	0.58	0.60	0.64
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1/N	0.57	0.56	0.55	0.45	0.36	0.58

The lower turnover of the SSH portfolio suggests that its good performance is more likely to survive the effect of transaction costs than the alternative mean-variance portfolios. We measure the average transaction cost of portfolio p as

$$\bar{C}_p = \kappa \frac{1}{\tau_{\text{oos}}} \sum_{t=1}^{\tau_{\text{oos}}} \sum_{i=1}^{N} \left| w_{p,i,t} - w_{p,i,t^+} \right|, \tag{33}$$

where $w_{p,i,t}$ is the weight in asset i at time t, w_{p,i,t^+} is the weight in asset i at time t before rebalancing, κ is a proportional transaction-cost parameter, and τ_{oos} is the number of out-of-sample observations. We then subtract $\bar{C}_p$ from the average portfolio return to compute the Sharpe ratio net of proportional transaction costs. Following Barroso and Saxena (2022), we consider $\kappa = 10$,

Table 4: Portfolio turnover

This table reports the average monthly turnover of the proposed SSH portfolio and seven benchmarks. The first four benchmarks are mean-variance portfolios constructed with the shrinkage covariance matrices of Ledoit and Wolf (2004) (MV-ID), Ledoit and Wolf (2003) (MV-MKT), Ledoit and Wolf (2020) (MV-NL), and Chen and Yuan (2016) (MV-PCA). The remaining ones are the MAXSER portfolio of Ao, Li, and Zheng (2019), the reward-to-risk timing strategy of Kirby and Ostdiek (2012) (RTR), and the 1/N portfolio. The six datasets are described in Section 2.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
SSH	0.26	0.41	0.34	0.38	1.66	0.46
MV-ID	2.00	2.10	1.40	1.68	12.49	3.64
MV-MKT	2.66	2.98	1.46	2.41	15.09	3.71
MV-NL	2.55	2.51	1.47	2.50	8.69	2.64
MV-PCA	0.55	0.91	0.33	3.24	14.36	0.59
MAXSER	1.16	1.49	1.02	1.58	23.41	1.67
RTR	0.04	0.04	0.04	0.05	0.04	0.04
1/N	0.04	0.04	0.04	0.04	0.04	0.04

30, and 50 basis points. Table 5 shows that the impact of transaction costs on the performance of the SSH portfolio is limited. For $\kappa = 30$ basis points, for instance, SSH achieves net-of-cost Sharpe ratios of 0.78, 0.85, 0.88, 1.04, 1.91, and 0.99 for the 25SBETA, 25SBTM, 25OPINV, 16ANOM, 46ANOM, and 100SINV datasets, respectively. In relative terms, the average outperformance of the SSH portfolio across datasets is 65%, 113%, 67%, 42%, and 27% for the MV-ID, MV-MKT, MV-NL, MV-PCA, and MAXSER benchmark portfolios, respectively. Moreover, even though RTR and 1/N require a low turnover, SSH outperforms them by 79% and 134% on average, respectively.

Table 6 reports the annualized out-of-sample volatility of the SSH portfolio and the seven benchmark strategies. For comparability, all portfolio returns are normalized to match the expected return of the 1/N benchmark. Consistent with the Sharpe ratio and turnover evidence, SSH also delivers superior performance in terms of volatility. Specifically, SSH achieves volatilities of

Table 5: Net-of-cost out-of-sample Sharpe ratios

Panels A, B, and C report the annualized out-of-sample Sharpe ratio net of proportional transaction costs of 10, 30, and 50 basis points, respectively, of the proposed SSH portfolio and seven benchmarks. The first four benchmarks are mean-variance portfolios constructed with the shrinkage covariance matrices of Ledoit and Wolf (2004) (MV-ID), Ledoit and Wolf (2003) (MV-MKT), Ledoit and Wolf (2020) (MV-NL), and Chen and Yuan (2016) (MV-PCA). The remaining ones are the MAXSER portfolio of Ao, Li, and Zheng (2019), the reward-to-risk timing strategy of Kirby and Ostdiek (2012) (RTR), and the 1/N portfolio. We also report in parentheses the two-sided p-values of the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The six datasets are described in Section 2. The numbers in bold font identify the portfolio strategies with the highest Sharpe ratio net of proportional transaction costs.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: costs of 10bps						
SSH	**0.81** (0.10)	**0.89** (0.05)	**0.92** (0.01)	**1.08** (0.00)	2.00 (0.00)	1.05 (0.00)
MV-ID	0.48 (0.75)	0.43 (0.62)	0.63 (0.67)	0.96 (0.06)	1.95 (0.00)	0.78 (0.34)
MV-MKT	0.38 (0.45)	0.30 (0.29)	0.59 (0.79)	0.89 (0.09)	1.80 (0.00)	0.75 (0.43)
MV-NL	0.46 (0.67)	0.42 (0.62)	0.61 (0.75)	0.88 (0.10)	2.02 (0.00)	0.91 (0.12)
MV-PCA	0.61 (0.81)	0.70 (0.49)	0.78 (0.17)	0.83 (0.15)	1.18 (0.01)	0.82 (0.18)
MAXSER	0.58 (0.90)	0.64 (0.67)	0.69 (0.51)	0.93 (0.08)	**2.09** (0.00)	**1.21** (0.00)
RTR	0.64 (0.00)	0.61 (0.00)	0.61 (0.00)	0.58 (0.00)	0.59 (0.00)	0.63 (0.00)
1/N	0.57	0.56	0.55	0.45	0.36	0.58
Panel B: costs of 30bps						
SSH	**0.78** (0.14)	**0.85** (0.07)	**0.88** (0.02)	**1.04** (0.00)	**1.91** (0.00)	0.99 (0.00)
MV-ID	0.39 (0.48)	0.35 (0.40)	0.55 (0.97)	0.87 (0.11)	1.85 (0.00)	0.60 (0.90)
MV-MKT	0.28 (0.22)	0.21 (0.14)	0.51 (0.89)	0.79 (0.20)	1.66 (0.00)	0.57 (0.98)
MV-NL	0.36 (0.35)	0.33 (0.35)	0.53 (0.97)	0.79 (0.20)	**1.91** (0.00)	0.76 (0.37)
MV-PCA	0.57 (0.95)	0.64 (0.68)	0.74 (0.24)	0.72 (0.30)	1.02 (0.04)	0.76 (0.28)
MAXSER	0.50 (0.80)	0.57 (0.88)	0.62 (0.69)	0.85 (0.13)	1.90 (0.00)	**1.07** (0.01)
RTR	0.63 (0.00)	0.60 (0.00)	0.61 (0.00)	0.57 (0.00)	0.59 (0.00)	0.63 (0.00)
1/N	0.56	0.55	0.54	0.45	0.35	0.58
Panel C: costs of 50bps						
SSH	**0.75** (0.22)	**0.81** (0.10)	**0.84** (0.05)	**1.00** (0.00)	**1.83** (0.00)	**0.94** (0.01)
MV-ID	0.29 (0.22)	0.27 (0.23)	0.46 (0.75)	0.78 (0.18)	1.74 (0.00)	0.41 (0.44)
MV-MKT	0.17 (0.07)	0.12 (0.05)	0.43 (0.61)	0.69 (0.34)	1.51 (0.00)	0.38 (0.38)
MV-NL	0.25 (0.15)	0.23 (0.17)	0.44 (0.66)	0.69 (0.35)	1.79 (0.00)	0.60 (0.88)
MV-PCA	0.54 (0.94)	0.58 (0.85)	0.69 (0.35)	0.61 (0.53)	0.84 (0.13)	0.71 (0.41)
MAXSER	0.41 (0.48)	0.49 (0.84)	0.55 (0.95)	0.76 (0.23)	1.69 (0.00)	0.92 (0.07)
RTR	0.63 (0.00)	0.60 (0.00)	0.60 (0.00)	0.57 (0.00)	0.58 (0.00)	0.62 (0.00)
1/N	0.56	0.55	0.53	0.44	0.34	0.57

Table 6: Out-of-sample volatilities

This table reports the annualized out-of-sample volatility of the proposed SSH portfolio and seven benchmarks. For better comparison, the out-of-sample portfolio returns are normalized to match the expected return of 1/N. The first four benchmarks are mean-variance portfolios constructed with the shrinkage covariance matrices of Ledoit and Wolf (2004) (MV-ID), Ledoit and Wolf (2003) (MV-MKT), Ledoit and Wolf (2020) (MV-NL), and Chen and Yuan (2016) (MV-PCA). The remaining ones are the MAXSER portfolio of Ao, Li, and Zheng (2019), the reward-to-risk timing strategy of Kirby and Ostdiek (2012) (RTR), and the 1/N portfolio. The six datasets are described in Section 2. The numbers in bold font identify the portfolio strategies with the lowest volatility.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
SSH	**0.120**	**0.108**	**0.092**	**0.072**	0.032	0.096
MV-ID	0.188	0.211	0.129	0.079	0.033	0.118
MV-MKT	0.231	0.287	0.137	0.084	0.035	0.123
MV-NL	0.194	0.211	0.134	0.085	0.031	0.104
MV-PCA	0.160	0.136	0.108	0.090	0.052	0.122
MAXSER	0.160	0.146	0.121	0.081	**0.030**	**0.080**
RTR	0.155	0.162	0.141	0.135	0.109	0.162
1/N	0.174	0.176	0.158	0.174	0.182	0.176

0.120, 0.108, 0.092, 0.072, 0.032, and 0.096 for the 25SBETA, 25SBTM, 25OPINV, 16ANOM, 46ANOM, and 100SINV datasets, respectively. In relative terms, SSH reduces volatility substantially across datasets, with average reductions of 24%, 31%, 23%, 23%, 10%, 41%, and 50% relative to the MV-ID, MV-MKT, MV-NL, MV-PCA, MAXSER, RTR, and 1/N benchmarks, respectively.

4.2.3 Sources of Empirical Performance Gains of the SSH Portfolio

To understand why the proposed SSH portfolio delivers a good out-of-sample performance, we now study its performance across sentiment regimes, and its exposure to the 1/N component conditional on sentiment.[39]

[39] In Section T of the Internet Appendix, we show that SSH performance is not driven by contractionary economic periods.

In Figure 9, we depict the out-of-sample gross Sharpe ratio across sentiment regimes of the 1/N and alpha components, the SSH portfolio, and, for comparison purposes, the MV-ID portfolio exploiting the shrinkage covariance matrix of Ledoit and Wolf (2004).

We draw several insights from Figure 9, which are consistent with the simulation analysis in Section 4.1. First, over the whole sample, SSH performs at least equally well, and in general substantially better, than both the 1/N and alpha portfolios. Second, the performance of the 1/N and alpha portfolios varies substantially across sentiment regimes. On average, 1/N delivers a Sharpe ratio of 0.24 (0.87) in high-sentiment (low-sentiment) periods, while the alpha portfolio delivers a Sharpe ratio of 1.37 (0.54) in high-sentiment (low-sentiment) regimes. Third, due to its strategic exposure to the 1/N and alpha components, SSH performs consistently well across sentiment periods with an average Sharpe ratio of 1.46 (0.88) in high-sentiment (low-sentiment) periods, which is 9.5% (51%) higher than that of MV-ID. Therefore, the substantial economic gains delivered by the SSH portfolio relative to benchmark mean-variance portfolios that ignore the role of sentiment in the performance of their 1/N and alpha components come from the stronger performance during low-sentiment regimes.

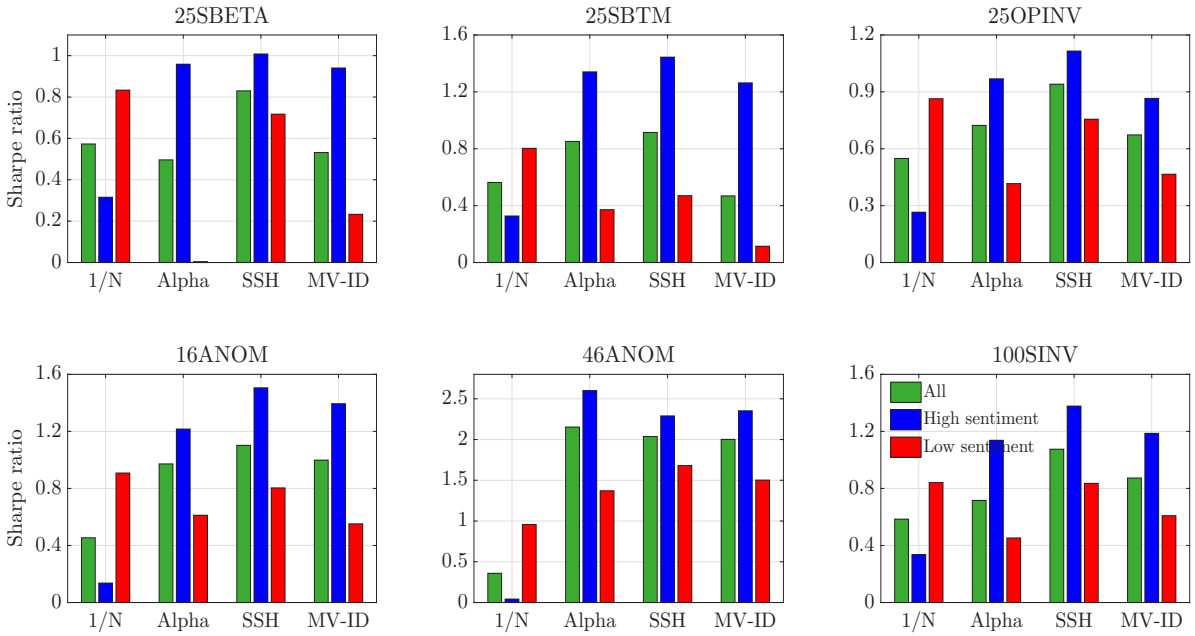
We now study the performance of the SSH portfolio using a conditional time-series regression. In particular, we consider the one-factor model in (21) and regress the out-of-sample returns of the SSH portfolio on 1/N portfolio returns, allowing for the slope coefficient to vary over time with investor sentiment, i.e.,

$$r_{ssh,t+1} = \alpha + (\beta_0 + \beta_1 \,\text{Sentiment}_t) r_{ew,t+1} + \epsilon_{t+1}. \tag{34}$$

Given that the sentiment index is standardized to have zero mean and unit variance, β_0 represents the average beta. Table 7 reports the intercept, slope coefficient, and t-statistics from the conditional

Figure 9: Sentiment and out-of-sample Sharpe ratios

This figure depicts the out-of-sample annualized gross Sharpe ratios of the 1/N portfolio, the alpha portfolio, the SSH portfolio, and the MV-ID portfolio that exploits the shrinkage covariance matrix in Ledoit and Wolf (2004). The six datasets are described in Section 2. The out-of-sample Sharpe ratios are evaluated either on the whole out-of-sample period or separately in high and low-sentiment regimes using the Huang et al. (2015) sentiment index. Like Barroso and Detzel (2021), we define high-sentiment regimes as those years for which the sentiment index at the end of the prior year is above its historical median. The SSH portfolio delivers consistently strong performance across both sentiment regimes, with its gains relative to benchmark mean-variance portfolios stemming primarily from its substantial outperformance during low-sentiment periods.
Alt text: Six panels compare portfolio Sharpe ratios across sentiment regimes. Sentiment-based shrinkage outperforms the identity-shrinkage benchmark during low sentiment in every dataset; differences are smaller during high sentiment.



time-series regression across our six datasets. For comparison purposes, we also report the results for the MV-ID portfolio.

We observe that the exposure of the SSH portfolio to the 1/N component is negatively related with sentiment, i.e., $\beta_1 < 0$ in (34), and β_1 is statistically significant across all datasets. Similarly, the conditional beta of the MV-ID portfolio also decreases with sentiment; however its statistical significance is, in general, lower. As discussed in Section 3.4, the correlation of any mean-variance

Table 7: Conditional exposure to the 1/N component

This table reports the intercept, slope coefficients, and Newey-West t-statistics (in square brackets) from the conditional time-series regressions in Equation (34). We run time-series regressions for the out-of-sample returns of the SSH portfolio and the mean-variance portfolio that exploits the Ledoit and Wolf (2004) shrinkage covariance matrix (MV-ID) across the six datasets described in Section 2.

	25SBETA		25SBTM		25OPINV		16ANOM		46ANOM		100SINV	
	SSH	MV-ID	SSH	MV-ID	SSH	MV-ID	SSH	MV-ID	SSH	MV-ID	SSH	MV-ID
α	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.03	0.08	0.50	0.01	0.03
	[3.59]	[3.10]	[4.28]	[2.29]	[4.27]	[3.30]	[5.22]	[4.36]	[7.92]	[5.34]	[5.35]	[5.13]
β_0	0.57	0.28	0.78	0.60	0.73	0.51	0.74	0.37	0.23	−1.76	0.72	0.34
	[9.50]	[1.37]	[13.88]	[2.51]	[12.96]	[3.66]	[9.11]	[2.29]	[1.44]	[−2.22]	[13.57]	[2.39]
β_1	−0.23	−0.18	−0.37	−1.12	−0.26	−0.66	−0.53	−0.92	−0.61	−2.34	−0.28	−0.52
	[−3.02]	[−0.87]	[−4.10]	[−2.17]	[−3.00]	[−3.27]	[−4.89]	[−4.03]	[−3.62]	[−2.13]	[−5.00]	[−3.56]

portfolio with 1/N returns should decrease when 1/N returns are smaller, which is typically the case in high-sentiment periods.

We also observe that SSH keeps a higher exposure to the 1/N component relative to MV-ID. Indeed, in Table 7, the average beta (β_0) of SSH is always larger than that of MV-ID. This can be explained because over time, as the size of the expanding window increases and parameter uncertainty in the covariance matrix decreases, MV-ID shrinks less intensively and becomes less exposed to 1/N. In contrast, SSH does not only shrink as a function of estimation errors in the covariance matrix. First, as explained in Section 3.3, SSH shrinks as a function of the estimation risk affecting the alpha component, and thus, as a function of parameter uncertainty in the covariance matrix, but also the vector of means, which can be substantial even for large sample sizes. Second, SSH shrinks as a function of sentiment, and thus, can shrink intensively when sentiment is low to keep a larger exposure to the 1/N component, even if the sample size is large.

The higher average exposure of the SSH portfolio to 1/N is also an important distinction relative to the benchmark mean-variance portfolios because it reduces transaction costs, as evidenced by the substantially lower turnover of our strategy in Table 4.

4.2.4 Alternative Channels for Sentiment-based Shrinkage

Our baseline SSH implementation uses sentiment to determine the shrinkage intensity of the covariance matrix. To assess whether the economic gains depend on this particular channel, we construct two alternatives. Mean-SSH applies sentiment-based shrinkage to the mean vector by shrinking the sample means toward their cross-sectional average, whereas weight-SSH applies it directly to the portfolio weights by shrinking the mean-variance portfolio toward its scaled 1/N component. In each case, the same predicted probability p_t used by baseline SSH calibrates a channel-specific shrinkage intensity through the correlation-matching criterion in Section 3.3. Section E of the Internet Appendix provides the theoretical derivations.

Panel A of Table 8 shows that the three implementations have similar gross performance: mean-SSH and weight-SSH both attain an average Sharpe ratio of 1.16, compared with 1.15 for SSH. Panel C shows, however, that SSH's average turnover reduction across datasets is 53% relative to mean-SSH and 60% relative to weight-SSH. Consequently, Panel B shows that SSH delivers an average net-of-cost Sharpe ratio of 1.08, compared with 0.98 for mean-SSH and weight-SSH, and has the higher net-of-cost Sharpe ratio in four and five of the six datasets, respectively. The tests against 1/N also generally yield smaller p-values for SSH.

These results show that investor sentiment is useful for timing exposure to the 1/N and alpha components beyond the covariance-matrix implementation. They also show why covariance

Table 8: Alternative channels for sentiment-based shrinkage

This table reports the annualized out-of-sample gross Sharpe ratio (Panel A), Sharpe ratio net of 30-bps transaction costs (Panel B), and turnover (Panel C) of sentiment-based shrinkage applied to the covariance matrix (SSH), the mean vector (mean-SSH), and portfolio weights (weight-SSH). The table reports in parentheses the two-sided p-values for tests of equality between each portfolio's Sharpe ratio and that of 1/N. The six datasets are described in Section 2. The numbers in bold identify the portfolio strategies with the highest Sharpe ratio in each dataset.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Gross Sharpe ratio						
SSH	**0.83**	**0.91**	**0.94**	1.10	2.04	1.08
	(0.08)	(0.03)	(0.01)	(0.00)	(0.00)	(0.00)
mean-SSH	0.72	0.75	0.93	**1.19**	2.25	**1.14**
	(0.47)	(0.38)	(0.02)	(0.00)	(0.00)	(0.00)
weight-SSH	0.78	0.85	0.92	1.14	**2.26**	1.01
	(0.20)	(0.11)	(0.01)	(0.00)	(0.00)	(0.00)
Panel B: Net-of-cost Sharpe ratio (30bps)						
SSH	**0.78**	**0.85**	**0.88**	1.04	1.91	**0.99**
	(0.14)	(0.07)	(0.02)	(0.00)	(0.00)	(0.00)
mean-SSH	0.57	0.58	0.84	**1.07**	1.94	0.90
	(0.94)	(0.89)	(0.07)	(0.01)	(0.00)	(0.06)
weight-SSH	0.62	0.69	0.82	1.02	**1.97**	0.74
	(0.72)	(0.44)	(0.06)	(0.01)	(0.00)	(0.28)
Panel C: Turnover						
SSH	0.26	0.41	0.34	0.38	1.66	0.46
mean-SSH	0.78	1.14	0.48	0.66	3.39	1.34
weight-SSH	0.90	1.06	0.52	0.75	5.21	1.86

shrinkage is our preferred implementation: it produces substantially more stable portfolio weights and therefore superior net-of-cost performance.[40]

4.2.5 Shrinkage Estimator of the Mean Vector

In this section, we test whether the outperformance of SSH versus the benchmarks is explained by the use of the sample mean estimator instead of a shrinkage mean estimator. To implement this

[40] Section U of the Internet Appendix additionally shows that a direct sentiment-based combination of the scaled 1/N and alpha components has higher turnover and lower gross and net-of-cost Sharpe ratios than SSH.

robustness test, we employ the shrinkage mean estimator of Asimit, Chen, and Lassance (2026, Theorem 4). Asimit, Chen, and Lassance (2026) develop simple distribution-free shrinkage estimators of the mean vector that are consistent in high-dimensional settings. Given a sample of N assets and M_t observations, and given the sample mean and covariance matrix $(\hat{\boldsymbol{\mu}}_t, \widehat{\boldsymbol{\Sigma}}_t)$, the estimator in Asimit, Chen, and Lassance (2026, Theorem 4) shrinks $\hat{\boldsymbol{\mu}}_t$ toward the cross-sectional mean $\hat{\mu}_{ew,t} = \boldsymbol{w}_{ew}^\top \hat{\boldsymbol{\mu}}_t$. The shrinkage intensity is defined to minimize the mean squared error, which yields

$$\begin{aligned} &\hat{\boldsymbol{\mu}}_{acl,t} = (1 - \hat{\delta}_{acl,t})\hat{\boldsymbol{\mu}}_t + \hat{\delta}_{acl,t}\hat{\mu}_{ew,t}\mathbf{1} \\ &\text{with} \quad \hat{\delta}_{acl,t} = \frac{N\text{tr}(\widehat{\boldsymbol{\Sigma}}_t) - \mathbf{1}^\top\widehat{\boldsymbol{\Sigma}}_t\mathbf{1}}{N\text{tr}(\widehat{\boldsymbol{\Sigma}}_t) - \mathbf{1}^\top\widehat{\boldsymbol{\Sigma}}_t\mathbf{1} + M_t\left[N\hat{\boldsymbol{\mu}}_t^\top\hat{\boldsymbol{\mu}}_t - (\mathbf{1}^\top\hat{\boldsymbol{\mu}}_t)^2\right]}. \end{aligned} \tag{35}$$

Table 9 reports the gross and net-of-cost annualized out-of-sample Sharpe ratio, and the turnover, of the SSH portfolio and the seven benchmarks using the shrinkage mean estimator of Asimit, Chen, and Lassance (2026) defined in Equation (35).[41]

The main insight from Table 9 is that SSH continues to dominate across benchmarks when the sample mean is shrunk. The SSH portfolio achieves gross Sharpe ratios of 0.89, 0.94, 0.96, 1.14, 2.14, and 1.12 for the 25SBETA, 25SBTM, 25OPINV, 16ANOM, 46ANOM, and 100SINV datasets, respectively, which are all greater than those under the sample mean reported in Table 3. In relative terms, the average outperformance of SSH across the six datasets is 23%, 44%, 27%, 14%, and 20% versus MV-ID, MV-MKT, MV-NL, MV-PCA, and MAXSER, respectively. Moreover, SSH delivers significantly lower turnover than these benchmarks, resulting in greater relative outperformance under transaction costs of 30 basis points. In particular, the SSH portfolio achieves net-of-cost Sharpe ratios of 0.83, 0.88, 0.90, 1.08, 2.01, and 1.04 for the 25SBETA, 25SBTM,

[41] We do not apply the shrinkage mean estimator to MAXSER because it is not a plug-in estimator.

Table 9: Out-of-sample performance with shrinkage of the mean vector

This table reports the annualized out-of-sample Sharpe ratio, and the turnover, of the proposed SSH portfolio and seven benchmarks. The first four benchmarks are mean-variance portfolios constructed with the shrinkage covariance matrices of Ledoit and Wolf (2004) (MV-ID), Ledoit and Wolf (2003) (MV-MKT), Ledoit and Wolf (2020) (MV-NL), and Chen and Yuan (2016) (MV-PCA). The remaining ones are the MAXSER portfolio of Ao, Li, and Zheng (2019), the reward-to-risk timing strategy of Kirby and Ostdiek (2012) (RTR), and the 1/N portfolio. All portfolios except MAXSER and 1/N use the shrinkage mean vector of Asimit, Chen, and Lassance (2026, Theorem 4). The table also reports in parentheses the two-sided p-values for the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The Sharpe ratios are either gross or net of proportional transaction costs of 30 basis points. The six datasets are described in Section 2. The numbers in bold font identify the portfolio strategies with the highest Sharpe ratio.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Gross Sharpe ratio						
SSH	**0.89**	**0.94**	**0.96**	**1.14**	2.14	1.12
	(0.04)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
MV-ID	0.69	0.56	0.78	1.06	2.08	1.05
	(0.57)	(0.98)	(0.22)	(0.02)	(0.00)	(0.02)
MV-MKT	0.58	0.37	0.76	1.02	1.90	1.04
	(0.92)	(0.50)	(0.27)	(0.03)	(0.00)	(0.02)
MV-NL	0.65	0.47	0.79	1.02	**2.27**	1.17
	(0.69)	(0.77)	(0.22)	(0.03)	(0.00)	(0.00)
MV-PCA	0.85	0.81	0.82	1.12	1.75	0.90
	(0.08)	(0.08)	(0.04)	(0.00)	(0.00)	(0.02)
MAXSER	0.62	0.68	0.72	0.98	2.18	**1.28**
	(0.75)	(0.55)	(0.41)	(0.05)	(0.00)	(0.00)
RTR	0.63	0.59	0.58	0.52	0.56	0.62
	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)
1/N	0.57	0.56	0.55	0.45	0.36	0.58
Panel B: Net-of-cost Sharpe ratio (30bps)						
SSH	**0.83**	**0.88**	**0.90**	**1.08**	2.01	1.04
	(0.06)	(0.03)	(0.01)	(0.00)	(0.00)	(0.00)
MV-ID	0.53	0.41	0.65	0.93	1.92	0.78
	(0.92)	(0.59)	(0.55)	(0.06)	(0.00)	(0.28)
MV-MKT	0.40	0.22	0.63	0.87	1.72	0.78
	(0.44)	(0.18)	(0.62)	(0.10)	(0.00)	(0.29)
MV-NL	0.47	0.32	0.66	0.86	**2.09**	0.95
	(0.69)	(0.36)	(0.52)	(0.11)	(0.00)	(0.04)
MV-PCA	0.80	0.75	0.77	0.96	1.52	0.84
	(0.12)	(0.17)	(0.07)	(0.04)	(0.00)	(0.05)
MAXSER	0.50	0.57	0.62	0.85	1.90	**1.07**
	(0.80)	(0.88)	(0.69)	(0.13)	(0.00)	(0.01)
RTR	0.62	0.58	0.57	0.51	0.55	0.61
	(0.00)	(0.03)	(0.03)	(0.01)	(0.00)	(0.00)
1/N	0.56	0.55	0.54	0.45	0.35	0.58
Panel C: Turnover						
SSH	0.24	0.36	0.30	0.36	1.63	0.40
MV-ID	1.08	1.69	1.07	1.35	12.36	2.28
MV-MKT	1.48	2.31	1.09	1.90	13.77	2.21
MV-NL	1.31	1.96	1.05	1.98	7.51	1.55
MV-PCA	0.15	0.26	0.17	1.27	9.58	0.27
MAXSER	1.16	1.49	1.02	1.58	23.41	1.67
RTR	0.04	0.04	0.04	0.05	0.04	0.04
1/N	0.04	0.04	0.04	0.04	0.04	0.04

25OPINV, 16ANOM, 46ANOM, and 100SINV datasets, respectively. In relative terms, this represents an average outperformance across the six datasets of 44%, 87%, 53%, 18%, and 33% for the MV-ID, MV-MKT, MV-NL, MV-PCA, and MAXSER benchmarks, respectively. Finally, SSH also significantly outperforms the low-turnover RTR and 1/N benchmarks.

5. Conclusion

We introduce a conditional shrinkage method for portfolio selection that exploits investor sentiment. Our approach relies on a decomposition of the mean-variance portfolio as the sum of a 1/N component and an orthogonal alpha component and shrinks the covariance matrix conditional on investor sentiment to tilt the mean-variance portfolio toward either of these two components. We show through simulations that our shrinkage methodology is flexible and can capture the directional effect of sentiment-based mispricing. For instance, if investors are optimistic and the overall market is overpriced, our methodology shrinks *against* sentiment. On the other hand, if investors are pessimistic and the overall market is underpriced, our methodology shrinks *with* sentiment. Empirically, our methodology favors shrinking against sentiment, i.e., tilting the mean-variance portfolio toward 1/N—and away from the alpha component—when sentiment is low. Our sentiment-based shrinkage covariance matrix allows mean-variance investors to attain significant economic gains relative to alternative benchmark mean-variance portfolios at a substantially lower turnover.

Finally, applying the same sentiment-based shrinkage method to the mean vector or directly to the portfolio weights delivers comparable gross performance, indicating that the gains stem from the timing of the 1/N and alpha components rather than from the specific shrinkage channel.

Shrinking the covariance matrix remains preferable because it generally yields lower turnover and higher performance net of transaction costs.

Data availability

Replication materials are available in the Review of Finance Dataverse: https://doi.org/10.7910/DVN/RKJ9KT.

The replication package accompanying this article documents the programs and frozen input files used for the main-text results. Portfolio returns are based on Kenneth French's Data Library and the anomaly portfolios of Novy-Marx and Velikov. Investor-sentiment inputs are obtained from Jeffrey Wurgler's data page. The additional conditioning variables use the Fama–French daily factor data, Robert Shiller's stock-market data, and the June 2026 vintage of FRED-MD. Source links, file formats, sample alignment, and software requirements are documented in the replication README. The replication code and included inputs cover the main text; the online appendix is provided separately.

References

Ao, Mengmeng, Yingying Li, and Xinghua Zheng 2019. "Approaching mean-variance efficiency for large portfolios." *The Review of Financial Studies* 32, 2890–2919. https://doi.org/10.1093/rfs/hhy105

Asimit, Vali, Ziwei Chen, and Nathan Lassance 2026. "Distribution-free shrinkage of high-dimensional mean vector." *Journal of Business & Economic Statistics* (forthcoming). https://doi.org/10.1080/07350015.2026.2638490

Bakalli, Gaetan, Stéphane Guerrier, and Olivier Scaillet 2023. "A penalized two-pass regression to predict stock returns with time-varying risk premia." *Journal of Econometrics* 237. https://doi.org/10.1016/j.jeconom.2022.12.004

Baker, Malcolm, and Jeffrey Wurgler 2006. "Investor sentiment and the cross-section of stock returns." *The Journal of Finance* 61, 1645–1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x

Baker, Malcolm, and Jeffrey Wurgler 2007. "Investor sentiment in the stock market." *Journal of Economic Perspectives* 21, 129–152. https://doi.org/10.1257/jep.21.2.129

[dataset] Baker, Malcolm, and Jeffrey Wurgler n.d. "Investor sentiment data.", New York University. Frozen sentiment inputs through June 2022. https://pages.stern.nyu.edu/~jwurgler/

Barroso, Pedro, and Andrew Detzel 2021. "Do limits to arbitrage explain the benefits of volatility-managed portfolios?" *Journal of Financial Economics* 140, 744–767. https://doi.org/10.1016/j.jfineco.2021.02.009

Barroso, Pedro, Andrew L. Detzel, and Paulo Maio 2025. "The volatility puzzle of the beta anomaly." *Journal of Financial Economics* 165, 103994. https://doi.org/10.1016/j.jfineco.2025.103994

Barroso, Pedro, and Konark Saxena 2022. "Lest we forget: Learn from out-of-sample forecast errors when optimizing portfolios." *The Review of Financial Studies* 35, 1222–1278. https://doi.org/10.1093/rfs/hhab041

Britten-Jones, Mark 1999. "The sampling error in estimates of mean-variance efficient portfolio weights." *The Journal of Finance* 54, 655–671. https://doi.org/10.1111/0022-1082.00120

Bryzgalova, Svetlana, Markus Pelger, and Jason Zhu 2025. "Forest through the trees: Building cross-sections of stock returns." *The Journal of Finance* 80, 2447–2506. https://doi.org/10.1111/jofi.13477

Cai, Zhaokun, Zhenyu Cui, Nathan Lassance, and Majeed Simaan 2024. "The economic value of mean squared error: Evidence from portfolio selection." *SSRN working paper*. https://doi.org/10.2139/ssrn.4856139

Campbell, John, and Samuel Thompson 2008. "Predicting excess stock returns out of sample: Can anything beat the historical average?" *The Review of Financial Studies* 21, 1509–1531. https://doi.org/10.1093/rfs/hhm055

Chen, Jiaqin, and Ming Yuan 2016. "Efficient portfolio selection in a large market." *Journal of Financial Econometrics* 14, 496–524. https://doi.org/10.1093/jjfinec/nbw003

Cochrane, John H. 2009. *Asset Pricing: Revised Edition* (Princeton University Press).

DeMiguel, Victor, Lorenzo Garlappi, Francisco Nogales, and Raman Uppal 2009. "A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms." *Management Science* 55, 798–812. https://doi.org/10.1287/mnsc.1080.0986

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal 2009. "Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?" *The Review of Financial Studies* 22, 1915–1953. https://doi.org/10.1093/rfs/hhm075

DeMiguel, Victor, Alberto Martín-Utrera, and Raman Uppal 2024. "A multifactor perspective on volatility-managed portfolios." *The Journal of Finance* 79, 3859–3891. https://doi.org/10.1111/jofi.13395

DeMiguel, Victor, Francisco J Nogales, and Raman Uppal 2014. "Stock return serial dependence and out-of-sample portfolio performance." *The Review of Financial Studies* 27, 1031–1073. https://doi.org/10.1093/rfs/hhu002

Dumas, Bernard, Alexander Kurshev, and Raman Uppal 2009. "Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility." *The Journal of Finance* 64, 579–629. https://doi.org/10.1111/j.1540-6261.2009.01444.x

[dataset] Federal Reserve Bank of St. Louis 2026. "FRED-MD: Monthly database for macroeconomic research.", June 2026 vintage. https://www.stlouisfed.org/research/economists/mccracken/fred-databases

[dataset] French, Kenneth R. n.d. "Data library: Portfolio returns and research factors.", Dartmouth College. Frozen files and vintages are listed in the replication README. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken 1989. "A test of the efficiency of a given portfolio." *Econometrica* 57, 1121–1152. https://doi.org/10.2307/1913625

Gu, Shihao, Bryan Kelly, and Dacheng Xiu 2020. "Empirical asset pricing via machine learning." *The Review of Financial Studies* 33, 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Guijarro-Ordonez, Jorge, Markus Pelger, and Greg Zanotti 2024. "Deep learning statistical arbitrage." *Management Science* (forthcoming). https://doi.org/10.1287/mnsc.2022.03132

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman 2009. *The Elements of Statistical Learning: Data Mining, Inference and Prediction* (Springer), second edition. https://doi.org/10.1007/978-0-387-84858-7

Hong, Harrison, and David Sraer 2016. "Speculative betas." *The Journal of Finance* 71, 2095–2144. https://doi.org/10.1111/jofi.12431

Huang, Dashan, Fuwei Jiang, Jun Tu, and Guofu Zhou 2015. "Investor sentiment aligned: A powerful predictor of stock returns." *The Review of Financial Studies* 28, 791–837. https://doi.org/10.1093/rfs/hhu080

Kirby, Chris, and Barbara Ostdiek 2012. "It's all in the timing: Simple active portfolio strategies that outperform naïve diversification." *Journal of Financial and Quantitative Analysis* 47, 437–467. https://doi.org/10.1017/s0022109012000117

Ledoit, Olivier, and Michael Wolf 2003. "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection." *Journal of Empirical Finance* 10, 603–621. https://doi.org/10.1016/s0927-5398(03)00007-0

Ledoit, Olivier, and Michael Wolf 2004. "A well-conditioned estimator for large-dimensional covariance matrices." *Journal of Multivariate Analysis* 88, 365–411. https://doi.org/10.1016/s0047-259x(03)00096-4

Ledoit, Olivier, and Michael Wolf 2008. "Robust performance hypothesis testing with the Sharpe ratio." *Journal of Empirical Finance* 15, 850–859. https://doi.org/10.1016/j.jempfin.2008.03.002

Ledoit, Olivier, and Michael Wolf 2017. "Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks." *The Review of Financial Studies* 30, 4349–4388. https://doi.org/10.1093/rfs/hhx052

Ledoit, Olivier, and Michael Wolf 2020. "Analytical nonlinear shrinkage of large-dimensional covariance matrices." *The Annals of Statistics* 48, 3043–3065. https://doi.org/10.1214/19-aos1921

Maenhout, Pascal J. 2004. "Robust portfolio rules and asset pricing." *The Review of Financial Studies* 17, 951–983. https://doi.org/10.1093/rfs/hhh003

Markowitz, Harry 1952. "Portfolio selection." *The Journal of Finance* 7, 77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Miller, Edward 1977. "Risk, uncertainty, and divergence of opinion." *The Journal of Finance* 32, 1151–1168. https://doi.org/10.1111/j.1540-6261.1977.tb03317.x

Moreira, Alan, and Tyler Muir 2017. "Volatility-managed portfolios." *The Journal of Finance* 72, 1611–1644. https://doi.org/10.1111/jofi.12513

Novy-Marx, Robert, and Mihail Velikov 2016. "A taxonomy of anomalies and their trading costs." *The Review of Financial Studies* 29, 104–147. https://doi.org/10.1093/rfs/hhv063

[dataset] Novy-Marx, Robert, and Mihail Velikov n.d. "Anomaly portfolio returns.", Robert Novy-Marx data library. Frozen portfolio files through December 2013. https://rnm.simon.rochester.edu/data_lib/

Pástor, Ľuboš 2000. "Portfolio selection and asset pricing models." *The Journal of Finance* 55, 179–223. https://doi.org/10.1111/0022-1082.00204

Pettenuzzo, Davide, Allan Timmermann, and Rossen Valkanov 2014. "Forecasting stock returns under economic constraints." *Journal of Financial Economics* 114, 517–553. https://doi.org/10.1016/j.jfineco.2014.07.015

Politis, Dimitris, and Joseph Romano 1994. "The stationary bootstrap." *Journal of the American Statistical Association* 89, 1303–1313. https://doi.org/10.1080/01621459.1994.10476870

Raponi, Valentina, Raman Uppal, and Paolo Zaffaroni 2023. "Robust portfolio choice." *SSRN working paper*. https://doi.org/10.2139/ssrn.3933063

Ross, Stephen A 1976. "The arbitrage theory of capital asset pricing." *Journal of Economic Theory* 13, 341–360. https://doi.org/10.1016/0022-0531(76)90046-6

Shen, Junyan, Jianfeng Yu, and Shen Zhao 2017. "Investor sentiment and economic forces." *Journal of Monetary Economics* 86, 1–21. https://doi.org/10.1016/j.jmoneco.2017.01.001

[dataset] Shiller, Robert J. n.d. "Stock market data.", Yale University. Frozen ie_data.xls input supplied with the replication package. http://www.econ.yale.edu/~shiller/data.htm

Stambaugh, Robert, Jianfeng Yu, and Yu Yuan 2012. "The short of it: Investor sentiment and anomalies." *Journal of Financial Economics* 104, 288–302. https://doi.org/10.1016/j.jfineco.2011.12.001

Yu, Jianfeng, and Yu Yuan 2011. "Investor sentiment and the mean–variance relation." *Journal of Financial Economics* 100, 367–381. https://doi.org/10.1016/j.jfineco.2010.10.011

Internet Appendix to

A Conditional Shrinkage Method for Portfolio Selection

Nathan Lassance Alberto Martín-Utrera

Section A analytically characterizes the out-of-sample performance loss of the alpha portfolio due to estimation risk. Section B imposes a cap on portfolio turnover. Section C reports the empirical performance of our method using industry portfolios. Section D studies the relation between the shrinkage mean-variance portfolio and its 1/N component when the shrinkage intensity $\delta = 1$. Section E develops the theoretical foundations for applying SSH to the mean vector and portfolio weights. Section F uses alternative probability models to calibrate the SSH shrinkage intensity. Section G examines the sensitivity of SSH to the estimation-window split ratio. Section H assesses the robustness of SSH to estimating the logistic regression using a rolling window. Section I assesses the robustness of SSH to the construction of the sentiment index, considering a value-weighted market-return PLS index and the raw Baker and Wurgler (2006) sentiment index. Section J tests the robustness of our results to re-estimating the logistic model infrequently. Section K extends our model of the economy to include short-sale-constrained investors. Section L studies the link between model-implied and observable sentiment. Section M examines the performance of SSH when the logistic probability model is augmented with additional macroeconomic conditioning variables. Section N considers alternative calibrations of the risk-aversion coefficient. Section O replaces the sample covariance matrix with a conditional EWMA covariance estimator and applies either SSH or Ledoit and Wolf (2004) shrinkage to alleviate estimation error. Section P reports pairwise Ledoit and Wolf (2008) Sharpe-ratio tests between SSH and the benchmarks. Section Q constructs volatility-managed versions of SSH and the mean-variance benchmarks. Section R decomposes the turnover gap between SSH and each mean-variance benchmark into a benchmark-to-permuted-SSH component and a temporal-ordering component. Section S evaluates the impact of shorting costs and limits to capacity. Section T asks whether contractionary economic periods

explain the performance of our SSH portfolio. Section U tests the performance of a combination of the 1/N and alpha portfolios. Section V details the proofs of all theoretical results.

A. The Alpha Portfolio's Out-of-sample Loss

In this section, we characterize analytically the impact that estimation risk has on the out-of-sample Sharpe ratio of the alpha portfolio. Given a sample of size T, the sample vector of means and covariance matrix of asset returns at time t are defined as

$$\hat{\boldsymbol{\mu}}_t = \frac{1}{T}\sum_{m=t-T+1}^{t} \boldsymbol{r}_m \quad \text{and} \quad \widehat{\boldsymbol{\Sigma}}_t = \frac{1}{T}\sum_{m=t-T+1}^{t} (\boldsymbol{r}_m - \hat{\boldsymbol{\mu}}_t)(\boldsymbol{r}_m - \hat{\boldsymbol{\mu}}_t)^\top. \tag{A1}$$

Then, the in-sample estimate of the alpha portfolio is

$$\hat{\boldsymbol{w}}_\alpha = \frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\hat{\boldsymbol{\mu}}_t - \hat{\lambda}_t \boldsymbol{w}_{ew} \quad \text{with} \quad \hat{\lambda}_t = \frac{\boldsymbol{w}_{ew}^\top \hat{\boldsymbol{\mu}}_t}{\gamma \boldsymbol{w}_{ew}^\top \widehat{\boldsymbol{\Sigma}}_t \boldsymbol{w}_{ew}}. \tag{A2}$$

Moreover, we define the out-of-sample Sharpe ratio loss of $\hat{\boldsymbol{w}}_\alpha$ as

$$\mathrm{L}(\hat{\boldsymbol{w}}_\alpha) = \frac{\mathrm{SR}(\boldsymbol{w}_\alpha) - \mathrm{SR}(\hat{\boldsymbol{w}}_\alpha)}{\mathrm{SR}(\boldsymbol{w}_\alpha)} \geq 0, \tag{A3}$$

where the out-of-sample Sharpe ratio of $\hat{\boldsymbol{w}}_\alpha$ is defined as

$$\mathrm{SR}(\hat{\boldsymbol{w}}_\alpha) = \frac{\mathrm{E}\big[\hat{\boldsymbol{w}}_\alpha^\top \boldsymbol{r}_{t+1}\big]}{\sqrt{\mathrm{Var}\big[\hat{\boldsymbol{w}}_\alpha^\top \boldsymbol{r}_{t+1}\big]}}. \tag{A4}$$

The following proposition provides an analytical characterization of the alpha component's out-of-sample Sharpe ratio.

Proposition A.1 *Let the sample of asset returns* $(\boldsymbol{r}_{t-T+1}, \ldots, \boldsymbol{r}_{t+1})$ *be i.i.d. multivariate normally distributed. Then, the following holds:*

1. *For $T > N + 4$, the out-of-sample Sharpe ratio of $\hat{\boldsymbol{w}}_\alpha$, $\mathrm{SR}(\hat{\boldsymbol{w}}_\alpha)$ in (A4), is given by*

$$\frac{\mathrm{SR}^2(\boldsymbol{w}_\alpha) + \frac{N-1}{T-3}\mathrm{SR}^2(\boldsymbol{w}_{ew})}{\sqrt{a\mathrm{SR}^4(\boldsymbol{w}_\alpha) + b\mathrm{SR}^2(\boldsymbol{w}_\alpha)\mathrm{SR}^2(\boldsymbol{w}_{ew}) + c\mathrm{SR}^4(\boldsymbol{w}_{ew}) + d\mathrm{SR}^2(\boldsymbol{w}_\alpha) + e\mathrm{SR}^2(\boldsymbol{w}_{ew}) + f}}, \tag{A5}$$

   where $a, b, c, d, e, f \geq 0$ are functions of N and T given by Equations (A90)–(A95).

2. *In the high-dimensional asymptotic regime where $N \to \infty$, $T \to \infty$, and $N/T \to \rho \in [0, 1)$, we have*

$$\mathrm{SR}(\hat{\boldsymbol{w}}_\alpha) \to \frac{\sqrt{1-\rho}\big(\mathrm{SR}^2(\boldsymbol{w}_\alpha) + \rho\mathrm{SR}^2(\boldsymbol{w}_{ew})\big)}{\sqrt{\mathrm{SR}^2(\boldsymbol{w}_\alpha) + \rho(1+\rho-\rho^2)\mathrm{SR}^2(\boldsymbol{w}_{ew}) + \rho}}. \tag{A6}$$

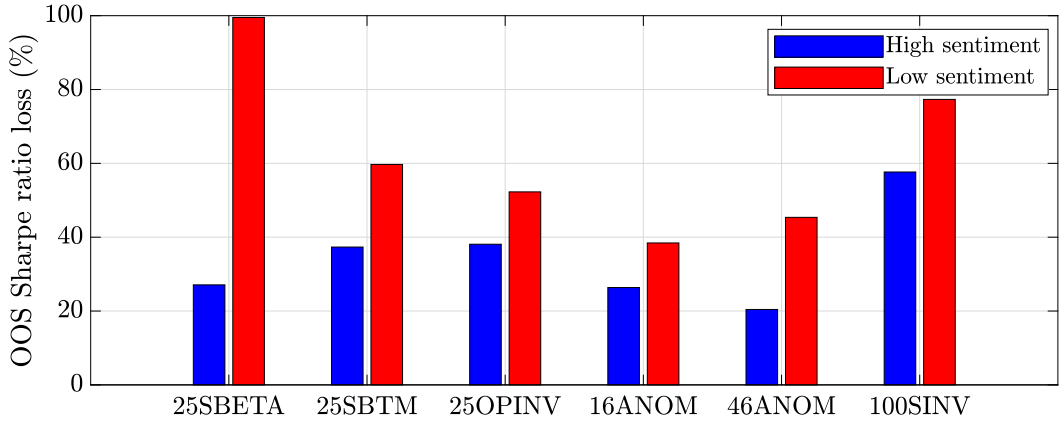
3. *In the high-dimensional asymptotic regime, the out-of-sample Sharpe ratio loss, $\mathrm{L}(\hat{\boldsymbol{w}}_\alpha)$ in (A3), decreases with $\mathrm{SR}(\boldsymbol{w}_\alpha)$ under two sufficient mild conditions: $\mathrm{SR}^2(\boldsymbol{w}_{ew}) < (1-\rho)/(1+2\rho^2-\rho^3)$ and $\mathrm{SR}^2(\boldsymbol{w}_\alpha) \geq \mathrm{SR}^2(\boldsymbol{w}_{ew})$.*

Proposition A.1 implies that the out-of-sample performance loss of the alpha portfolio decreases with the population Sharpe ratio of the alpha portfolio. We now assess the empirical validity of the results presented in this section. Figure A.1 depicts the out-of-sample Sharpe ratio loss of the alpha portfolio across sentiment regimes. In particular, for each dataset, we generate out-of-sample alpha portfolio returns using expanding windows with an initial sample size of $T = 120$,[1] and report the performance loss relative to the in-sample Sharpe ratio. We observed that, on average across datasets, the out-of-sample Sharpe ratio loss of the alpha portfolio is 62.1% following low-sentiment periods but only 34.5% following high-sentiment periods. This result is consistent with Proposition A.1, which shows that the out-of-sample loss is larger when the alpha portfolio's Sharpe ratio is lower, as is typically the case during low-sentiment periods. These results suggest

[1]The insights from this analysis are similar if we use rolling windows with a fixed sample size.

Figure A.1: Out-of-sample Sharpe ratio loss of the alpha portfolio

This figure depicts the out-of-sample Sharpe ratio loss for the alpha portfolio $\boldsymbol{w}_\alpha$ in (4), in percentage points, across the six datasets described in Section 2. For each dataset, we estimate the alpha portfolio using an expanding window with an initial size of $T = 120$ months, evaluate its out-of-sample return on the next month, and apply this method recursively to obtain a time series of out-of-sample alpha portfolio returns. We also compute the time series of in-sample alpha portfolio returns over the same time period. We then split these two time series in high and low-sentiment regimes using the Huang et al. (2015) sentiment index, and evaluate the out-of-sample Sharpe ratio loss in each regime. Like Barroso and Detzel (2021), we define high-sentiment regimes as those years for which the sentiment index at the end of the prior year is above the historical median. On average across datasets, the out-of-sample Sharpe ratio loss of the alpha portfolio is 62.1% following low-sentiment periods but only 34.5% following high-sentiment periods, indicating that estimation risk is more damaging when the alpha portfolio's population Sharpe ratio is lower.
Alt text: Paired bars show larger alpha-portfolio Sharpe ratio losses following low sentiment than high sentiment in all six datasets. The largest loss occurs for size and beta portfolios.



that it is particularly valuable to explicitly account for the impact of estimation risk in the mean-variance portfolio's alpha component when the number of assets is large relative to the sample size, but also when the alpha portfolio's Sharpe ratio is low.

B. Cap on Portfolio Turnover

In principle, one could construct low-turnover versions of any mean-variance portfolio strategy. It is therefore important to assess whether SSH's outperformance persists when all strategies are

subject to a common turnover constraint. Accordingly, we compare the out-of-sample Sharpe ratios of SSH and the benchmark strategies after imposing a common monthly turnover cap. Specifically, at the end of month t, let $w_{i,t+1}$ denote the desired weight on asset i for month $t+1$, and let w_{i,t^+} denote the pre-rebalancing weight at time t. The turnover required to implement the desired portfolio is

$$\text{TO}_{t+1} = \sum_{i=1}^{N} |w_{i,t+1} - w_{i,t^+}|. \tag{A7}$$

Let TOmax denote a constant monthly turnover cap. The adjusted portfolio $\tilde{\boldsymbol{w}}_{t+1}$ is then defined as

$$\tilde{\boldsymbol{w}}_{t+1} = \begin{cases} \boldsymbol{w}_{t+1}, & \text{if } \text{TO}_{t+1} \leq \text{TOmax}, \\ \boldsymbol{w}_{t^+} + \dfrac{\text{TOmax}}{\text{TO}_{t+1}}(\boldsymbol{w}_{t+1} - \boldsymbol{w}_{t^+}), & \text{if } \text{TO}_{t+1} > \text{TOmax}, \end{cases} \tag{A8}$$

so that realized turnover equals TOmax whenever the desired turnover exceeds the cap.

Table A.1 reports gross and net-of-cost out-of-sample Sharpe ratios, as well as turnover, for SSH and the seven benchmark strategies under a monthly cap of TOmax $= 15\%$. This turnover cap corresponds with a monthly reshuffling of about 7.5%, which is economically reasonable in light of observed trading activity in actively managed mutual funds. For example, van Binsbergen et al. (2024) report that the median fund in their sample exhibits annual portfolio reshuffling of approximately 150–160%, corresponding to roughly 12–13% per month.[2]

Table A.1 shows that SSH continues to deliver strong performance even when all strategies are subject to a common 15% monthly turnover cap. In the net-of-cost results reported in Panel B,

[2]Portfolio reshuffling in van Binsbergen et al. (2024) is defined as the minimum of aggregate purchases and sales over a period divided by total net assets, thereby avoiding double counting. By contrast, our turnover measure sums absolute changes in portfolio weights and counts both increases and decreases. For example, in a \$100 portfolio that decreases one position by \$7.5 and increases another by \$7.5 within a month, portfolio reshuffling equals 7.5%, whereas our turnover measure equals 15%. Thus, a 15% monthly turnover cap under our definition corresponds to roughly 7.5% reshuffling per month (about 90% annually), which is conservative relative to the median fund in van Binsbergen et al. (2024).

SSH attains Sharpe ratios of 0.82, 0.80, 0.86, 1.00, 0.59, and 0.95 for the 25SBETA, 25SBTM, 25OPINV, 16ANOM, 46ANOM, and 100SINV datasets, respectively. In each dataset, SSH delivers a higher net-of-cost Sharpe ratio than the mean-variance benchmark portfolios. This ranking persists under a common turnover cap, indicating that SSH's performance is not driven by differences in trading intensity.

Panel C shows that the turnover cap substantially compresses dispersion in trading activity across the mean-variance strategies. Nevertheless, equalizing turnover does not eliminate SSH's advantage. Even when MV-ID, MV-MKT, MV-NL, MV-PCA, and MAXSER operate at the same turnover level, their Sharpe ratios remain below that of SSH in most datasets. Overall, the evidence suggests that SSH's performance reflects differences in portfolio construction rather than simply lower turnover.

C. Performance on Industry Portfolios

One potential concern is that the economic gains of our sentiment-based shrinkage (SSH) methodology may arise solely in the context of characteristic-sorted portfolios, given the documented relation between sentiment and anomaly performance (Stambaugh, Yu, and Yuan 2012). While it is true that our approach allows mean-variance portfolios to take on greater exposure to the alpha component when sentiment is high, it also restricts its influence in the mean-variance portfolio when sentiment is low—shifting weight toward the 1/N portfolio to capture its higher premia. As documented in Section 4.2.3 of the manuscript, the economic gains generated by our shrinkage portfolio are primarily driven by this increased exposure to the 1/N component during low-sentiment periods. Therefore, we expect our SSH methodology to perform well even in alternative

Table A.1: Out-of-sample performance with monthly turnover capped at 15%

This table reports the annualized out-of-sample Sharpe ratio, and the turnover, of the proposed SSH portfolio and seven benchmarks when the monthly turnover is capped at 15% following the procedure outlined in Section B. The first four benchmarks are mean-variance portfolios constructed with the shrinkage covariance matrices of Ledoit and Wolf (2004) (MV-ID), Ledoit and Wolf (2003) (MV-MKT), Ledoit and Wolf (2020) (MV-NL), and Chen and Yuan (2016) (MV-PCA). The remaining ones are the MAXSER portfolio of Ao, Li, and Zheng (2019), the reward-to-risk timing strategy of Kirby and Ostdiek (2012) (RTR), and the 1/N portfolio. The table also reports in parentheses the two-sided p-values for the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The Sharpe ratios are either gross or net of proportional transaction costs of 30 basis points. The six datasets are described in Section 2. The numbers in bold font identify the portfolio strategies with the highest Sharpe ratio.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Gross Sharpe ratio						
SSH	**0.84**	**0.82**	**0.88**	**1.02**	0.59	**0.97**
	(0.08)	(0.12)	(0.03)	(0.00)	(0.06)	(0.00)
MV-ID	0.28	0.27	0.52	0.54	0.50	0.30
	(0.02)	(0.05)	(0.84)	(0.61)	(0.26)	(0.05)
MV-MKT	0.19	0.26	0.52	0.53	0.53	0.32
	(0.01)	(0.05)	(0.81)	(0.64)	(0.18)	(0.06)
MV-NL	0.24	0.26	0.53	0.53	0.49	0.33
	(0.02)	(0.05)	(0.87)	(0.65)	(0.28)	(0.06)
MV-PCA	0.36	0.36	0.76	0.53	0.55	0.72
	(0.18)	(0.15)	(0.24)	(0.66)	(0.15)	(0.41)
MAXSER	0.31	0.24	0.63	0.57	0.45	0.50
	(0.07)	(0.03)	(0.62)	(0.47)	(0.43)	(0.51)
RTR	0.64	0.61	0.62	0.59	**0.60**	0.64
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1/N	0.57	0.56	0.55	0.46	0.36	0.59
Panel B: Net-of-cost Sharpe ratio (30bps)						
SSH	**0.82**	**0.80**	**0.86**	**1.00**	**0.59**	**0.95**
	(0.10)	(0.12)	(0.03)	(0.04)	(0.00)	(0.01)
MV-ID	0.28	0.27	0.52	0.54	0.50	0.30
	(0.03)	(0.06)	(0.91)	(0.57)	(0.21)	(0.06)
MV-MKT	0.19	0.26	0.52	0.53	0.53	0.32
	(0.01)	(0.06)	(0.86)	(0.62)	(0.16)	(0.07)
MV-NL	0.24	0.26	0.53	0.53	0.49	0.33
	(0.01)	(0.05)	(0.92)	(0.60)	(0.24)	(0.08)
MV-PCA	0.36	0.36	0.73	0.53	0.55	0.71
	(0.18)	(0.17)	(0.27)	(0.62)	(0.15)	(0.42)
MAXSER	0.31	0.24	0.62	0.57	0.45	0.50
	(0.09)	(0.04)	(0.61)	(0.41)	(0.38)	(0.56)
RTR	0.63	0.60	0.61	0.58	**0.59**	0.63
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
1/N	0.57	0.56	0.54	0.45	0.35	0.58
Panel C: Turnover						
SSH	0.13	0.14	0.14	0.14	0.15	0.15
MV-ID	0.15	0.15	0.15	0.15	0.15	0.15
MV-MKT	0.15	0.15	0.15	0.15	0.15	0.15
MV-NL	0.15	0.15	0.15	0.15	0.15	0.15
MV-PCA	0.15	0.15	0.14	0.15	0.15	0.15
MAXSER	0.15	0.15	0.15	0.15	0.15	0.15
RTR	0.04	0.04	0.04	0.05	0.04	0.04
1/N	0.04	0.04	0.04	0.04	0.04	0.04

settings where test assets are not constructed based on anomaly characteristics. To test this, we now evaluate the performance of our approach using industry portfolios.

Table A.2 reports the annualized out-of-sample Sharpe ratios without transaction costs (Panel A) and with transaction costs of 30 basis points (Panel B) of the proposed SSH portfolio and seven benchmarks for two datasets of 10 industry portfolios (10IND) and 49 industry portfolios (49IND), spanning July 1965 to June 2022 like our other datasets. Across both datasets, the SSH portfolio consistently outperforms all five mean-variance benchmark portfolios. Notably, for the 49IND dataset, the improvement in performance is statistically significant relative to the MV-ID portfolio, which relies on the shrinkage estimator of the covariance matrix toward the identity of Ledoit and Wolf (2004), and thus, represents the benchmark methodologically closest to ours. Moreover, the SSH portfolio performs similarly to 1/N, which is a notoriously difficult benchmark among industry portfolios. These findings support the use of sentiment in the construction of shrinkage covariance matrices for portfolio selection, even in settings where sentiment does not have a direct or well-established relation to the performance of the underlying test assets.

D. Relation to the 1/N Component when $\delta = 1$

When the shrinkage intensity $\delta = 1$ in the shrinkage covariance matrix $\boldsymbol{\Sigma}_s = (1-\delta)\boldsymbol{\Sigma} + \delta\bar{\sigma}^2\boldsymbol{I}$, the shrinkage mean-variance portfolio becomes

$$\boldsymbol{w}_{smv} = \frac{1}{\gamma}\boldsymbol{\Sigma}_s^{-1}\boldsymbol{\mu} = \frac{1}{\gamma\bar{\sigma}^2}\boldsymbol{\mu}, \tag{A9}$$

which is proportional to $\boldsymbol{\mu}$. This result is consistent with MacKinlay and Pástor (2000, p.889), who show that imposing a one-factor structure and a constant diagonal covariance matrix for the error term "implies that the tangency portfolio weights are proportional to the vector of mean returns." In

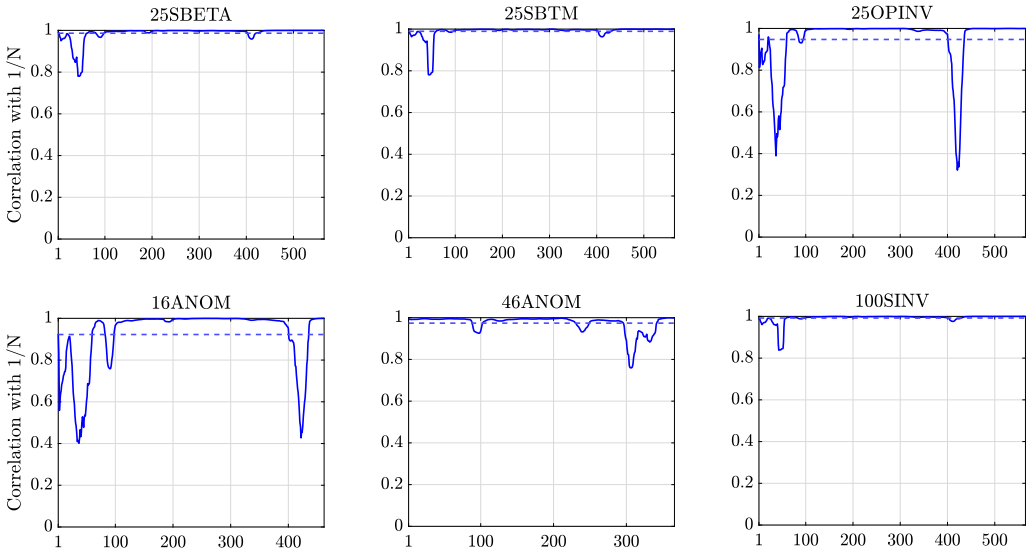
Table A.2: Out-of-sample Sharpe ratios with industry portfolios

This table reports the annualized out-of-sample Sharpe ratios without transaction costs (Panel A) and with transaction costs of 30 basis points (Panel B) of the proposed SSH portfolio and seven benchmarks for two datasets of 10 industry portfolios (10IND) and 49 industry portfolios (49IND), spanning July 1965 to June 2022. The first four benchmarks are mean-variance portfolios constructed with the shrinkage covariance matrices of Ledoit and Wolf (2004) (MV-ID), Ledoit and Wolf (2003) (MV-MKT), Ledoit and Wolf (2020) (MV-NL), and Chen and Yuan (2016) (MV-PCA). The remaining ones are the MAXSER portfolio by Ao, Li, and Zheng (2019), the reward-to-risk timing strategy of Kirby and Ostdiek (2012) (RTR), and the 1/N portfolio. The table also reports two-sided p-values of the significance test of the difference of each portfolio's Sharpe ratio with that of MV-ID. Both datasets of industry portfolios are downloaded from Kenneth French's website.

Strategy	10IND	49IND
Panel A: Gross Sharpe ratio		
SSH	0.85	0.79
	(0.18)	(0.00)
MV-ID	0.78	0.40
MV-MKT	0.77	0.49
	(0.49)	(0.02)
MV-NL	0.77	0.41
	(0.36)	(0.40)
MV-PCA	0.77	0.63
	(0.43)	(0.16)
MAXSER	0.80	0.64
	(0.66)	(0.05)
RTR	0.89	0.82
	(0.36)	(0.02)
1/N	0.87	0.78
	(0.48)	(0.04)
Panel B: Net-of-cost Sharpe ratio (30 bps)		
SSH	0.82	0.74
	(0.10)	(0.00)
MV-ID	0.73	0.27
MV-MKT	0.72	0.37
	(0.26)	(0.01)
MV-NL	0.72	0.29
	(0.15)	(0.21)
MV-PCA	0.71	0.59
	(0.20)	(0.04)
MAXSER	0.77	0.55
	(0.52)	(0.01)
RTR	0.88	0.81
	(0.24)	(0.00)
1/N	0.86	0.77
	(0.31)	(0.01)

Figure A.2: Correlation with the 1/N component when $\delta = 1$

This figure depicts the return correlation between the shrinkage mean-variance portfolio with a shrinkage intensity $\delta = 1$ ($\boldsymbol{w}_{smv}$ in (A9)) and the 1/N component ($\boldsymbol{w}_{ew}$ in (3)). The six datasets are described in Section 2. The correlations are obtained using rolling windows of $T = 120$ months and we depict their moving average over 12 months.
Alt text: Six panels show correlations between fully shrunk and equal-weighted portfolio returns. Correlations remain close to one for most windows, interrupted by occasional sharp downward spikes.



the next proposition, we study how the portfolio in (A9) relates to the 1/N portfolio, $\boldsymbol{w}_{ew} = \mathbf{1}/N$. In particular, we demonstrate an equivalence between $\boldsymbol{w}_{smv}$ with $\delta = 1$ and $\boldsymbol{w}_{ew}$ when there is no cross-sectional variation in average returns, i.e., when $\boldsymbol{\mu} = \bar{\mu}\mathbf{1}$ with $\bar{\mu} = \frac{1}{N}\sum_{i=1}^{N}\mu_i$. For that purpose, we define

$$\boldsymbol{\eta} = \boldsymbol{\mu} - \bar{\mu}\mathbf{1}. \tag{A10}$$

Proposition A.2 *When the shrinkage intensity $\delta = 1$, the following relations hold between the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ and the 1/N portfolio $\boldsymbol{w}_{ew}$:*

1. *The squared correlation between their returns is*

$$\text{Corr}^2(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}) = 1 - \boldsymbol{x}^\top \boldsymbol{\Sigma} \boldsymbol{\eta}, \tag{A11}$$

where

$$\boldsymbol{x} = \frac{(\mathbf{1}^\top \boldsymbol{\Sigma} \mathbf{1})\boldsymbol{\eta} - (\mathbf{1}^\top \boldsymbol{\Sigma} \boldsymbol{\eta})\mathbf{1}}{(\mathbf{1}^\top \boldsymbol{\Sigma} \mathbf{1})(\boldsymbol{\mu}^\top \boldsymbol{\Sigma} \boldsymbol{\mu})}. \tag{A12}$$

In particular, the correlation is equal to one when $\boldsymbol{\eta} = \mathbf{0}$.

2. *The squared Euclidean norm of their weight difference is*

$$\left|\left| \boldsymbol{w}_{ew} - \frac{\gamma \bar{\sigma}^2}{N \bar{\mu}} \boldsymbol{w}_{smv} \right|\right|^2 = \frac{\bar{\eta}_2 + 2\bar{\mu}\bar{\eta}_1}{N \bar{\mu}^2}, \tag{A13}$$

where $\bar{\eta}_k = \frac{1}{N} \sum_{i=1}^N \eta_i^k$. *In particular, this norm is equal to zero when* $\boldsymbol{\eta} = \mathbf{0}$.

In Figure A.2, we depict, across the six datasets described in Section 2, the return correlation between the shrinkage mean-variance portfolio with a shrinkage intensity $\delta = 1$, $\boldsymbol{w}_{smv}$ in (A9), and its 1/N component $\boldsymbol{w}_{ew}$, which we obtain from (A11). We compute the correlations over time using rolling windows of $T = 120$ months and we depict their moving averages over 12 months for visibility. The figure shows that the correlation is close to one on average for all datasets and remains close to one in most time periods.

E. SSH for the Mean Vector and Portfolio Weights

In the main body of the manuscript, our proposed sentiment-based SSH methodology shrinks the covariance matrix to control the relative importance of 1/N and the alpha component in the mean-

variance portfolio. In this section, we show how to adapt the SSH methodology to shrinking the mean vector and mean-variance portfolio weights in Sections E.1 and E.2, respectively.

E.1 SSH for the Mean Vector

In this section, we study a sentiment-based shrinkage estimator where we shrink the mean vector $\boldsymbol{\mu}$ toward the cross-sectional average of mean returns:

$$\boldsymbol{\mu}_s = (1-\delta)\boldsymbol{\mu} + \delta\mu_{ew}\mathbf{1}. \tag{A14}$$

The resulting shrinkage mean-variance portfolio is given by

$$\boldsymbol{w}_{smv} = \frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_s = (1-\delta)\boldsymbol{w}_{mv} + \frac{\delta\mu_{ew}}{\gamma}\boldsymbol{\Sigma}^{-1}\mathbf{1}. \tag{A15}$$

To apply our sentiment-based shrinkage methodology, we derive in Proposition A.3 the squared correlation between the return of the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ and that of its 1/N component $\boldsymbol{w}_{ew}$ and the shrinkage alpha component $\boldsymbol{w}_{s\alpha} = \boldsymbol{w}_{smv} - \lambda\boldsymbol{w}_{ew}$.

Proposition A.3 *Let $R_{\sigma^2} = \sigma_g^2/\sigma_{ew}^2 \leq 1$ be the ratio between the variance of the GMV and 1/N portfolios, and $R_\mu = \mu_g/\mu_{ew}$ the ratio between their mean return. Then, the squared correlations between the return of the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ in (A15) and that of its 1/N component $\boldsymbol{w}_{ew}$ and shrinkage alpha component $\boldsymbol{w}_{s\alpha} = \boldsymbol{w}_{smv} - \lambda\boldsymbol{w}_{ew}$ are given by*

$$\text{Corr}^2(\boldsymbol{w}_{smv}^\top\boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top\boldsymbol{r}_{t+1}) = \frac{\text{SR}^2(\boldsymbol{w}_{ew})}{\text{SR}^2(\boldsymbol{w}_{ew})(1+\eta(\delta)) + (1-\delta)^2\text{SR}^2(\boldsymbol{w}_\alpha)}, \tag{A16}$$

$$\text{Corr}^2\big(\boldsymbol{w}_{smv}^\top\boldsymbol{r}_{t+1}, \boldsymbol{w}_{s\alpha}^\top\boldsymbol{r}_{t+1}\big) = 1 - \text{Corr}^2\big(\boldsymbol{w}_{smv}^\top\boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top\boldsymbol{r}_{t+1}\big), \tag{A17}$$

where $\eta(\delta) = [(1+R_{\sigma^2}-2R_\mu)\delta^2+2(R_\mu-R_{\sigma^2})\delta]/R_{\sigma^2}$. *When the first eigenvector* $\boldsymbol{v}_1 = \mathbf{1}/\sqrt{N}$, *we have* $\eta(\delta) = 0$*; in this case, a higher* δ *increases the squared correlation with the 1/N component and decreases that with the alpha component.*

Proposition A.3 shows that shrinking the mean vector also allows us to control the mean-variance portfolio exposure to its 1/N and alpha components. Using this proposition, we can find the SSH shrinkage intensity following the methodology in Section 3.3. We first use the predicted probability p_t that 1/N outperforms the alpha portfolio out of sample to interpolate the squared correlation between the shrinkage mean-variance portfolio return and that of 1/N:

$$\text{Corr}^2(p_t) = (1-p_t)\,\text{Corr}_0^2 + p_t\,\text{Corr}_1^2 = (1-p_t)\frac{\text{SR}^2(\boldsymbol{w}_{ew})}{\text{SR}^2(\boldsymbol{w}_{mv})} + p_t R_{\sigma^2}. \tag{A18}$$

Then, we determine the sentiment-based shrinkage intensity δ_{ssh} such that the squared correlation with 1/N in (A16) matches the target squared correlation $\text{Corr}^2(p_t)$ in (A18).

The SSH portfolio at time t is then given by

$$\hat{\boldsymbol{w}}_{ssh,t} = \frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\hat{\boldsymbol{\mu}}_{s,t}, \tag{A19}$$

where $\hat{\boldsymbol{\mu}}_{s,t} = (1-\hat{\delta}_{ssh,t})\hat{\boldsymbol{\mu}}_t + \hat{\delta}_{ssh,t}\hat{\mu}_{ew,t}\mathbf{1}$, $\hat{\mu}_{ew,t} = \boldsymbol{w}_{ew}^\top\hat{\boldsymbol{\mu}}_t$, $\hat{\boldsymbol{\mu}}_t$ is the sample mean, $\widehat{\boldsymbol{\Sigma}}_t$ is the sample covariance matrix, and $\hat{\delta}_{ssh,t}$ is the sentiment-based shrinkage intensity for the vector of means.

Section 4.2.4 of the main manuscript reports the empirical performance of mean-SSH together with the corresponding covariance- and weight-shrinkage implementations.

E.2 SSH for Portfolio Weights

In this section, we implement an alternative specification in which sentiment-based shrinkage is applied directly to the vector of portfolio weights rather than through the covariance matrix. Specif-

ically, we define the shrinkage mean-variance portfolio as

$$\boldsymbol{w}_{smv} = (1-\delta)\boldsymbol{w}_{mv} + \delta\lambda\boldsymbol{w}_{ew} = \lambda\boldsymbol{w}_{ew} + (1-\delta)\boldsymbol{w}_{\alpha}, \tag{A20}$$

where $\lambda = \mu_{ew}/(\gamma\sigma_{ew}^2)$. To extend the SSH methodology to portfolio weights, we introduce Proposition A.4, which derives the squared correlation between the return of the shrinkage mean-variance portfolio in (A20) and those of its 1/N and alpha components.

Proposition A.4 *The squared correlations between the return of the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ in (A20) and that of its 1/N component $\boldsymbol{w}_{ew}$ and alpha component $\boldsymbol{w}_{\alpha}$ are given by*

$$\text{Corr}^2(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}) = \frac{\text{SR}^2(\boldsymbol{w}_{ew})}{\text{SR}^2(\boldsymbol{w}_{ew}) + (1-\delta)^2\text{SR}^2(\boldsymbol{w}_{\alpha})}, \tag{A21}$$

$$\text{Corr}^2\big(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{s\alpha}^\top \boldsymbol{r}_{t+1}\big) = 1 - \text{Corr}^2\big(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}\big). \tag{A22}$$

Moreover, a higher δ increases the squared correlation with the 1/N component and decreases that with the alpha component.

Using this proposition, we derive the shrinkage intensity applied to the vector of portfolio weights following the methodology in Section 3.3. We first use the time-t predicted probability p_t that the 1/N portfolio outperforms the alpha portfolio out of sample to interpolate the squared correlation between the shrinkage mean-variance portfolio return and that of 1/N:

$$\text{Corr}^2(p_t) = (1-p_t)\,\text{Corr}_0^2 + p_t\,\text{Corr}_1^2 = (1-p_t)\frac{\text{SR}^2(\boldsymbol{w}_{ew})}{\text{SR}^2(\boldsymbol{w}_{mv})} + p_t. \tag{A23}$$

We then determine the shrinkage intensity $\delta_{ssh,t}$ such that the squared correlation with 1/N in (A21) equals the target squared correlation $\text{Corr}^2(p_t)$ in (A23). This yields the following closed-

form expression:

$$\delta_{ssh,t} = 1 - \sqrt{\frac{(1-p_t)\text{SR}^2(\boldsymbol{w}_{ew})}{\text{SR}^2(\boldsymbol{w}_{ew}) + p_t\text{SR}^2(\boldsymbol{w}_{\alpha})}}. \tag{A24}$$

Unlike the covariance-shrinkage implementation used in the main body of the manuscript, the shrinkage intensity applied directly to the vector of portfolio weights admits a closed-form expression for $\delta_{ssh,t}$. Consistent with intuition, $\delta_{ssh,t} = 1$ when $p_t = 1$ and $\delta_{ssh,t} = 0$ when $p_t = 0$. Accordingly, the sentiment-based shrinkage portfolio at time t is given by

$$\hat{\boldsymbol{w}}_{ssh,t} = (1 - \hat{\delta}_{ssh,t})\frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\hat{\boldsymbol{\mu}}_t + \hat{\delta}_{ssh,t}\hat{\lambda}_t\boldsymbol{w}_{ew}, \tag{A25}$$

where

$$\hat{\lambda}_t = \frac{\boldsymbol{w}_{ew}^{\top}\hat{\boldsymbol{\mu}}_t}{\gamma\boldsymbol{w}_{ew}^{\top}\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}},$$

$\hat{\boldsymbol{\mu}}_t$ denotes the sample mean vector, $\widehat{\boldsymbol{\Sigma}}_t$ is the sample covariance matrix, and $\hat{\delta}_{ssh,t}$ is the sentiment-based shrinkage intensity for the portfolio weights.

Section 4.2.4 of the main manuscript reports the empirical performance of weight-SSH together with the corresponding covariance- and mean-shrinkage implementations.

F. Alternative Probability Models to Calibrate the SSH Shrinkage Intensity

In the main body of the manuscript, we make a specific choice on how to estimate the probability p_t that the 1/N portfolio outperforms the alpha portfolio out of sample. Now, we conduct extensive analyses to examine the robustness of our results to alternative model specifications. In particular, we estimate the probability p_t using three alternative probability models:

1. A second-order polynomial in sentiment, defined as sign(Sent)$\times$Sent^2: using the same logistic specification as in the main body of the manuscript, we replace the linear sentiment index with this transformed measure, thereby allowing for nonlinear effects of investor sentiment on asset returns.

2. Alternative functional form (Probit): instead of using the logistic function in the main body of the manuscript, we now use a probit function to model the probability of 1/N outperforming the alpha portfolio.

3. Alternative sentiment measure (PCA sent.): instead of the partial least squares method of Huang et al. (2015) used in the main body of the manuscript, we test the robustness of our approach to considering the PCA method of Baker and Wurgler (2006), who extract the latent sentiment index using the first principal component of the sentiment proxies: (1) the closed-end fund discount, (2) the number of IPOs, (3) the first-day returns on IPOs, (4) the equity share in new issues, and (5) the dividend premium.

Table A.3 reports the annualized out-of-sample Sharpe ratios for each specification. The results demonstrate that the SSH portfolio's performance is robust across all three alternative models. Specifically, the Sent^2 specification achieves Sharpe ratios ranging from 0.82 to 1.90 across the six datasets, the Probit specification achieves Sharpe ratios ranging from 0.83 to 2.03, and the PCA sentiment specification achieves Sharpe ratios ranging from 0.84 to 2.12. These values are comparable to those for our baseline model specification in the main body of the manuscript, where SSH achieves Sharpe ratios ranging from 0.83 to 2.04.

Table A.4 reports the average monthly turnover for each specification. The turnover levels are comparable across specifications, with the Probit and PCA sentiment models delivering slightly lower turnover than the Sent^2 specification in most datasets. These turnovers also compare favor-

Table A.3: Out-of-sample Sharpe ratios for different SSH models

This table reports the annualized out-of-sample Sharpe ratio of the proposed SSH portfolio, where the shrinkage intensity is estimated under three alternative probability models to estimate the probability p_t that the 1/N portfolio outperforms the alpha portfolio used in the construction of the shrinkage covariance matrix in Equation (20). In particular, we estimate p_t using the square of our partial least squares sentiment index while keeping its sign, a probit function instead of a logit function that models the probability of 1/N outperforming the alpha portfolio, and using the first principal component of the sentiment proxies as in Baker and Wurgler (2006) to extract the latent sentiment index instead of the partial least squares method we use in the main body of the manuscript following Huang et al. (2015). The table also reports two-sided p-values for the significance test of the difference between each portfolio's Sharpe ratio with that of 1/N. The six datasets are described in Section 2.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
sign(Sent)×Sent^2	0.82	0.89	0.93	1.10	1.90	1.06
	(0.09)	(0.04)	(0.01)	(0.00)	(0.00)	(0.00)
Probit	0.83	0.91	0.94	1.10	2.03	1.06
	(0.08)	(0.03)	(0.01)	(0.00)	(0.00)	(0.00)
PCA sent.	0.84	0.92	0.95	1.10	2.12	1.05
	(0.08)	(0.03)	(0.01)	(0.00)	(0.00)	(0.00)

Table A.4: Turnover for different SSH models

This table reports the average monthly turnover of the proposed SSH portfolio, where the shrinkage intensity is estimated under three different models to estimate the probability p_t that the 1/N portfolio outperforms the alpha portfolio used in the construction of the shrinkage covariance matrix in Equation (20). In particular, we estimate p_t using the square of our partial least squares sentiment index while keeping its sign, a probit function instead of a logit function that models the probability of 1/N outperforming the alpha portfolio, and using the first principal component of the sentiment proxies as in Baker and Wurgler (2006) to extract the latent sentiment index instead of the partial least squares method we use in the main body of the manuscript following Huang et al. (2015). The six datasets are described in Section 2.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
sign(Sent)×Sent^2	0.28	0.47	0.37	0.41	1.86	0.51
Probit	0.26	0.42	0.34	0.38	1.69	0.52
PCA sent.	0.25	0.39	0.33	0.36	1.59	0.53

Table A.5: Out-of-sample Sharpe ratios net of transaction costs for different SSH models

This table reports the annualized out-of-sample Sharpe ratio net of transaction costs of 30 basis points of the proposed SSH portfolio, where the shrinkage intensity is estimated under three different models to estimate the probability p_t that the 1/N portfolio outperforms the alpha portfolio used in the construction of the shrinkage covariance matrix in Equation (20). In particular, we estimate p_t using the square of our partial least squares sentiment index while keeping its sign, a probit function instead of a logit function that models the probability of 1/N outperforming the alpha portfolio, and using the first principal component of the sentiment proxies as in Baker and Wurgler (2006) to extract the latent sentiment index instead of the partial least squares method we use in the main body of the manuscript following Huang et al. (2015). The table also reports two-sided p-values for the significance test of the difference between each portfolio's Sharpe ratio with that of 1/N. The six datasets are described in Section 2.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
sign(Sent)×Sent^2	0.77	0.82	0.86	1.03	1.77	0.97
	(0.17)	(0.10)	(0.03)	(0.00)	(0.00)	(0.01)
Probit	0.78	0.85	0.88	1.04	1.90	0.97
	(0.14)	(0.07)	(0.02)	(0.00)	(0.00)	(0.00)
PCA sent.	0.79	0.85	0.89	1.04	2.00	0.96
	(0.12)	(0.07)	(0.02)	(0.00)	(0.00)	(0.00)

ably relative to those required by the mean-variance portfolio benchmarks in Table 4. As a result of this lower turnover, the strong performance of SSH reported in Table A.3 under these three alternative model specifications persists after accounting for transaction costs of 30 basis points, as we see in Table A.5.

Together, these results confirm that our SSH methodology is robust to alternative functional forms and sentiment measures. The choice between logit and probit has minimal impact on performance, and using the Baker and Wurgler (2006) PCA-based index yields similar results to our baseline PLS-based sentiment index.

Table A.6: Sensitivity of SSH to the estimation-window split ratio

This table reports the annualized out-of-sample gross Sharpe ratio (Panel A), turnover (Panel B), and Sharpe ratio net of 30-bps transaction costs (Panel C) of the SSH portfolio for estimation-window split ratios of 50/50 (baseline), 60/40, and 40/60, where the first figure is the fraction of the window used to estimate the mean-variance components and the second the fraction used for the logistic regression. All ratios are evaluated over identical out-of-sample months. The six datasets are described in Section 2.

	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Gross Sharpe ratio						
50/50 (baseline)	0.83	0.91	0.94	1.10	2.04	1.08
60/40	0.87	0.92	0.92	1.13	1.83	1.09
40/60	0.83	0.89	0.92	1.08	1.90	1.09
Panel B: Turnover						
50/50 (baseline)	0.26	0.41	0.34	0.38	1.66	0.46
60/40	0.26	0.43	0.35	0.42	3.81	0.55
40/60	0.25	0.37	0.30	0.37	1.65	0.51
Panel C: Net-of-cost Sharpe ratio						
50/50 (baseline)	0.78	0.85	0.88	1.04	1.91	0.99
60/40	0.82	0.86	0.85	1.06	1.64	1.00
40/60	0.78	0.84	0.86	1.02	1.77	0.99

G. Sensitivity to the Estimation-window Split Ratio

We now assess the robustness of the baseline 50/50 split in the estimation window (Section 3.3) to construct the SSH portfolio to 60/40 and 40/60 split alternatives, where the first figure is the fraction of the window used to estimate the mean-variance components and the second the fraction used for the logistic regression. All other choices are unchanged, and all three ratios are evaluated over identical out-of-sample months. Table A.6 shows that the gross and net-of-cost Sharpe ratios under the alternative splits are close to the baseline, and the main performance conclusions are unchanged. The largest sensitivity arises in 46ANOM, where the 60/40 split raises turnover and lowers the net-of-cost Sharpe ratio from 1.91 to 1.64.

H. Rolling-window Estimation of Logistic Regression

In the main body of the manuscript, we use an expanding window to estimate the logistic regression that produces p_t, the probability that the scaled $1/N$ portfolio outperforms the alpha portfolio out of sample conditional on investor sentiment; see Figure 3. The expanding-window specification provides a larger sample for estimating the logistic-regression coefficients, β_0 and β_1, and helps reduce estimation error. However, it may respond more slowly to changes in the relation between sentiment and subsequent portfolio performance than a rolling-window specification. In this section, we evaluate how estimating the logistic regression with rolling windows instead of expanding windows affects the out-of-sample performance and turnover of the SSH portfolio.

Table A.7 reports results for rolling windows of 60 and 120 months. Panel A shows that rolling-window estimation increases the gross out-of-sample Sharpe ratio in some datasets and reduces it in others, with no systematic improvement over the baseline expanding-window specification. Panel C shows that both rolling-window specifications increase portfolio turnover in every dataset, consistent with the greater sensitivity of the estimated logistic-regression coefficients to recent observations. Consequently, after accounting for transaction costs, Panel B shows that the expanding-window specification delivers higher net-of-cost Sharpe ratios than either rolling-window alternative across all six datasets. Overall, these findings support the use of expanding windows in the implementation of SSH.

I. Robustness to the Sentiment Index

This section assesses the robustness of SSH to the construction of the sentiment index.

Table A.7: Out-of-sample performance of SSH with rolling-window estimation of logistic regression

This table reports the annualized out-of-sample Sharpe ratio and turnover of SSH when the logistic regression is estimated using an expanding window or rolling windows of 60 and 120 months. The table also reports in parentheses the two-sided p-values for the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The six datasets are described in Section 2. The numbers in bold font identify the portfolio strategy with the highest Sharpe ratio.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Gross Sharpe ratio						
SSH (expanding)	**0.83**	0.91	**0.94**	1.10	**2.04**	1.08
	(0.08)	(0.03)	(0.01)	(0.00)	(0.00)	(0.00)
SSH (rolling, 60 months)	**0.83**	0.71	0.83	0.97	1.72	1.03
	(0.09)	(0.46)	(0.07)	(0.00)	(0.00)	(0.00)
SSH (rolling, 120 months)	0.82	**0.92**	0.88	**1.12**	1.68	**1.09**
	(0.09)	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)
Panel B: Net-of-cost Sharpe ratio (30bps)						
SSH (expanding)	**0.78**	**0.85**	**0.88**	**1.04**	**1.91**	**0.99**
	(0.14)	(0.07)	(0.02)	(0.00)	(0.00)	(0.00)
SSH (rolling, 60 months)	0.56	0.38	0.60	0.77	1.08	0.57
	(0.97)	(0.40)	(0.72)	(0.05)	(0.00)	(0.96)
SSH (rolling, 120 months)	0.71	0.80	0.77	1.02	1.43	0.81
	(0.32)	(0.09)	(0.10)	(0.00)	(0.00)	(0.08)
Panel C: Turnover						
SSH (expanding)	0.26	0.41	0.34	0.38	1.66	0.46
SSH (rolling, 60 months)	1.38	2.44	1.32	1.27	5.04	2.71
SSH (rolling, 120 months)	0.55	0.71	0.58	0.66	3.17	1.52

I.1 Using the Value-weighted Market Return in the PLS Step

Our baseline PLS sentiment index uses the 1/N return as the target return. We now follow the Huang et al. (2015) approach and use the value-weighted market return. We therefore reconstruct the index at each rebalancing date using the value-weighted market return in excess of the risk-free rate and downloaded from Kenneth French's website, while leaving the rest of the SSH method unchanged. Table A.8 shows that performance is generally similar across the two constructions;

Table A.8: SSH with a value-weighted market-return PLS sentiment index

This table reports the annualized out-of-sample Sharpe ratio, turnover, and Sharpe ratio net of 30-bps transaction costs of the SSH portfolio when the PLS sentiment index targets the 1/N return (Panel A, baseline) or the Fama–French value-weighted market excess return (Panel B). The six datasets are described in Section 2.

	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Baseline SSH (1/N-return PLS)						
Sharpe ratio	0.83	0.91	0.94	1.10	2.04	1.08
Turnover	0.26	0.41	0.34	0.38	1.66	0.46
Net-of-cost Sharpe ratio	0.78	0.85	0.88	1.04	1.91	0.99
Panel B: SSH with value-weighted market-return PLS						
Sharpe ratio	0.82	0.92	0.94	1.10	1.80	1.06
Turnover	0.27	0.42	0.34	0.38	2.22	0.52
Net-of-cost Sharpe ratio	0.77	0.85	0.87	1.04	1.67	0.97

the main exception is 46ANOM, where the market-return index yields a slightly lower gross Sharpe ratio (1.80 vs 2.04).

I.2 Using the Raw Baker–Wurgler Index

We next replace the PLS sentiment index with the raw Baker and Wurgler (2006) index obtained from Jeffrey Wurgler's website, keeping SSH otherwise unchanged. Table A.9 shows that the resulting performance is similar to the baseline across datasets.

J. Infrequent Re-estimation of the Logistic Model

We now test the robustness of our main results to infrequent re-estimation of the logistic model. In particular, we re-estimate the logistic model only once every five years, rather than at each monthly rebalancing date. This specification closely mimics an investor who learns the mapping from sentiment to shrinkage intensity infrequently and then applies it prospectively over extended

Table A.9: SSH with the raw Baker–Wurgler sentiment index

This table reports the annualized out-of-sample Sharpe ratio, turnover, and Sharpe ratio net of 30-bps transaction costs of the SSH portfolio when the conditioning variable is the baseline PLS sentiment index (Panel A) or the raw Baker and Wurgler (2006) index (Panel B), obtained from Jeffrey Wurgler's website. The raw index spans July 1965 to December 2023, covering the out-of-sample period of all six datasets described in Section 2, so both panels are evaluated over the same months.

	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Baseline SSH (PLS sentiment index)						
Sharpe ratio	0.83	0.91	0.94	1.10	2.04	1.08
Turnover	0.26	0.41	0.34	0.38	1.66	0.46
Net-of-cost Sharpe ratio	0.78	0.85	0.88	1.04	1.91	0.99
Panel B: SSH with raw Baker–Wurgler sentiment						
Sharpe ratio	0.84	0.88	0.95	1.11	2.09	1.04
Turnover	0.27	0.44	0.33	0.38	1.64	0.57
Net-of-cost Sharpe ratio	0.79	0.82	0.89	1.05	1.97	0.94

periods. By substantially reducing the re-estimation frequency, this approach limits the scope for overfitting to recent episodes and alleviates concerns associated with learning from overlapping return windows.

Tables A.10, A.11, and A.12 report the results for this robustness test. Comparing these results with the baseline results in Tables 3, 4, and Panel B of Table 5, we find that the performance of the SSH portfolio is largely unchanged under infrequent re-estimation.

In terms of gross Sharpe ratios, Table A.10 shows that the five-year re-estimation scheme yields values of 0.82, 0.90, 0.91, 0.96, 2.16, and 1.05 for the 25SBETA, 25SBTM, 25OPINV, 16ANOM, 46ANOM, and 100SINV datasets, respectively—very close to the corresponding baseline values of 0.83, 0.91, 0.94, 1.10, 2.04, and 1.08 reported in Table 3. The SSH portfolio continues to outperform the benchmark strategies in most datasets. Turnover, moreover, decreases notably: Table A.11 reports SSH turnover of 0.22, 0.31, 0.27, 0.32, 0.80, and 0.32 across the six datasets, compared with baseline values of 0.26, 0.41, 0.34, 0.38, 1.66, and 0.46 in Table 4. The reduction is

particularly striking for the 46ANOM dataset, where turnover falls from 1.66 to 0.80. As a result, net-of-cost performance is largely preserved. After accounting for proportional transaction costs of 30 basis points, Table A.12 shows that the SSH portfolio delivers net Sharpe ratios of 0.78, 0.85, 0.86, 0.91, 2.09, and 0.99—values that are very close to, and in some cases exceed, the baseline net Sharpe ratios of 0.78, 0.85, 0.88, 1.04, 1.91, and 0.99 reported in Panel B of Table 5.

Importantly, under the five-year re-estimation scheme, the SSH portfolio continues to significantly outperform the 1/N benchmark, as indicated by the reported p-values in Tables A.10 and A.12. Moreover, SSH continues to dominate the alternative mean-variance benchmarks (MV-ID, MV-MKT, MV-NL, and MV-PCA) in terms of net-of-cost Sharpe ratios across nearly all datasets.

Taken together, these results indicate that our findings are not driven by small-sample overfitting or look-ahead bias. The expanding-window estimation mitigates concerns related to model instability, and the strong performance under infrequent re-estimation demonstrates that the effectiveness of our sentiment-based shrinkage approach does not rely on short-horizon learning. Instead, the results are consistent with an investor who learns the sentiment–shrinkage mapping infrequently and applies it prospectively.

K. Model of the Economy with Short-sale Constrained Investors

In this section, we extend the model of the economy in Assumption 1 and Proposition 5 following Miller (1977) and Hong and Sraer (2016). Specifically, we consider an economy with three groups of investors: optimistic short-sale-constrained sentiment investors, pessimistic short-sale-constrained sentiment investors, and rational unconstrained investors. In addition, the economy satisfies the following characteristics:

Table A.10: Out-of-sample Sharpe ratios for SSH logistic model estimated only once every five years

This table reports the annualized out-of-sample Sharpe ratio of the proposed SSH portfolio, where the probability p_t that the 1/N portfolio outperforms the alpha portfolio obtained from the logistic model in Equation (18), used to determine the shrinkage intensity, is estimated only once every five years. The table also reports in parentheses the two-sided p-values for the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The six datasets are described in Section 2. The numbers in bold font identify the portfolio strategies with the highest Sharpe ratio.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
SSH	**0.82**	**0.90**	**0.91**	0.96	2.16	1.05
	(0.09)	(0.05)	(0.02)	(0.00)	(0.00)	(0.00)
MV-ID	0.53	0.47	0.67	**1.00**	2.00	0.87
	(0.93)	(0.74)	(0.52)	(0.04)	(0.00)	(0.15)
MV-MKT	0.43	0.34	0.63	0.94	1.86	0.84
	(0.61)	(0.36)	(0.66)	(0.06)	(0.00)	(0.24)
MV-NL	0.51	0.47	0.65	0.93	2.08	0.99
	(0.84)	(0.75)	(0.62)	(0.08)	(0.00)	(0.05)
MV-PCA	0.62	0.73	0.80	0.88	1.26	0.84
	(0.76)	(0.42)	(0.14)	(0.11)	(0.00)	(0.14)
MAXSER	0.62	0.68	0.72	0.98	**2.18**	**1.28**
	(0.76)	(0.57)	(0.42)	(0.06)	(0.00)	(0.00)
RTR	0.64	0.61	0.62	0.58	0.60	0.64
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1/N	0.57	0.56	0.55	0.45	0.36	0.58
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)

Table A.11: Turnover for SSH logistic model estimated only once every five years

This table reports the average monthly turnover of the proposed SSH portfolio, where the probability p_t that the 1/N portfolio outperforms the alpha portfolio obtained from the logistic model in Equation (18), used to determine the shrinkage intensity, is estimated only once every five years. The six datasets are described in Section 2.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
SSH	0.22	0.31	0.27	0.32	0.80	0.32
MV-ID	2.00	2.10	1.40	1.68	12.49	3.64
MV-MKT	2.66	2.98	1.46	2.41	15.09	3.71
MV-NL	2.55	2.51	1.47	2.50	8.69	2.64
MV-PCA	0.55	0.91	0.33	3.24	14.36	0.59
MAXSER	1.16	1.49	1.02	1.58	23.41	1.67
RTR	0.04	0.04	0.04	0.05	0.04	0.04
1/N	0.04	0.04	0.04	0.04	0.04	0.04

Assumption A.1 *The economy satisfies the following characteristics.*

1. *Assets live for two periods, t and $t+1$.*

2. *There are N risky assets that pay random dividends at time $t+1$,*

$$\boldsymbol{d}_{t+1} = \boldsymbol{\mu_d} + \boldsymbol{u}, \tag{A26}$$

   where $\boldsymbol{\mu_d} \in \mathbb{R}^N$ is the vector of average dividends and $\boldsymbol{u} \in \mathbb{R}^N$ follows a zero-mean multivariate distribution with diagonal covariance matrix $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \ldots, \sigma_N^2)$.

3. *All investors have mean-variance preferences with risk-aversion coefficient $\gamma > 0$.*

4. *There are three groups of investors:*

   - *A mass $\phi^+ \in [0, 1)$ of optimistic short-sale-constrained sentiment investors who believe that $\text{E}[\boldsymbol{d}_{t+1}] = \boldsymbol{\mu_d} + \boldsymbol{b}$ with $b_i \geq 0$ for all i and at least one $b_i > 0$.*

Table A.12: Out-of-sample Sharpe ratios net of transaction costs for SSH logistic model estimated only once every five years

This table reports the annualized out-of-sample Sharpe ratio net of transaction costs of 30 basis points of the proposed SSH portfolio, where the probability p_t that the 1/N portfolio outperforms the alpha portfolio obtained from the logistic model in Equation (18), used to determine the shrinkage intensity, is estimated only once every five years. The table also reports in parentheses the two-sided p-values for the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The six datasets are described in Section 2. The numbers in bold font identify the portfolio strategies with the highest Sharpe ratio.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
SSH	**0.78**	**0.85**	**0.86**	**0.91**	**2.09**	0.99
	(0.14)	(0.07)	(0.03)	(0.01)	(0.00)	(0.00)
MV-ID	0.39	0.35	0.55	0.87	1.85	0.60
	(0.46)	(0.40)	(0.95)	(0.11)	(0.00)	(0.90)
MV-MKT	0.28	0.21	0.51	0.79	1.66	0.57
	(0.20)	(0.13)	(0.89)	(0.18)	(0.00)	(0.97)
MV-NL	0.36	0.33	0.53	0.79	1.91	0.76
	(0.35)	(0.34)	(0.94)	(0.19)	(0.00)	(0.37)
MV-PCA	0.57	0.64	0.74	0.72	1.02	0.76
	(0.95)	(0.66)	(0.23)	(0.29)	(0.04)	(0.26)
MAXSER	0.50	0.57	0.62	0.85	1.90	**1.07**
	(0.78)	(0.90)	(0.67)	(0.13)	(0.00)	(0.01)
RTR	0.63	0.60	0.61	0.57	0.59	0.63
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
1/N	0.56	0.55	0.54	0.45	0.35	0.58
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)

- *A mass* $\phi^- \in [0, 1)$ *of pessimistic short-sale-constrained sentiment investors who believe that* $\mathrm{E}[\boldsymbol{d}_{t+1}] = \boldsymbol{\mu_d} - \boldsymbol{b}$.
- *A mass* $1 - \phi^+ - \phi^- \in (0, 1]$ *of unconstrained rational investors who believe that* $\mathrm{E}[\boldsymbol{d}_{t+1}] = \boldsymbol{\mu_d}$.

In this economy, the parameter $b_i \geq 0$ measures the degree of disagreement between optimistic and pessimistic sentiment investors regarding the expected cashflows on asset i. In the next proposition, we present the counterpart of Proposition 5 for this economy, focusing on the case where

the market portfolio is the 1/N portfolio for conciseness. Specifically, we derive the equilibrium vector of expected returns $\boldsymbol{\mu}$ and the Sharpe ratios of the 1/N and alpha portfolios.

Proposition A.5 *Let Assumption A.1 hold and the equilibrium market portfolio be the 1/N portfolio, $\boldsymbol{w}_{ew}$. Then, the following holds at the market equilibrium:*

1. *When $b_i < \gamma(\sigma_i^2/N)/(1+\phi^+-\phi^-) = b_i^\star$, both optimistic and pessimistic investors hold positive quantities of asset i and its expected return is $\mu_i = \gamma\sigma_i^2/N - (\phi^+-\phi^-)b_i$. If $b_i < b_i^\star$ for all i, the 1/N and alpha Sharpe ratios are*

$$\text{SR}(\boldsymbol{w}_{ew}) = \gamma\sigma_{ew} - \frac{(\phi^+-\phi^-)\bar{b}}{\sigma_{ew}} \quad \textit{and} \quad \text{SR}(\boldsymbol{w}_\alpha) = |\phi^+-\phi^-|\sqrt{\boldsymbol{b}^\top \boldsymbol{Q}\boldsymbol{b}}. \tag{A27}$$

   For fixed (ϕ^+,ϕ^-), $\text{SR}(\boldsymbol{w}_{ew})$ decreases (increases) with $\bar{b}$ if optimism (pessimism) dominates, i.e., $\phi^+ > \phi^-$ ($\phi^- > \phi^+$), and, assuming $\boldsymbol{b} = \bar{b}\mathbf{1}$, $\text{SR}(\boldsymbol{w}_\alpha)$ increases with $\bar{b}$.

2. *When $b_i \geq b_i^\star$, only optimistic investors hold positive quantities of asset i and its expected return is $\mu_i = (\gamma\sigma_i^2/N - \phi^+ b_i)/(1-\phi^-)$. If $b_i \geq b_i^\star$ for all i, the 1/N and alpha Sharpe ratios are*

$$\text{SR}(\boldsymbol{w}_{ew}) = \frac{\gamma\sigma_{ew}^2 - \phi^+\bar{b}}{(1-\phi^-)\sigma_{ew}} \quad \textit{and} \quad \text{SR}(\boldsymbol{w}_\alpha) = \frac{\phi^+}{1-\phi^-}\sqrt{\boldsymbol{b}^\top \boldsymbol{Q}\boldsymbol{b}}. \tag{A28}$$

   For fixed (ϕ^+,ϕ^-), $\text{SR}(\boldsymbol{w}_{ew})$ decreases with $\bar{b}$ and, assuming $\boldsymbol{b} = \bar{b}\mathbf{1}$, $\text{SR}(\boldsymbol{w}_\alpha)$ increases with $\bar{b}$.

Proposition A.5 shows that when disagreement between optimistic and pessimistic sentiment investors about asset i is sufficiently small, i.e., $b_i < b_i^\star$, both types of investors hold positive quantities of the asset. When this condition holds for all assets, the Sharpe ratios of the $1/N$ and alpha portfolios are similar to those in the baseline model in Equation (29), which abstracts from short-sale constraints.

When optimism dominates among sentiment investors, i.e., $\phi^+ > \phi^-$, prices are pushed above fundamental values, so that the Sharpe ratio of the $1/N$ (alpha) portfolio decreases (increases) with the average bias among sentiment investors, $\bar{b}$. In this case, an investor should optimally shrink less toward the $1/N$ portfolio as the mass of sentiment investors increases. In contrast, when pessimism dominates, i.e., $\phi^- > \phi^+$, both the $1/N$ and alpha Sharpe ratios increase with the magnitude of the bias, $\bar{b}$, making the direction of optimal shrinkage ambiguous.

When disagreement is sufficiently large, i.e., $b_i \geq b_i^\star$, pessimistic investors are sidelined due to short-sale constraints, and only optimistic investors hold positive quantities of the asset, consistent with Miller (1977) and Hong and Sraer (2016). As a result, only the optimistic component of sentiment shows up in prices, which helps rationalize why it is empirically optimal to shrink toward $1/N$ when sentiment is low. This is because the Sharpe ratio of the $1/N$ (alpha) portfolio decreases (increases) with the mass of optimistic sentiment investors.

L. Link Between Model-implied and Observable Sentiment

We now study whether the predictions from the theoretical model in Section 4 are consistent with the observed empirical relation between investor sentiment and asset returns.

L.1 Empirical Evidence on the Logistic Slope Coefficient

We now assess whether the estimated logistic coefficients are qualitatively consistent with the model's comparative statics. First, we note that our conditional shrinkage approach is data-driven, and thus, the sign of the relation between investor sentiment and the relative performance of the scaled 1/N and alpha components is estimated, not imposed. Figure A.3 depicts boxplots of the

slope coefficient β_1 from the logistic regression in Equation (18), estimated using the procedure described in Figure 3. The median coefficient is negative in every dataset, consistent with higher sentiment being associated with a lower probability that the scaled 1/N component outperforms the alpha component. We complement this evidence with the time series of the estimated coefficients and the correlations between the implemented SSH shrinkage intensities and sentiment.

Figure A.3: Distribution of the estimated slope coefficient in the logistic model

This figure depicts boxplots for the slope coefficient β_1 from the logistic regression in Equation (18), estimated using an expanding-window procedure. For each dataset and each out-of-sample period, the logistic model is estimated using an expanding window of past observations, where the dependent variable is an indicator for whether the scaled 1/N portfolio return exceeds the alpha portfolio return, and the predictor is the sentiment index constructed from only past observable sentiment proxies to avoid look-ahead biases. At each point in time, the estimated coefficients (β_0, β_1) are obtained using only information available up to that date, and are then used to form the out-of-sample probability forecast. The figure summarizes the distribution of the estimated β_1 coefficients across all expanding windows for each dataset. The central line in each box indicates the median, the box represents the interquartile range, and the whiskers extend to 1.5 times the interquartile range, with outliers shown individually.
Alt text: Six boxplots show negative median sentiment slopes, with the one hundred asset dataset closest to zero. Positive and negative outliers indicate occasional departures from typical estimates.

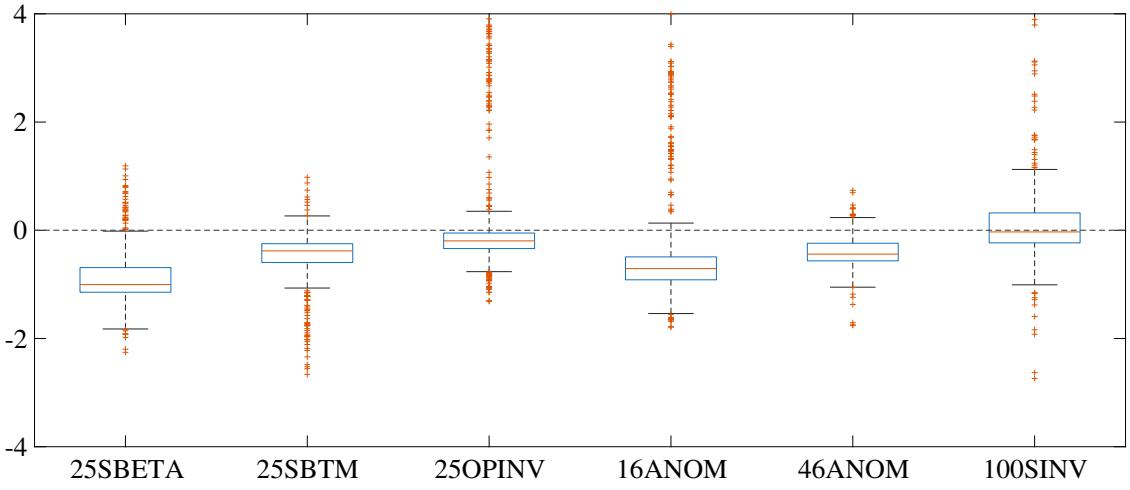


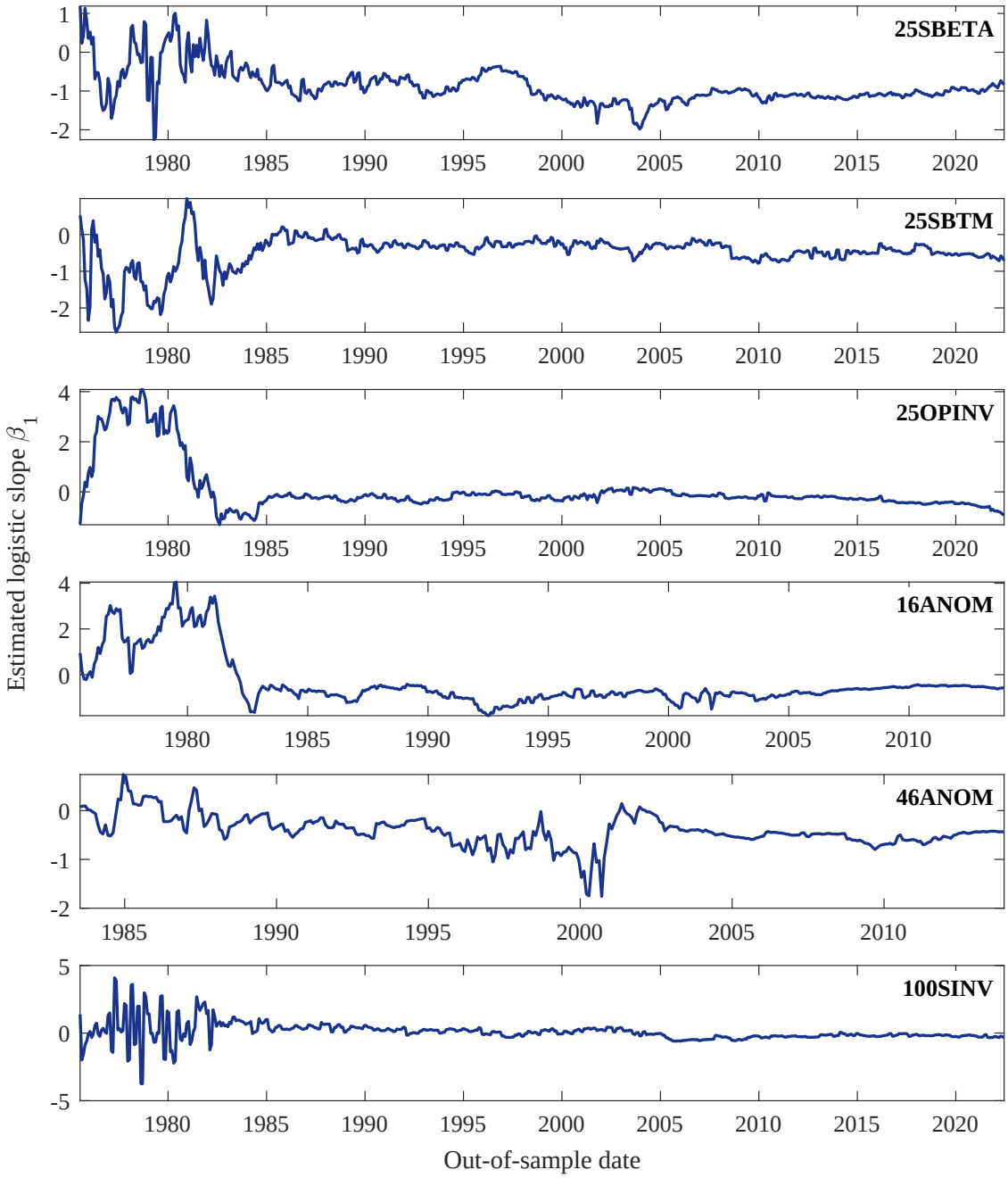
Figure A.4 plots the expanding-window estimates of β_1 used in the out-of-sample analysis. The coefficient is negative in 93.6%, 95.2%, 79.1%, 83.5%, and 91.0% of the expanding windows

for 25SBETA, 25SBTM, 25OPINV, 16ANOM, and 46ANOM, respectively. There is substantial time variation early in the out-of-sample period, when the first logistic regression uses only 60 monthly observations. As the estimation window lengthens, the estimates generally vary less and remain predominantly negative. At the final estimation date of each dataset, the estimated coefficients are -0.84, -0.71, -0.94, -0.56, -0.44, and -0.33 for 25SBETA, 25SBTM, 25OPINV, 16ANOM, 46ANOM, and 100SINV, respectively. Moreover, the coefficient is negative in every month of the final ten years for the first five datasets and in 98.3% of these months for 100SINV. Thus, although 100SINV has a negative coefficient in 52.5% of all expanding windows and a full-period median of -0.03, its later estimates also clearly favor shrinking against sentiment.

We next examine whether the implemented shrinkage intensity moves against sentiment. Because the shrinkage intensity increases with the fitted probability p_t in Equation (18), and p_t decreases with sentiment when $\hat{\beta}_1 < 0$, the implemented intensity is, other things equal, a decreasing function of sentiment during any period in which $\hat{\beta}_1$ remains negative, suggesting a negative correlation between the SSH shrinkage intensity and sentiment. If instead the estimated sign of $\hat{\beta}_1$ changed within the period, the intensity would decrease with sentiment in some months and increase in others, potentially pulling their correlation toward zero. The correlation between the implemented shrinkage intensity and sentiment therefore summarizes the direction and strength of their comovement over the period. Table A.13 reports the correlations between the implemented shrinkage intensity and the recursive sentiment index used at the same portfolio-formation date, over the full out-of-sample period and the final 20, 15, and 10 years of each dataset, corresponding to its final 240, 180, and 120 monthly observations. At each formation date, the sentiment index is reconstructed using only the information available then, and the shrinkage intensity is the one implemented by the baseline SSH procedure.

Figure A.4: Time series of the estimated logistic slope coefficient β_1

This figure plots the time series of the estimated slope coefficient β_1 from the logistic regression in Equation (18), re-estimated at each out-of-sample date using only information available up to that date; the sign of β_1 is not imposed. Each panel corresponds to one of the six datasets described in Section 2.
Alt text: Six time-series panels show sentiment slope estimates fluctuating sharply early in the sample and generally becoming more stable and negative later. Patterns differ across the six datasets.



Over the full out-of-sample period, the correlations are negative for 25SBETA, 25SBTM, 25OPINV, and 16ANOM, and close to zero for 46ANOM and 100SINV. The near-zero full-period correlations for these two datasets do not indicate that the shrinkage intensity is unrelated to sentiment; they reflect the changes in the sign of $\hat{\beta}_1$ during the early years of the out-of-sample period, when the estimation windows are short. As the estimation windows lengthen, the correlations become uniformly negative. Over the final 20 years, they are negative for five datasets and close to zero for 46ANOM, and over the final 15 and 10 years they are negative in every dataset, ranging from -0.94 to -0.14 and from -0.96 to -0.45, respectively. These results are not specific to the horizons reported in the table: the correlation is negative in every dataset for every terminal window between 5 and 16 years, and it is negative for every terminal window between 5 and 25 years for all datasets other than 46ANOM. Taken together, the negative full-period correlations for four datasets and the uniformly negative correlations once the estimation windows are sufficiently long indicate that the estimated direction of the relation between sentiment and the relative performance of the scaled 1/N and alpha portfolios is stable, and that shrinking against sentiment remains a feature of the implemented strategy.

L.2 Empirical Link Between Model-implied and Observable Sentiment

We now empirically document the link between the model-implied sentiment and the observable sentiment index used in our baseline specification. Motivated by the theoretical model in Section 4.1, we construct a time-varying measure of aggregate sentiment-induced pricing pressure based on the vector of equilibrium expected returns in Equation (27). In the model, expected returns satisfy

$$\boldsymbol{\mu} = \gamma \boldsymbol{\Sigma} \boldsymbol{w}_{ew} - \phi \boldsymbol{b}, \tag{A29}$$

Table A.13: Correlation between implemented shrinkage intensity and sentiment

This table reports the sample correlations between the implemented SSH shrinkage intensity and the recursive sentiment index used at the same portfolio-formation date. The index is reconstructed separately for each dataset and each date using only the information available then. All dates refer to portfolio-return months. The full out-of-sample period begins in July 1975, except for 46ANOM, which begins in July 1983. The final 20, 15, and 10 years use the last 240, 180, and 120 monthly observations of each dataset, respectively. For 25SBETA, 25SBTM, 25OPINV, and 100SINV, these periods begin in July 2002, July 2007, and July 2012 and end in June 2022. For 16ANOM and 46ANOM, they begin in January 1994, January 1999, and January 2004 and end in December 2013.

	Full period	Final 20 years	Final 15 years	Final 10 years
25SBETA	−0.477	−0.896	−0.937	−0.958
25SBTM	−0.259	−0.547	−0.638	−0.863
25OPINV	−0.229	−0.671	−0.847	−0.961
16ANOM	−0.371	−0.747	−0.774	−0.490
46ANOM	0.002	0.012	−0.142	−0.447
100SINV	0.064	−0.596	−0.664	−0.822

where $\boldsymbol{w}_{ew}$ denotes the (market) equal-weighted portfolio, γ is the risk aversion parameter, $\phi \in [0, 1]$ is the mass of investors with biased beliefs about cash flows, and $\boldsymbol{b}$ is the vector of cash-flow biases so that $\phi\boldsymbol{b}$ captures the pricing distortion induced by sentiment investors.

Premultiplying both sides of Equation (A29) by $\boldsymbol{w}_{ew}$ yields the expected return on the 1/N portfolio:

$$\mu_{ew} = \boldsymbol{w}_{ew}^{\top}\boldsymbol{\mu} = \gamma\boldsymbol{w}_{ew}^{\top}\boldsymbol{\Sigma}\boldsymbol{w}_{ew} - \phi\boldsymbol{w}_{ew}^{\top}\boldsymbol{b}. \tag{A30}$$

Defining $\sigma_{ew}^2 = \boldsymbol{w}_{ew}^{\top}\boldsymbol{\Sigma}\boldsymbol{w}_{ew}$ and $\bar{b} = \boldsymbol{w}_{ew}^{\top}\boldsymbol{b}$, Equation (A30) can be written as

$$\mu_{ew} = \gamma\sigma_{ew}^2 - \phi\bar{b}. \tag{A31}$$

Rearranging Equation (A31) implies that the model-implied aggregate sentiment pressure is given by

$$\phi\bar{b} = \gamma\sigma_{ew}^2 - \mu_{ew}. \tag{A32}$$

This motivates our empirical model-implied sentiment proxy

$$z_t = \gamma \hat{\sigma}^2_{ew,t} - \hat{\mu}_{ew,t}, \tag{A33}$$

where $\hat{\mu}_{ew,t}$ and $\hat{\sigma}^2_{ew,t}$ are estimated using past returns. At each point in time, we compute $\hat{\mu}_{ew,t}$ as the sample mean of the 1/N portfolio return and $\hat{\sigma}^2_{ew,t}$ as its sample variance, using only information available up to that date.[3]

By construction, z_t corresponds to the composite term $\phi_t \bar{b}_t$ in Equation (A32), which combines the mass of sentiment investors and the average bias in their beliefs. Importantly, the model does not separately identify ϕ_t and $\bar{b}_t$, but implies that equilibrium expected returns depend on their product. We therefore interpret z_t as the model-implied aggregate sentiment pressure. This mapping provides a direct link between observable return moments and the theoretical object driving mispricing in the model. We then estimate the logistic specification in Equation (18) using the procedure described in Figure 3, replacing the empirical sentiment index with z_t. This yields a second set of fitted probabilities, denoted as $\hat{p}_t^{(Z)}$, which we compare to the baseline fitted probabilities $\hat{p}_t^{(S)}$ obtained using the sentiment index.

Table A.14 reports statistics about the fitted probabilities from the two specifications. Across all datasets, the fitted probabilities are positively related. The correlation between $\hat{p}_t^{(S)}$ and $\hat{p}_t^{(Z)}$ ranges from 0.29 to 0.83 across datasets, and is 0.73 in the stacked sample. Moreover, the regression of $\hat{p}_t^{(S)}$ on $\hat{p}_t^{(Z)}$ yields a slope of 0.82 with a t-statistic of 59.95 in the stacked sample, indicating a strong and statistically significant association between the two probability measures. Because both specifications map their respective sentiment measures to the probability that the 1/N portfolio outperforms the alpha component, the positive association indicates that the observable

[3] We set $\gamma = 1$ as in the out-of-sample analysis and an initial estimation window of $T = 120$ monthly observations.

Table A.14: Comparison of fitted probabilities from observed and model-implied sentiment

This table compares the fitted probabilities from two logistic specifications. The first specification is based on the observed sentiment index, and the second specification is based on the model-implied sentiment variable $z_t = \gamma\hat{\sigma}^2_{ew,t} - \hat{\mu}_{ew,t}$, as explained in Section L. For each dataset, the first column reports the dataset acronym, the second column reports the number of observations used in each regression, the third column reports the correlation between the two fitted probabilities, and the remaining columns report the intercept, slope, and slope t-statistic for the regression of $\hat{p}_t^{(S)}$ on $\hat{p}_t^{(Z)}$.

Dataset	Observations	Correlation	Intercept	Slope	t(Slope)
25SBETA	564	0.29	0.25	0.50	7.21
25SBTM	564	0.65	0.06	0.88	20.23
25OPINV	564	0.62	0.17	0.60	18.77
16ANOM	462	0.46	0.19	0.49	11.13
46ANOM	366	0.83	0.05	0.77	28.27
100SINV	564	0.77	0.14	0.75	28.38
Stacked	3084	0.73	0.08	0.82	59.95

sentiment index and the model-implied variable z_t capture similar time-series variation in the underlying economic mechanism defining the probability of 1/N outperforming the alpha component.

Finally, Table A.15 reports the empirical out-of-sample Sharpe ratio, and the turnover, of our SSH portfolio constructed with a logistic probability fitted to the observed sentiment index ($\hat{p}_t^{(S)}$) or the model-implied sentiment variable ($\hat{p}_t^{(Z)}$). The out-of-sample Sharpe ratios and the turnover are similar across datasets for the two specifications. This result means that the out-of-sample performance gains delivered by the SSH portfolio we document in the manuscript are robust to using a model-implied sentiment proxy based purely on historical returns.

The analysis in this section shows that observable sentiment measures are closely related to the model-implied sentiment pressure in Section 4.1, eventually leading to similar portfolio performance, and that the estimated logistic coefficients are quantitatively consistent with the model's comparative statics.

Table A.15: Out-of-sample performance for SSH based on observed and model-implied sentiment

This table reports the annualized out-of-sample Sharpe ratio, and the turnover, of the SSH portfolio. In the logistic regression, we use either the observed sentiment index or the model-implied sentiment variable $z_t = \gamma\hat{\sigma}^2_{ew,t} - \hat{\mu}_{ew,t}$, as explained in Section L. The table also reports in parentheses the two-sided p-values for the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The six datasets are described in Section 2.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Gross Sharpe ratio						
SSH (observed)	0.83	0.91	0.94	1.10	2.04	1.08
	(0.08)	(0.03)	(0.01)	(0.00)	(0.00)	(0.00)
SSH (model)	0.85	0.89	0.94	1.04	2.13	1.07
	(0.05)	(0.04)	(0.01)	(0.00)	(0.00)	(0.00)
Panel B: Net-of-cost Sharpe ratio (30bps)						
SSH (observed)	0.78	0.85	0.88	1.04	1.91	0.99
	(0.14)	(0.07)	(0.02)	(0.00)	(0.00)	(0.00)
SSH (model)	0.81	0.84	0.88	0.99	2.01	1.00
	(0.11)	(0.08)	(0.03)	(0.00)	(0.00)	(0.00)
Panel C: Turnover						
SSH (observed)	0.26	0.41	0.34	0.38	1.66	0.46
SSH (model)	0.21	0.33	0.32	0.34	1.51	0.45

M. Exploiting Additional Conditioning Variables

Table 2 in the main body of the manuscript shows that investor sentiment is an informative conditioning variable for predicting the relative performance of the 1/N and alpha components. We now examine whether SSH performance can be improved by augmenting the baseline sentiment specification with each of the additional conditioning variables considered in Table 2, one at a time. Table A.16 shows that adding an additional conditioning variable has only a limited effect on gross portfolio performance. Across the augmented specifications, the average gross Sharpe ratio ranges from 1.13 to 1.17, compared with 1.15 for the sentiment-only specification. However, turnover increases, and the average net-of-cost Sharpe ratios are generally similar to or slightly lower than that of the baseline.

N. Alternative Calibrations of the Risk-aversion Coefficient

In the main body of the manuscript, the SSH portfolio and the considered benchmarks described in Section 4.2.2 aim to estimate the mean-variance portfolio $\boldsymbol{w}_{mv} = \gamma^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ with $\gamma = 1$. We now consider two alternatives: $\gamma = 3$ and $\gamma = (\boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})^{1/2}/\sigma$, which amounts to targeting a volatility σ. We set $\sigma = 0.3/\sqrt{12}$. Table A.17 reports the net-of-cost out-of-sample Sharpe ratios using proportional transaction costs of 30 basis points, and the average monthly turnover, for the eight portfolio strategies we consider.

The conclusions in the main body of the manuscript obtained for $\gamma = 1$ are robust to using different calibrations of γ. Specifically, the average outperformance of the SSH portfolio relative to the five benchmark mean-variance portfolios is 24% and 16% for $\gamma = 3$ and $\gamma = (\boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})^{1/2}/\sigma$, respectively. The SSH portfolio also largely outperforms the simple RTR and 1/N portfolios. More-

Table A.16: SSH with additional conditioning variables

This table reports the annualized out-of-sample gross Sharpe ratio (Panel A), turnover (Panel B), and Sharpe ratio net of 30-bps transaction costs (Panel C) of the baseline SSH portfolio and of five specifications that add one macroeconomic variable to sentiment in the logistic regression for p_t in Equation (18). The macroeconomic variables are defined and lagged as in Table 2; all specifications share the same out-of-sample months. The six datasets are described in Section 2.

	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Gross Sharpe ratio						
Sentiment only	0.83	0.91	0.94	1.10	2.04	1.08
+ Market volatility	0.83	0.91	0.94	1.11	1.95	1.09
+ Credit spread	0.80	1.00	0.96	1.11	2.07	1.07
+ Dividend yield	0.85	0.95	0.95	1.14	1.84	1.04
+ Industrial production growth	0.83	0.89	0.95	1.12	2.00	1.06
+ Initial claims growth	0.83	0.91	0.95	1.11	2.01	1.10
Panel B: Turnover						
Sentiment only	0.26	0.41	0.34	0.38	1.66	0.46
+ Market volatility	0.31	0.52	0.44	0.50	3.26	0.65
+ Credit spread	0.30	0.53	0.37	0.46	1.70	0.69
+ Dividend yield	0.28	0.57	0.37	0.43	2.10	0.65
+ Industrial production growth	0.32	0.63	0.49	0.55	2.73	0.86
+ Initial claims growth	0.41	0.67	0.60	0.54	3.09	1.02
Panel C: Net-of-cost Sharpe ratio						
Sentiment only	0.78	0.85	0.88	1.04	1.91	0.99
+ Market volatility	0.77	0.83	0.86	1.03	1.72	0.97
+ Credit spread	0.75	0.92	0.89	1.05	1.92	0.94
+ Dividend yield	0.80	0.86	0.88	1.07	1.70	0.93
+ Industrial production growth	0.77	0.80	0.86	1.03	1.79	0.91
+ Initial claims growth	0.75	0.81	0.84	1.03	1.78	0.92

over, the benchmark mean-variance portfolios require a turnover that is on average 343% and 236% larger than that of the SSH portfolio for the two calibrations of γ, respectively.

Table A.17: Sharpe ratios and turnover for alternative calibrations of γ

This table reports the annualized out-of-sample Sharpe ratios net of proportional transaction costs of 30 basis points, and the turnover, of the proposed SSH portfolio and seven benchmarks. The first four benchmarks are mean-variance portfolios constructed with the shrinkage covariance matrices of Ledoit and Wolf (2004) (MV-ID), Ledoit and Wolf (2003) (MV-MKT), Ledoit and Wolf (2020) (MV-NL), and Chen and Yuan (2016) (MV-PCA). The remaining ones are the MAXSER portfolio by Ao, Li, and Zheng (2019), the reward-to-risk timing strategy of Kirby and Ostdiek (2012) (RTR), and the 1/N portfolio. Panels A and B consider a risk-aversion coefficient $\gamma = 3$ and $\gamma = (\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})^{1/2}/\sigma$ with $\sigma = 0.3/\sqrt{12}$, respectively. We also report two-sided p-values of the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The six datasets are described in Section 2.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: $\gamma = 3$						
Net Sharpe ratio						
SSH	**0.79**	0.88	**0.89**	**1.07**	**2.16**	0.97
	(0.12)	(0.03)	(0.02)	(0.00)	(0.00)	(0.01)
MV-ID	0.58	0.76	0.68	1.00	2.09	0.62
	(0.92)	(0.35)	(0.46)	(0.03)	(0.00)	(0.82)
MV-MKT	0.47	0.58	0.67	0.94	1.93	0.60
	(0.66)	(0.88)	(0.54)	(0.05)	(0.00)	(0.89)
MV-NL	0.53	0.69	0.71	0.96	1.86	0.81
	(0.84)	(0.54)	(0.41)	(0.04)	(0.00)	(0.24)
MV-PCA	0.69	0.73	0.73	0.90	1.38	0.80
	(0.46)	(0.34)	(0.25)	(0.08)	(0.00)	(0.20)
MAXSER	0.59	**0.94**	0.78	0.99	2.02	**1.10**
	(0.84)	(0.05)	(0.22)	(0.03)	(0.00)	(0.01)
RTR	0.63	0.60	0.61	0.57	0.59	0.63
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1/N	0.56	0.55	0.54	0.45	0.35	0.58
Turnover						
SSH	0.24	0.29	0.27	0.27	1.01	0.52
MV-ID	1.10	1.19	1.08	1.05	6.60	3.48
MV-MKT	1.52	1.56	1.10	1.44	7.35	3.42
MV-NL	1.36	1.41	1.02	1.29	3.31	2.37
MV-PCA	0.32	0.54	0.22	1.76	5.84	0.38
MAXSER	0.94	1.00	0.74	1.15	9.30	1.58
RTR	0.04	0.04	0.04	0.05	0.04	0.04
1/N	0.04	0.04	0.04	0.04	0.04	0.04
Panel B: $\gamma = (\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})^{1/2}/\sigma$ **with** $\sigma = 0.3/\sqrt{12}$						
Net Sharpe ratio						
SSH	**0.80**	0.89	**0.90**	**1.07**	**1.97**	0.99
	(0.11)	(0.03)	(0.02)	(0.00)	(0.00)	(0.00)
MV-ID	0.60	0.89	0.75	1.02	1.94	0.83
	(0.88)	(0.11)	(0.29)	(0.02)	(0.00)	(0.17)
MV-MKT	0.52	0.78	0.74	0.98	**1.97**	0.84
	(0.83)	(0.29)	(0.32)	(0.04)	(0.00)	(0.17)
MV-NL	0.53	0.84	0.75	0.96	1.77	0.91
	(0.84)	(0.18)	(0.29)	(0.04)	(0.00)	(0.07)
MV-PCA	0.72	0.72	0.75	0.94	1.42	0.80
	(0.40)	(0.36)	(0.23)	(0.06)	(0.00)	(0.17)
MAXSER	0.60	**0.95**	0.73	1.00	1.85	**1.11**
	(0.85)	(0.07)	(0.35)	(0.02)	(0.00)	(0.01)
RTR	0.63	0.60	0.61	0.57	0.59	0.63
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1/N	0.56	0.55	0.54	0.45	0.35	0.58
Turnover						
SSH	0.23	0.28	0.29	0.27	0.41	0.48
MV-ID	0.92	0.98	0.87	0.73	1.44	1.99
MV-MKT	1.15	1.19	0.87	0.90	1.56	1.91
MV-NL	1.11	1.12	0.83	0.88	1.05	1.47
MV-PCA	0.31	0.51	0.31	1.05	1.79	0.39
MAXSER	1.03	1.01	0.83	0.97	2.46	1.31
RTR	0.04	0.04	0.04	0.05	0.04	0.04
1/N	0.04	0.04	0.04	0.04	0.04	0.04

O. Autocorrelated Covariance Matrix

In this section, we assess the performance of mean-variance portfolios constructed using a dynamic covariance matrix that exploits the serial autocorrelation of variances and covariances. This conditional specification allows the covariance estimate to incorporate recent return information more directly than the unconditional sample covariance matrix. Specifically, given an estimation window of T months, we estimate the covariance matrix using the exponentially weighted moving average model:

$$\widehat{\boldsymbol{\Sigma}}_{ewma} = \frac{1-\lambda}{1-\lambda^T} \sum_{t=1}^{T} \lambda^{T-t} (\boldsymbol{r}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{r}_t - \hat{\boldsymbol{\mu}})^\top, \tag{A34}$$

where $\lambda = 0.94$ as recommended by RiskMetrics. Then, we shrink $\widehat{\boldsymbol{\Sigma}}_{ewma}$ toward the identity matrix using either the SSH shrinkage intensity or that of Ledoit and Wolf (2004) (MV-ID). Thus, MV-ID is also a conditional and regularized benchmark in this exercise: the EWMA specification captures time variation in covariances, while Ledoit–Wolf shrinkage alleviates estimation error. In Table A.18, we report the annualized Sharpe ratios, gross and net of proportional transaction costs of 30 basis points, and the turnover of the two portfolios. Panels A, B, and C consider the three calibrations of the risk-aversion coefficient γ considered in this manuscript, i.e., $\gamma = 1$, $\gamma = 3$, and $\gamma = (\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})^{1/2}/\sigma$ with $\sigma = 0.3/\sqrt{12}$, respectively. Two results stand out. First, the SSH portfolio continues to substantially outperform the MV-ID portfolio and to deliver a much smaller turnover. Second, our proposed SSH portfolio still delivers a good performance when applied to dynamic autocorrelated covariance matrices.

Table A.18: Out-of-sample performance with autocorrelated covariance matrix

This table reports the annualized out-of-sample Sharpe ratios, and the turnover, of the SSH portfolio and the mean-variance portfolio constructed with the shrinkage covariance matrix of Ledoit and Wolf (2004) (MV-ID). The Sharpe ratios are either gross or net of proportional transaction costs of 30 basis points. The covariance matrix is estimated using the exponentially weighted moving average model in (A34). Panels A, B, and C consider a risk-aversion coefficient $\gamma = 1$, $\gamma = 3$ and $\gamma = (\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})^{1/2}/\sigma$ with $\sigma = 0.3/\sqrt{12}$, respectively. The six datasets are described in Section 2.

Strategy		25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: $\gamma = 1$							
SSH	Gross SR	0.68	0.73	0.89	0.89	2.00	0.92
	Net SR (30 bps)	0.61	0.64	0.80	0.82	1.86	0.79
	Turnover	0.39	0.59	0.45	0.42	1.20	0.77
MV-ID	Gross SR	0.50	0.31	0.31	0.56	1.10	0.40
	Net SR (30 bps)	0.14	−0.07	−0.01	0.28	0.70	−0.06
	Turnover	3.41	5.72	6.96	6.34	28.1	19.0
Panel B: $\gamma = 3$							
SSH	Gross SR	0.76	0.81	0.90	0.91	2.01	0.92
	Net SR (30 bps)	0.69	0.73	0.81	0.85	1.88	0.79
	Turnover	0.32	0.47	0.38	0.30	1.05	0.72
MV-ID	Gross SR	0.67	0.51	0.44	0.71	1.63	0.54
	Net SR (30 bps)	0.30	0.14	0.11	0.44	1.21	0.19
	Turnover	2.60	3.37	3.99	3.41	15.0	9.08
Panel C: $\gamma = (\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})^{1/2}/\sigma$ **with** $\sigma = 0.3/\sqrt{12}$							
SSH	Gross SR	0.79	0.84	0.89	0.92	2.15	0.95
	Net SR (30 bps)	0.72	0.76	0.81	0.86	2.05	0.84
	Turnover	0.32	0.46	0.39	0.30	0.53	0.60
MV-ID	Gross SR	0.79	0.74	0.55	0.77	1.99	0.83
	Net SR (30 bps)	0.39	0.34	0.22	0.50	1.57	0.39
	Turnover	2.33	2.81	2.93	2.01	5.04	4.29

P. Pairwise Sharpe-ratio Tests

We complement the statistical analysis in the main body of the manuscript, which compares the Sharpe ratio of each portfolio policy with that of 1/N, by conducting pairwise one-sided Sharpe-ratio tests across all strategies. We apply the studentized circular-block bootstrap of Ledoit and Wolf (2008) to portfolio returns net of 30-bps transaction costs. For each dataset, we test

$$H_0 : \text{SR}_i \leq \text{SR}_j \qquad \text{versus} \qquad H_1 : \text{SR}_i > \text{SR}_j.$$

The entry in row i and column j reports the corresponding one-sided bootstrap p-value. Thus, a small p-value in the SSH row provides evidence that SSH has a higher Sharpe ratio than the strategy in the corresponding column. We use a bootstrap block size of five, as in the main-text tests.[4]

Table A.19 reports the pairwise p-value matrices. SSH has a significantly higher net-of-cost Sharpe ratio than all five mean-variance benchmarks in 25SBETA and 25SBTM, and than four of the five benchmarks in 25OPINV. It also has a higher Sharpe ratio than 1/N at the 5% significance level in five of the six datasets; the evidence is only marginally weaker in 25SBETA, where the exact p-value is 0.052. SSH has a significantly higher Sharpe ratio than RTR in five of the six datasets. In 16ANOM, 46ANOM, and 100SINV, several mean-variance benchmarks also attain high net-of-cost Sharpe ratios, so some pairwise differences are not statistically significant. Across the six datasets, SSH has a significantly higher net-of-cost Sharpe ratio in 29 of the 42 comparisons with the seven benchmarks at the 5% level. The median one-sided p-value across these comparisons is 0.014, substantially below the next-lowest median p-value of 0.133 for MAXSER.

[4]The main body of the manuscript reports two-sided p-values. Here, we report one-sided p-values to facilitate interpretation of the tables: a small entry in row i and column j always provides evidence that the strategy in row i has a higher Sharpe ratio than the strategy in column j. This simple comparison rule would not hold if we reported two-sided p-values.

Table A.19: Pairwise Sharpe-ratio tests, net of costs

Each panel reports one-sided Ledoit and Wolf (2008) bootstrap p-values for the null hypothesis $H_0 : \text{SR}_i \leq \text{SR}_j$ against $H_1 : \text{SR}_i > \text{SR}_j$, using returns net of 30-bps transaction costs and a studentized circular-block bootstrap with block size equal to five. The entry in row i and column j reports the corresponding one-sided bootstrap p-value. The final column ("Avg.") is the average of the off-diagonal p-values in each row; diagonal entries are omitted. The six datasets are described in Section 2.

	SSH	MV-ID	MV-MKT	MV-NL	MV-PCA	MAXSER	RTR	1/N	Avg.
Panel A: 25SBETA									
SSH	–	0.01	0.00	0.00	0.02	0.01	0.11	0.05	0.03
MV-ID	1.00	–	0.00	0.10	0.94	0.94	0.86	0.76	0.66
MV-MKT	1.00	1.00	–	1.00	0.98	0.99	0.93	0.88	0.97
MV-NL	1.00	0.90	0.00	–	0.95	0.98	0.88	0.80	0.79
MV-PCA	0.98	0.07	0.02	0.06	–	0.23	0.60	0.45	0.35
MAXSER	0.99	0.06	0.01	0.02	0.78	–	0.74	0.61	0.46
RTR	0.89	0.14	0.07	0.13	0.41	0.27	–	0.00	0.27
1/N	0.94	0.24	0.12	0.20	0.55	0.39	1.00	–	0.49
Panel B: 25SBTM									
SSH	–	0.00	0.00	0.00	0.02	0.03	0.03	0.02	0.01
MV-ID	1.00	–	0.00	0.31	0.98	1.00	0.84	0.79	0.70
MV-MKT	1.00	1.00	–	1.00	1.00	1.00	0.95	0.92	0.98
MV-NL	1.00	0.69	0.00	–	0.99	1.00	0.88	0.81	0.77
MV-PCA	0.98	0.01	0.00	0.01	–	0.31	0.40	0.34	0.30
MAXSER	0.97	0.00	0.00	0.00	0.69	–	0.52	0.47	0.38
RTR	0.97	0.16	0.05	0.13	0.60	0.47	–	0.00	0.34
1/N	0.97	0.21	0.07	0.18	0.67	0.55	1.00	–	0.52
Panel C: 25OPINV									
SSH	–	0.00	0.00	0.00	0.07	0.01	0.02	0.01	0.02
MV-ID	1.00	–	0.00	0.14	0.86	0.75	0.53	0.42	0.53
MV-MKT	1.00	1.00	–	0.85	0.91	0.87	0.61	0.49	0.82
MV-NL	1.00	0.87	0.16	–	0.90	0.84	0.60	0.46	0.69
MV-PCA	0.93	0.14	0.09	0.11	–	0.24	0.17	0.11	0.26
MAXSER	0.99	0.25	0.14	0.16	0.78	–	0.43	0.33	0.44
RTR	0.98	0.46	0.38	0.42	0.83	0.58	–	0.00	0.52
1/N	0.99	0.59	0.50	0.55	0.90	0.69	1.00	–	0.74

Table A.19: Pairwise Sharpe-ratio tests, net of costs (continued)

Continuation of the previous table for the remaining three datasets.

	SSH	MV-ID	MV-MKT	MV-NL	MV-PCA	MAXSER	RTR	1/N	Avg.
Panel D: 16ANOM									
SSH	–	0.11	0.06	0.06	0.04	0.09	0.00	0.00	0.05
MV-ID	0.88	–	0.01	0.01	0.00	0.23	0.11	0.05	0.18
MV-MKT	0.94	0.99	–	0.30	0.00	0.89	0.19	0.09	0.49
MV-NL	0.94	0.99	0.70	–	0.00	0.93	0.19	0.10	0.55
MV-PCA	0.96	1.00	1.00	1.00	–	0.99	0.27	0.15	0.77
MAXSER	0.91	0.77	0.11	0.07	0.01	–	0.13	0.06	0.30
RTR	1.00	0.89	0.80	0.80	0.72	0.87	–	0.00	0.73
1/N	1.00	0.94	0.91	0.91	0.85	0.94	1.00	–	0.94
Panel E: 46ANOM									
SSH	–	0.38	0.09	0.51	0.00	0.52	0.00	0.00	0.21
MV-ID	0.62	–	0.02	0.71	0.00	0.76	0.00	0.00	0.30
MV-MKT	0.92	0.98	–	0.99	0.00	0.99	0.00	0.00	0.55
MV-NL	0.48	0.27	0.01	–	0.00	0.49	0.00	0.00	0.18
MV-PCA	1.00	1.00	1.00	1.00	–	1.00	0.10	0.03	0.73
MAXSER	0.48	0.25	0.01	0.51	0.00	–	0.00	0.00	0.18
RTR	1.00	1.00	1.00	1.00	0.91	1.00	–	0.00	0.84
1/N	1.00	1.00	1.00	1.00	0.97	1.00	1.00	–	1.00
Panel F: 100SINV									
SSH	–	0.01	0.01	0.10	0.02	0.83	0.01	0.01	0.14
MV-ID	0.99	–	0.07	0.99	0.87	1.00	0.53	0.44	0.70
MV-MKT	0.99	0.94	–	1.00	0.91	1.00	0.59	0.49	0.85
MV-NL	0.91	0.00	0.00	–	0.55	1.00	0.25	0.19	0.41
MV-PCA	0.98	0.13	0.09	0.45	–	0.99	0.18	0.11	0.42
MAXSER	0.17	0.00	0.00	0.00	0.01	–	0.01	0.01	0.03
RTR	0.99	0.48	0.41	0.75	0.82	0.99	–	0.00	0.64
1/N	1.00	0.56	0.51	0.82	0.87	0.99	1.00	–	0.82

Q. Volatility-managed SSH and Mean-variance Benchmarks

We construct volatility-managed versions of the mean-variance benchmarks by scaling each benchmark's weights using lagged market volatility as in DeMiguel, Martín-Utrera, and Uppal (2024). For each mean-variance benchmark p, the volatility-managed weights are

$$\tilde{\boldsymbol{w}}_{p,t} = c_{p,t} \frac{\bar{\sigma}^{RV}_{MKT,t}}{\sigma^{RV}_{MKT,t}} \boldsymbol{w}_{p,t}, \qquad c_{p,t} = \frac{\sigma_{MKT,t}}{\sigma_{p,t}}, \tag{A35}$$

where $\boldsymbol{w}_{p,t}$ is the baseline mean-variance portfolio formed at the end of month t; $\sigma_{MKT,t}$ and $\sigma_{p,t}$ are the standard deviations of the monthly market return and the return of portfolio p, respectively, estimated over the expanding estimation window ending in month t; $\sigma^{RV}_{MKT,t}$ is the realized market volatility in month t, computed as the standard deviation of daily market returns within that month; and $\bar{\sigma}^{RV}_{MKT,t}$ is the arithmetic average of monthly realized market volatility over the expanding estimation window. All these quantities are therefore observable when the portfolio is formed at the end of month t, and the resulting weights are applied to returns in month $t+1$. The normalization factor $c_{p,t}$ rescales the baseline portfolio's in-sample volatility to that of the market before the volatility-timing adjustment is applied. The ratio $\bar{\sigma}^{RV}_{MKT,t}/\sigma^{RV}_{MKT,t}$ then increases exposure following months of relatively low realized market volatility and decreases exposure following months of relatively high realized market volatility.

Table A.20 reports the results for the five volatility-managed mean-variance benchmarks and for a volatility-managed version of SSH. Volatility management raises the benchmarks' gross Sharpe ratios in every dataset and generally increases their turnover. Nevertheless, the net-of-cost Sharpe ratio of every volatility-managed benchmark remains below that of the baseline SSH portfolio in every dataset. Applying the same volatility-management procedure to SSH raises its gross Sharpe ratio in every dataset but lowers its net-of-cost Sharpe ratio in five of the six datasets.

Table A.20: Volatility-managed SSH and mean-variance benchmarks

This table reports the annualized out-of-sample gross Sharpe ratio (Panel A), turnover (Panel B), and Sharpe ratio net of 30-bps transaction costs (Panel C) of the volatility-managed portfolios constructed as in Equation (A35). "SSH (baseline)" is the unmanaged SSH portfolio; the remaining rows are volatility-managed versions of SSH and of the mean-variance benchmarks. The six datasets are described in Section 2.

	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Gross Sharpe ratio						
SSH (baseline)	0.83	0.91	0.94	1.10	2.04	1.08
SSH (vol-managed)	0.91	0.94	1.01	1.28	2.47	1.13
MV-ID (vol-managed)	0.84	1.04	0.90	1.16	2.35	1.07
MV-MKT (vol-managed)	0.79	0.98	0.89	1.11	2.33	1.07
MV-NL (vol-managed)	0.80	1.01	0.89	1.10	2.31	1.09
MV-PCA (vol-managed)	0.75	0.79	0.82	1.07	2.02	0.85
MAXSER (vol-managed)	0.89	1.10	0.93	1.21	2.33	1.31
Panel B: Turnover						
SSH (baseline)	0.26	0.41	0.34	0.38	1.66	0.46
SSH (vol-managed)	1.00	1.06	1.24	1.01	1.89	1.58
MV-ID (vol-managed)	3.00	2.87	2.57	2.05	4.20	5.57
MV-MKT (vol-managed)	3.33	3.16	2.54	2.37	4.32	5.37
MV-NL (vol-managed)	3.33	3.16	2.49	2.51	4.92	4.57
MV-PCA (vol-managed)	1.25	1.34	1.19	2.68	2.75	1.24
MAXSER (vol-managed)	1.79	1.77	1.39	1.72	3.57	2.39
Panel C: Net-of-cost Sharpe ratio						
SSH (baseline)	0.78	0.85	0.88	1.04	1.91	0.99
SSH (vol-managed)	0.66	0.70	0.74	1.01	2.07	0.78
MV-ID (vol-managed)	0.23	0.47	0.42	0.71	1.69	0.20
MV-MKT (vol-managed)	0.13	0.37	0.42	0.60	1.65	0.21
MV-NL (vol-managed)	0.13	0.39	0.42	0.56	1.56	0.30
MV-PCA (vol-managed)	0.48	0.47	0.54	0.51	1.47	0.57
MAXSER (vol-managed)	0.50	0.73	0.65	0.82	1.76	0.89

R. Decomposing SSH's Turnover Advantage

To understand why the turnover of the mean-variance benchmarks is, on average, substantially higher than that of SSH, we decompose the turnover gap into a component associated with the magnitude and distribution of SSH shrinkage intensities and a component associated with their sentiment-driven temporal ordering. Specifically, for each mean-variance benchmark B, we define its turnover gap with SSH as

$$\mathrm{TO}_B - \mathrm{TO}_{\mathrm{SSH}} = (\mathrm{TO}_B - \mathrm{E}[\mathrm{TO}_{\mathrm{perm}}]) + (\mathrm{E}[\mathrm{TO}_{\mathrm{perm}}] - \mathrm{TO}_{\mathrm{SSH}}), \tag{A36}$$

where $\mathrm{TO}_{\mathrm{perm}}$ is the turnover of SSH after randomly permuting its realized monthly shrinkage intensities. To estimate $\mathrm{E}[\mathrm{TO}_{\mathrm{perm}}]$, we generate 1,000 random orderings of the entire realized out-of-sample intensity sequence, with each intensity appearing exactly once in every reordered path. For each path, we recompute monthly SSH weights while holding all other inputs fixed and calculate average monthly turnover. $\mathrm{E}[\mathrm{TO}_{\mathrm{perm}}]$ is the arithmetic mean across the 1,000 path-specific turnover estimates. Thus, the procedure preserves the mean and empirical distribution of shrinkage intensities while removing their observed temporal ordering.

Hence, the first term captures the part of the benchmark-SSH turnover gap remaining after temporal ordering is removed and is therefore associated with the magnitude and distribution of shrinkage intensities. The second component captures the reduction in turnover obtained by restoring the observed sentiment-driven temporal ordering. A larger first (second) component therefore indicates a greater role for the magnitude (temporal ordering) of shrinkage intensities.

The Aggregate column of Table A.21 shows that temporal ordering accounts for 67.4%, 53.9%, 79.3%, 81.1%, and 49.8% of the average turnover gap relative to MV-ID, MV-MKT, MV-NL, MV-PCA, and MAXSER, respectively, and for 63.8% relative to the mean benchmark—the

Table A.21: Decomposition of the benchmark-SSH turnover gap

This table reports the percentage contributions of the two components in Equation (A36). Each Panel A (Panel B) entry equals 100 times the first (second) component divided by $\mathrm{TO}_B - \mathrm{TO}_{\mathrm{SSH}}$, so the two panels sum to 100% for every benchmark-dataset pair. Negative contributions, or contributions above 100%, indicate that one component offsets the other. Expected permuted-SSH turnover is estimated from 1,000 random permutations of the realized SSH shrinkage-intensity path, holding all other SSH inputs unchanged. For each benchmark, the Aggregate column reports the component summed across the six datasets as a percentage of the corresponding total benchmark-SSH turnover gap summed across the six datasets. Mean benchmark is the average of the five mean-variance benchmarks. The six datasets are described in Section 2.

	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV	Aggregate
Panel A: Benchmark-to-permuted-SSH share (%)							
MV-ID	27.7	−15.5	−27.2	−50.9	54.5	40.0	32.6
MV-MKT	47.5	24.0	−20.2	3.2	63.3	41.5	46.1
MV-NL	45.1	6.8	−19.2	7.2	29.9	12.6	20.7
MV-PCA	−330.2	−291.8	14117.5	31.2	61.2	−1442.7	18.9
MAXSER	−39.7	−81.6	−97.5	−63.4	77.3	−57.1	50.2
Mean benchmark	17.5	−23.1	−69.1	−3.3	62.5	4.2	36.2
Panel B: SSH temporal-ordering share (%)							
MV-ID	72.3	115.5	127.2	150.9	45.5	60.0	67.4
MV-MKT	52.5	76.0	120.2	96.8	36.7	58.5	53.9
MV-NL	54.9	93.2	119.2	92.8	70.1	87.4	79.3
MV-PCA	430.2	391.8	−14017.5	68.8	38.8	1542.7	81.1
MAXSER	139.7	181.6	197.5	163.4	22.7	157.1	49.8
Mean benchmark	82.5	123.1	169.1	103.3	37.5	95.8	63.8

equal-weight average of the five mean-variance benchmarks—versus 36.2% for the other component.[5] Thus, although both components contribute materially, SSH's lower turnover primarily reflects the slow temporal variation in sentiment-driven shrinkage rather than its shrinkage magnitude alone.

[5]Some percentages for the MV-PCA benchmark are very large because MV-PCA and SSH have nearly identical turnover in those datasets, making the denominator of the percentage decomposition close to zero. These values reflect two large offsetting components rather than unusually large turnover effects. The Aggregate column is computed as the turnover gap summed across datasets and therefore does not average the unstable dataset-level percentages.

S. Shorting Costs and Limits to Capacity

Beyond the proportional transaction costs we consider in the main body of the manuscript, a realistic evaluation of implementability should also account for frictions associated with short positions and the strategy's capacity. We now discuss in detail the implementability of the SSH strategy relative to benchmark mean-variance strategies and explicitly address these concerns through additional analyses and robustness tests.

We first note that an important advantage of the SSH portfolio is its substantially lower turnover relative to benchmark mean-variance portfolios, which already improves implementability. As shown in Table 4 of the manuscript, SSH requires dramatically less trading than the benchmark mean-variance strategies. This difference in turnover is a key reason why the performance advantage of SSH becomes even more pronounced once trading costs are taken into account.

We extend our net-of-cost performance evaluation by explicitly incorporating shorting costs. In particular, in addition to proportional transaction costs based on portfolio turnover, we introduce a reduced-form shorting cost proportional to the portfolio's total short exposure. Specifically, we measure the average shorting cost of portfolio p as

$$\bar{S}_p = \lambda \frac{1}{\tau_{\text{oos}}} \sum_{t=1}^{\tau_{\text{oos}}} \sum_{i=1}^{N} \max(-w_{p,i,t}, 0), \tag{A37}$$

where $w_{p,i,t}$ denotes the weight on asset i at time t, λ is the shorting-cost parameter, and τ_{oos} is the number of out-of-sample observations. We then subtract $\bar{S}_p$ from the average portfolio return net of proportional transaction costs to compute the Sharpe ratio.

In the analysis, we consider proportional transaction costs of $\kappa = 30$ basis points per unit of portfolio turnover and a monthly shorting cost of $\lambda = 10$ basis points, which corresponds to approximately 120 basis points per year. This value is conservative and consistent with the empirical

Table A.22: Out-of-sample performance net of transaction and shorting costs

This table reports the annualized out-of-sample Sharpe ratios net of both transaction costs of 30 basis points and shorting costs of 10 basis points, for the proposed SSH portfolio and seven benchmarks. The first four benchmarks are mean-variance portfolios constructed with the shrinkage covariance matrices of Ledoit and Wolf (2004) (MV-ID), Ledoit and Wolf (2003) (MV-MKT), Ledoit and Wolf (2020) (MV-NL), and Chen and Yuan (2016) (MV-PCA). The remaining ones are the MAXSER portfolio of Ao, Li, and Zheng (2019), the reward-to-risk timing strategy of Kirby and Ostdiek (2012) (RTR), and the 1/N portfolio. The table also reports in parentheses the two-sided p-values for the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The six datasets are described in Section 2. The numbers in bold font identify the portfolio strategies with the highest Sharpe ratio.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
SSH	**0.70**	**0.75**	**0.76**	**0.93**	**1.88**	0.83
	(0.37)	(0.23)	(0.13)	(0.01)	(0.00)	(0.08)
MV-ID	0.16	0.12	0.31	0.64	1.63	0.14
	(0.06)	(0.07)	(0.26)	(0.47)	(0.00)	(0.05)
MV-MKT	0.02	−0.03	0.28	0.53	1.34	0.12
	(0.01)	(0.01)	(0.22)	(0.76)	(0.00)	(0.05)
MV-NL	0.08	0.05	0.29	0.51	1.53	0.31
	(0.02)	(0.03)	(0.26)	(0.85)	(0.00)	(0.22)
MV-PCA	0.49	0.50	0.61	0.43	0.84	0.64
	(0.69)	(0.81)	(0.66)	(0.94)	(0.13)	(0.71)
MAXSER	0.36	0.43	0.51	0.67	1.55	**0.89**
	(0.36)	(0.64)	(0.92)	(0.38)	(0.00)	(0.12)
RTR	0.63	0.60	0.61	0.57	0.59	0.63
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
1/N	0.56	0.55	0.54	0.45	0.35	0.58
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)

distribution of stock-loan fees documented in the securities-lending literature (Engelberg, Reed, and Ringgenberg 2018). Using this calibration, we report in Table A.22 the out-of-sample Sharpe ratios net of both transaction and shorting costs.

The results show that the strong performance of SSH persists after accounting for both transaction and shorting costs. In absolute terms, SSH achieves net-of-cost Sharpe ratios of 0.70, 0.75, 0.76, 0.93, 1.88, and 0.83 for the 25SBETA, 25SBTM, 25OPINV, 16ANOM, 46ANOM, and 100SINV datasets, respectively. These levels remain economically large and are substantially

higher than those delivered by most benchmark mean-variance portfolios once implementation frictions are incorporated.

In relative terms, SSH continues to outperform the benchmark mean-variance strategies by a wide margin. Across the six datasets, the average net Sharpe ratio of SSH is markedly higher than that of MV-ID, MV-MKT, and MV-NL, which often deliver Sharpe ratios close to zero once shorting costs are imposed. Even compared to stronger competitors such as MV-PCA and MAXSER, SSH remains competitive or dominant in most datasets. The main exception is the 100SINV dataset, where MAXSER attains a slightly higher net Sharpe ratio; however, this comes at the expense of substantially higher turnover.

Moreover, SSH continues to outperform the low-turnover benchmarks RTR and 1/N in most datasets, despite the fact that these strategies involve no short exposure. Overall, these results indicate that the performance advantage of SSH is not an artifact of ignoring short-sale frictions. Rather, SSH combines strong gross performance with substantially lower turnover and short exposure than the benchmark mean-variance portfolios, yielding superior net-of-cost Sharpe ratios under realistic implementation costs.

To further assess potential limits to capacity, Table A.23 reports the monthly average cross-sectional maximum and minimum portfolio weights for SSH and the benchmark strategies. Compared to the benchmark mean-variance portfolios, SSH systematically exhibits smaller maximum and minimum weights across all datasets, indicating a much lower reliance on extreme long or short positions. In contrast, standard mean-variance strategies such as MV-ID, MV-MKT, and MV-NL exhibit far more extreme weights, several times larger than those of SSH.

These relative differences also hold when comparing SSH to MAXSER. While MAXSER can occasionally deliver competitive performance, it does so with markedly larger extreme positions

Table A.23: Cross-sectional maximum and minimum portfolio weights

This table reports the monthly average for the cross-sectional maximum and minimum portfolio weights of the proposed SSH portfolio and the seven benchmarks. The first four benchmarks are mean-variance portfolios constructed with the shrinkage covariance matrices of Ledoit and Wolf (2004) (MV-ID), Ledoit and Wolf (2003) (MV-MKT), Ledoit and Wolf (2020) (MV-NL), and Chen and Yuan (2016) (MV-PCA). The remaining ones are the MAXSER portfolio of Ao, Li, and Zheng (2019), the reward-to-risk timing strategy of Kirby and Ostdiek (2012) (RTR), and the 1/N portfolio. The six datasets are described in Section 2.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Maximum weight						
SSH	0.28	0.54	0.38	0.48	1.07	0.10
MV-ID	2.85	3.18	1.60	2.59	6.56	1.12
MV-MKT	4.06	4.94	1.78	4.20	6.54	1.27
MV-NL	3.60	3.73	1.79	3.98	2.88	0.71
MV-PCA	0.75	1.01	0.31	6.07	3.25	0.12
MAXSER	1.50	2.21	0.96	1.82	7.34	0.25
RTR	0.06	0.06	0.07	0.10	0.03	0.02
1/N	0.04	0.04	0.04	0.06	0.02	0.01
Panel B: Minimum weight						
SSH	−0.20	−0.45	−0.33	−0.35	−1.04	−0.08
MV-ID	−2.63	−2.80	−1.55	−2.72	−6.39	−1.10
MV-MKT	−3.91	−4.34	−1.74	−4.74	−6.53	−1.23
MV-NL	−3.47	−3.50	−1.71	−4.20	−2.82	−0.69
MV-PCA	−0.70	−0.87	−0.26	−6.43	−3.17	−0.10
MAXSER	−1.29	−1.75	−0.93	−2.36	−7.00	−0.25
RTR	0.03	0.03	0.01	0.03	0.01	0.01
1/N	0.04	0.04	0.04	0.06	0.02	0.01

on both the long and short sides, suggesting more severe capacity constraints. By comparison, SSH achieves strong out-of-sample performance with significantly more balanced portfolios, which alleviates concerns related to scalability and short-sale constraints.[6]

[6]For the purpose of assessing capacity, what matters are relative differences in portfolio weights rather than absolute levels. In unconstrained mean-variance settings, portfolio weights can always be rescaled to achieve a desired overall level of exposure; however, such rescaling does not alter the relative concentration of positions across assets, which is the relevant determinant of implementability and capacity.

T. Can Contractionary Economic Periods Explain SSH?

To assess whether the profitability of our SSH strategy is driven by a small number of extreme economic episodes, we conduct a panel regression analysis that pools monthly observations across our six datasets described in Section 2, while explicitly controlling for the salient macro-financial stress episodes in 1973-75, 2000-02, and 2008-09, as well as broader recessionary conditions. Specifically, we estimate the following pooled panel regression:

$$r_{i,t}^{SSH} = \alpha + \beta_1 r_{i,t}^{1/N} + \beta_2 r_{i,t}^{MV\text{-}ID} + \gamma_{74} D_t^{74} + \gamma_{00} D_t^{00} + \gamma_{08} D_t^{08} + \gamma_{\text{NBER}} \text{NBER}_{t-1} + \sum_{j=2}^{6} \mu_j \mathbf{1}\{i=j\} + \varepsilon_{i,t}. \tag{A38}$$

Here, $r_{i,t}^{SSH}$ denotes the time-t return of the SSH portfolio in dataset i, $r_{i,t}^{MV\text{-}ID}$ is the time-t return on the MV-ID portfolio in dataset i, and $r_{i,t}^{1/N}$ is the time-t return on the 1/N portfolio in dataset i. The episode indicators D_t^{74}, D_t^{00}, and D_t^{08} equal one during the three contractionary economic periods 1973:01–1975:12, 2000:01–2002:12, and 2008:01–2009:12, respectively, and zero otherwise. We additionally include the official NBER recession indicator lagged by one month, NBER_{t-1}, to control for additional recessionary conditions. We include dataset fixed effects via dataset dummies $\{\mathbf{1}\{i=j\}\}_{j=2}^{6}$ and estimate Equation (A38) using OLS with heteroskedasticity-robust standard errors. Our primary object of interest is the intercept α, which captures the abnormal return of the SSH portfolio after controlling for exposures to the 1/N and MV-ID benchmark returns, as well as extreme economic episodes.

Table A.24 shows that the SSH portfolio delivers a statistically significant monthly abnormal return of $\alpha = 0.56\%$, indicating the presence of a sizable average return component that is not explained by contemporaneous benchmark returns and is not confined to periods of extreme eco-

Table A.24: Panel regression evidence on whether SSH performance is driven by major stress episodes

This table reports estimates from the pooled (unbalanced) panel regression in Equation (A38), where the dependent variable is the monthly return of the SSH strategy across the six datasets described in Section 2. The regressors include contemporaneous returns on the 1/N and MV-ID benchmark portfolios, indicators for the stress episodes 1973:01–1975:12 (D^{74}), 2000:01–2002:12 (D^{00}), and 2008:01–2009:12 (D^{08}), and the NBER recession indicator lagged one month (NBER_{t-1}). Dataset fixed effects are included via dataset dummies (dataset 1 omitted). Standard errors are heteroskedasticity-robust, and the table reports t-statistics in parentheses.

	Coefficient
α	0.0056
	(3.84)
$r^{1/N}$	0.5967
	(23.78)
$r^{MV\text{-}ID}$	0.1192
	(15.88)
D^{74}	−0.0311
	(−3.11)
D^{00}	0.0094
	(1.80)
D^{08}	−0.0053
	(−1.20)
NBER_{t-1}	−0.0064
	(−1.76)
Observations	3078

nomic events. The table also shows that SSH returns are lower during the 1973–75 episode, while the coefficients associated with the 2000–02 and 2008–09 episodes, as well as NBER recessions, are comparatively small. Overall, the evidence indicates that SSH profitability is not driven by a small number of extreme episodes, but instead reflects a pervasive abnormal-return component that is present across time and datasets.

U. Combination of the 1/N and Alpha Portfolios

In the baseline SSH specification, sentiment affects the portfolio through shrinkage of the covariance matrix, which determines the relative weight placed on the 1/N and alpha components in the mean-variance decomposition. We therefore evaluate whether the performance of SSH arises specifically from the covariance-shrinkage channel or whether a sentiment-based combination of the 1/N and alpha portfolios delivers comparable performance. Accordingly, we construct a convex combination of 1/N and the alpha component. At each rebalancing date t, the resulting portfolio is defined as:

$$\hat{\boldsymbol{w}}_t^{mix} = p_t \hat{\lambda}_t \boldsymbol{w}_{ew} + (1 - p_t)\hat{\boldsymbol{w}}_{\alpha,t},$$

where $\hat{\lambda}_t = (\boldsymbol{w}_{ew}^\top \hat{\boldsymbol{\mu}}_t)/(\gamma \boldsymbol{w}_{ew}^\top \widehat{\boldsymbol{\Sigma}}_t \boldsymbol{w}_{ew})$, $\hat{\boldsymbol{w}}_{\alpha,t} = \gamma^{-1}\widehat{\boldsymbol{\Sigma}}_t^{-1}\hat{\boldsymbol{\mu}}_t - \hat{\lambda}_t \boldsymbol{w}_{ew}$, $\widehat{\boldsymbol{\Sigma}}_t$ is the Ledoit and Wolf (2004) covariance estimator, and p_t is the probability that $\hat{\lambda}_t \boldsymbol{w}_{ew}$ outperforms $\hat{\boldsymbol{w}}_{\alpha,t}$ out of sample estimated from a logistic model.

Table A.25 reports the gross and net-of-cost Sharpe ratios, and the turnover, for the sentiment-based combination of the 1/N and alpha components (EW-α). For comparison, we also report the performance of the SSH, MV-ID, and 1/N portfolios. We observe that, net of costs, EW-α outperforms 1/N and MV-ID in five out of the six datasets. However, EW-α consistently underperforms our SSH portfolio and has higher turnover. These findings indicate that implementing sentiment-based shrinkage through covariance estimation delivers superior performance and substantially lower turnover.

Table A.25: Out-of-sample performance of combination of the 1/N and alpha portfolios

This table reports the annualized out-of-sample Sharpe ratio, and the turnover, of the combination of the 1/N and alpha portfolios introduced in Section U of the Internet Appendix, our proposed SSH portfolio, the mean-variance portfolio constructed with the shrinkage covariance matrix of Ledoit and Wolf (2004) (MV-ID), and the 1/N portfolio. The table also reports in parentheses the two-sided p-values for the significance test of the difference of each portfolio's Sharpe ratio with that of 1/N. The Sharpe ratios are either gross or net of proportional transaction costs of 30 basis points. The six datasets are described in Section 2. The numbers in bold font identify the portfolio strategies with the highest Sharpe ratio.

Strategy	25SBETA	25SBTM	25OPINV	16ANOM	46ANOM	100SINV
Panel A: Gross Sharpe ratio						
EW-α	0.63	0.72	0.74	1.05	1.77	0.96
	(0.76)	(0.47)	(0.36)	(0.02)	(0.00)	(0.06)
SSH	**0.83**	**0.91**	**0.94**	**1.10**	**2.04**	**1.08**
	(0.08)	(0.03)	(0.01)	(0.00)	(0.00)	(0.00)
MV-ID	0.53	0.47	0.67	1.00	2.00	0.87
	(0.93)	(0.73)	(0.53)	(0.04)	(0.00)	(0.15)
1/N	0.57	0.56	0.55	0.45	0.36	0.58
Panel B: Net-of-cost Sharpe ratio (30bps)						
EW-α	0.47	0.57	0.62	0.93	1.60	0.65
	(0.64)	(0.93)	(0.69)	(0.05)	(0.00)	(0.71)
SSH	**0.78**	**0.85**	**0.88**	**1.04**	**1.91**	**0.99**
	(0.14)	(0.07)	(0.02)	(0.00)	(0.00)	(0.00)
MV-ID	0.39	0.35	0.55	0.87	1.85	0.60
	(0.48)	(0.40)	(0.97)	(0.11)	(0.00)	(0.90)
1/N	0.56	0.55	0.54	0.45	0.35	0.58
Panel C: Turnover						
EW-α	1.80	1.85	1.45	1.70	14.59	3.83
SSH	0.26	0.41	0.34	0.38	1.66	0.46
MV-ID	2.00	2.10	1.40	1.68	12.49	3.64
1/N	0.04	0.04	0.04	0.04	0.04	0.04

V. Proofs of All Theoretical Results

Proof of Proposition 1

The squared correlations with the 1/N and alpha components are given by

$$\text{Corr}^2(\boldsymbol{w}_{mv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}) = \frac{(\boldsymbol{w}_{mv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{ew})^2}{\sigma_{ew}^2(\boldsymbol{w}_{mv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{mv})} = \frac{\text{SR}^2(\boldsymbol{w}_{ew})}{\text{SR}^2(\boldsymbol{w}_{mv})}, \tag{A39}$$

$$\text{Corr}^2(\boldsymbol{w}_{mv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{\alpha}^\top \boldsymbol{r}_{t+1}) = \frac{(\boldsymbol{w}_{mv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{\alpha})^2}{(\boldsymbol{w}_{\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{\alpha})(\boldsymbol{w}_{mv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{mv})} = \frac{\text{SR}^2(\boldsymbol{w}_{\alpha})}{\text{SR}^2(\boldsymbol{w}_{mv})}, \tag{A40}$$

because $\boldsymbol{w}_{mv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{ew} = \mu_{ew}/\gamma$, $\boldsymbol{w}_{mv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{mv} = \text{SR}^2(\boldsymbol{w}_{mv})/\gamma^2$, and $\boldsymbol{w}_{mv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{\alpha} = \boldsymbol{w}_{\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{\alpha} = \text{SR}^2(\boldsymbol{w}_{\alpha})/\gamma^2$. Moreover, the two squared correlations sum up to one given (6), which concludes the proof.

Proof of Proposition 2

We start with the squared correlation between the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ in (12) and its 1/N component $\boldsymbol{w}_{ew}$, which is

$$\text{Corr}^2(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}) = \frac{\left(\boldsymbol{w}_{ew}^\top \boldsymbol{\Sigma} \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\mu}\right)^2}{\sigma_{ew}^2 \gamma^2 (\boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv})} = \frac{\sigma_{ew}^2 \lambda_s^2}{\boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv}}. \tag{A41}$$

Using Equations (A54)–(A55), we have that

$$\sigma_{ew}^2 \lambda_s^2 = \frac{\text{SR}^2(\boldsymbol{w}_{ew}) \times (1 + \varepsilon(\delta))^2}{\gamma^2 (1 - \delta + \delta \bar{\sigma}^2 / \sigma_{PC_1}^2)^2}, \tag{A42}$$

where $\varepsilon(\delta)$ is defined in (A53), which, combined with the expression for $\boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv}$ in (A56), yields the result in (15).

Then, we turn to the squared correlation between the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ and its alpha component $\boldsymbol{w}_{s\alpha}$ in (13). Using

$$\gamma \boldsymbol{w}_{s\alpha}^{\top} \boldsymbol{\Sigma} \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\mu} = \gamma^2 \boldsymbol{w}_{s\alpha}^{\top} \boldsymbol{\Sigma} \boldsymbol{w}_{s\alpha} = \gamma^2 \big(\boldsymbol{w}_{smv}^{\top} \boldsymbol{\Sigma} \boldsymbol{w}_{smv} - \sigma_{ew}^2 \lambda_s^2 \big), \tag{A43}$$

we find

$$\text{Corr}^2(\boldsymbol{w}_{smv}^{\top} \boldsymbol{r}_{t+1}, \boldsymbol{w}_{s\alpha}^{\top} \boldsymbol{r}_{t+1}) = \frac{\big(\boldsymbol{w}_{s\alpha}^{\top} \boldsymbol{\Sigma} \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\mu} \big)^2}{(\boldsymbol{w}_{s\alpha}^{\top} \boldsymbol{\Sigma} \boldsymbol{w}_{s\alpha})(\boldsymbol{w}_{smv}^{\top} \boldsymbol{\Sigma} \boldsymbol{w}_{smv})} = 1 - \frac{\sigma_{ew}^2 \lambda_s^2}{\boldsymbol{w}_{smv}^{\top} \boldsymbol{\Sigma} \boldsymbol{w}_{smv}}, \tag{A44}$$

which, given (A41), is equal to $1 - \text{Corr}^2(\boldsymbol{w}_{smv}^{\top} \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^{\top} \boldsymbol{r}_{t+1})$.

It is easy to check that $s_i(\delta) \in [0, 1]$. Moreover, $s_i(\delta)$ decreases with i if it increases with $\sigma_{PC_i}^2$, which is indeed the case:

$$\frac{\partial s_i(\delta)}{\partial \sigma_{PC_i}^2} = \frac{\kappa(\delta) \sigma_{PC_1}^2}{\big(\sigma_{PC_i}^2 + \kappa(\delta)(\sigma_{PC_1}^2 - \sigma_{PC_i}^2) \big)^2} \geq 0. \tag{A45}$$

Then, $s_i(\delta)$ decreases with δ if $\kappa(\delta)$ increases with δ, which is the case too:

$$\frac{\partial \kappa(\delta)}{\partial \delta} = \frac{\bar{\sigma}^2 \sigma_{PC_1}^2}{(\delta \bar{\sigma}^2 + (1 - \delta) \sigma_{PC_1}^2)^2} \geq 0. \tag{A46}$$

Finally, when $\boldsymbol{v}_1 = \mathbf{1}/\sqrt{N}$, we have that $\varepsilon(\delta) = 0$ in (A53) because $\bar{v}_i = 0$ for all $i \geq 2$. Therefore, the squared correlation with the 1/N component simplifies to

$$\text{Corr}^2(\boldsymbol{w}_{smv}^{\top} \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^{\top} \boldsymbol{r}_{t+1}) = \frac{\text{SR}^2(\boldsymbol{w}_{ew})}{\text{SR}^2(\boldsymbol{w}_{ew}) + \sum_{i=2}^{N} s_i^2(\delta) \text{SR}^2(\boldsymbol{v}_i)}. \tag{A47}$$

This squared correlation is increasing in δ because, as shown in Proposition 4, the shrinkage factor $s_i(\delta)$ is decreasing in δ. As a result, $\text{Corr}^2(\boldsymbol{w}_{smv}^{\top} \boldsymbol{r}_{t+1}, \boldsymbol{w}_{s\alpha}^{\top} \boldsymbol{r}_{t+1})$, is decreasing in δ because it is equal to $1 - \text{Corr}^2(\boldsymbol{w}_{smv}^{\top} \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^{\top} \boldsymbol{r}_{t+1})$, which concludes the proof.

Proof of Proposition 3

The OLS solution to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in the one-factor model (21) is $\boldsymbol{\alpha} = \boldsymbol{\mu} - \boldsymbol{\beta}\mu_{ew}$ and $\boldsymbol{\beta} = \boldsymbol{\Sigma}\boldsymbol{w}_{ew}/\sigma_{ew}^2$. Therefore, we can expand $\boldsymbol{\alpha}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha}$ as

$$\boldsymbol{\alpha}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha} = \boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \mu_{ew}^2\boldsymbol{\beta}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta} - 2\mu_{ew}\boldsymbol{\beta}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}. \tag{A48}$$

Using $\boldsymbol{\beta} = \boldsymbol{\Sigma}\boldsymbol{w}_{ew}/\sigma_{ew}^2$, we have $\mu_{ew}^2\boldsymbol{\beta}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta} = \mu_{ew}\boldsymbol{\beta}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \text{SR}^2(\boldsymbol{w}_{ew})$, and thus,

$$\boldsymbol{\alpha}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha} = \boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \text{SR}^2(\boldsymbol{w}_{ew}) = \text{SR}^2(\boldsymbol{w}_\alpha), \tag{A49}$$

which concludes the proof.

Proof of Proposition 4

The mean return of the shrinkage alpha portfolio, $\boldsymbol{w}_{s\alpha}$ in (13), is

$$\begin{aligned}
\gamma\boldsymbol{w}_{s\alpha}^\top\boldsymbol{\mu} &= \gamma\big(\boldsymbol{w}_{smv}^\top\boldsymbol{\mu} - \lambda_s\mu_{ew}\big) \\
&= \gamma\boldsymbol{w}_{smv}^\top\boldsymbol{\mu} - \text{SR}^2(\boldsymbol{w}_{ew})\sum_{i=1}^{N}\frac{\bar{v}_i}{1-\delta+\delta\bar{\sigma}^2/\sigma_{PC_i}^2}\frac{\mu_{PC_i}}{\mu_{ew}} \\
&= \gamma\boldsymbol{w}_{smv}^\top\boldsymbol{\mu} - \frac{\text{SR}^2(\boldsymbol{w}_{ew})}{1-\delta+\delta\bar{\sigma}^2/\sigma_{PC_1}^2}\sum_{i=1}^{N}s_i(\delta)\bar{v}_i\frac{\mu_{PC_i}}{\mu_{ew}},
\end{aligned} \tag{A50}$$

where $\bar{v}_i = \boldsymbol{w}_{ew}^\top\boldsymbol{v}_i$. Now, notice that $\mu_{ew} = \sum_{i=1}^{N}\bar{v}_i\mu_{PC_i}$. Therefore, using

$$\gamma\boldsymbol{w}_{smv}^\top\boldsymbol{\mu} = \boldsymbol{\mu}^\top\boldsymbol{\Sigma}_s^{-1}\boldsymbol{\mu} = \sum_{i=1}^{N}\frac{\text{SR}^2(\boldsymbol{v}_i)}{1-\delta+\delta\bar{\sigma}^2/\sigma_{PC_i}^2} = \frac{\sum_{i=1}^{N}s_i(\delta)\text{SR}^2(\boldsymbol{v}_i)}{1-\delta+\delta\bar{\sigma}^2/\sigma_{PC_1}^2}, \tag{A51}$$

Equation (A50) becomes

$$\gamma\boldsymbol{w}_{s\alpha}^\top\boldsymbol{\mu} = \frac{\text{SR}^2(\boldsymbol{v}_1) + \sum_{i=2}^{N}[s_i(\delta)\text{SR}^2(\boldsymbol{v}_i)] - \text{SR}^2(\boldsymbol{w}_{ew})(1+\varepsilon(\delta))}{1-\delta+\delta\bar{\sigma}^2/\sigma_{PC_1}^2}, \tag{A52}$$

where

$$\varepsilon(\delta) = \sum_{i=2}^{N} (s_i(\delta) - 1)\bar{v}_i \frac{\mu_{PC_i}}{\mu_{ew}}. \tag{A53}$$

Then, the return variance of $\boldsymbol{w}_{s\alpha}$ is

$$\begin{aligned}\gamma^2 \boldsymbol{w}_{s\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{s\alpha} &= \gamma^2 \left(\boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv} + \lambda_s^2 \sigma_{ew}^2 - 2\lambda_s \boldsymbol{w}_{ew}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv} \right) \\ &= \gamma^2 \boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv} - \text{SR}^2(\boldsymbol{w}_{ew}) \left(\sum_{i=1}^{N} \frac{\bar{v}_i}{1 - \delta + \delta \bar{\sigma}^2 / \sigma_{PC_i}^2} \frac{\mu_{PC_i}}{\mu_{ew}} \right)^2, \end{aligned} \tag{A54}$$

where we use $\lambda_s^2 \sigma_{ew}^2 = \lambda_s \boldsymbol{w}_{ew}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv}$. Using the development in (A50), we have that

$$\sum_{i=1}^{N} \frac{\bar{v}_i}{1 - \delta + \delta \bar{\sigma}^2 / \sigma_{PC_i}^2} \frac{\mu_{PC_i}}{\mu_{ew}} = \frac{1 + \varepsilon(\delta)}{1 - \delta + \delta \bar{\sigma}^2 / \sigma_{PC_1}^2}. \tag{A55}$$

Therefore, using

$$\gamma^2 \boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv} = \boldsymbol{\mu}^\top \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\Sigma} \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\mu} = \sum_{i=1}^{N} \frac{\text{SR}^2(\boldsymbol{v}_i)}{(1 - \delta + \delta \bar{\sigma}^2 / \sigma_{PC_i}^2)^2} = \frac{\sum_{i=1}^{N} s_i^2(\delta) \text{SR}^2(\boldsymbol{v}_i)}{(1 - \delta + \delta \bar{\sigma}^2 / \sigma_{PC_1}^2)^2}, \tag{A56}$$

Equation (A54) becomes

$$\gamma^2 \boldsymbol{w}_{s\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{s\alpha} = \frac{\text{SR}^2(\boldsymbol{v}_1) + \sum_{i=2}^{N} [s_i^2(\delta) \text{SR}^2(\boldsymbol{v}_i)] - \text{SR}^2(\boldsymbol{w}_{ew})(1 + \varepsilon(\delta))^2}{(1 - \delta + \delta \bar{\sigma}^2 / \sigma_{PC_1}^2)^2}. \tag{A57}$$

Finally, combining (A52) and (A57) delivers the squared Sharpe ratio of $\boldsymbol{w}_{s\alpha}$ as

$$\text{SR}^2(\boldsymbol{w}_{s\alpha}) = \frac{\left[\text{SR}^2(\boldsymbol{v}_1) + \sum_{i=2}^{N} s_i(\delta) \text{SR}^2(\boldsymbol{v}_i) - \text{SR}^2(\boldsymbol{w}_{ew})(1 + \varepsilon(\delta)) \right]^2}{\text{SR}^2(\boldsymbol{v}_1) + \sum_{i=2}^{N} s_i^2(\delta) \text{SR}^2(\boldsymbol{v}_i) - \text{SR}^2(\boldsymbol{w}_{ew})(1 + \varepsilon(\delta))^2}. \tag{A58}$$

Equation (A58) simplifies to (24) when $\boldsymbol{v}_1 = \mathbf{1}/\sqrt{N}$ because $\text{SR}^2(\boldsymbol{w}_{ew}) = \text{SR}^2(\boldsymbol{v}_1)$ and $\varepsilon(\delta) = 0$ in that case. Finally, the condition under which $\text{SR}^2(\boldsymbol{w}_{s\alpha})$ decreases with δ is

$$\frac{\partial \text{SR}^2(\boldsymbol{w}_{s\alpha})}{\partial \delta} \leq 0$$

$$\Leftrightarrow \left(\sum_{i=2}^{N} s_i'(\delta)\text{SR}^2(\boldsymbol{v}_i)\right)\left(\sum_{i=2}^{N} s_i(\delta)^2\text{SR}^2(\boldsymbol{v}_i)\right) \leq \left(\sum_{i=2}^{N} s_i(\delta)\text{SR}^2(\boldsymbol{v}_i)\right)\left(\sum_{i=2}^{N} s_i(\delta)s_i'(\delta)\text{SR}^2(\boldsymbol{v}_i)\right)$$
$$\Leftrightarrow \sum_{i=2}^{N}\sum_{j=2}^{N} s_i'(\delta)s_j(\delta)^2\text{SR}^2(\boldsymbol{v}_i)\text{SR}^2(\boldsymbol{v}_j) \leq \sum_{i=2}^{N}\sum_{j=2}^{N} s_i(\delta)s_i'(\delta)s_j(\delta)\text{SR}^2(\boldsymbol{v}_i)\text{SR}^2(\boldsymbol{v}_j)$$
$$\Leftrightarrow \sum_{i=2}^{N}(-s_i'(\delta))\text{SR}^2(\boldsymbol{v}_i)L_i \geq 0, \tag{A59}$$

where L_i is defined in (25). This concludes the proof.

Proof of Proposition 5

The optimality conditions of the sophisticated and biased investors give the following optimal mean-variance portfolios:

$$\boldsymbol{w}_l = \frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}(\text{E}_l[\boldsymbol{d}_{t+1}] - \boldsymbol{P}_t), \tag{A60}$$

where $\boldsymbol{P}_t$ is the vector of equilibrium prices and $\text{E}_l[\boldsymbol{d}_{t+1}]$ for $l = \{s, b\}$ is the vector of expected dividends for the sophisticated (s) and biased (b) investors. Under Assumption 1, the market clearing condition is

$$\frac{\phi}{\gamma}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu_d} + \boldsymbol{b} - \boldsymbol{P}_t) + \frac{1-\phi}{\gamma}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu_d} - \boldsymbol{P}_t) = \boldsymbol{w}_{mkt}. \tag{A61}$$

From the clearing condition (A61), the vector of equilibrium prices $\boldsymbol{P}_t$ is

$$\boldsymbol{P}_t = \boldsymbol{\mu_d} - \gamma\boldsymbol{\Sigma}\boldsymbol{w}_{mkt} + \phi\boldsymbol{b}. \tag{A62}$$

Accordingly, the vector of equilibrium expected returns $\boldsymbol{\mu}$ is

$$\boldsymbol{\mu} = \text{E}[\boldsymbol{r}_{t+1}] = \text{E}[\boldsymbol{d}_{t+1} - \boldsymbol{P}_t] = \boldsymbol{\mu_d} - \boldsymbol{P}_t = \gamma\boldsymbol{\Sigma}\boldsymbol{w}_{mkt} - \phi\boldsymbol{b}, \tag{A63}$$

which corresponds to (27) and means that the Sharpe ratio of the market portfolio is

$$\text{SR}(\boldsymbol{w}_{mkt}) = \frac{\boldsymbol{w}_{mkt}^\top \boldsymbol{\mu}}{\sigma_{mkt}} = \frac{\gamma\sigma_{mkt}^2 - \phi \boldsymbol{b}^\top \boldsymbol{w}_{mkt}}{\sigma_{mkt}}, \tag{A64}$$

and this corresponds to (28). Turning to the alpha portfolio orthogonal to $\boldsymbol{w}_{mkt}$, its weights are

$$\boldsymbol{w}_\alpha = \frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \frac{\boldsymbol{w}_{mkt}^\top \boldsymbol{\mu}}{\gamma\sigma_{mkt}^2}\boldsymbol{w}_{mkt} = -\frac{\phi}{\gamma}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{b} - \frac{\boldsymbol{b}^\top \boldsymbol{w}_{mkt}}{\sigma_{mkt}^2}\boldsymbol{\Sigma}\boldsymbol{w}_{mkt}\right). \tag{A65}$$

Therefore, the mean and variance of the alpha portfolio return are

$$\gamma \boldsymbol{w}_\alpha^\top \boldsymbol{\mu} = \gamma^2 \boldsymbol{w}_\alpha^\top \boldsymbol{\Sigma} \boldsymbol{w}_\alpha = \phi^2 \left(\boldsymbol{b}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{b} - \frac{(\boldsymbol{b}^\top \boldsymbol{w}_{mkt})^2}{\sigma_{mkt}^2}\right) = \phi^2 \boldsymbol{b}^\top \boldsymbol{Q} \boldsymbol{b}, \tag{A66}$$

where $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{w}_{mkt}\boldsymbol{w}_{mkt}^\top/\sigma_{mkt}^2$ is positive semidefinite because $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{I} - \boldsymbol{u}\boldsymbol{u}^\top)\boldsymbol{\Sigma}^{-1/2}$ with $\boldsymbol{u} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{w}_{mkt}/\sigma_{mkt}$, and $\boldsymbol{I} - \boldsymbol{u}\boldsymbol{u}^\top$ is positive semidefinite. As a result, the alpha portfolio's Sharpe ratio is

$$\text{SR}(\boldsymbol{w}_\alpha) = \frac{\boldsymbol{w}_\alpha^\top \boldsymbol{\mu}}{\sqrt{\boldsymbol{w}_\alpha^\top \boldsymbol{\Sigma} \boldsymbol{w}_\alpha}} = \phi\sqrt{\boldsymbol{b}^\top \boldsymbol{Q} \boldsymbol{b}}. \tag{A67}$$

Proof of Corollary 1

Plugging $\boldsymbol{w}_{mkt} = \boldsymbol{w}_{ew}$ into (28) yields the 1/N and alpha Sharpe ratios in (29) with $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{w}_{ew}\boldsymbol{w}_{ew}^\top/\sigma_{ew}^2$. The relations that $\text{SR}(\boldsymbol{w}_{ew})$ and $\text{SR}(\boldsymbol{w}_\alpha)$ have with ϕ and $\boldsymbol{b}$ are straightforward to check. This concludes the proof.

Proof of Proposition A.1

Part 1. Under the assumption that the asset returns $(\boldsymbol{r}_{t-T+1}, \ldots, \boldsymbol{r}_{t+1})$ are i.i.d. multivariate normally distributed, we have that $\hat{\boldsymbol{\mu}}_t \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/T)$, $T\widehat{\boldsymbol{\Sigma}}_t \sim \mathcal{W}_N(T-1, \boldsymbol{\Sigma})$, and $\hat{\boldsymbol{\mu}}_t \perp \widehat{\boldsymbol{\Sigma}}_t \perp \boldsymbol{r}_{t+1}$.

Then, $\mathrm{E}\big[\widehat{\boldsymbol{\Sigma}}_t^{-1}\big] = \frac{T}{T-N-2}\boldsymbol{\Sigma}^{-1}$ if $T > N+2$ and $\mathrm{E}\big[1/(\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew})\big] = \frac{T}{T-3}1/\sigma_{ew}^2$ if $T > 3$. Therefore, the expected out-of-sample return of $\hat{\boldsymbol{w}}_\alpha$ in (A2) is

$$\begin{aligned}\gamma\mathrm{E}\big[\hat{\boldsymbol{w}}_\alpha^\top\boldsymbol{r}_{t+1}\big] = \gamma\mathrm{E}\big[\hat{\boldsymbol{w}}_\alpha^\top\boldsymbol{\mu}\big] &= \frac{T}{T-N-2}\mathrm{SR}^2(\boldsymbol{w}_{mv}) - \frac{T}{T-3}\mathrm{SR}^2(\boldsymbol{w}_{ew}) \\ &= \frac{T}{T-N-2}\bigg(\mathrm{SR}^2(\boldsymbol{w}_\alpha) + \frac{N-1}{T-3}\mathrm{SR}^2(\boldsymbol{w}_{ew})\bigg). \end{aligned} \tag{A68}$$

Then, the variance of the out-of-sample return of $\hat{\boldsymbol{w}}_\alpha$ is

$$\mathrm{Var}\big[\hat{\boldsymbol{w}}_\alpha^\top\boldsymbol{r}_{t+1}\big] = \mathrm{E}\big[\hat{\boldsymbol{w}}_\alpha^\top\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_\alpha\big] + \mathrm{Var}\big[\hat{\boldsymbol{w}}_\alpha^\top\boldsymbol{\mu}\big], \tag{A69}$$

where

$$\gamma^2\mathrm{E}\big[\hat{\boldsymbol{w}}_\alpha^\top\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_\alpha\big] = \mathrm{E}\Big[\hat{\boldsymbol{\mu}}_t^\top\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\hat{\boldsymbol{\mu}}_t\Big] + \mathrm{E}\left[\frac{\sigma_{ew}^2(\boldsymbol{w}_{ew}^\top\hat{\boldsymbol{\mu}}_t)^2}{(\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew})^2}\right] - 2\mathrm{E}\left[\frac{(\boldsymbol{w}_{ew}^\top\hat{\boldsymbol{\mu}}_t)(\boldsymbol{w}_{ew}^\top\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\hat{\boldsymbol{\mu}}_t)}{\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}}\right], \tag{A70}$$

$$\begin{aligned}\gamma^2\mathrm{Var}\big[\hat{\boldsymbol{w}}_\alpha^\top\boldsymbol{\mu}\big] = \mathrm{E}\Big[(\hat{\boldsymbol{\mu}}_t^\top\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu})^2\Big] &+ \mathrm{E}\left[\frac{\mu_{ew}^2(\boldsymbol{w}_{ew}^\top\hat{\boldsymbol{\mu}}_t)^2}{(\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew})^2}\right] - 2\mathrm{E}\left[\frac{\mu_{ew}(\boldsymbol{w}_{ew}^\top\hat{\boldsymbol{\mu}}_t)(\hat{\boldsymbol{\mu}}_t^\top\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu})}{\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}}\right] \\ &- \gamma^2\mathrm{E}^2\big[\hat{\boldsymbol{w}}_\alpha^\top\boldsymbol{\mu}\big]. \end{aligned} \tag{A71}$$

We now treat the seven expectations in (A70)–(A71) separately using the independence between $\hat{\boldsymbol{\mu}}_t$ and $\widehat{\boldsymbol{\Sigma}}_t$ and the property $\mathrm{E}[\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}] = \mathrm{E}[\boldsymbol{x}]^\top\boldsymbol{A}\mathrm{E}[\boldsymbol{x}] + \mathrm{tr}(\mathrm{Var}[\boldsymbol{x}]\boldsymbol{A})$ when $\boldsymbol{x}$ is a random vector and $\boldsymbol{A}$ a constant matrix. For the first expectation in (A70), we have

$$\mathrm{E}\Big[\hat{\boldsymbol{\mu}}_t^\top\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\hat{\boldsymbol{\mu}}_t\Big] = \boldsymbol{\mu}^\top\mathrm{E}\big[\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\big]\boldsymbol{\mu} + \frac{1}{T}\mathrm{tr}\Big(\mathrm{E}\big[\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\big]\boldsymbol{\Sigma}\Big). \tag{A72}$$

From Kan and Zhou (2007), we have that $\mathrm{E}\big[\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\big] = \frac{T^2(T-2)}{(T-N-1)(T-N-2)(T-N-4)}\boldsymbol{\Sigma}^{-1}$ if $T > N+4$. Therefore, (A72) becomes

$$\mathrm{E}\Big[\hat{\boldsymbol{\mu}}_t^\top\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\hat{\boldsymbol{\mu}}_t\Big] = \frac{T^2(T-2)}{(T-N-1)(T-N-2)(T-N-4)}\left(\mathrm{SR}^2(\boldsymbol{w}_{mv}) + \frac{N}{T}\right). \tag{A73}$$

The second expectation in (A70) is

$$\mathrm{E}\left[\frac{\sigma_{ew}^2(\boldsymbol{w}_{ew}^\top\hat{\boldsymbol{\mu}}_t)^2}{(\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew})^2}\right] = \mathrm{E}\big[(\boldsymbol{w}_{ew}^\top\hat{\boldsymbol{\mu}}_t)^2\big] \times \mathrm{E}\left[\frac{\sigma_{ew}^2}{(\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew})^2}\right]. \tag{A74}$$

We have $\mathrm{E}\big[(\boldsymbol{w}_{ew}^\top\hat{\boldsymbol{\mu}}_t)^2\big] = \mu_{ew}^2 + \sigma_{ew}^2/T$ and, given that $T\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew} \sim \sigma_{ew}^2\chi_{T-1}^2$, we have

$$\mathrm{E}\Big[\big(\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}\big)^{-2}\Big] = \frac{T^2}{(T-3)(T-5)\sigma_{ew}^4} \tag{A75}$$

if $T > 5$. As a result, (A74) becomes

$$\mathrm{E}\left[\frac{\sigma_{ew}^2(\boldsymbol{w}_{ew}^\top\hat{\boldsymbol{\mu}}_t)^2}{(\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew})^2}\right] = \frac{T^2}{(T-3)(T-5)}\left(\mathrm{SR}^2(\boldsymbol{w}_{ew}) + \frac{1}{T}\right). \tag{A76}$$

The third expectation in (A70) is

$$\mathrm{E}\left[\frac{(\boldsymbol{w}_{ew}^\top\hat{\boldsymbol{\mu}}_t)(\boldsymbol{w}_{ew}^\top\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\hat{\boldsymbol{\mu}}_t)}{\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}}\right] = \mu_{ew}\boldsymbol{w}_{ew}^\top\boldsymbol{\Sigma}\mathrm{E}\left[\frac{\widehat{\boldsymbol{\Sigma}}_t^{-1}}{\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}}\right]\boldsymbol{\mu} + \frac{1}{T}\left(\boldsymbol{w}_{ew}^\top\boldsymbol{\Sigma}\mathrm{E}\left[\frac{\widehat{\boldsymbol{\Sigma}}_t^{-1}}{\boldsymbol{w}_{ew}^\top\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}}\right]\boldsymbol{\Sigma}\boldsymbol{w}_{ew}\right). \tag{A77}$$

We proceed with the following lemma, which we prove in the next section.

Lemma 1 *Let* $\boldsymbol{W} \sim \mathcal{W}_m(n, \boldsymbol{\Sigma})$, $\boldsymbol{a}$ *be an* $m \times 1$ *constant vector,* $n > 4$, *and* $n > m+1$. *Then,*

$$\mathrm{E}\left[\frac{\boldsymbol{W}^{-1}}{\boldsymbol{a}^\top\boldsymbol{W}\boldsymbol{a}}\right] = \frac{1}{(n-2)(n-m-1)}\left(\frac{\boldsymbol{\Sigma}^{-1}}{\boldsymbol{a}^\top\boldsymbol{\Sigma}\boldsymbol{a}} + \frac{2}{n-4}\frac{\boldsymbol{a}\boldsymbol{a}^\top}{(\boldsymbol{a}^\top\boldsymbol{\Sigma}\boldsymbol{a})^2}\right). \tag{A78}$$

Using Lemma 1 and $T\widehat{\boldsymbol{\Sigma}}_t \sim \mathcal{W}_N(T-1, \boldsymbol{\Sigma})$, we find that if $T > 5$ and $T > N+2$,

$$\mathrm{E}\left[\frac{\widehat{\boldsymbol{\Sigma}}_t^{-1}}{\boldsymbol{w}_{ew}^\top \widehat{\boldsymbol{\Sigma}}_t \boldsymbol{w}_{ew}}\right] = \frac{T^2}{(T-3)(T-N-2)}\left(\frac{\boldsymbol{\Sigma}^{-1}}{\sigma_{ew}^2} + \frac{2\boldsymbol{w}_{ew}\boldsymbol{w}_{ew}^\top}{(T-5)\sigma_{ew}^4}\right). \tag{A79}$$

Plugging (A79) into (A77), the third expectation in (A70) becomes

$$\mathrm{E}\left[\frac{(\boldsymbol{w}_{ew}^\top \hat{\boldsymbol{\mu}}_t)(\boldsymbol{w}_{ew}^\top \boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_t^{-1}\hat{\boldsymbol{\mu}}_t)}{\boldsymbol{w}_{ew}^\top \widehat{\boldsymbol{\Sigma}}_t \boldsymbol{w}_{ew}}\right] = \frac{T^2}{(T-5)(T-N-2)}\left(\mathrm{SR}^2(\boldsymbol{w}_{ew}) + \frac{1}{T}\right). \tag{A80}$$

The first expectation in (A71) is given by

$$\mathrm{E}\Big[(\hat{\boldsymbol{\mu}}_t^\top \widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu})^2\Big] = \boldsymbol{\mu}^\top \mathrm{E}\Big[\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^\top \widehat{\boldsymbol{\Sigma}}_t^{-1}\Big]\boldsymbol{\mu} + \frac{1}{T}\mathrm{tr}\Big(\mathrm{E}\Big[\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^\top \widehat{\boldsymbol{\Sigma}}_t^{-1}\Big]\boldsymbol{\Sigma}\Big). \tag{A81}$$

We then use the property that, given $\boldsymbol{W} \sim \mathcal{W}_m(n, \boldsymbol{\Sigma})$, a constant symmetric matrix $\boldsymbol{A}$, and $n > m+3$,

$$\mathrm{E}\big[\boldsymbol{W}^{-1}\boldsymbol{A}\boldsymbol{W}^{-1}\big] = \frac{\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{A})\boldsymbol{\Sigma}^{-1} + (n-m-1)\boldsymbol{\Sigma}^{-1}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}}{(n-m)(n-m-1)(n-m-3)}. \tag{A82}$$

This gives, provided $T > N+4$,

$$\mathrm{E}\Big[\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^\top \widehat{\boldsymbol{\Sigma}}_t^{-1}\Big] = \frac{T^2\big(\mathrm{SR}^2(\boldsymbol{w}_{mv})\boldsymbol{\Sigma}^{-1} + (T-N-2)\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\big)}{(T-N-1)(T-N-2)(T-N-4)}. \tag{A83}$$

Plugging (A83) into (A81), the first expectation in (A71) becomes

$$\mathrm{E}\Big[(\hat{\boldsymbol{\mu}}_t^\top \widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu})^2\Big] = \frac{T^2\mathrm{SR}^2(\boldsymbol{w}_{mv})\big((T-N-1)\mathrm{SR}^2(\boldsymbol{w}_{mv}) + \frac{T-2}{T}\big)}{(T-N-1)(T-N-2)(T-N-4)}. \tag{A84}$$

Then, from (A76), the second expectation in (A71) is

$$\mathrm{E}\left[\frac{\mu_{ew}^2(\boldsymbol{w}_{ew}^\top \hat{\boldsymbol{\mu}}_t)^2}{(\boldsymbol{w}_{ew}^\top \widehat{\boldsymbol{\Sigma}}_t \boldsymbol{w}_{ew})^2}\right] = \frac{T^2\mathrm{SR}^2(\boldsymbol{w}_{ew})\big(\mathrm{SR}^2(\boldsymbol{w}_{ew}) + \frac{1}{T}\big)}{(T-3)(T-5)}. \tag{A85}$$

For the third expectation in (A71), we have

$$\mathrm{E}\left[\frac{\mu_{ew}(\boldsymbol{w}_{ew}^{\top}\hat{\boldsymbol{\mu}}_t)(\hat{\boldsymbol{\mu}}_t^{\top}\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu})}{\boldsymbol{w}_{ew}^{\top}\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}}\right] = \mu_{ew}^2\boldsymbol{\mu}^{\top}\mathrm{E}\left[\frac{\widehat{\boldsymbol{\Sigma}}_t^{-1}}{\boldsymbol{w}_{ew}^{\top}\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}}\right]\boldsymbol{\mu} + \frac{\mu_{ew}}{T}\boldsymbol{\mu}^{\top}\mathrm{E}\left[\frac{\widehat{\boldsymbol{\Sigma}}_t^{-1}}{\boldsymbol{w}_{ew}^{\top}\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}}\right]\boldsymbol{\Sigma}\boldsymbol{w}_{ew}, \tag{A86}$$

which, using (A79), becomes

$$\mathrm{E}\left[\frac{\mu_{ew}(\boldsymbol{w}_{ew}^{\top}\hat{\boldsymbol{\mu}}_t)(\hat{\boldsymbol{\mu}}_t^{\top}\widehat{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu})}{\boldsymbol{w}_{ew}^{\top}\widehat{\boldsymbol{\Sigma}}_t\boldsymbol{w}_{ew}}\right] = \frac{T^2\mathrm{SR}^2(\boldsymbol{w}_{ew})}{(T-3)(T-N-2)}\left(\mathrm{SR}^2(\boldsymbol{w}_{mv}) + \frac{2\mathrm{SR}^2(\boldsymbol{w}_{ew})}{T-5} + \frac{T-3}{T(T-5)}\right). \tag{A87}$$

Finally, using (A68), the fourth expectation in (A71) is given by

$$\mathrm{E}^2\left[\hat{\boldsymbol{w}}_{\alpha}^{\top}\boldsymbol{\mu}\right] = \frac{T^2}{(T-N-2)^2}\left(\mathrm{SR}^4(\boldsymbol{w}_{\alpha}) + \frac{(N-1)^2}{(T-3)^2}\mathrm{SR}^4(\boldsymbol{w}_{ew}) + \frac{2(N-1)}{T-3}\mathrm{SR}^2(\boldsymbol{w}_{\alpha})\mathrm{SR}^2(\boldsymbol{w}_{ew})\right). \tag{A88}$$

Plugging (A73), (A76), (A80), (A84), (A85), (A87), and (A88) into (A70)–(A71) and letting $\mathrm{SR}^2(\boldsymbol{w}_{mv}) = \mathrm{SR}^2(\boldsymbol{w}_{\alpha}) + \mathrm{SR}^2(\boldsymbol{w}_{ew})$, we find that the variance of the out-of-sample return of $\hat{\boldsymbol{w}}_{\alpha}$ in (A69) becomes

$$\mathrm{Var}\left[\hat{\boldsymbol{w}}_{\alpha}^{\top}\boldsymbol{r}_{t+1}\right] = \frac{T^2\left(a\mathrm{SR}^4(\boldsymbol{w}_{\alpha}) + b\mathrm{SR}^2(\boldsymbol{w}_{\alpha})\mathrm{SR}^2(\boldsymbol{w}_{ew}) + c\mathrm{SR}^4(\boldsymbol{w}_{ew}) + d\mathrm{SR}^2(\boldsymbol{w}_{\alpha}) + e\mathrm{SR}^2(\boldsymbol{w}_{ew}) + f\right)}{\gamma^2(T-N-2)^2}, \tag{A89}$$

where $a, b, c, d, e, f \geq 0$ are the following functions of N and T:

$$a = \frac{2}{T-N-4}, \tag{A90}$$

$$b = \frac{4}{T-N-4}, \tag{A91}$$

$$c = \frac{2(N-1)[T(T-7) + N(T-N-3) + 13]}{(T-3)^2(T-5)(T-N-4)}, \tag{A92}$$

$$d = \frac{(T+1)(T-2)(T-N-2)}{T(T-N-1)(T-N-4)}, \tag{A93}$$

$$e = \frac{(N-1)(T+1)(T-N-2)[T(T-7)+N(T-N-2)+14]}{T(T-3)(T-5)(T-N-1)(T-N-4)}, \tag{A94}$$

$$f = \frac{(N-1)(T-N-2)[(T-1)(T-4)^2+N(T-N-2)]}{T(T-3)(T-5)(T-N-1)(T-N-4)}. \tag{A95}$$

Combining (A68) and (A89), the out-of-sample Sharpe ratio of $\hat{\boldsymbol{w}}_\alpha$ corresponds to (A5).

Part 2. As $N \to \infty$, $T \to \infty$, and $N/T \to \rho \in [0,1)$, we have $a, b, c \to 0$, $d \to 1/(1-\rho)$, $e \to \rho(1+\rho-\rho^2)/(1-\rho)$, and $f \to \rho/(1-\rho)$, which gives the asymptotic out-of-sample Sharpe ratio of $\hat{\boldsymbol{w}}_\alpha$ in (A6).

Part 3. As $N \to \infty$, $T \to \infty$, and $N/T \to \rho \in [0,1)$, we have from (A6) that the out-of-sample Sharpe ratio loss of $\hat{\boldsymbol{w}}_\alpha$ tends to

$$\mathrm{L}(\hat{\boldsymbol{w}}_\alpha) \to 1 - \frac{\sqrt{1-\rho}\big(\mathrm{SR}(\boldsymbol{w}_\alpha) + \rho\mathrm{SR}^2(\boldsymbol{w}_{ew})/\mathrm{SR}(\boldsymbol{w}_\alpha)\big)}{\sqrt{\mathrm{SR}^2(\boldsymbol{w}_\alpha) + \rho(1+\rho-\rho^2)\mathrm{SR}^2(\boldsymbol{w}_{ew}) + \rho}}. \tag{A96}$$

After some developments, we can show that the derivative of $\mathrm{L}(\hat{\boldsymbol{w}}_\alpha)$ with respect to $\mathrm{SR}(\boldsymbol{w}_\alpha)$ is negative if and only if

$$\left(1 - \rho\frac{\mathrm{SR}^2(\boldsymbol{w}_{ew})}{\mathrm{SR}^2(\boldsymbol{w}_\alpha)}\right)\big(\mathrm{SR}^2(\boldsymbol{w}_\alpha) + \rho(1+\rho-\rho^2)\mathrm{SR}^2(\boldsymbol{w}_{ew}) + \rho\big) \geq \mathrm{SR}^2(\boldsymbol{w}_\alpha) + \rho\mathrm{SR}^2(\boldsymbol{w}_{ew}). \tag{A97}$$

We find that if $(1-\rho+\rho^2)\mathrm{SR}^2(\boldsymbol{w}_{ew}) < 1$, which holds in particular if $\mathrm{SR}^2(\boldsymbol{w}_{ew}) < 1$, the inequality (A97) becomes

$$\frac{\mathrm{SR}^2(\boldsymbol{w}_\alpha)}{\mathrm{SR}^2(\boldsymbol{w}_{ew})} \geq \rho \times \frac{1 + (1+\rho-\rho^2)\mathrm{SR}^2(\boldsymbol{w}_{ew})}{1 - (1-\rho+\rho^2)\mathrm{SR}^2(\boldsymbol{w}_{ew})}. \tag{A98}$$

We can show that this lower bound is lower than one if $\text{SR}^2(\boldsymbol{w}_{ew}) \leq (1-\rho)/(1+2\rho^2-\rho^3)$, in which case the condition becomes $\text{SR}^2(\boldsymbol{w}_{\alpha}) \geq \text{SR}^2(\boldsymbol{w}_{ew})$. The former condition and $\text{SR}^2(\boldsymbol{w}_{ew}) < 1$ can be combined in the condition $\text{SR}^2(\boldsymbol{w}_{ew}) < (1-\rho)/(1+2\rho^2-\rho^3)$. This concludes the proof.

Proof of Lemma 1

Denote

$$\boldsymbol{f}(\boldsymbol{a}) = \text{E}\left[\frac{\boldsymbol{W}^{-1}}{\boldsymbol{a}^\top \boldsymbol{W} \boldsymbol{a}}\right]. \tag{A99}$$

Let $\boldsymbol{S} = \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{W} \boldsymbol{\Sigma}^{-\frac{1}{2}} \sim \mathcal{W}_m(n, \boldsymbol{I}_m)$. We have

$$\boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{\Sigma}^{-\frac{1}{2}} \text{E}\left[\frac{\boldsymbol{S}^{-1}}{\tilde{\boldsymbol{a}}^\top \boldsymbol{S} \tilde{\boldsymbol{a}}}\right] \boldsymbol{\Sigma}^{-\frac{1}{2}}, \tag{A100}$$

where $\tilde{\boldsymbol{a}} = \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{a}$. Consider an $m \times m$ orthonormal matrix $\boldsymbol{P}$ whose first column is equal to $\tilde{\boldsymbol{a}}/(\tilde{\boldsymbol{a}}^\top \tilde{\boldsymbol{a}})^{\frac{1}{2}}$. Define

$$\tilde{\boldsymbol{S}} = \boldsymbol{P}^\top \boldsymbol{S} \boldsymbol{P} \sim \mathcal{W}_m(n, \boldsymbol{I}_m). \tag{A101}$$

Then, we can write

$$\boldsymbol{f}(\boldsymbol{a}) = \frac{1}{\tilde{\boldsymbol{a}}^\top \tilde{\boldsymbol{a}}} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}^\top \text{E}\left[\frac{\tilde{\boldsymbol{S}}^{-1}}{\tilde{S}_{11}}\right] \boldsymbol{P} \boldsymbol{\Sigma}^{-\frac{1}{2}} = \frac{1}{\boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{a}} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}^\top \text{E}\left[\frac{\tilde{\boldsymbol{S}}^{-1}}{\tilde{S}_{11}}\right] \boldsymbol{P} \boldsymbol{\Sigma}^{-\frac{1}{2}}, \tag{A102}$$

where

$$\tilde{S}_{11} = \boldsymbol{e}_1^\top \tilde{\boldsymbol{S}} \boldsymbol{e}_1 \tag{A103}$$

is the $(1,1)$ element of $\tilde{\boldsymbol{S}}$. Therefore, we can focus on obtaining $\text{E}[\tilde{\boldsymbol{S}}^{-1}/\tilde{S}_{11}]$. By symmetry, this expectation is equal to

$$\text{E}\left[\frac{\tilde{\boldsymbol{S}}^{-1}}{\tilde{S}_{11}}\right] = \begin{bmatrix} c_1 & \boldsymbol{0}_{m-1}^\top \\ \boldsymbol{0}_{m-1} & c_2 \boldsymbol{I}_{m-1} \end{bmatrix} = (c_1 - c_2) \boldsymbol{e}_1 \boldsymbol{e}_1^\top + c_2 \boldsymbol{I}_m, \tag{A104}$$

where c_1 and c_2 are two constants to be determined. This gives us

$$\boldsymbol{f}(\boldsymbol{a}) = \frac{c_2 \boldsymbol{\Sigma}^{-1}}{\boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{a}} + (c_1 - c_2) \frac{\boldsymbol{a}\boldsymbol{a}^\top}{(\boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{a})^2}. \tag{A105}$$

Using the formula for the inverse of a partitioned matrix, we obtain

$$\tilde{\boldsymbol{S}}^{-1} = \begin{bmatrix} \tilde{S}^{11} & \tilde{\boldsymbol{S}}^{12} \\ \tilde{\boldsymbol{S}}^{21} & \tilde{\boldsymbol{S}}^{22} \end{bmatrix}, \tag{A106}$$

where

$$\tilde{S}^{11} = (\tilde{S}_{11} - \tilde{\boldsymbol{S}}_{12} \tilde{\boldsymbol{S}}_{22}^{-1} \tilde{\boldsymbol{S}}_{21})^{-1} \equiv \tilde{S}_{11 \cdot 2}^{-1}, \tag{A107}$$

$$\tilde{\boldsymbol{S}}^{12} = -\tilde{S}_{11 \cdot 2}^{-1} \tilde{\boldsymbol{S}}_{12} \tilde{\boldsymbol{S}}_{22}^{-1}, \tag{A108}$$

$$\tilde{\boldsymbol{S}}^{22} = \tilde{\boldsymbol{S}}_{22}^{-1} + \tilde{\boldsymbol{S}}_{22}^{-1} \tilde{\boldsymbol{S}}_{21} \tilde{S}_{11 \cdot 2}^{-1} \tilde{\boldsymbol{S}}_{12} \tilde{\boldsymbol{S}}_{22}^{-1}. \tag{A109}$$

It is well known that $\tilde{S}_{11 \cdot 2} \sim \chi^2_{n-m+1}$, $\tilde{\boldsymbol{S}}_{22} \sim \mathcal{W}_{m-1}(n, \boldsymbol{I}_{m-1})$, $\boldsymbol{z} \equiv \tilde{\boldsymbol{S}}_{22}^{-\frac{1}{2}} \tilde{\boldsymbol{S}}_{21} \sim \mathcal{N}(\boldsymbol{0}_{m-1}, \boldsymbol{I}_{m-1})$, and they are mutually independent. It follows that

$$c_1 = \mathrm{E}\left[\frac{\tilde{S}_{11 \cdot 2}^{-1}}{\tilde{S}_{11}} \right]. \tag{A110}$$

Note that we can write

$$\tilde{S}_{11} = \tilde{S}_{11 \cdot 2} + \tilde{\boldsymbol{S}}_{12} \tilde{\boldsymbol{S}}_{22}^{-1} \tilde{\boldsymbol{S}}_{21} = \tilde{\boldsymbol{S}}_{11 \cdot 2} + \boldsymbol{z}^\top \boldsymbol{z} \equiv x_1 + x_2, \tag{A111}$$

where $x_1 = \tilde{S}_{11 \cdot 2} \sim \chi^2_{n-m+1}$, $x_2 = \boldsymbol{z}^\top \boldsymbol{z} \sim \chi^2_{m-1}$, and they are mutually independent. As a result, we have

$$c_1 = \mathrm{E}\left[\frac{1}{x_1 (x_1 + x_2)} \right] = \mathrm{E}\left[\left(\frac{x_1 + x_2}{x_1} \right) \frac{1}{(x_1 + x_2)^2} \right]. \tag{A112}$$

Since $b = x_1/(x_1 + x_2) \sim \text{Beta}((n-m+1)/2, (m-1)/2)$ is independent of $x_1 + x_2 \sim \chi^2_n$, we have

$$c_1 = \text{E}[b^{-1}]\text{E}[(x_1+x_2)^{-2}] = \left(\frac{n-2}{n-m-1}\right)\frac{1}{(n-2)(n-4)} = \frac{1}{(n-4)(n-m-1)}, \quad \text{(A113)}$$

assuming $n > 4$ and $n > m+1$. For the $(2,2)$ block of $\text{E}[\tilde{\boldsymbol{S}}^{-1}/\tilde{S}_{11}]$, we have

$$\begin{aligned}\text{E}\left[\frac{1}{(x_1+\boldsymbol{z}^\top\boldsymbol{z})}\tilde{\boldsymbol{S}}_{22}^{-\frac{1}{2}}\left(\boldsymbol{I}_{m-1}+\frac{\boldsymbol{z}\boldsymbol{z}^\top}{x_1}\right)\tilde{\boldsymbol{S}}_{22}^{-\frac{1}{2}}\right] &= \text{E}\left[\frac{\tilde{\boldsymbol{S}}_{22}^{-1}}{x_1+\boldsymbol{z}^\top\boldsymbol{z}}\right] + \text{E}\left[\tilde{\boldsymbol{S}}_{22}^{-\frac{1}{2}}\frac{\boldsymbol{z}\boldsymbol{z}^\top}{x_1(x_1+\boldsymbol{z}^\top\boldsymbol{z})}\tilde{\boldsymbol{S}}_{22}^{-\frac{1}{2}}\right] \\ &= \frac{\boldsymbol{I}_{m-1}}{(n-2)(n-m)} + \text{E}\left[\tilde{\boldsymbol{S}}_{22}^{-\frac{1}{2}}\frac{\boldsymbol{z}\boldsymbol{z}^\top}{x_1(x_1+\boldsymbol{z}^\top\boldsymbol{z})}\tilde{\boldsymbol{S}}_{22}^{-\frac{1}{2}}\right].\end{aligned} \quad \text{(A114)}$$

We can show that

$$\text{E}\left[\frac{\boldsymbol{z}\boldsymbol{z}^\top}{x_1(x_1+\boldsymbol{z}^\top\boldsymbol{z})}\right] = \frac{\boldsymbol{I}_{m-1}}{(n-2)(n-m-1)}, \quad \text{(A115)}$$

and thus, we have

$$\text{E}\left[\tilde{\boldsymbol{S}}_{22}^{-\frac{1}{2}}\frac{\boldsymbol{z}\boldsymbol{z}^\top}{x_1(x_1+\boldsymbol{z}^\top\boldsymbol{z})}\tilde{\boldsymbol{S}}_{22}^{-\frac{1}{2}}\right] = \frac{\text{E}[\tilde{\boldsymbol{S}}_{22}^{-1}]}{(n-2)(n-m-1)} = \frac{\boldsymbol{I}_{m-1}}{(n-2)(n-m-1)(n-m)}. \quad \text{(A116)}$$

Finally, this means that

$$c_2 = \frac{1}{(n-2)(n-m)} + \frac{1}{(n-2)(n-m-1)(n-m)} = \frac{1}{(n-2)(n-m-1)}, \quad \text{(A117)}$$

which concludes the proof after plugging c_1 and c_2 into (A105).

Proof of Proposition A.2

Part 1. The squared correlation between the return of the 1/N portfolio, $\boldsymbol{w}_{ew} \propto \mathbf{1}$, and that of the shrinkage mean-variance portfolio with $\delta = 1$, $\boldsymbol{w}_{smv} \propto \boldsymbol{\mu}$, is

$$\text{Corr}^2(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}) = \frac{(\mathbf{1}^\top \boldsymbol{\Sigma} \boldsymbol{\mu})^2}{(\mathbf{1}^\top \boldsymbol{\Sigma} \mathbf{1})(\boldsymbol{\mu}^\top \boldsymbol{\Sigma} \boldsymbol{\mu})} = 1 - \frac{(\mathbf{1}^\top \boldsymbol{\Sigma} \mathbf{1})(\boldsymbol{\mu}^\top \boldsymbol{\Sigma} \boldsymbol{\mu}) - (\mathbf{1}^\top \boldsymbol{\Sigma} \boldsymbol{\mu})^2}{(\mathbf{1}^\top \boldsymbol{\Sigma} \mathbf{1})(\boldsymbol{\mu}^\top \boldsymbol{\Sigma} \boldsymbol{\mu})}. \quad \text{(A118)}$$

Replacing $\mu = \eta + \bar{\mu}\mathbf{1}$ in (A118) gives, after some developments,

$$\text{Corr}^2(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}) = 1 - \frac{(\mathbf{1}^\top \boldsymbol{\Sigma} \mathbf{1})(\boldsymbol{\eta}^\top \boldsymbol{\Sigma} \boldsymbol{\eta}) - (\mathbf{1}^\top \boldsymbol{\Sigma} \boldsymbol{\eta})^2}{(\mathbf{1}^\top \boldsymbol{\Sigma} \mathbf{1})(\boldsymbol{\mu}^\top \boldsymbol{\Sigma} \boldsymbol{\mu})}, \quad \text{(A119)}$$

which corresponds to $1 - \boldsymbol{x}^\top \boldsymbol{\Sigma} \boldsymbol{\eta}$ with $\boldsymbol{x}$ defined in (A12), as desired.

Part 2. The squared Euclidean norm in (A13) is

$$\left|\left| \boldsymbol{w}_{ew} - \frac{1}{N\bar{\mu}} \boldsymbol{\mu} \right|\right|^2 = \boldsymbol{w}_{ew}^\top \boldsymbol{w}_{ew} - \frac{2\boldsymbol{w}_{ew}^\top \boldsymbol{\mu}}{N\bar{\mu}} + \frac{\boldsymbol{\mu}^\top \boldsymbol{\mu}}{N^2 \bar{\mu}^2} = \frac{\boldsymbol{\mu}^\top \boldsymbol{\mu}}{N^2 \bar{\mu}^2} - \frac{1}{N}. \quad \text{(A120)}$$

Now, using $\boldsymbol{\mu} = \boldsymbol{\eta} + \bar{\mu}\mathbf{1}$, we have that

$$\frac{\boldsymbol{\mu}^\top \boldsymbol{\mu}}{N} = \frac{\boldsymbol{\eta}^\top \boldsymbol{\eta} + N\bar{\mu}^2 + 2\bar{\mu}\mathbf{1}^\top \boldsymbol{\eta}}{N} = \bar{\eta}_2 + \bar{\mu}^2 + 2\bar{\mu}\bar{\eta}_1. \quad \text{(A121)}$$

Replacing (A121) into (A120) gives the desired result in (A13), which concludes the proof.

Proof of Proposition A.3

We start with the squared correlation between the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ in (A15) and its 1/N component $\boldsymbol{w}_{ew}$, which is

$$\text{Corr}^2(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}) = \frac{(\boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{ew})^2}{\sigma_{ew}^2 (\boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv})}. \quad \text{(A122)}$$

Using $\sigma_g^2 = (\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1})^{-1}$ and $\mu_g/\sigma_g^2 = \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$, we have

$$\gamma \boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{ew} = (1-\delta)\mu_{ew} + \delta \mu_{ew} = \mu_{ew}, \tag{A123}$$

$$\begin{aligned}\gamma^2 \boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv} &= (1-\delta)^2(\text{SR}^2(\boldsymbol{w}_{ew}) + \text{SR}^2(\boldsymbol{w}_\alpha)) + \delta^2 \frac{\mu_{ew}^2}{\sigma_g^2} + 2\delta(1-\delta)\frac{\mu_{ew}\mu_g}{\sigma_g^2} \\ &= \text{SR}^2(\boldsymbol{w}_{ew})(1+\eta(\delta)) + (1-\delta)^2 \text{SR}^2(\boldsymbol{w}_\alpha),\end{aligned} \tag{A124}$$

where $\eta(\delta)$ is defined in the proposition statement. Plugging (A123)–(A124) into (A122) yields the squared correlation in (A16).

Then, we turn to the squared correlation between the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ and its alpha component $\boldsymbol{w}_{s\alpha} = \boldsymbol{w}_{smv} - \lambda \boldsymbol{w}_{ew}$, which is

$$\text{Corr}^2(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{s\alpha}^\top \boldsymbol{r}_{t+1}) = \frac{(\boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{s\alpha})^2}{(\boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv})(\boldsymbol{w}_{s\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{s\alpha})}, \tag{A125}$$

where $\boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv}$ is given by (A124). Regarding $\boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{s\alpha}$ and $\boldsymbol{w}_{s\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{s\alpha}$, we have

$$\begin{aligned}\gamma^2 \boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{s\alpha} = \gamma^2 \boldsymbol{w}_{s\alpha}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{s\alpha} &= \gamma^2 \boldsymbol{w}_{smv}^\top \boldsymbol{\Sigma} \boldsymbol{w}_{smv} - \text{SR}^2(\boldsymbol{w}_{ew}) \\ &= \eta(\delta)\text{SR}^2(\boldsymbol{w}_{ew}) + (1-\delta)^2 \text{SR}^2(\boldsymbol{w}_\alpha).\end{aligned} \tag{A126}$$

Plugging (A124) and (A126) into (A125) yields $\text{Corr}^2\big(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{s\alpha}^\top \boldsymbol{r}_{t+1}\big) = 1 - \text{Corr}^2\big(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1}\big)$, as in (A17).

Finally, when the first eigenvector $\boldsymbol{v}_1 = \mathbf{1}/\sqrt{N}$, we have $\boldsymbol{\Sigma}^{-1}\mathbf{1} \propto \mathbf{1}$, and thus, the GMV and 1/N portfolios are equal: $\boldsymbol{w}_g = (\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1})^{-1}\boldsymbol{\Sigma}^{-1}\mathbf{1} = \boldsymbol{w}_{ew}$. As a result, we have $R_{\sigma^2} = R_\mu = 1$, and thus, $\eta(\delta) = 0$. When $\eta(\delta) = 0$, it is clear that the squared correlation $\text{Corr}^2(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1})$ in (A16) increases with δ. Therefore, $\text{Corr}^2(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{s\alpha}^\top \boldsymbol{r}_{t+1})$ is decreasing in δ because it is equal to $1 - \text{Corr}^2(\boldsymbol{w}_{smv}^\top \boldsymbol{r}_{t+1}, \boldsymbol{w}_{ew}^\top \boldsymbol{r}_{t+1})$, which concludes the proof.

Proof of Proposition A.4

Proposition A.4 is a special case of Proposition A.3 when the first eigenvector $\boldsymbol{v}_1 = \mathbf{1}/\sqrt{N}$. Indeed, in that case, the shrinkage mean-variance portfolio $\boldsymbol{w}_{smv}$ with shrinkage of weights in (A20) corresponds to that with shrinkage of the mean in (A15). This is because, when $\boldsymbol{v}_1 = \mathbf{1}/\sqrt{N}$, we have $\boldsymbol{\Sigma}^{-1}\mathbf{1} = \boldsymbol{w}_{ew}/\sigma^2_{ew}$, and thus, $(\mu_{ew}/\gamma)\boldsymbol{\Sigma}^{-1}\mathbf{1} = \lambda \boldsymbol{w}_{ew}$. Therefore, Proposition A.4 corresponds to Proposition A.3 with $\eta(\delta) = 0$, which is because $\eta(\delta) = 0$ when $\boldsymbol{v}_1 = \mathbf{1}/\sqrt{N}$ as shown in the proof of Proposition A.3. This concludes the proof.

Proof of Proposition A.5

Because the covariance matrix $\boldsymbol{\Sigma}$ is diagonal, we can separate the Karush-Kuhn-Tucker (KKT) conditions of the three investor groups across individual assets. For asset i, the KKT condition for the rational, optimistic, and pessimistic investors is, respectively,

$$\mu_{d,i} - P_{i,t} - \gamma\sigma_i^2 w_i^r = 0, \tag{A127}$$

$$\mu_{d,i} + b_i - P_{i,t} - \gamma\sigma_i^2 w_i^+ + \lambda_i^+ = 0, \tag{A128}$$

$$\mu_{d,i} - b_i - P_{i,t} - \gamma\sigma_i^2 w_i^- + \lambda_i^- = 0, \tag{A129}$$

where $\lambda_i^+ \geq 0$ and $\lambda_i^- \geq 0$ are the Lagrange multipliers, $\mu_{d,i}$ is the ith entry of the vector of average dividends $\boldsymbol{\mu_d}$, and $P_{i,t}$ is the equilibrium price of asset i. The solution to the optimal weight on asset i for the three groups of investors is, respectively,

$$w_i^r = \frac{\mu_{d,i} - P_{i,t}}{\gamma\sigma_i^2}, \tag{A130}$$

$$w_i^+ = \max\left(0, \frac{\mu_{d,i} + b_i - P_{i,t}}{\gamma\sigma_i^2}\right), \tag{A131}$$

$$w_i^- = \max\left(0, \frac{\mu_{d,i} - b_i - P_{i,t}}{\gamma\sigma_i^2}\right). \tag{A132}$$

Under Assumption A.1, the market clearing condition is

$$(1 - \phi^+ - \phi^-)w_i^r + \phi^+ w_i^+ + \phi^- w_i^- = 1/N. \tag{A133}$$

Part 1. Pessimistic investors hold positive quantities of asset i if $w_i^- > 0$, which implies that $\lambda_i^- = 0$, and thus, $w_i^+ > 0$ and $\lambda_i^+ = 0$. Therefore, the market clearing condition becomes

$$(1 - \phi^+ - \phi^-)\frac{\mu_{d,i} - P_{i,t}}{\gamma\sigma_i^2} + \phi^+\frac{\mu_{d,i} + b_i - P_{i,t}}{\gamma\sigma_i^2} + \phi^-\frac{\mu_{d,i} - b_i - P_{i,t}}{\gamma\sigma_i^2} = 1/N, \tag{A134}$$

which gives an equilibrium price $P_{i,t} = \mu_{d,i} + (\phi^+ - \phi^-)b_i - \gamma\sigma_i^2/N$. The condition $w_i^- > 0$ is equivalent to $(\mu_{d,i} - b_i - P_{i,t})/(\gamma\sigma_i^2) > 0$, which gives $b_i < \gamma(\sigma_i^2/N)/(1 + \phi^+ - \phi^-) = b_i^\star$. The expected return of asset i is then $\mu_i = \mu_{d,i} - P_{i,t} = \gamma\sigma_i^2/N - (\phi^+ - \phi^-)b_i$. If $b_i < b_i^\star$ for all i, the vector of expected returns becomes $\boldsymbol{\mu} = \gamma\boldsymbol{\Sigma}\boldsymbol{w}_{ew} - (\phi^+ - \phi^-)\boldsymbol{b}$, and thus, the 1/N Sharpe ratio is

$$\text{SR}(\boldsymbol{w}_{ew}) = \gamma\sigma_{ew} - \frac{(\phi^+ - \phi^-)\bar{b}}{\sigma_{ew}}, \tag{A135}$$

where $\bar{b} = \boldsymbol{b}^\top\boldsymbol{w}_{ew}$ is the average bias among optimistic investors. Regarding the alpha Sharpe ratio, it becomes

$$\text{SR}(\boldsymbol{w}_\alpha) = \sqrt{\boldsymbol{\mu}^\top\boldsymbol{Q}\boldsymbol{\mu}} = |\phi^+ - \phi^-|\sqrt{\boldsymbol{b}^\top\boldsymbol{Q}\boldsymbol{b}}, \tag{A136}$$

where $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{w}_{ew}\boldsymbol{w}_{ew}^\top/\sigma_{ew}^2$. Finally, the relations that $\text{SR}(\boldsymbol{w}_{ew})$ and $\text{SR}(\boldsymbol{w}_\alpha)$ have with $\boldsymbol{b}$ are straightforward to check.

Part 2. When $b_i \geq b_i^\star$, we have $w_i^- = 0$ and $\lambda_i^- \geq 0$. Therefore, the market clearing condition becomes

$$(1-\phi^+-\phi^-)\frac{\mu_{d,i}-P_{i,t}}{\gamma\sigma_i^2}+\phi^+\max\left(0,\frac{\mu_{d,i}+b_i-P_{i,t}}{\gamma\sigma_i^2}\right)=1/N, \tag{A137}$$

which can be satisfied only if $\mu_{d,i}+b_i-P_{i,t}\geq 0$, meaning that $w_i^+\geq 0$ and optimistic investors hold positive quantities of asset i. Therefore, dropping the max operator in (A137) gives an equilibrium price $P_{i,t}=\mu_{d,i}+(\phi^+b_i-\gamma\sigma_i^2/N)/(1-\phi^-)$ and an expected return $\mu_i=\mu_{d,i}-P_{i,t}=(\gamma\sigma_i^2/N-\phi^+b_i)/(1-\phi^-)$. If $b_i\geq b_i^\star$ for all i, the vector of expected returns becomes $\boldsymbol{\mu}=(\gamma\boldsymbol{\Sigma}\boldsymbol{w}_{ew}-\phi^+\boldsymbol{b})/(1-\phi^-)$, and thus, the 1/N Sharpe ratio is

$$\text{SR}(\boldsymbol{w}_{ew})=\frac{\gamma\sigma_{ew}^2-\phi^+\bar{b}}{(1-\phi^-)\sigma_{ew}}. \tag{A138}$$

Regarding the alpha Sharpe ratio, it becomes

$$\text{SR}(\boldsymbol{w}_\alpha)=\sqrt{\boldsymbol{\mu}^\top\boldsymbol{Q}\boldsymbol{\mu}}=\frac{\phi^+}{1-\phi^-}\sqrt{\boldsymbol{b}^\top\boldsymbol{Q}\boldsymbol{b}}. \tag{A139}$$

Finally, the relations that $\text{SR}(\boldsymbol{w}_{ew})$ and $\text{SR}(\boldsymbol{w}_\alpha)$ have with $\boldsymbol{b}$ are straightforward to check, which concludes the proof.

References

Ao, Mengmeng, Yingying Li, and Xinghua Zheng 2019. "Approaching mean-variance efficiency for large portfolios." *The Review of Financial Studies* 32, 2890–2919. https://doi.org/10.1093/rfs/hhy105

Baker, Malcolm, and Jeffrey Wurgler 2006. "Investor sentiment and the cross-section of stock returns." *The Journal of Finance* 61, 1645–1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x

Barroso, Pedro, and Andrew Detzel 2021. "Do limits to arbitrage explain the benefits of volatility-managed portfolios?" *Journal of Financial Economics* 140, 744–767. https://doi.org/10.1016/j.jfineco.2021.02.009

Chen, Jiaqin, and Ming Yuan 2016. "Efficient portfolio selection in a large market." *Journal of Financial Econometrics* 14, 496–524. https://doi.org/10.1093/jjfinec/nbw003

DeMiguel, Victor, Alberto Martín-Utrera, and Raman Uppal 2024. "A multifactor perspective on volatility-managed portfolios." *The Journal of Finance* 79, 3859–3891. https://doi.org/10.1111/jofi.13395

Engelberg, Joseph E, Adam V Reed, and Matthew C Ringgenberg 2018. "Short-selling risk." *The Journal of Finance* 73, 755–786. https://doi.org/10.1111/jofi.12601

Hong, Harrison, and David Sraer 2016. "Speculative betas." *The Journal of Finance* 71, 2095–2144. https://doi.org/10.1111/jofi.12431

Huang, Dashan, Fuwei Jiang, Jun Tu, and Guofu Zhou 2015. "Investor sentiment aligned: A powerful predictor of stock returns." *The Review of Financial Studies* 28, 791–837. https://doi.org/10.1093/rfs/hhu080

Kan, Raymond, and Guofu Zhou 2007. "Optimal portfolio choice with parameter uncertainty." *Journal of Financial and Quantitative Analysis* 42, 621–656. https://doi.org/10.1017/s0022109000004129

Kirby, Chris, and Barbara Ostdiek 2012. "It's all in the timing: Simple active portfolio strategies that outperform naïve diversification." *Journal of Financial and Quantitative Analysis* 47, 437–467. https://doi.org/10.1017/s0022109012000117

Ledoit, Olivier, and Michael Wolf 2003. "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection." *Journal of Empirical Finance* 10, 603–621.

https://doi.org/10.1016/s0927-5398(03)00007-0

Ledoit, Olivier, and Michael Wolf 2004. "A well-conditioned estimator for large-dimensional covariance matrices." *Journal of Multivariate Analysis* 88, 365–411. https://doi.org/10.1016/s0047-259x(03)00096-4

Ledoit, Olivier, and Michael Wolf 2008. "Robust performance hypothesis testing with the Sharpe ratio." *Journal of Empirical Finance* 15, 850–859. https://doi.org/10.1016/j.jempfin.2008.03.002

Ledoit, Olivier, and Michael Wolf 2020. "Analytical nonlinear shrinkage of large-dimensional covariance matrices." *The Annals of Statistics* 48, 3043–3065. https://doi.org/10.1214/19-aos1921

MacKinlay, A. Craig, and Ľuboš Pástor 2000. "Asset pricing models: Implications for expected returns and portfolio selection." *The Review of Financial Studies* 13, 883–916. https://doi.org/10.1093/rfs/13.4.883

Miller, Edward 1977. "Risk, uncertainty, and divergence of opinion." *The Journal of Finance* 32, 1151–1168. https://doi.org/10.1111/j.1540-6261.1977.tb03317.x

Stambaugh, Robert, Jianfeng Yu, and Yu Yuan 2012. "The short of it: Investor sentiment and anomalies." *Journal of Financial Economics* 104, 288–302. https://doi.org/10.1016/j.jfineco.2011.12.001

van Binsbergen, Jules, Jungsuk Han, Hongxun Ruan, and Ran Xing 2024. "A horizon-based decomposition of mutual fund value added using transactions." *The Journal of Finance* 79, 1831–1882.