MAXIMIZING THE OUT-OF-SAMPLE SHARPE RATIO

Nathan Lassance

REPRINT • 2026 / 02



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

Maximizing the Out-of-Sample Sharpe Ratio*

Nathan Lassance

August 10, 2026

Operations Research, forthcoming

Abstract

We derive analytical expressions for the expectation and variance of the out-of-sample Sharpe ratio of sample mean-variance portfolios, which allows us to design portfolio strategies with an optimal Sharpe ratio under estimation risk. This problem is mathematically more challenging than maximizing the out-of-sample utility, the standard objective in the literature on portfolio choice with estimation risk. When a risk-free asset is available, we show that maximizing the expected out-of-sample utility approximately–but not exactly–optimizes the expected out-of-sample Sharpe ratio. The two objectives, however, differ in the common setting with a full investment in risky assets. In this case, we derive the fully invested mean-variance portfolio that maximizes the expected out-of-sample Sharpe ratio, along with a robust version that penalizes the Sharpe ratio variance. Using simulated and empirical data, we find that our portfolio rules generally deliver a superior out-of-sample Sharpe ratio relative to the benchmarks. In particular, across different empirical configurations, an equally weighted combination of our four strategies delivers on average an annualized out-of-sample Sharpe ratio of 0.891 and 0.847 before and after transaction costs, respectively, compared with 0.760 and 0.732 for the global minimum-variance portfolio with nonlinear shrinkage, and 0.802 and 0.728 for a fully invested MAXSER portfolio.

Keywords: estimation risk, mean-variance portfolio, out-of-sample performance, parameter uncertainty, shrinkage estimator.

JEL classification: C44, G11.

*Affiliation: LFIN/LIDAM, UCLouvain, Belgium. Email address: nathan.lassance@uclouvain.be. I am grateful to Victor DeMiguel, Raymond Kan, Felix Kircher, Alberto Martín-Utrera, Majeed Simaan, Xiaolu Wang, and Guofu Zhou for their feedback. This work was supported by the Fonds de la Recherche Scientifique (F.R.S.-FNRS) [grant number J.0135.25].

1 Introduction

The literature on portfolio choice with estimation risk warns investors not to blindly apply sample estimates of asset-return parameters in their portfolios. In particular, optimal mean-variance portfolios using sample estimates of the mean and covariance matrix deliver a notoriously poor out-of-sample performance.[1] As a remedy, researchers have developed different methods to alleviate the impact of estimation risk. Many of these methods improve out-of-sample performance *indirectly*, for example using improved estimators of the mean and covariance matrix (Ledoit and Wolf, 2004; Barroso and Saxena, 2021; Ding, Li, and Zheng, 2021), penalties on portfolio weights (DeMiguel, Garlappi, and Uppal, 2009; Olivares-Nadal and DeMiguel, 2018), principal and independent component analysis (Zhao, Chakrabarti, and Muthuraman, 2019; Lassance, DeMiguel, and Vrins, 2022), or minimax uncertainty-averse portfolios (Nguyen et al., 2024; Wu et al., 2025).

In this paper, we *directly* optimize the *out-of-sample Sharpe ratio*, using an analytical approach. When it comes to directly optimizing the out-of-sample performance, an important contribution is Kan and Zhou (2007) who derive analytical expressions for the expected out-of-sample *utility* (i.e., certainty-equivalent return) of mean-variance portfolios and design optimal two-fund and three-fund rules accordingly. Since then, many papers have used the expected out-of-sample utility (or variance) as an objective.[2] Recent developments include the case without a risk-free asset (Kan, Wang, and Zhou, 2021) and studying higher-order moments of out-of-sample performance (Lassance, Martín-Utrera, and Simaan, 2024; Lassance, Kan, and Wang, 2025), which are analytically challenging.

Despite these developments, no paper has addressed the important problem of finding analytical expressions for the expected out-of-sample Sharpe ratio and designing optimal portfolios accordingly. This is surprising because the Sharpe ratio is a pervasive objective among investors. As Kelly, Malamud, and Zhou (2024, p. 468) put it: *"Academics and real-world investors rely nearly universally on the [...] Sharpe ratio when evaluating empirical trading strategies."* One reason for this focus on the expected out-of-sample utility instead of Sharpe ratio is that the latter is analytically more challenging to evaluate. As Tu and Zhou (2011, p.208) write, *"[...] we focus on maximizing the expected utility as it is easier to solve than maximizing the Sharpe ratio because the latter is to maximize a highly nonlinear function of the portfolio weights and there are no closed-form solutions available in the presence of estimation errors."* Our paper addresses this challenge.

Although previous research has developed investment strategies maximizing the expected out-of-sample Sharpe ratio, we are the first to provide an exact finite-sample analytical approach, to the best of our knowledge.[3] DeMiguel, Martín-Utrera, and Nogales (2013) propose the expected out-of-

sample Sharpe ratio among other criteria for optimizing shrinkage portfolios, but they only use an approximation and do not provide analytical solutions for the optimal portfolios. Yuan and Zhou (2024) combine mean-variance and 1/N portfolios to maximize the out-of-sample Sharpe ratio, but in an asymptotic regime rather than the exact finite-sample setting. Chakrabarti (2021) and Kircher and Rosch (2021) maximize the expected out-of-sample Sharpe ratio using data-driven approaches based on robust optimization and bootstrap, respectively, which differ from our analytical approach.

We make four main contributions to the portfolio choice literature with estimation risk. First, in Section 3, we consider the setting with a risk-free asset investment. We show that combining portfolios to trade off between expected out-of-sample mean return and variance is mathematically similar to classical portfolio theory. In particular, *maximizing the expected out-of-sample utility optimizes the expected Sharpe ratio as well*, under a Taylor-series approximation of the latter. We illustrate this result for the three-fund rule of Kan and Zhou (2007) combining the risk-free asset, the sample tangency portfolio, and the sample global minimum-variance (GMV) portfolio. We also derive an exact analytical expression for the expected out-of-sample Sharpe ratio of this three-fund rule and find that the resulting optimal combination coefficients are close to those under the approximation.

Second, in Section 4, we consider the common setting of a full investment in risky assets. In that case, the combination coefficients on different portfolios sum to one, and thus, the expected out-of-sample Sharpe ratio delivers a unique optimal portfolio combination different from that under the expected out-of-sample utility. As a special case, we consider portfolios located on the mean-variance efficient frontier as in Kan, Wang, and Zhou (2021) and Lassance, Martín-Utrera, and Simaan (2024), which are combinations of the sample tangency and GMV portfolios, and derive their approximate and exact expected out-of-sample Sharpe ratio. When the covariance matrix is known, the optimal combination of the sample tangency and GMV portfolios under the expected out-of-sample utility is the same as under the approximate expected out-of-sample Sharpe ratio, and almost the same as under the exact one. However, this is no longer true when the covariance matrix is unknown, in which case the expected out-of-sample Sharpe ratio implies a smaller weight on the GMV portfolio. Therefore, a naive application of the results in Kan, Wang, and Zhou (2021), which pertain to the expected utility, does not deliver an optimal Sharpe ratio out of sample.

Third, in Section 5, we derive a Taylor-series approximation and exact analytical expression for the variance of the out-of-sample Sharpe ratio of fully invested mean-variance efficient portfolios. Following Lassance, Martín-Utrera, and Simaan (2024), we then find the combination of the sample tangency and GMV portfolios that delivers an optimal tradeoff between the expectation and variance of the out-of-sample Sharpe ratio. Penalizing the variance increases the combination

coefficient allocated to the sample GMV portfolio, which alleviates the impact of estimation risk in the combination coefficient. This more intensive shrinkage toward GMV is valuable empirically.

Finally, in Section 6, we evaluate the out-of-sample Sharpe ratio of the four fully invested mean-variance portfolios we introduce in the paper, across four datasets. We first consider simulated data and show that, in most cases, our portfolios outperform the 1/N portfolio, the portfolio of Kan, Wang, and Zhou (2021) maximizing the expected out-of-sample utility, and the two combined portfolios, i.e., the tangency and GMV portfolios. This outperformance holds under the sample covariance matrix and the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020). Moreover, our robust portfolios that penalize the variance of the out-of-sample Sharpe ratio are useful when there is high estimation risk, i.e., when the sample size is small, the sample covariance matrix is used, or when the return signal in the dataset is limited, e.g., in industry portfolios.

We turn next to empirical data and consider a larger set of benchmarks. Our mean-variance portfolios deliver larger out-of-sample Sharpe ratios than GMV in most cases, even under the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020). Moreover, our portfolios generally outperform the benchmarks, including a fully invested version of the MAXSER portfolio of Ao, Li, and Zheng (2019). In particular, under the Ledoit-Wolf covariance estimator, our four strategies deliver on average an annualized out-of-sample Sharpe ratio of 0.871 and 0.829 before and after transaction costs, respectively, versus 0.760 and 0.732 for the GMV portfolio and 0.802 and 0.728 for MAXSER. Compared to the data-driven approach of Kircher and Rosch (2021) that maximizes the expected out-of-sample Sharpe ratio via bootstrap, we obtain better Sharpe ratios overall, particularly among industry portfolios, along with a significantly lower turnover. Finally, relative to maximizing the expectation only, our robust portfolios that penalize the Sharpe ratio variance often help improve the Sharpe ratio, particularly net of costs and in industry portfolios where the return signal is limited, and deliver a lower turnover by shrinking more intensively toward GMV.

2 Preliminaries

2.1 Notation, assumptions, and definitions

We consider an investment universe of $N \geq 2$ assets. At the end of period t, the investor cares about her next-period portfolio excess return, $w'r_{t+1}$, where w is the vector of portfolio weights on the risky assets and r_{t+1} is the random vector of asset excess returns with mean vector μ and positive-definite covariance matrix $\mathbf{\Sigma}$. We assume that μ and $\mathbf{\Sigma}$ are constant over time and solve the unconditional portfolio choice problem. For a given w, a proportion $w'\iota$ is invested in risky

assets and $1 - w'\iota$ in the risk-free asset, where ι is an N-dimensional vector of ones. When the portfolio is fully invested in risky assets as in Section 4, we have $w'\iota = 1$.

The parameters μ and $\mathbf{\Sigma}$ are unknown. Given a sample of T historical return observations, $(r_{t-T+1}, \ldots, r_t)$, assumed i.i.d., the sample estimators of μ and $\mathbf{\Sigma}$ at time t are

$$\hat{\mu} = \frac{1}{T}\sum_{m=1}^{T} r_{t-T+m} \quad \text{and} \quad \hat{\mathbf{\Sigma}} = \frac{1}{T}\sum_{m=1}^{T}(r_{t-T+m} - \hat{\mu})(r_{t-T+m} - \hat{\mu})'. \tag{1}$$

In our theoretical results, we consider the covariance matrix $\mathbf{\Sigma}$ to be either known (and replaced by the Ledoit and Wolf (2020) nonlinear shrinkage estimator empirically) or estimated by $\hat{\mathbf{\Sigma}}$.

To derive some of our analytical results, but not all of them, we assume that returns are i.i.d. multivariate normal. Some comments follow. First, this is a pervasive assumption in this literature; see Kan and Lassance (2025) for a review. Second, as Kan, Wang, and Zheng (2024) argue, this assumption allows to isolate the effect of estimation risk from other factors explaining the empirical gap between in-sample and out-of-sample performance such as market frictions, data mining, outliers, and structural breaks. Third, this assumption allows analytical insights that would otherwise be difficult to obtain, which lead to the strategies we propose. Fourth, in our empirical analysis of Section 6.4, the 'Expected Sharpe (approx)' and 'Robust Sharpe (approx)' strategies implemented with the Ledoit and Wolf (2020) covariance matrix, which are not built upon the multivariate normality assumption, deliver among the largest out-of-sample Sharpe ratios; see Table 4.

Finally, we introduce some population parameters. The tangency portfolio is

$$w_{sr} = \frac{1}{\gamma_{sr}}\mathbf{\Sigma}^{-1}\mu \quad \text{with} \quad \gamma_{sr} = \iota'\mathbf{\Sigma}^{-1}\mu, \tag{2}$$

where γ_{sr} is the risk-aversion coefficient implied by w_{sr}. We make the mild assumption that $\gamma_{sr} > 0$, in which case the tangency portfolio delivers the maximum achievable Sharpe ratio:

$$\theta = \frac{w_{sr}'\mu}{\sqrt{w_{sr}'\mathbf{\Sigma}w_{sr}}} = \sqrt{\mu'\mathbf{\Sigma}^{-1}\mu}. \tag{3}$$

In the unlikely case that $\gamma_{sr} < 0$, w_{sr} delivers the lowest Sharpe ratio $-\theta$, and the maximum-Sharpe-ratio portfolio differs from w_{sr}; see Maller and Turkington (2003). We omit this case for simplicity. The global minimum-variance (GMV) portfolio is

$$w_g = \frac{\mathbf{\Sigma}^{-1}\iota}{\iota'\mathbf{\Sigma}^{-1}\iota}, \tag{4}$$

and we denote the return mean, variance, and Sharpe ratio of the GMV portfolio, as well as the difference between the maximum squared Sharpe ratio and that of GMV, as

$$\mu_g = \frac{\iota'\boldsymbol{\Sigma}^{-1}\mu}{\iota'\boldsymbol{\Sigma}^{-1}\iota}, \quad \sigma_g^2 = \frac{1}{\iota'\boldsymbol{\Sigma}^{-1}\iota}, \quad \theta_g = \frac{\mu_g}{\sigma_g}, \quad \text{and} \quad \psi^2 = \theta^2 - \theta_g^2. \tag{5}$$

2.2 The expected out-of-sample Sharpe ratio

The maximum Sharpe ratio, θ in (3), is unattainable because μ and $\boldsymbol{\Sigma}$ are unknown. This calls for a criterion accounting for the impact of estimation risk on the out-of-sample Sharpe ratio. Given an estimated portfolio $\hat{w}$, its out-of-sample return $\hat{w}'r_{t+1}$ has mean $\hat{w}'\mu$ and variance $\hat{w}'\boldsymbol{\Sigma}\hat{w}$ conditional on $\hat{w}$. Then, as in, e.g., Kircher and Rosch (2021) and Kan, Wang, and Zheng (2024), we define the *expected out-of-sample Sharpe ratio* of $\hat{w}$ as

$$\mathcal{S}(\hat{w}) = \mathbb{E}\left[\frac{\hat{w}'\mu}{\sqrt{\hat{w}'\boldsymbol{\Sigma}\hat{w}}}\right], \tag{6}$$

where the expectation is evaluated under the true distribution of asset returns with mean μ and covariance matrix $\boldsymbol{\Sigma}$. In addition to the exact expected out-of-sample Sharpe ratio $\mathcal{S}(\hat{w})$, we also consider a zero-order Taylor-series approximation of $\mathcal{S}(\hat{w})$, which is also used by DeMiguel, Martín-Utrera, and Nogales (2013) for tractability. This approximation is given by

$$\widetilde{\mathcal{S}}(\hat{w}) = \frac{\mathbb{E}[\hat{w}'\mu]}{\sqrt{\mathbb{E}[\hat{w}'\boldsymbol{\Sigma}\hat{w}]}}, \tag{7}$$

and it satisfies $\mathcal{S}(\hat{w}) = \widetilde{\mathcal{S}}(\hat{w}) + O(1/T)$. That is, asymptotically as $T \to \infty$ (with N fixed or $N/T \to c \in [0, 1)$), we have $\widetilde{\mathcal{S}}(\hat{w}) \to \mathcal{S}(\hat{w})$. Compared to the exact $\mathcal{S}(\hat{w})$, the approximation $\widetilde{\mathcal{S}}(\hat{w})$ is easier to evaluate analytically and allows more theoretical insights.

3 Expected out-of-sample Sharpe ratio with a risk-free asset

In this section, a risk-free asset investment is allowed. We show that combining portfolios to maximize the expected out-of-sample utility and Sharpe ratio is mathematically similar to combining assets in classical portfolio theory. In particular, when the combination coefficients are unconstrained, maximizing the expected out-of-sample utility optimizes the approximate expected out-of-sample Sharpe ratio $\widetilde{\mathcal{S}}(\hat{w})$ as well. We then focus on the special case of the three-fund rule of Kan and Zhou (2007), for which we derive the exact and approximate expected out-of-sample Sharpe ratio.

3.1 Expected out-of-sample Sharpe ratio versus utility

Consider a combination of the risk-free asset and K estimated portfolios of risky assets,

$$\hat{w}(\kappa) = \hat{w}(\kappa_1, \ldots, \kappa_K) = \sum_{i=1}^{K} \kappa_i \hat{w}_i, \tag{8}$$

where κ_i is the ith combination coefficient and the weight on the risk-free asset is $1 - \hat{w}(\kappa)'\iota$. Then, the expected out-of-sample mean return and variance of this portfolio combination are

$$\mathbb{E}[\hat{w}(\kappa)'\mu] = \kappa'\eta \quad \text{with} \quad \eta_i = \mathbb{E}[\hat{w}_i'\mu], \tag{9}$$

$$\mathbb{E}[\hat{w}(\kappa)'\boldsymbol{\Sigma}\hat{w}(\kappa)] = \kappa'\boldsymbol{S}\kappa \quad \text{with} \quad \boldsymbol{S}_{ij} = \mathbb{E}[\hat{w}_i'\boldsymbol{\Sigma}\hat{w}_j], \tag{10}$$

assuming η and $\boldsymbol{S}$ exist, where $\boldsymbol{S}$ is positive definite if the $\hat{w}_i$'s are linearly independent. (9)–(10) shows that combining portfolios to trade off between the expected out-of-sample mean return and variance, $\kappa'\eta$ and $\kappa'\boldsymbol{S}\kappa$, is similar to combining assets to trade off between the mean return and variance, $w'\mu$ and $w'\boldsymbol{\Sigma}w$. In particular, when a risk-free asset is available and the combination coefficients are unconstrained, we recover the classic result that maximizing utility maximizes the Sharpe ratio, under the approximation $\widetilde{\mathcal{S}}(\hat{w})$. Indeed, let the expected out-of-sample utility of $\hat{w}$ be

$$\mathcal{U}(\hat{w}) = \mathbb{E}[\hat{w}'\mu] - \frac{\gamma}{2}\mathbb{E}[\hat{w}'\boldsymbol{\Sigma}\hat{w}], \tag{11}$$

where $\gamma > 0$ is the risk-aversion coefficient. Then, the vector of combination coefficients κ maximizing the expected out-of-sample utility of the portfolio combination $\hat{w}(\kappa)$ in (8) is given by

$$\kappa_\gamma = \arg\max_{\kappa \in \mathbb{R}^K} \mathcal{U}(\hat{w}(\kappa)) = \arg\max_{\kappa \in \mathbb{R}^K} \kappa'\eta - \frac{\gamma}{2}\kappa'\boldsymbol{S}\kappa = \frac{1}{\gamma}\boldsymbol{S}^{-1}\eta, \tag{12}$$

and it delivers the maximum approximate expected out-of-sample Sharpe ratio:

$$\max_{\kappa \in \mathbb{R}^K} \widetilde{\mathcal{S}}(\hat{w}(\kappa)) = \max_{\kappa \in \mathbb{R}^K} \frac{\kappa'\eta}{\sqrt{\kappa'\boldsymbol{S}\kappa}} = \sqrt{\eta'\boldsymbol{S}^{-1}\eta} = \widetilde{\mathcal{S}}(\hat{w}(\kappa_\gamma)). \tag{13}$$

We formalize this result in the next proposition; proofs are available in the Electronic Companion.

Proposition 1. *Consider $\hat{w}(\kappa)$ in (8) that combines the risk-free asset and K estimated portfolios of risky assets. Assume η and $\boldsymbol{S}$ in (9)–(10) exist and the combination coefficients $\kappa \in \mathbb{R}^K$ are unconstrained. Then, κ_γ in (12) maximizing the expected out-of-sample utility, $\mathcal{U}(\hat{w}(\kappa))$ in (11), also maximizes the approximate expected out-of-sample Sharpe ratio, $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ in (7).*

The assumption that the combination coefficients are unconstrained is key. For example, Tu and Zhou (2011) combine the risk-free asset, the sample tangency portfolio, and the 1/N portfolio to maximize the expected out-of-sample utility. They constrain the two combination coefficients on the tangency and 1/N portfolios to sum to one, and thus, their portfolio does not maximize the expected out-of-sample Sharpe ratio, which leads them to write (p.208): *"When the true parameters are known, maximizing the Sharpe ratio and maximizing the expected utility are equivalent, a well-known fact. However, once the true parameters are unknown, the two are different."* Proposition 1 shows that, under $\widetilde{\mathcal{S}}(\hat{w})$, this result no longer holds with unconstrained combination coefficients.

To illustrate Proposition 1, consider the popular three-fund rule of Kan and Zhou (2007) that combines the risk-free asset, the sample tangency portfolio, and the sample GMV portfolio:

$$\hat{w}(\kappa) = \kappa_1 \hat{w}_1 + \kappa_2 \hat{w}_2, \quad \text{where} \quad \hat{w}_1 = \hat{\boldsymbol{\Sigma}}^{-1}\hat{\mu} \quad \text{and} \quad \hat{w}_2 = \hat{\boldsymbol{\Sigma}}^{-1}\iota. \tag{14}$$

Assuming that asset returns are i.i.d. multivariate normal, the vector η in (9) exists for $T \geq N+3$ and the matrix $\boldsymbol{S}$ in (10) for $T \geq N+5$, and they are given by

$$\eta = \frac{T}{T-N-2}\begin{pmatrix} \theta^2 \\ \mu_g/\sigma_g^2 \end{pmatrix}, \quad \boldsymbol{S} = \frac{T^2(T-2)}{(T-N-1)(T-N-2)(T-N-4)}\begin{pmatrix} \theta^2 + \frac{N}{T} & \mu_g/\sigma_g^2 \\ \mu_g/\sigma_g^2 & 1/\sigma_g^2 \end{pmatrix}. \tag{15}$$

Then, the combination coefficients in (12) maximizing the expected out-of-sample utility are

$$\kappa_\gamma = \frac{1}{\gamma}\boldsymbol{S}^{-1}\eta = \frac{(T-N-1)(T-N-4)}{\gamma T(T-2)}\left(\frac{\psi^2}{\psi^2 + N/T}, \frac{\mu_g \frac{N}{T}}{\psi^2 + \frac{N}{T}}\right)', \tag{16}$$

and they maximize the approximate expected out-of-sample Sharpe ratio:

$$\max_{\kappa \in \mathbb{R}^2} \widetilde{\mathcal{S}}(\hat{w}(\kappa)) = \widetilde{\mathcal{S}}(\hat{w}(\kappa_\gamma)) = \sqrt{\eta' \boldsymbol{S}^{-1}\eta} = \sqrt{\frac{(T-N-1)(T-N-4)}{(T-2)(T-N-2)}\left(\frac{\psi^2\theta^2 + \frac{N}{T}\theta_g^2}{\psi^2 + N/T}\right)}. \tag{17}$$

3.2 Exact versus approximate expected out-of-sample Sharpe ratio

The equivalence between the expected out-of-sample utility and Sharpe ratio holds under the approximation $\widetilde{\mathcal{S}}(\hat{w})$ in (7). We now evaluate the accuracy of this approximation and the optimal combination coefficients it delivers. For that purpose, we focus on the three-fund rule of Kan and Zhou (2007) in (14), an important analytically tractable class of strategies. We derive its exact expected out-of-sample Sharpe ratio in the next proposition, which is not available in the literature.

Proposition 2. *Let asset returns be i.i.d. multivariate normally distributed. Consider the three-fund rule $\hat{w}(\kappa)$ in (14) and denote $\omega = \kappa_2/\kappa_1$, $\kappa_1 > 0$.*

1. *Given $T \geq N + 5$, the approximate expected out-of-sample Sharpe ratio of $\hat{w}(\kappa)$ is given by*

$$\widetilde{\mathcal{S}}(\hat{w}(\kappa)) = \sqrt{\frac{(T-N-1)(T-N-4)}{(T-2)(T-N-2)}}\left(\frac{\theta^2 + \omega\mu_g/\sigma_g^2}{\sqrt{\theta^2 + \frac{N}{T} + 2\omega\mu_g/\sigma_g^2 + \omega^2/\sigma_g^2}}\right), \tag{18}$$

and it is maximized by

$$\tilde{\omega}_{sr} = \arg\max_{\omega\in\mathbb{R}} \widetilde{\mathcal{S}}(\hat{w}(\kappa)) = \frac{N\mu_g}{T\psi^2}. \tag{19}$$

2. *The exact expected out-of-sample Sharpe ratio of $\hat{w}(\kappa)$ is given by*

$$\mathcal{S}(\hat{w}(\kappa)) = C(N,T)\left(\theta^2 + \frac{\omega\mu_g}{\sigma_g^2}\right){}_1F_1\left(\frac{1}{2}; \frac{N+2}{2}; -\frac{T}{2}\left[\theta^2 + \frac{2\omega\mu_g}{\sigma_g^2} + \frac{\omega^2}{\sigma_g^2}\right]\right), \tag{20}$$

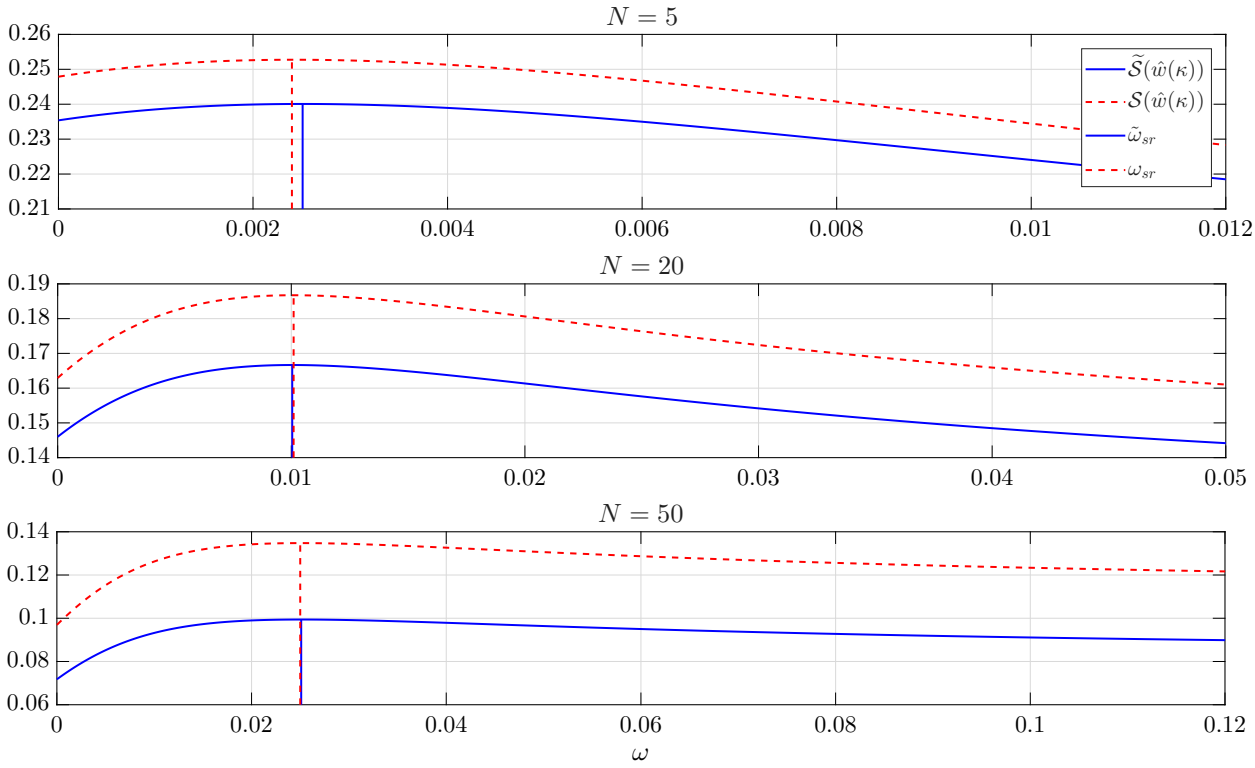
where ${}_1F_1$ is the confluent hypergeometric function and

$$C(N,T) = \sqrt{\frac{T}{2}}\frac{\Gamma(\frac{T}{2})\Gamma(\frac{T-N+2}{2})\Gamma(\frac{N+1}{2})}{\Gamma(\frac{T+1}{2})\Gamma(\frac{T-N+1}{2})\Gamma(\frac{N+2}{2})}. \tag{21}$$

We now offer several comments. First, the exact expected out-of-sample Sharpe ratio in (20)–(21) extends upon that in Kan, Wang, and Zheng (2024) for $\kappa_1 = 1$ and $\kappa_2 = 0$, i.e., the portfolio $\hat{\mathbf{\Sigma}}^{-1}\hat{\mu}$. It also extends upon Yuan and Zhou (2024) who derive its high-dimensional asymptotic limit. Second, because the Sharpe ratio is scaling-invariant, both $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ and $\mathcal{S}(\hat{w}(\kappa))$ only depend on $\omega = \kappa_2/\kappa_1$, not κ_1 and κ_2 separately. Third, the approximate expected out-of-sample Sharpe ratio allows a closed-form expression for the optimal ω, given by $\tilde{\omega}_{sr}$ in (19), which delivers the maximum $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ in (17). Fourth, a relevant by-product of the proof of Proposition 2 is a stochastic representation for the exact *distribution* of the out-of-sample Sharpe ratio of $\hat{w}(\kappa)$.[4]

In Figure 1, we compare the exact expected out-of-sample Sharpe ratio, $\mathcal{S}(\hat{w}(\kappa))$ in (20)–(21), with its approximation, $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ in (18). We use realistic values of $\mu_g = 0.05/12$, $\sigma_g = 0.1/\sqrt{12}$, $\theta = 0.3$, $T = 120$, $N \in \{5, 20, 50\}$, and we depict $\mathcal{S}(\hat{w}(\kappa))$ and $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ as a function of ω. We observe that the approximation $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ is conservative, in the sense that it underestimates the exact expected Sharpe ratio $\mathcal{S}(\hat{w}(\kappa))$, and the difference is not negligible. However, the approximate and exact optimal ω, $\tilde{\omega}_{sr}$ and ω_{sr}, are very similar, a result that is robust to different values of the parameters. There is a slightly larger difference between $\tilde{\omega}_{sr}$ and ω_{sr} when $N = 5$ rather than 20

Figure 1: Exact versus approximate expected out-of-sample Sharpe ratio of the three-fund rule



Notes. This figure depicts the expected out-of-sample Sharpe ratio of the three-fund rule, $\hat{w}(\kappa)$ in (14), as a function of $\omega = \kappa_2/\kappa_1$. We depict the approximate Sharpe ratio, $\tilde{\mathcal{S}}(\hat{w}(\kappa))$ in (18), and the exact one, $\mathcal{S}(\hat{w}(\kappa))$ in (20)–(21), as well as the resulting optimal $\tilde{\omega}_{sr}$ and ω_{sr}. We set $\mu_g = 0.05/12$, $\sigma_g = 0.1/\sqrt{12}$, $\theta = 0.3$, $T = 120$, and $N \in \{5, 20, 50\}$.

or 50, which is expected given the asymptotic equivalence between $\mathcal{S}(\hat{w}(\kappa))$ and $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$.

Overall, Figure 1 suggests that maximizing the approximate expected out-of-sample Sharpe ratio $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ almost maximizes the exact Sharpe ratio $\mathcal{S}(\hat{w}(\kappa))$ too. Therefore, the equivalence between maximizing the expected out-of-sample utility and the approximate expected out-of-sample Sharpe ratio established in Section 3.1 essentially applies to the exact Sharpe ratio as well.

4 Expected out-of-sample Sharpe ratio without a risk-free asset

In this section, there is no risk-free investment, we combine fully invested portfolios, the combination coefficients are constrained to sum to one, and there is no longer an equivalence between maximizing the expected out-of-sample utility and Sharpe ratio. This is similar to classical portfolio theory where there is a single tangency portfolio on the fully invested efficient frontier that maximizes the Sharpe ratio. As a special case, we combine portfolios on the sample mean-variance efficient frontier as in Kan, Wang, and Zhou (2021) and Lassance, Martín-Utrera, and Simaan (2024).

4.1 Expected out-of-sample utility versus Sharpe ratio

Consider a combination of K estimated portfolios fully invested in risky assets,

$$\hat{w}(\kappa) = \hat{w}(\kappa_1, \ldots, \kappa_K) = \sum_{i=1}^{K} \kappa_i \hat{w}_i, \quad \hat{w}_i'\iota = 1 \ \forall i. \tag{22}$$

To ensure that $\hat{w}(\kappa)'\iota = 1$, we need $\kappa'\iota = 1$. Due to this constraint on κ, there are two main changes relative to the case with a risk-free asset in Section 3.1. First, the combination coefficients maximizing the approximate expected out-of-sample Sharpe ratio are unique and given by

$$\tilde{\kappa}_{sr} = \arg \max_{\substack{\kappa \in \mathbb{R}^K \\ \kappa'\iota = 1}} \widetilde{\mathcal{S}}(\hat{w}(\kappa)) = \frac{\boldsymbol{S}^{-1}\eta}{\iota'\boldsymbol{S}^{-1}\eta}, \tag{23}$$

provided $\iota'\boldsymbol{S}^{-1}\eta > 0$, where η and $\boldsymbol{S}$ are defined in (9)–(10). Second, provided η is not a constant vector, the combination coefficients maximizing the expected out-of-sample utility in (11),

$$\kappa_\gamma = \arg \max_{\substack{\kappa \in \mathbb{R}^K \\ \kappa'\iota = 1}} \mathcal{U}(\hat{w}(\kappa)) = \frac{\boldsymbol{S}^{-1}\iota}{\iota'\boldsymbol{S}^{-1}\iota} + \frac{1}{\gamma}\left(\boldsymbol{S}^{-1} - \frac{\boldsymbol{S}^{-1}\iota\iota'\boldsymbol{S}^{-1}}{\iota'\boldsymbol{S}^{-1}\iota}\right)\eta, \tag{24}$$

are different from $\tilde{\kappa}_{sr}$ in (23), except if $\gamma = \iota'\boldsymbol{S}^{-1}\eta$. We now formalize these results.

Proposition 3. *Consider $\hat{w}(\kappa)$ in (22) that combines K estimated portfolios fully invested in risky assets. Assume η and $\boldsymbol{S}$ in (9)–(10) exist, $\iota'\boldsymbol{S}^{-1}\eta > 0$, and the combination coefficients $\kappa \in \mathbb{R}^K$ satisfy $\kappa'\iota = 1$. Then, the unique κ maximizing the approximate expected out-of-sample Sharpe ratio, $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ in (7), is given by $\tilde{\kappa}_{sr}$ in (23). Moreover, if η is not a constant vector, $\tilde{\kappa}_{sr}$ differs from κ_γ in (24) maximizing the expected out-of-sample utility, $\mathcal{U}(\hat{w}(\kappa))$ in (11), except if $\gamma = \iota'\boldsymbol{S}^{-1}\eta$.*

4.2 Investing on the sample mean-variance efficient frontier

We now use the expected out-of-sample Sharpe ratio to invest on the fully invested sample mean-variance efficient frontier. We first introduce some definitions. Given a risk-aversion coefficient $\gamma > 0$,

$$w_\gamma = \arg \max_{\substack{w \in \mathbb{R}^N \\ w'\iota = 1}} \mathcal{U}(w) = w_g + \frac{1}{\gamma} w_z \tag{25}$$

is the utility-optimal portfolio on the efficient frontier, where w_g in (4) is the GMV portfolio and

$$w_z = \boldsymbol{B}\mu \quad \text{with} \quad \boldsymbol{B} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{I}_N - \iota w_g') \tag{26}$$

is a zero-cost portfolio, i.e., $w_z'\iota = 0$. The tangency portfolio w_{sr} in (2), which maximizes the Sharpe ratio, is obtained for $\gamma = \gamma_{sr}$ in (2), where we assume $\gamma_{sr} > 0$. That is,

$$w_{sr} = \arg \max_{\substack{w \in \mathbb{R}^N \\ w'\iota = 1}} \mathcal{S}(w) = \frac{\boldsymbol{\Sigma}^{-1}\mu}{\iota'\boldsymbol{\Sigma}^{-1}\mu} = w_g + \frac{1}{\gamma_{sr}} w_z = w_{\gamma_{sr}}. \tag{27}$$

4.2.1 Uncertainty in mean returns only

We first consider the case where the covariance matrix $\boldsymbol{\Sigma}$ is known and estimation errors only step from the sample mean-return vector $\hat{\mu}$ as in, e.g., Garlappi, Uppal, and Wang (2007). This assumption allows us to derive the optimal portfolio when estimation errors in the covariance matrix are negligible relative to those in mean returns. Empirically, we then replace $\boldsymbol{\Sigma}$ with the nonlinear shrinkage estimator of Ledoit and Wolf (2020).

We invest on the sample efficient frontier by combining the GMV portfolio w_g with the mean-variance portfolio having unit exposure to the sample zero-cost portfolio, $\hat{w}_z^{\sigma} = \boldsymbol{B}\hat{\mu}$. This gives

$$\hat{w}^{\sigma}(\kappa) = (1-\kappa)w_g + \kappa(w_g + \hat{w}_z^{\sigma}) = w_g + \kappa \boldsymbol{B}\hat{\mu}, \tag{28}$$

where the superscript σ refers to $\boldsymbol{\Sigma}$ being known. Approximating the universe of portfolios with combinations of two portfolios makes our problem analytically tractable, but also means a loss in out-of-sample efficiency. We call $\gamma_{sr}\kappa$ the *shrinkage intensity* because it quantifies the intensity with which the GMV portfolio is shrunk toward the tangency portfolio:

$$\hat{w}^{\sigma}(\kappa) = (1 - \gamma_{sr}\kappa)w_g + \gamma_{sr}\kappa\hat{w}_{sr}, \tag{29}$$

where $\hat{w}_{sr} = \gamma_{sr}^{-1}\boldsymbol{\Sigma}^{-1}\hat{\mu}$. From Kan, Wang, and Zhou (2021), we have that the combination coefficient κ maximizing the expected out-of-sample utility is given by

$$\kappa_{\gamma}^{\sigma} = \arg\max_{\kappa \in \mathbb{R}} \mathcal{U}(\hat{w}^{\sigma}(\kappa)) = \frac{1}{\gamma}\left(\frac{\psi^2}{\psi^2 + \frac{N-1}{T}}\right). \tag{30}$$

In the next proposition, we derive analytical expressions for the approximate expected out-of-sample Sharpe ratio of $\hat{w}^{\sigma}(\kappa)$, the resulting optimal κ, and the exact expected out-of-sample Sharpe ratio, which are new in the literature. The approximate Sharpe ratio does not require a distributional assumption beyond returns being i.i.d., whereas the exact one assumes i.i.d. multivariate normality.

Proposition 4. *Let the asset returns be i.i.d. and the covariance matrix* $\boldsymbol{\Sigma}$ *be known. Consider*

the fully invested portfolio combination $\hat{w}^{\sigma}(\kappa)$ *in* (28).

1. *The approximate expected out-of-sample Sharpe ratio of* $\hat{w}^{\sigma}(\kappa)$ *is given by*

$$\widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\kappa)) = \frac{\mu_g + \kappa\psi^2}{\sqrt{\sigma_g^2 + \kappa^2(\psi^2 + \frac{N-1}{T})}}. \tag{31}$$

The combination coefficient $\kappa \in \mathbb{R}$ *maximizing* $\widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\kappa))$ *is*

$$\tilde{\kappa}_{sr}^{\sigma} = \frac{1}{\gamma_{sr}}\left(\frac{\psi^2}{\psi^2 + \frac{N-1}{T}}\right), \tag{32}$$

and it delivers the following approximate expected out-of-sample Sharpe ratio:

$$\widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\tilde{\kappa}_{sr}^{\sigma})) = \sqrt{\frac{\psi^2\theta^2 + \frac{N-1}{T}\theta_g^2}{\psi^2 + \frac{N-1}{T}}}. \tag{33}$$

2. *Let the asset returns be multivariate normally distributed. The exact expected out-of-sample Sharpe ratio of* $\hat{w}^{\sigma}(\kappa)$*, provided the integral exists,*[5] *is given by*

$$\mathcal{S}(\hat{w}^{\sigma}(\kappa)) = \int_0^{\infty}\left(1 + \frac{2\kappa^2 t}{T}\right)^{-\frac{N+1}{2}}\left[\left(1 + \frac{2\kappa^2 t}{T}\right)\mu_g + \kappa\psi^2\right]\exp\left\{-t\left(\sigma_g^2 + \frac{\kappa^2 T\psi^2}{T + 2\kappa^2 t}\right)\right\}(\pi t)^{-\frac{1}{2}}\mathrm{d}t. \tag{34}$$

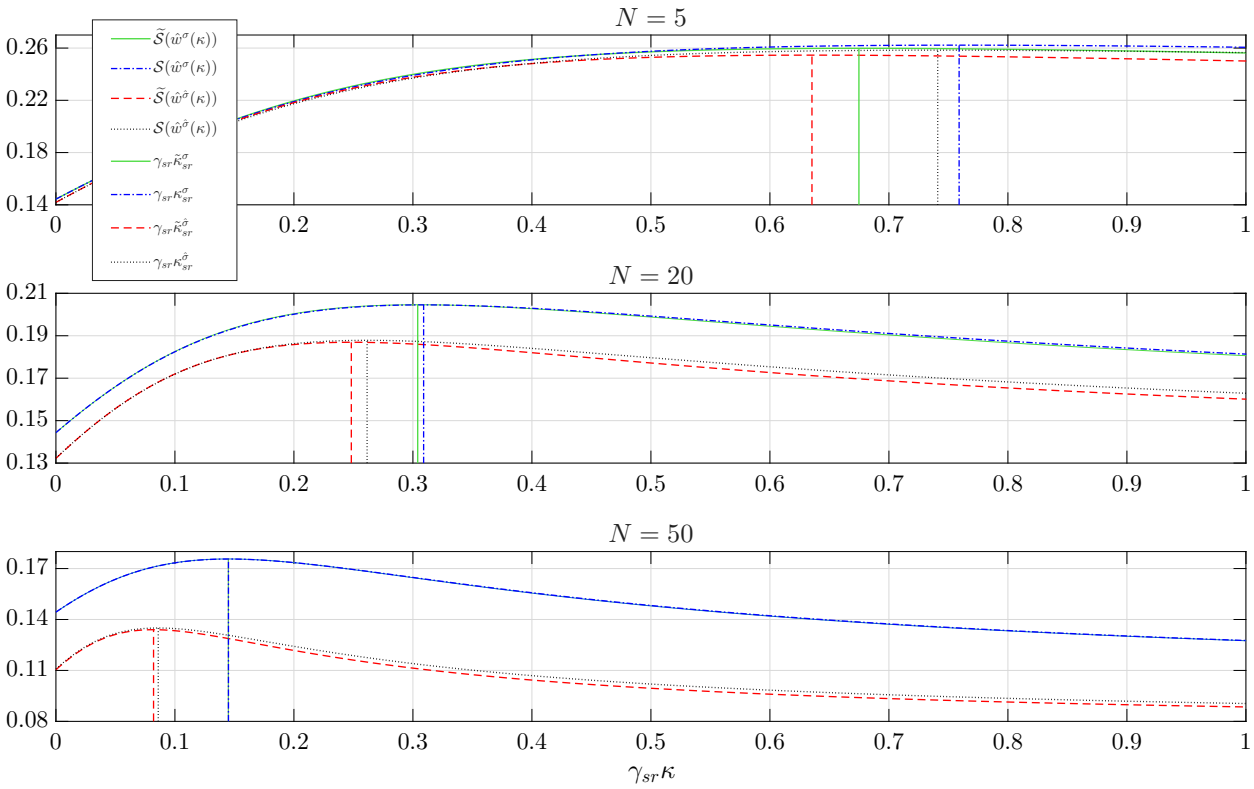
A number of comments follow. First, to maximize the approximate expected out-of-sample Sharpe ratio $\widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\kappa))$, one can naively apply the combination coefficient maximizing the expected out-of-sample utility in (30) and replace γ by γ_{sr} in (2). That is, $\tilde{\kappa}_{sr}^{\sigma} = \kappa_{\gamma_{sr}}^{\sigma}$. Referring to the result in Proposition 3, this is because $\iota'\boldsymbol{S}^{-1}\eta = \gamma_{sr}$ for the considered portfolio combination in (28).

Second, to maximize the approximate expected out-of-sample Sharpe ratio $\widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\kappa))$, the GMV portfolio is shrunk toward the sample tangency portfolio with an intensity $\psi^2/(\psi^2 + (N-1)/T)$.

Third, a relevant by-product of the proof of Proposition 4 is a stochastic representation for the exact distribution of the out-of-sample Sharpe ratio of $\hat{w}^{\sigma}(\kappa)$.[6]

In Figure 2, we depict the approximate and exact expected out-of-sample Sharpe ratio, $\widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\kappa))$ in (31) and $\mathcal{S}(\hat{w}^{\sigma}(\kappa))$ in (34), as a function of the shrinkage intensity $\gamma_{sr}\kappa$. We also depict the resulting optimal shrinkage intensities, $\gamma_{sr}\kappa_{sr}^{\sigma}$ and $\gamma_{sr}\tilde{\kappa}_{sr}^{\sigma}$. We use the same values of N, T, μ_g, σ_g, and θ as in Figure 1. The figure depicts results for the case where $\boldsymbol{\Sigma}$ is known as in this section, but also under the sample covariance matrix $\hat{\boldsymbol{\Sigma}}$ as in Section 4.2.2. First, $\widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\kappa))$ provides a remarkably

Figure 2: Expected out-of-sample Sharpe ratio of fully invested sample mean-variance portfolios



Notes. This figure depicts the expected out-of-sample Sharpe ratio of fully invested portfolio combinations, $\hat{w}^{\sigma}(\kappa)$ in (28) when $\boldsymbol{\Sigma}$ is known and $\hat{w}^{\hat{\sigma}}(\kappa)$ in (35) under the sample estimator $\hat{\boldsymbol{\Sigma}}$ in (1), as a function of the shrinkage intensity $\gamma_{sr}\kappa$. We depict the approximate Sharpe ratios, $\tilde{\mathcal{S}}(\hat{w}^{\sigma}(\kappa))$ in (31) and $\tilde{\mathcal{S}}(\hat{w}^{\hat{\sigma}}(\kappa))$ in (39), the exact ones, $\mathcal{S}(\hat{w}^{\sigma}(\kappa))$ in (34) and $\mathcal{S}(\hat{w}^{\hat{\sigma}}(\kappa))$ in (40), and the resulting optimal $\gamma_{sr}\kappa$. We set $\mu_g = 0.05/12$, $\sigma_g = 0.1/\sqrt{12}$, $\theta = 0.3$, $T = 120$, and $N \in \{5, 20, 50\}$.

accurate approximation to the exact $\mathcal{S}(\hat{w}^{\sigma}(\kappa))$. Second, with the right shrinkage intensity, the investor can obtain a significantly larger expected out-of-sample Sharpe ratio than both the GMV and tangency portfolios. Third, compared to $\theta = 0.3$, the maximum exact expected out-of-sample Sharpe ratio is 0.262, 0.205, and 0.176 for $N = 5$, 20, and 50, respectively. Finally, the optimal shrinkage intensity of the GMV portfolio toward the sample tangency portfolio decreases with N, and it is similar under the approximate and exact expected out-of-sample Sharpe ratio except for $N = 5$. The approximation is conservative and implies a higher weight on GMV, i.e., $\tilde{\kappa}_{sr}^{\sigma} < \kappa_{sr}^{\sigma}$.

4.2.2 Uncertainty in mean returns and the covariance matrix

We now turn to the case where the covariance matrix $\boldsymbol{\Sigma}$ is estimated using the sample estimator $\hat{\boldsymbol{\Sigma}}$ in (1). In this case, given a combination coefficient κ, the fully invested portfolio combination is

$$\hat{w}^{\hat{\sigma}}(\kappa) = (1-\kappa)\hat{w}_g + \kappa(\hat{w}_g + \hat{w}_z^{\hat{\sigma}}) = \hat{w}_g + \kappa\hat{\boldsymbol{B}}\hat{\mu}. \tag{35}$$

From Kan, Wang, and Zhou (2021), we have under i.i.d. multivariate normality and $T \geq N+4$ that the combination coefficient κ maximizing the expected out-of-sample utility is given by

$$\kappa_\gamma^{\hat{\sigma}} = \arg\max_{\kappa \in \mathbb{R}} \mathcal{U}(\hat{w}^{\hat{\sigma}}(\kappa)) = \frac{(T-N)(T-N-3)}{\gamma T(T-2)} \left(\frac{\psi^2}{\psi^2 + \frac{N-1}{T}} \right). \tag{36}$$

In the next proposition, we derive analytical expressions under the same assumptions for the approximate expected out-of-sample Sharpe ratio of $\hat{w}^{\hat{\sigma}}(\kappa)$, the resulting optimal κ, and the exact expected out-of-sample Sharpe ratio, which are new in the literature.

Proposition 5. *Let the asset returns be i.i.d. multivariate normally distributed. Consider the fully invested portfolio combination $\hat{w}^{\hat{\sigma}}(\kappa)$ in (35) using the sample covariance matrix $\hat{\mathbf{\Sigma}}$ in (1).*

1. *Let $T \geq N+4$. The approximate expected out-of-sample Sharpe ratio of $\hat{w}^{\hat{\sigma}}(\kappa)$ is given by*

$$\widetilde{\mathcal{S}}(\hat{w}^{\hat{\sigma}}(\kappa)) = \frac{\mu_g + \frac{\kappa T \psi^2}{T-N-1}}{\sqrt{\frac{T-2}{T-N-1}\sigma_g^2 + \frac{\kappa^2 T(T-2)(T\psi^2+N-1)}{(T-N)(T-N-1)(T-N-3)}}}. \tag{37}$$

The combination coefficient $\kappa \in \mathbb{R}$ maximizing $\widetilde{\mathcal{S}}(\hat{w}^{\hat{\sigma}}(\kappa))$ is

$$\tilde{\kappa}_{sr}^{\hat{\sigma}} = \frac{(T-N)(T-N-3)}{\gamma_{sr} T(T-N-1)} \left(\frac{\psi^2}{\psi^2 + \frac{N-1}{T}} \right), \tag{38}$$

which is smaller than that when the covariance matrix is known in (32), and it delivers the following approximate expected out-of-sample Sharpe ratio:

$$\widetilde{\mathcal{S}}(\hat{w}^{\hat{\sigma}}(\tilde{\kappa}_{sr}^{\hat{\sigma}})) = \sqrt{\frac{T-N-1}{T-2}\theta_g^2 + \frac{(T-N)(T-N-3)}{(T-2)(T-N-1)} \left(\frac{\psi^2}{\psi^2 + \frac{N-1}{T}} \right) \psi^2}, \tag{39}$$

which is smaller than that when the covariance matrix is known in (33).

2. *Let $T \geq N+1$. The exact expected out-of-sample Sharpe ratio of $\hat{w}^{\hat{\sigma}}(\kappa)$ is given by*

$$\begin{aligned} \mathcal{S}(\hat{w}^{\hat{\sigma}}(\kappa)) = \frac{\Gamma(\frac{T}{2})\Gamma(\frac{T-N+3}{2})}{\Gamma(\frac{T+1}{2})\Gamma(\frac{T-N+2}{2})} \int_0^\infty \int_0^\infty (\pi v t)^{-\frac{1}{2}} (1+2\alpha)^{-\frac{N-1}{2}} \exp\left(-\frac{\alpha T \psi^2}{1+2\alpha} - vt \right) \\ \times \left(\frac{\theta_g v}{(1+2t)^{\frac{1}{2}}} + \frac{\kappa T \psi^2}{\sigma_g (1+2\alpha)(1+2t)^{\frac{3}{2}}} \right) f_v(v) \mathrm{d}t \mathrm{d}v, \end{aligned} \tag{40}$$

provided the integral exists, where f_v is the pdf of $v \sim \chi^2_{T-N+1}$ and $\alpha = \kappa^2 T t / [\sigma_g^2 v(1+2t)]$.

We note the following relevant comments. First, to maximize the approximate expected out-of-

sample Sharpe ratio, one *cannot* naively use the combination coefficient of Kan, Wang, and Zhou (2021) maximizing the expected out-of-sample utility in (36) and replace γ by γ_{sr}, unlike when the covariance matrix is known in Section 4.2.1. Instead, to obtain $\tilde{\kappa}_{sr}^{\hat{\sigma}}$ in (38), γ must be replaced by $\gamma_{sr}\frac{T-N-1}{T-2}$, which is the value of $\iota'\boldsymbol{S}^{-1}\eta$ in Proposition 3. In particular, one must allocate less weight to the sample GMV portfolio to maximize the approximate expected out-of-sample Sharpe ratio instead of the expected out-of-sample utility.

Second, introducing parameter uncertainty in the covariance matrix reduces the maximum Sharpe ratio that is achievable out of sample, which can be seen by comparing (39) with (33).

Third, relative to the case where the covariance matrix is known, a more intensive shrinkage toward the GMV portfolio is optimal in the presence of estimation errors in the covariance matrix. For example, when $T = 120$ and $N = 20$ in Figure 2, the optimal weight $1 - \gamma_{sr}\kappa$ on the GMV portfolio is 0.739 under the exact expected out-of-sample Sharpe ratio when the covariance matrix is unknown, versus 0.691 when it is known. Figure 2 also shows that the approximation to the expected out-of-sample Sharpe ratio is accurate, conservative, and delivers a slightly larger weight on the sample GMV portfolio and a good approximation for the optimal κ.

Fourth, a relevant by-product of the proof of Proposition 4 is a stochastic representation for the exact distribution of the out-of-sample Sharpe ratio of $\hat{w}^{\hat{\sigma}}(\kappa)$ obtained from Lassance, Kan, and Wang (2025, Theorem 1); see Equations (EC70)–(EC71) of the Electronic Companion.

5 Penalizing the variance of the out-of-sample Sharpe ratio

In our simulation and empirical analysis in Section 6, we focus on fully invested portfolios as in Section 4. This is because, as shown in Section 3, when a risk-free investment is allowed there is essentially no difference between maximizing the expected out-of-sample Sharpe ratio or the expected out-of-sample utility as done in prior literature. Now, when applying our optimal fully invested portfolio combination rules to the data, we need to estimate the combination coefficients κ that maximize the expected out-of-sample Sharpe ratio, i.e., κ_{sr}^{σ} $\tilde{\kappa}_{sr}^{\sigma}$, $\kappa_{sr}^{\hat{\sigma}}$, and $\tilde{\kappa}_{sr}^{\hat{\sigma}}$, which are oracles because they depend on the population parameters μ_g, σ_g^2, γ_{sr}, and ψ^2.[7] Estimation errors in κ introduce an additional layer of estimation risk that has a negative impact on performance.

To alleviate the impact of estimation errors in optimal combination coefficients, Kan and Wang (2023) and Lassance, Martín-Utrera, and Simaan (2024) show that shrinking the sample-optimal portfolio more intensively helps. In this section, we follow Lassance, Martín-Utrera, and Simaan (2024) who consider the portfolio combination $\hat{w}^{\hat{\sigma}}(\kappa)$ in (35) and find the κ that trades off between

the expectation and variance of the out-of-sample utility. Their idea is that, because the expected out-of-sample utility depends on the population μ and $\mathbf{\Sigma}$, we replace it with an estimator and construct an interval around it whose length is a multiple of the standard deviation of the out-of-sample utility. They can then find the κ maximizing the lower bound of this interval, resulting in a worst-case-robust portfolio. Similarly, we propose to find the combination coefficient κ that trades off between the expectation and variance of the out-of-sample Sharpe ratio. We show that it amounts to shrinking more intensively toward the GMV portfolio relative to maximizing the expected out-of-sample Sharpe ratio alone. We call the resulting portfolios as robust. Our empirical analysis in Section 6.4 indicates that the robust portfolio combination rules often improve the out-of-sample Sharpe ratio while reducing turnover due to the more intensive shrinkage.

In Sections 5.1 and 5.2, we derive exact and approximate formulas, respectively, for the variance of the out-of-sample Sharpe ratio of fully invested portfolio combination rules.

5.1 Exact variance of the out-of-sample Sharpe ratio

Following Lassance, Martín-Utrera, and Simaan (2024), we find the exact robust κ in $\hat{w}^{\sigma}(\kappa) = w_g + \kappa \boldsymbol{B}\hat{\mu}$ for a known $\mathbf{\Sigma}$, and in $\hat{w}^{\hat{\sigma}}(\kappa) = \hat{w}_g + \kappa \hat{\boldsymbol{B}}\hat{\mu}$ for the sample estimator of $\mathbf{\Sigma}$, as follows:

$$\kappa_{\lambda}^{\sigma} = \arg\max_{\kappa \in \mathbb{R}} \mathcal{S}(\hat{w}^{\sigma}(\kappa)) - \lambda \sqrt{\mathcal{V}(\hat{w}^{\sigma}(\kappa))}, \tag{41}$$

$$\kappa_{\lambda}^{\hat{\sigma}} = \arg\max_{\kappa \in \mathbb{R}} \mathcal{S}(\hat{w}^{\hat{\sigma}}(\kappa)) - \lambda \sqrt{\mathcal{V}(\hat{w}^{\hat{\sigma}}(\kappa))}, \tag{42}$$

where $\lambda \geq 0$, $\mathcal{S}(\hat{w})$ is the exact expected out-of-sample Sharpe ratio, and

$$\mathcal{V}(\hat{w}) = \mathbb{V}\left[\frac{\hat{w}'\mu}{\sqrt{\hat{w}'\mathbf{\Sigma}\hat{w}}}\right] \tag{43}$$

is the exact variance of the out-of-sample Sharpe ratio. We derive $\mathcal{S}(\hat{w}^{\sigma}(\kappa))$ and $\mathcal{S}(\hat{w}^{\hat{\sigma}}(\kappa))$ in (34) and (40), respectively. In the next proposition, we derive $\mathcal{V}(\hat{w}^{\sigma}(\kappa))$ and $\mathcal{V}(\hat{w}^{\hat{\sigma}}(\kappa))$ analytically.

Proposition 6. *The following holds about the exact variance of the out-of-sample Sharpe ratio:*

1. *Let the asset returns be i.i.d. and the covariance matrix $\mathbf{\Sigma}$ be known. Consider the fully invested portfolio combination $\hat{w}^{\sigma}(\kappa)$ in (28). We have $\mathcal{V}(\hat{w}^{\sigma}(\kappa)) = \mathbb{E}\big[(\hat{w}^{\sigma}(\kappa)'\mu)^2/\hat{w}^{\sigma}(\kappa)'\mathbf{\Sigma}\hat{w}^{\sigma}(\kappa)\big] - \mathcal{S}^2(\hat{w}^{\sigma}(\kappa))$ with $\mathcal{S}(\hat{w}^{\sigma}(\kappa))$ given by (34) and, provided the integral exists,*

$$\mathbb{E}\left[\frac{(\hat{w}^{\sigma}(\kappa)'\mu)^2}{\hat{w}^{\sigma}(\kappa)'\mathbf{\Sigma}\hat{w}^{\sigma}(\kappa)}\right] = \int_0^{\infty} \left(1 + \frac{2\kappa^2 t}{T}\right)^{-\frac{N+3}{2}} \left\{\left[\left(1 + \frac{2\kappa^2 t}{T}\right)\mu_g + \kappa\psi^2\right]^2 + \frac{\kappa^2\psi^2(T + 2\kappa^2 t)}{T^2}\right\}$$

$$\times \exp\left\{-t\left(\sigma_g^2 + \frac{\kappa^2\psi^2 T}{T+2\kappa^2 t}\right)\right\}\mathrm{d}t. \tag{44}$$

2. *Let the asset returns be i.i.d. multivariate normally distributed and* $T \geq N+1$. *Consider the fully invested portfolio combination* $\hat{w}^{\hat{\sigma}}(\kappa)$ *in* (35) *using* $\hat{\boldsymbol{\Sigma}}$ *in* (1). *We have* $\mathcal{V}(\hat{w}^{\hat{\sigma}}(\kappa)) = \mathbb{E}[(\hat{w}^{\hat{\sigma}}(\kappa)'\mu)^2/\hat{w}^{\hat{\sigma}}(\kappa)'\boldsymbol{\Sigma}\hat{w}^{\hat{\sigma}}(\kappa)] - \mathcal{S}^2(\hat{w}^{\hat{\sigma}}(\kappa))$ *with* $\mathcal{S}(\hat{w}^{\hat{\sigma}}(\kappa))$ *given by* (40) *and*

$$\mathbb{E}\left[\frac{(\hat{w}^{\hat{\sigma}}(\kappa)'\mu)^2}{\hat{w}^{\hat{\sigma}}(\kappa)'\boldsymbol{\Sigma}\hat{w}^{\hat{\sigma}}(\kappa)}\right] = \frac{(N-2)\psi^2}{(N-1)T}{}_1F_1\left(1;\frac{N+1}{2};-\frac{T\psi^2}{2}\right) + \frac{T-N+2}{T}\int_0^\infty\int_0^\infty (1+2\alpha)^{-\frac{N-1}{2}}$$
$$\times \exp\left(-\frac{T\psi^2\alpha}{1+2\alpha} - vt\right)\left\{\frac{\theta_g^2 v}{(1+2t)^{\frac{1}{2}}} + \frac{2\kappa T\theta_g\psi^2}{\sigma_g(1+2\alpha)(1+2t)^{\frac{3}{2}}} + \frac{\kappa^2 T\psi^2(1+2\alpha+T\psi^2)}{\sigma_g^2 v(1+2\alpha)^2(1+2t)^{\frac{5}{2}}}\right.$$
$$\left. + \frac{\psi^2}{(1+2t)^{\frac{3}{2}}}\left[1 - \frac{N-2}{N-1}{}_1F_1\left(1;\frac{N+1}{2};-\frac{T\psi^2}{2(1+2\alpha)}\right)\right]\right\} f_v(v)\mathrm{d}t\mathrm{d}v, \tag{45}$$

provided the integral exists, where $f_v(\cdot)$ *is the pdf of* $v \sim \chi^2_{T-N+1}$ *and* $\alpha = \kappa^2 Tt/[\sigma_g^2 v(1+2t)]$.

5.2 Approximate variance of the out-of-sample Sharpe ratio

We now employ a more tractable approximation to the variance of the out-of-sample Sharpe ratio of a portfolio $\hat{w}$, which is based on the following approximation for the second moment:

$$\widetilde{\mathbb{E}}\left[\left(\frac{\hat{w}'\mu}{\sqrt{\hat{w}'\boldsymbol{\Sigma}\hat{w}}}\right)^2\right] = \frac{\mathbb{E}[(\hat{w}'\mu)^2]}{\mathbb{E}[\hat{w}'\boldsymbol{\Sigma}\hat{w}]} = \mathbb{E}\left[\left(\frac{\hat{w}'\mu}{\sqrt{\hat{w}'\boldsymbol{\Sigma}\hat{w}}}\right)^2\right] + O(1/T). \tag{46}$$

The resulting approximation for $\mathcal{V}(\hat{w})$ is

$$\widetilde{\mathcal{V}}(\hat{w}) = \widetilde{\mathbb{E}}\left[\left(\frac{\hat{w}'\mu}{\sqrt{\hat{w}'\boldsymbol{\Sigma}\hat{w}}}\right)^2\right] - \widetilde{\mathcal{S}}^2(\hat{w}) = \frac{\mathbb{V}[\hat{w}'\mu]}{\mathbb{E}[\hat{w}'\boldsymbol{\Sigma}\hat{w}]}, \tag{47}$$

where $\widetilde{\mathcal{S}}(\hat{w})$ is defined in (7). The resulting approximate robust κ in $\hat{w}^{\sigma}(\kappa) = w_g + \kappa\boldsymbol{B}\hat{\mu}$ for a known $\boldsymbol{\Sigma}$, and in $\hat{w}^{\hat{\sigma}}(\kappa) = \hat{w}_g + \kappa\hat{\boldsymbol{B}}\hat{\mu}$ for the sample estimator of $\boldsymbol{\Sigma}$, are obtained as

$$\tilde{\kappa}_\lambda^{\sigma} = \arg\max_{\kappa\in\mathbb{R}} \widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\kappa)) - \lambda\sqrt{\widetilde{\mathcal{V}}(\hat{w}^{\sigma}(\kappa))}, \tag{48}$$

$$\tilde{\kappa}_\lambda^{\hat{\sigma}} = \arg\max_{\kappa\in\mathbb{R}} \widetilde{\mathcal{S}}(\hat{w}^{\hat{\sigma}}(\kappa)) - \lambda\sqrt{\widetilde{\mathcal{V}}(\hat{w}^{\hat{\sigma}}(\kappa))}. \tag{49}$$

We derive $\widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\kappa))$ and $\widetilde{\mathcal{S}}(\hat{w}^{\hat{\sigma}}(\kappa))$ in (31) and (37), respectively. In the next proposition, we derive $\widetilde{\mathcal{V}}(\hat{w}^{\sigma}(\kappa))$ and $\widetilde{\mathcal{V}}(\hat{w}^{\hat{\sigma}}(\kappa))$ analytically, and we characterize the resulting $\tilde{\kappa}_\lambda^{\sigma}$ and $\tilde{\kappa}_\lambda^{\hat{\sigma}}$.

Proposition 7. *The following holds about the approximate variance of out-of-sample Sharpe ratio:*

1. *Let the asset returns be i.i.d. and the covariance matrix* $\mathbf{\Sigma}$ *be known. Consider the fully invested portfolio combination* $\hat{w}^{\sigma}(\kappa)$ *in* (28)*. We have*

$$\widetilde{\mathcal{V}}(\hat{w}^{\sigma}(\kappa)) = \frac{1}{T}\left(\frac{\kappa^2\psi^2}{\sigma_g^2 + \kappa^2(\psi^2 + \frac{N-1}{T})}\right). \tag{50}$$

The resulting robust combination coefficient $\tilde{\kappa}_{\lambda}^{\sigma}$ *in* (48) *is given by*

$$\tilde{\kappa}_{\lambda}^{\sigma} = \begin{cases} \tilde{\kappa}_{sr}^{\sigma}\left(1 - \frac{\lambda}{\sqrt{T}\psi}\right) & \text{if } \lambda \leq \sqrt{T}\psi, \\ 0 & \text{if } \lambda \geq \sqrt{T}\psi, \end{cases} \tag{51}$$

where $\tilde{\kappa}_{sr}^{\sigma}$ *in* (32) *maximizes the approximate expected out-of-sample Sharpe ratio.*

2. *Let the asset returns be i.i.d. multivariate normally distributed and* $T \geq N + 4$*. Consider the fully invested portfolio combination* $\hat{w}^{\hat{\sigma}}(\kappa)$ *in* (35) *using* $\hat{\mathbf{\Sigma}}$ *in* (1)*. We have*

$$\widetilde{\mathcal{V}}(\hat{w}^{\hat{\sigma}}(\kappa)) = \frac{1}{T-N-1}\left(\frac{\sigma_g^2\psi^2 + \frac{\kappa^2\psi^2[T(T-2)(T-N-1)+2T^2(T-N)\psi^2]}{(T-N)(T-N-1)(T-N-3)}}{\frac{T-2}{T-N-1}\sigma_g^2 + \frac{\kappa^2T(T-2)(T\psi^2+N-1)}{(T-N)(T-N-1)(T-N-3)}}\right). \tag{52}$$

The resulting robust combination coefficient $\tilde{\kappa}_{\lambda}^{\hat{\sigma}}$ *in* (49) *satisfies* $\tilde{\kappa}_{\lambda}^{\hat{\sigma}} = \tilde{\kappa}_{sr}^{\hat{\sigma}}$ *if* $\lambda = 0$*,* $\tilde{\kappa}_{\lambda}^{\hat{\sigma}} = 0$ *if* $\lambda \to \infty$*, and* $0 \leq \tilde{\kappa}_{\lambda}^{\hat{\sigma}} \leq \tilde{\kappa}_{sr}^{\hat{\sigma}}$ *otherwise with strict inequalities if* $\psi^2 > 0$*, where* $\tilde{\kappa}_{sr}^{\hat{\sigma}}$ *in* (38) *maximizes the approximate expected out-of-sample Sharpe ratio.*

Proposition 7 shows that penalizing the approximate variance of the out-of-sample Sharpe ratio amounts to shrinking more intensively toward the GMV portfolio: $\tilde{\kappa}_{\lambda}^{\sigma} \leq \tilde{\kappa}_{sr}^{\sigma}$ and $\tilde{\kappa}_{\lambda}^{\hat{\sigma}} \leq \tilde{\kappa}_{sr}^{\hat{\sigma}}$.[8] The impact of the variance penalization on the optimal κ is different if $\mathbf{\Sigma}$ is known or unknown. When $\mathbf{\Sigma}$ is known, Equation (51) shows that the optimal $\kappa = 0$ if $\lambda \geq \sqrt{T}\psi$, in which case we recover the GMV portfolio. However, when using the sample covariance matrix, $\tilde{\kappa}_{\lambda}^{\hat{\sigma}} = 0$ only as $\lambda \to \infty$. This difference can be explained because, when $\mathbf{\Sigma}$ is known, the variance of the out-of-sample Sharpe ratio of the GMV portfolio is zero, and thus, it is optimal to fully shrink toward GMV once λ is large enough. In contrast, when $\mathbf{\Sigma}$ is estimated, the out-of-sample Sharpe ratio of GMV has a non-zero variance if T is finite, and thus, it is not optimal to fully shrink toward GMV.

6 Simulation and empirical analysis

In this section, we first run a simulation analysis where returns are i.i.d. multivariate normal, which, as Kan, Wang, and Zheng (2024) argue, allows us to isolate estimation risk from other factors that

can explain the empirical gap between in-sample and out-of-sample performance. Then, we conduct a more challenging and realistic empirical analysis. Throughout this section, we focus on fully invested portfolios because, when a risk-free investment is allowed, there is essentially no difference between maximizing the expected out-of-sample Sharpe ratio or the expected out-of-sample utility as done in prior literature, as shown in Section 3.

6.1 Datasets

In the simulation and empirical analysis, we consider four datasets of monthly excess returns downloaded from Kenneth French's website that are commonly used in the literature: (i) 10 portfolios sorted on momentum (10MOM), (ii) 25 portfolios sorted on size and book-to-market (25SBM), (iii) 25 portfolios sorted on operating profitability and investment (25OPIN), and (iv) 30 industry portfolios (30IND). The time period for each dataset spans July 1963 to May 2025.

6.2 Portfolio strategies

We consider 11 fully invested portfolio strategies, composed of five novel strategies introduced in this paper and six benchmark strategies, which we detail in Table 1.

The five novel strategies are the following. First and second, 'Expected Sharpe (approx)' and 'Expected Sharpe (exact)' are the mean-variance portfolios maximizing the approximate and exact expected out-of-sample Sharpe ratio, respectively, as studied in Section 4. Third and fourth, 'Robust Sharpe (approx)' and 'Robust Sharpe (exact)' are the mean-variance portfolios maximizing the tradeoff between the expectation and variance of the out-of-sample Sharpe ratio using the exact and approximate moments, respectively, as studied in Section 5. We use a coefficient $\lambda = 2$ as in Lassance, Martín-Utrera, and Simaan (2024).[9] Fifth, the equally weighted combination of these four portfolio strategies. For these five strategies, we use the optimal combination coefficients κ found under the sample covariance matrix $\hat{\boldsymbol{\Sigma}}$ in (1), or a known covariance matrix that we replace with the nonlinear shrinkage estimator of Ledoit and Wolf (2020).

The six benchmark strategies are the following. First, 'Bootstrap Sharpe' is the bootstrap methodology of Kircher and Rosch (2021) for finding the mean-variance portfolio maximizing the expected out-of-sample Sharpe ratio, which we detail in Section EC.1 of the Electronic Companion. Second, 'Expected utility' is the mean-variance portfolio maximizing the expected out-of-sample utility developed by Kan, Wang, and Zhou (2021) and obtained using Equations (30) and (36) under a known covariance matrix or its sample estimator, respectively, for which we use a risk-aversion coefficient $\gamma = 5$.[10] Third, 'GMV' is the global minimum-variance portfolio in (4). Fourth,

'Tangency' is the tangency portfolio in (27). Fifth, 'MAXSER' is the portfolio strategy of Ao, Li, and Zheng (2019) that we adapt to the setting without a risk-free asset; see Section EC.2 of the Electronic Companion. Sixth, '$1/N$' is the equally weighted portfolio.[11] All benchmark strategies except MAXSER and $1/N$ need an estimate of the covariance matrix, and we consider the sample estimator, $\hat{\mathbf{\Sigma}}$ in (1), or the nonlinear shrinkage estimator of Ledoit and Wolf (2020).

In Section EC.3 of the Electronic Companion, we explain how we estimate the parameters μ_g, σ_g^2, and ψ^2 in (5), as well as γ_{sr} in (2), on which many of the portfolio strategies depend.

6.3 Simulation analysis

6.3.1 Methodology for simulation analysis

The simulated data are i.i.d. and drawn from the multivariate normal distribution.[12] Specifically, we calibrate the population μ and $\mathbf{\Sigma}$ to the full sample of each of the four datasets in Section 6.1 and then randomly draw $M = 10{,}000$ times T independent vector of monthly excess returns from a multivariate normal distribution. For a given portfolio strategy, we obtain estimated portfolios $\hat{w}_m$ for $m = 1, \ldots, M$ and evaluate the annualized expected out-of-sample Sharpe ratio as

$$\frac{\sqrt{12}}{M} \sum_{m=1}^{M} \frac{\hat{w}_m' \mu}{\sqrt{\hat{w}_m' \mathbf{\Sigma} \hat{w}_m}}. \tag{53}$$

We consider a sample size $T \in \{60, 120, 240, 480\}$, i.e., 5, 10, 20, and 40 years of returns. Among the portfolio strategies listed in Table 1, we restrict to Expected Sharpe, Robust Sharpe, Expected utility, GMV, Tangency, and $1/N$ in this simulation analysis for brevity. We estimate the covariance matrix either with the Ledoit and Wolf (2020) covariance matrix or the sample covariance matrix.

6.3.2 Discussion of results for simulation analysis

In Tables 2 and 3, we report the annualized expected out-of-sample Sharpe ratio, and the average shrinkage intensity of the GMV portfolio toward the tangency portfolio (i.e., $\hat{\gamma}_{sr}\kappa$), for the Ledoit-Wolf and sample covariance matrix, respectively. We also report θ in (3) and γ_{sr} in (2) for each dataset, which are the maximum Sharpe ratio without estimation risk and the risk-aversion coefficient implied by the tangency portfolio, respectively. We make a number of observations.

First, although the $1/N$ portfolio can sometimes outperform the tangency portfolio when the sample size $T = 60$ or 120, it is consistently outperformed by all the other considered strategies. Thus, when asset returns are multivariate normally distributed and there are no transaction costs,

it is not optimal to fully ignore the information in the mean and covariance matrix of returns.

Second, it is generally useful to combine the GMV and tangency portfolios. Specifically, our four novel portfolio combinations mostly deliver a larger Sharpe ratio than the GMV and tangency portfolios. This result holds even for the 30IND dataset, which is known to be more challenging for strategies exploiting mean returns because there is less cross-sectional variation in mean returns across industry portfolios than characteristic-sorted portfolios. For instance, when $T = 120$, the Expected Sharpe (exact) portfolio delivers an annualized expected out-of-sample Sharpe ratio in the 30IND dataset of 0.538 and 0.487 under the Ledoit-Wolf and sample covariance matrix, respectively, versus 0.499 and 0.447 for the GMV portfolio and 0.420 and 0.355 for the tangency portfolio.

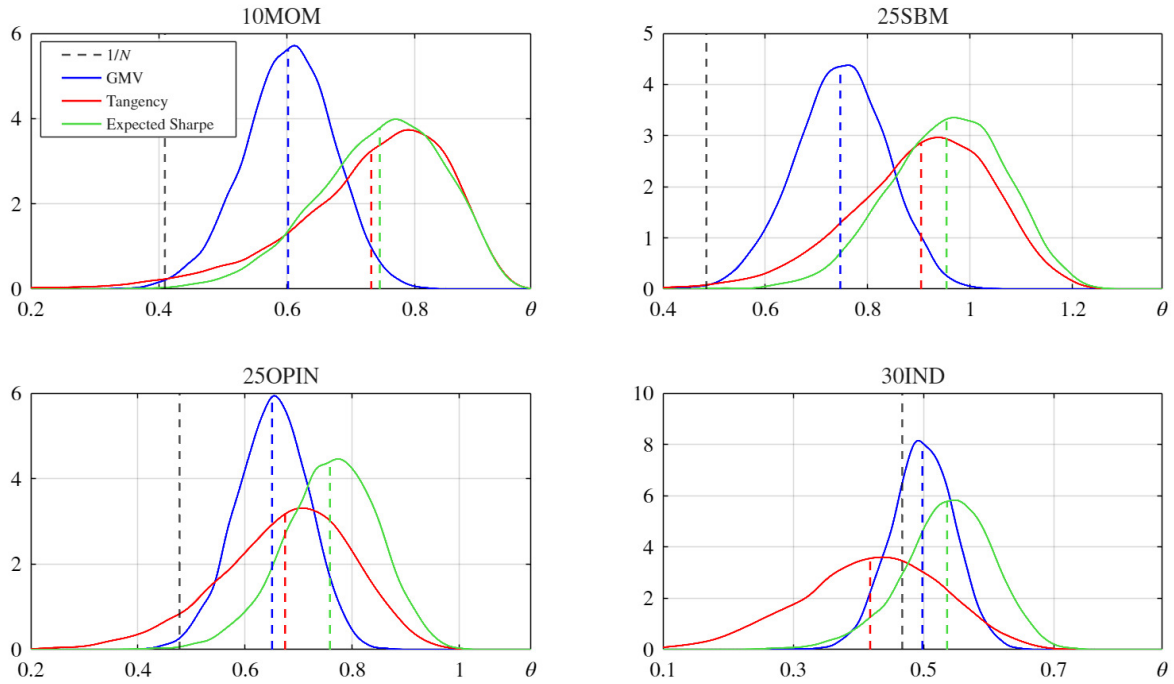
In Figure 3, we illustrate these first two observations by depicting the density of the out-of-sample Sharpe ratio obtained with the GMV, Tangency, and Expected Sharpe (exact) portfolios, as well as the constant $1/N$ Sharpe ratio, under the Ledoit-Wolf covariance matrix and $T = 120$. The figure shows that the out-of-sample Sharpe ratio of the proposed Expected Sharpe portfolio is not only larger than that of the tangency portfolio on average but is also less volatile because shrinking toward the GMV portfolio reduces estimation risk. Moreover, in this figure, the proposed Expected Sharpe portfolio delivers the largest expected out-of-sample Sharpe ratio for all datasets.

Third, comparing the proposed mean-variance portfolios maximizing the expected out-of-sample Sharpe ratios with those maximizing the expected out-of-sample utility as is standard in the literature, the latter can deliver a slightly larger Sharpe ratio than the former when, by chance, the investor's risk-aversion coefficient γ is close to the risk-aversion coefficient implied by the tangency portfolio, i.e., γ_{sr}. In our case, this happens when maximizing the expected utility with $\gamma = 5$, because γ_{sr} varies between 4.55 and 6.72 across datasets. The explanation is that fixing γ entails less estimation risk than having to estimate γ_{sr}. However, when the investor's γ is further away from γ_{sr}, such as the $\gamma = 3$ we consider, our proposed portfolios deliver a larger Sharpe ratio.

Fourth, the proposed Robust Sharpe portfolios that penalize the variance of the out-of-sample Sharpe ratio can outperform the Expected Sharpe portfolios when estimation risk is large, i.e., when $T = 60$ and the sample covariance matrix is used, and for the 30IND dataset that contains a weaker return signal. For instance, when $T = 60$, the Robust Sharpe (approx) portfolio delivers the largest expected out-of-sample Sharpe ratio in the 30IND dataset, which is equal to 0.503 and 0.384 for the Ledoit-Wolf and sample covariance matrix, respectively, versus 0.485 and 0.371 for the Exact Sharpe (approx) portfolio. Otherwise, when the sample size is larger and when the Ledoit-Wolf covariance matrix is used, the Robust Sharpe strategy underperforms Expected Sharpe.

Fifth, we evaluate the impact of using the exact or approximate expectation and variance of

Figure 3: Density of the out-of-sample Sharpe ratio in simulated data with Ledoit-Wolf covariance matrix and $T = 120$



Notes. This figure depicts the density of the out-of-sample Sharpe ratio delivered by the $1/N$ portfolio (dotted black vertical line), the GMV portfolio (in blue), the tangency portfolio (in red), and the Expected Sharpe (exact) portfolio (in green). These four strategies are described in Section 6.2 and Table 1. We estimate the covariance matrix with the nonlinear shrinkage estimator of Ledoit and Wolf (2020). The densities are obtained across 10,000 datasets of $T = 120$ monthly returns drawn from a multivariate normal distribution, with the population mean and covariance matrix calibrated to the full sample of the 10MOM, 25SBM, 25OPIN, and 30IND datasets described in Section 6.1. The dotted vertical lines depict the mean of each density. On the x-axis labels, θ in (3) is the maximum attainable Sharpe ratio without estimation risk.

the out-of-sample Sharpe ratio. For the Expected Sharpe strategy, there is little difference between the exact and approximate strategy, consistent with Figure 2. The main difference is that, under the sample covariance matrix, the approximate strategy shrinks more intensively toward the GMV portfolio, which is valuable for $T = 60$ and 120. Therefore, for Expected Sharpe, there is an incentive to use the approximate expected out-of-sample Sharpe ratio rather than the exact one that involves a single or a double integral. For the Robust Sharpe strategy, there are larger differences between using the exact or approximate strategy. In particular, the approximate strategy shrinks significantly more toward the GMV portfolio, which improves the Sharpe ratio only when $T = 60$ under the sample covariance matrix, and also for the 30IND dataset. Therefore, for Robust Sharpe, it is generally valuable to use the exact expectation and variance of the out-of-sample Sharpe ratio, at the cost of computational slowdown as they involve single or double integrals.

Sixth, in 84% of the cases, using the Ledoit-Wolf covariance matrix in Table 2 delivers a larger

expected out-of-sample Sharpe ratio than using the sample covariance matrix in Table 3. Therefore, it is generally preferable to use the Ledoit-Wolf covariance matrix and the wrong combination coefficients that assume the covariance matrix is known, rather than using the sample covariance matrix and the correct combination coefficients.

6.4 Empirical analysis

6.4.1 Methodology for empirical analysis

We consider the four datasets in Section 6.1 and all 11 portfolio strategies described in Section 6.2 and Table 1. As before, we estimate the covariance matrix in these different portfolio strategies with the nonlinear shrinkage estimator of Ledoit and Wolf (2020) or the sample estimator. In Section EC.4 of the Electronic Companion, we test the robustness of empirical results to imposing lower and upper-bound constraints on the portfolio weights.

For each portfolio strategy, we evaluate the out-of-sample Sharpe ratio using rolling windows; see, e.g., DeMiguel, Garlappi, and Uppal (2009). Specifically, given a sample size T, we estimate the portfolio on the estimation window composed of the first T months of the sample and compute the out-of-sample portfolio return on the next month. Then, we roll forward the estimation window by one month, rebalance the portfolio using data from the new estimation window, and compute again the next-month out-of-sample return. We proceed iteratively until the end of the sample. Rebalancing the portfolio every month incurs transaction costs, and we consider both the case with zero transaction costs and the case where the out-of-sample portfolio returns are adjusted to turnover using proportional transaction costs of 10 basis points, in line with recent values of average bid-ask spreads in Engle, Ferstenberg, and Russell (2012) and Frazzini, Israel, and Moskowitz (2018); our conclusions are robust to using 20 basis points. We use a sample size $T = 120$ and 240 months.

With the time series of out-of-sample returns, we compare the 12 portfolio strategies in terms of annualized gross and net-of-cost out-of-sample Sharpe ratio, which is the objective function that we consider in this paper. We also test statistically whether our four novel portfolio strategies that use information in sample mean returns outperform the GMV portfolio that ignores this information. Specifically, we compute two-sided p-values for the difference in out-of-sample Sharpe ratios using the block bootstrap approach of Ledoit and Wolf (2008), with 1,000 bootstrap resamples and a block size of five as in DeMiguel et al. (2009). In addition to the Sharpe ratio, we also report for each strategy the average monthly turnover induced by rebalancing. Finally, for those portfolio strategies that combine the GMV and tangency portfolios (i.e., all those in Table 1 except MAXSER

and $1/N$), we report the average weight allocated to the tangency portfolio across all estimation windows. As explained in Section 4.2.1, given a combination coefficient κ, this weight is $\hat{\gamma}_{sr}\kappa$. It is also the intensity with which the GMV portfolio is shrunk toward the tangency portfolio.

6.4.2 Discussion of results for empirical analysis

In Tables 4 and 5, we report, for the four datasets and 12 portfolio strategies, the annualized gross and net-of-cost out-of-sample Sharpe ratio, the average monthly turnover, and the average shrinkage intensity of the GMV portfolio toward the tangency portfolio (i.e., $\hat{\gamma}_{sr}\kappa$) for the Ledoit-Wolf and sample covariance matrix, respectively. We make a number of observations.

First, we find that the $1/N$ portfolio performs reasonably well both before and after transaction costs. In particular, on average across all configurations (i.e., four datasets, two sample sizes, and two covariance matrix estimators), the out-of-sample Sharpe ratio of $1/N$ is 0.510 and 0.507 before and after transaction costs, respectively. However, our four novel portfolio rules largely outperform $1/N$ even after transaction costs, therefore capturing the information in asset returns in a manner that is sufficiently effective for outperforming a myopic approach.

Second, the four optimal portfolio combinations we propose deliver better out-of-sample Sharpe ratios overall than the two combined portfolios, i.e., GMV and tangency. Relative to the tangency portfolio, all four of our novel portfolio rules systematically outperform with and without transaction costs, which highlights the clear value of combining the tangency portfolio with GMV. Relative to the GMV portfolio, and looking at the 10MOM, 25SBM, and 25OPIN datasets, all four proposed portfolio rules deliver a larger Sharpe ratio with and without transaction costs. As discussed in Section 6.3.2, the 30IND dataset is more challenging for strategies exploiting sample mean returns. In this case, when $T = 120$, our Robust Sharpe (approx) portfolio that shrinks most toward GMV among the four strategies we develop delivers a Sharpe ratio of 0.498 and 0.459 on average before and after transaction costs, respectively, versus 0.495 and 0.459 for GMV. Overall, across all configurations for the Ledoit-Wolf covariance matrix, the Combination Sharpe portfolio that equally combines the four strategies delivers on average an out-of-sample Sharpe ratio of 0.891 and 0.847 before and after costs, respectively, versus 0.760 and 0.732 for GMV, as written in the abstract.

Third, the mean-variance portfolio maximizing the expected out-of-sample utility with a risk-aversion coefficient $\gamma = 3$ or 5 delivers almost systematically a smaller out-of-sample Sharpe ratio than that of the four proposed portfolio rules designed to maximize it. However, in unreported results, we find that it delivers a larger out-of-sample utility, which highlights the importance of optimizing the combination coefficient using the investor's specific objective function.

Fourth, the Robust Sharpe portfolio that penalizes the variance of the out-of-sample Sharpe ratio shrinks more intensively toward the GMV portfolio. For example, the average weight across all configurations allocated to the tangency portfolio (i.e., $\gamma_{sr}\kappa$) with Robust Sharpe (exact) is 27.7% smaller than the average weight with Expected Sharpe (exact). This more intensive shrinkage has two benefits. First, it can help improve performance, particularly under transaction costs. Across all configurations, Robust Sharpe (exact) and Robust Sharpe (approx) deliver on average a net-of-cost out-of-sample Sharpe ratio of 0.809 and 0.820, respectively, versus 0.798 and 0.802 for Expected Sharpe (exact) and Expected Sharpe (approx). The improvement is even larger for the 30IND dataset, with respective average net-of-cost Sharpe ratios of 0.516 and 0.524 versus 0.484 and 0.485. The second benefit is that the Robust Sharpe strategies deliver a lower turnover than the Expected Sharpe strategies. Across all configurations, the Robust Sharpe (exact) and Robust Sharpe (approx) strategies have a turnover that is on average 10.5% and 24.7% smaller than the Expected Sharpe (exact) and Expected Sharpe (approx) strategies. Note that the approximate portfolio rules perform overall as well as the exact rules, while delivering a lower turnover and being computationally more efficient because they do not involve single or double integrals.

Fifth, the Bootstrap Sharpe strategy of Kircher and Rosch (2021), and the fully invested version of the MAXSER strategy of Ao, Li, and Zheng (2019), achieve their objective of maximizing the out-of-sample Sharpe ratio; they are sometimes the best considered strategy. However, overall, they underperform our four novel portfolio rules and deliver a larger turnover. In particular, across all configurations for the Ledoit-Wolf covariance matrix, the Combination Sharpe portfolio that equally combines our four novel strategies delivers on average an out-of-sample Sharpe ratio of 0.891 and 0.847 before and after costs, respectively, versus 0.866 and 0.819 for Bootstrap Sharpe and 0.802 and 0.728 for MAXSER, as written in the abstract. In terms of turnover, Robust Sharpe is significantly more stable. In particular, across all configurations for the Ledoit-Wolf covariance matrix, Robust Sharpe (exact) and Robust Sharpe (approx) deliver a turnover 15.7% and 31.8% smaller on average than Boostrap Sharpe, respectively, and 51.9% and 62.1% smaller than MAXSER.

We close with a discussion on the i.i.d. multivariate normality assumption for the asset returns underlying most of our proposed portfolio strategies. It is notable how well our proposed strategies perform empirically given that observed returns typically deviate from this assumption. Under i.i.d. normality, the out-of-sample performance loss originates solely from estimation errors in the mean vector and covariance matrix of asset returns; we abstract away from the effect of fat tails and serial dependence. The great empirical performance delivered by our theoretically optimal out-of-sample portfolios on the sample mean-variance efficient frontier under i.i.d. normality means

that abstracting away from other sources of estimation risk than those stemming from the mean vector and covariance matrix already provides a good approximation to the optimal out-of-sample solution. This result suggests that our analytical expressions for the expected out-of-sample Sharpe ratio are robust to moderate deviations from i.i.d. normality, like those we can observe with monthly-frequency returns. Accounting for fat tails and serial dependence would in principle provide further out-of-sample Sharpe ratio gains (see Kan and Lassance (2025) for the effect of fat tails on out-of-sample performance), but not to the point that ignoring them destroys the performance. However, extending the theory to accommodate fat tails and serial dependence would be at the cost of some loss of analytical tractability, which goes against our primary objective of studying analytically the out-of-sample Sharpe ratio of sample mean-variance portfolios. One can incorporate other sources of estimation risk that matter empirically using data-driven methods like bootstrap, following Kircher and Rosch (2021) who maximize the out-of-sample Sharpe ratio using a bootstrap-based shrinkage approach. However, as previously discussed, our portfolio strategies generally outperform this benchmark, which suggests that accounting for empirical sources of estimation risk beyond those captured under i.i.d. normality is not trivial to do in a way that generalizes well out of sample.

7 Conclusion

The Sharpe ratio is a widespread portfolio performance measure. What is particularly relevant for investors is to maximize the out-of-sample Sharpe ratio under estimation risk. However, so far, no exact analytical approach exists for finding the optimal mean-variance portfolio, with or without a risk-free investment, maximizing the expected out-of-sample Sharpe ratio. This paper addresses this challenge. We focus mostly on the setting without a risk-free asset and derive exact and approximate analytical expressions for the expectation and variance of the out-of-sample Sharpe ratio of sample mean-variance efficient portfolios. This allows us to develop portfolio strategies maximizing the expected out-of-sample Sharpe ratio, and robust versions penalizing the Sharpe ratio variance following Lassance, Martín-Utrera, and Simaan (2024). The intensity with which the sample-optimal portfolio is shrunk toward the global minimum-variance portfolio to maximize the expected out-of-sample Sharpe ratio differs from the intensity maximizing the expected out-of-sample utility in Kan, Wang, and Zhou (2021), which is the standard criterion in the literature on portfolio choice with estimation risk. Empirically, our strategies are overall superior to six relevant benchmarks in terms of out-of-sample Sharpe ratio and turnover. In future research, our stochastic representations for the distribution of the out-of-sample Sharpe ratio of sample mean-variance

portfolios, with and without a risk-free asset, may be used for finite-sample and asymptotic inference and may prove useful in adjacent applications where the Sharpe ratio is a decision criterion.

Code and data

The code and data required to replicate the results are published alongside the paper.

Notes

[1]We focus on mean-variance portfolios, which are analytically more tractable and less prone to estimation risk than higher-moment strategies (Simaan, 2014; Khashanah, Simaan, and Simaan, 2022; Lassance, 2022)

[2]See DeMiguel, Garlappi, and Uppal (2009), Frahm and Memmel (2010), Tu and Zhou (2011), DeMiguel, Martín-Utrera, and Nogales (2013, 2015), Branger, Lučivjanská, and Weissensteiner (2019), Kan, Wang, and Zhou (2021), Bodnar, Okhrin, and Parolya (2023), Kan and Wang (2023), Kan and Lassance (2025), Lassance, Martín-Utrera, and Simaan (2024), Lassance, Vanderveken, and Vrins (2024), and Kelly et al. (2025).

[3]Kan, Wang, and Zheng (2024) derive the exact distribution (and expectation) of the out-of-sample Sharpe ratio of sample mean-variance portfolios with a risk-free asset. However, unlike our objective of constructing optimal portfolios, they focus on studying the out-of-sample performance of multi-factor asset-pricing models.

[4]Let $b \sim \text{Beta}((T-N+1)/2,(N-1)/2)$, $r \sim n_1/(n'n)^{\frac{1}{2}}$ with $n \sim \mathcal{N}(0_{N-1}, \boldsymbol{I}_{N-1})$, $q \sim y/(y^2+u)^{\frac{1}{2}}$ with $y \sim \mathcal{N}\big(\sqrt{T}(\theta^2+\omega\mu_g/\sigma_g^2)/\theta, 1\big)$ and $u \sim \chi^2_{N-1}\big(T\omega^2\psi^2/(\sigma_g^2\theta^2)\big)$, where (b,r,q) are independent of each other. Then, the out-of-sample Sharpe ratio $\hat{w}(\kappa)'\mu/(\hat{w}(\kappa)'\boldsymbol{\Sigma}\hat{w}(\kappa))^{\frac{1}{2}} \sim \theta\big(\sqrt{b}q+\sqrt{(1-b)(1-q^2)}r\big)$.

[5]The expected out-of-sample Sharpe ratio, like any expectation of a continuous random variable, is an integral that in principle may not exist. We add this condition for caution given that the existence condition is difficult to study.

[6]Let $z \sim \mathcal{N}(\sqrt{T}\psi, 1)$, $u \sim \chi^2_{N-2}$, and they are independent of each other. Then, the out-of-sample Sharpe ratio $\hat{w}^\sigma(\kappa)'\mu/[\hat{w}^\sigma(\kappa)'\boldsymbol{\Sigma}\hat{w}^\sigma(\kappa)]^{\frac{1}{2}} \sim (\mu_g+\kappa\psi z/\sqrt{T})/[\sigma_g^2+\kappa^2(z^2+u)/T]^{\frac{1}{2}}$.

[7]In Section EC.3 of the Electronic Companion, we explain how we estimate the parameters μ_g, σ_g^2, γ_{sr}, and ψ^2.

[8]Kan and Wang (2023, Section 3.2) account for estimation errors in their optimal combination coefficients by shrinking more intensively toward the risk-free asset, and they calibrate the optimal intensity via cross-validation taking into account that combination coefficients are estimated. Similarly, we shrink more intensively toward the GMV portfolio, and the intensity depends on the coefficient λ that can also be calibrated using cross-validation.

[9]In unreported results, we also consider $\lambda = 1$ and calibrating λ via cross-validation and find similar conclusions.

[10]We have also considered a risk-aversion coefficient $\gamma = 3$ but omit these results because it delivers a systematically smaller out-of-sample Sharpe ratio than $\gamma = 5$, thus providing a weaker benchmark.

[11]We have also considered the reward-to-risk timing strategy in Kirby and Ostdiek (2012, Equation (13)) but omit these results because it performs similarly to $1/N$.

[12]In unreported results, we also consider the multivariate t-distribution and find similar conclusions.

Tables

Table 1: List of portfolio strategies considered in the simulation and empirical analysis

Name	Description
Novel portfolio strategies	
Expected Sharpe (approx)	Mean-variance portfolio maximizing the approximate expected out-of-sample Sharpe ratio estimated with the Ledoit-Wolf covariance matrix ($\tilde{\kappa}^{\sigma}_{sr}$ in (32)) or the sample covariance matrix ($\tilde{\kappa}^{\hat{\sigma}}_{sr}$ in (38)).
Expected Sharpe (exact)	Mean-variance portfolio maximizing the exact expected out-of-sample Sharpe ratio estimated with the Ledoit-Wolf covariance matrix (κ^{σ}_{sr} maximizing (34)) or the sample covariance matrix ($\kappa^{\hat{\sigma}}_{sr}$ in (40)).
Robust Sharpe (approx)	Mean-variance portfolio maximizing the tradeoff between the approximate expectation and variance of the out-of-sample Sharpe ratio estimated with the Ledoit-Wolf covariance matrix ($\tilde{\kappa}^{\sigma}_{\lambda}$ in (51)) or the sample covariance matrix ($\tilde{\kappa}^{\hat{\sigma}}_{\lambda}$ in (49)) using a coefficient $\lambda = 2$.
Robust Sharpe (exact)	Mean-variance portfolio maximizing the tradeoff between the exact expectation and variance of the out-of-sample Sharpe ratio estimated with the Ledoit-Wolf covariance matrix ($\kappa^{\sigma}_{\lambda}$ in (41)) or the sample covariance matrix ($\kappa^{\hat{\sigma}}_{\lambda}$ in (42)) using a coefficient $\lambda = 2$.
Combination Sharpe	Equally weighted combination of Expected Sharpe (approx), Expected Sharpe (exact), Robust Sharpe (approx), and Robust Sharpe (exact).
Benchmark portfolio strategies	
Bootstrap Sharpe	Mean-variance portfolio of Kircher and Rosch (2021) maximizing the expected out-of-sample Sharpe ratio using a bootstrap methodology estimated with the Ledoit-Wolf covariance matrix (κ^{σ}_{boot}) or the sample covariance matrix ($\kappa^{\hat{\sigma}}_{boot}$). See Section EC.1 of the Electronic Companion.
Expected utility	Mean-variance portfolio of Kan, Wang, and Zhou (2021) maximizing the expected out-of-sample utility estimated with the Ledoit-Wolf covariance matrix (κ^{σ}_{γ} in (30)) or the sample covariance matrix ($\kappa^{\hat{\sigma}}_{\gamma}$ in (36)) using a risk-aversion coefficient $\gamma = 5$.
GMV	Global minimum-variance portfolio, w_g in (4) obtained with $\kappa = 0$, estimated with the Ledoit-Wolf or sample covariance matrix.
Tangency	Tangency portfolio, w_{sr} in (27) with $\kappa = 1/\gamma_{sr}$, estimated with the Ledoit-Wolf or sample covariance matrix.
MAXSER	Fully invested version of the MAXSER portfolio of Ao, Li, and Zheng (2019). See Section EC.2 of the Electronic Companion.
$1/N$	Equally weighted portfolio.

Notes. This table describes the 11 fully invested portfolio strategies considered in the simulation and empirical analysis of Section 6. The strategies named Expected Sharpe, Robust Sharpe, Combination Sharpe, Expected utility, and Tangency depend on a combination coefficient κ whose estimation is described in Section EC.3 of the Electronic Companion. The portfolio strategies named Bootstrap Sharpe and MAXSER are described in Sections EC.1 and EC.2 of the Electronic Companion. The Ledoit-Wolf covariance matrix is the nonlinear shrinkage estimator of Ledoit and Wolf (2020).

Table 2: Results with simulated data and Ledoit-Wolf covariance matrix

	Annualized expected OOS Sharpe ratio				Average of $\hat{\gamma}_{sr}\kappa$			
10MOM ($\theta = 0.981$, $\gamma_{sr} = 4.55$)	$T = 60$	$T = 120$	$T = 240$	$T = 480$	$T = 60$	$T = 120$	$T = 240$	$T = 480$
Novel portfolios								
Expected Sharpe (approx)	0.681	0.746	0.815	0.878	0.334	0.404	0.529	0.686
Expected Sharpe (exact)	0.680	0.746	0.816	0.879	0.345	0.417	0.545	0.705
Robust Sharpe (approx)	0.627	0.673	0.740	0.825	0.070	0.113	0.217	0.386
Robust Sharpe (exact)	0.638	0.699	0.782	0.868	0.152	0.232	0.406	0.640
Benchmark portfolios								
Expected Utility ($\gamma = 5$)	0.680	0.749	0.819	0.881	0.458	0.489	0.592	0.732
GMV	0.585	0.600	0.610	0.617	0	0	0	0
Tangency	0.631	0.733	0.825	0.889	1	1	1	1
$1/N$	0.409	0.409	0.409	0.409	/	/	/	/
25SBM ($\theta = 1.376$, $\gamma_{sr} = 6.72$)	$T = 60$	$T = 120$	$T = 240$	$T = 480$	$T = 60$	$T = 120$	$T = 240$	$T = 480$
Novel portfolios								
Expected Sharpe (approx)	0.828	0.954	1.078	1.190	0.286	0.349	0.481	0.653
Expected Sharpe (exact)	0.827	0.954	1.078	1.190	0.290	0.353	0.488	0.662
Robust Sharpe (approx)	0.788	0.896	1.028	1.162	0.115	0.164	0.284	0.462
Robust Sharpe (exact)	0.795	0.916	1.056	1.185	0.169	0.238	0.400	0.616
Benchmark portfolios								
Expected Utility ($\gamma = 5$)	0.815	0.957	1.081	1.187	0.556	0.577	0.726	0.926
GMV	0.698	0.746	0.778	0.796	0	0	0	0
Tangency	0.729	0.904	1.059	1.181	1	1	1	1
$1/N$	0.485	0.485	0.485	0.485	/	/	/	/
25OPIN ($\theta = 1.134$, $\gamma_{sr} = 5.53$)	$T = 60$	$T = 120$	$T = 240$	$T = 480$	$T = 60$	$T = 120$	$T = 240$	$T = 480$
Novel portfolios								
Expected Sharpe (approx)	0.684	0.759	0.843	0.938	0.258	0.294	0.386	0.543
Expected Sharpe (exact)	0.683	0.759	0.843	0.938	0.262	0.298	0.391	0.550
Robust Sharpe (approx)	0.672	0.723	0.799	0.904	0.094	0.119	0.196	0.344
Robust Sharpe (exact)	0.672	0.732	0.817	0.926	0.139	0.174	0.282	0.477
Benchmark portfolios								
Expected Utility ($\gamma = 5$)	0.665	0.757	0.847	0.942	0.463	0.428	0.497	0.652
GMV	0.634	0.650	0.669	0.685	0	0	0	0
Tangency	0.550	0.674	0.805	0.925	1	1	1	1
$1/N$	0.478	0.478	0.478	0.478	/	/	/	/
30IND ($\theta = 0.867$, $\gamma_{sr} = 4.55$)	$T = 60$	$T = 120$	$T = 240$	$T = 480$	$T = 60$	$T = 120$	$T = 240$	$T = 480$
Novel portfolios								
Expected Sharpe (approx)	0.485	0.538	0.591	0.652	0.231	0.239	0.288	0.391
Expected Sharpe (exact)	0.484	0.538	0.591	0.652	0.234	0.241	0.291	0.395
Robust Sharpe (approx)	0.503	0.530	0.565	0.619	0.087	0.090	0.128	0.217
Robust Sharpe (exact)	0.495	0.530	0.572	0.634	0.123	0.127	0.179	0.298
Benchmark portfolios								
Expected Utility ($\gamma = 5$)	0.461	0.531	0.591	0.654	0.403	0.336	0.350	0.436
GMV	0.496	0.499	0.503	0.507	0	0	0	0
Tangency	0.329	0.420	0.523	0.628	1	1	1	1
$1/N$	0.468	0.468	0.468	0.468	/	/	/	/

Notes. This table reports the annualized expected out-of-sample Sharpe ratio, and the average intensity of the GMV portfolio toward the tangency portfolio ($\hat{\gamma}_{sr}\kappa$), for the simulation analysis of Section 6.3. We use four datasets of monthly excess returns described in Section 6.1. We estimate the covariance matrix with the nonlinear shrinkage estimator of Ledoit and Wolf (2020). The eight portfolio strategies are described in Section 6.2 and Table 1. The expected out-of-sample Sharpe ratio is obtained as an average across 10,000 datasets simulated from a multivariate normal distribution, with the population mean and covariance matrix calibrated to the full sample of each dataset. The sample size used to estimate the portfolios is $T = 60$, 120, 240, and 480 months. The parameters θ in (3) and γ_{sr} in (2) reported besides the dataset name are the maximum Sharpe ratio attainable without parameter uncertainty and the risk-aversion coefficient implied by the tangency portfolio, respectively.

Table 3: Results with simulated data and sample covariance matrix

	Annualized expected OOS Sharpe ratio				Average of $\hat{\gamma}_{sr}\kappa$			
10MOM ($\theta = 0.981$, $\gamma_{sr} = 4.55$)	$T = 60$	$T = 120$	$T = 240$	$T = 480$	$T = 60$	$T = 120$	$T = 240$	$T = 480$
Novel portfolios								
Expected Sharpe (approx)	0.654	0.734	0.805	0.875	0.266	0.364	0.499	0.669
Expected Sharpe (exact)	0.650	0.733	0.806	0.876	0.304	0.391	0.524	0.695
Robust Sharpe (approx)	0.645	0.705	0.762	0.834	0.134	0.185	0.271	0.412
Robust Sharpe (exact)	0.647	0.722	0.792	0.869	0.229	0.303	0.439	0.639
Benchmark portfolios								
Expected Utility ($\gamma = 5$)	0.655	0.736	0.808	0.877	0.308	0.405	0.536	0.701
GMV	0.575	0.600	0.611	0.617	0	0	0	0
Tangency	0.570	0.706	0.811	0.885	1	1	1	1
$1/N$	0.409	0.409	0.409	0.409	/	/	/	/
25SBM ($\theta = 1.376$, $\gamma_{sr} = 6.72$)	$T = 60$	$T = 120$	$T = 240$	$T = 480$	$T = 60$	$T = 120$	$T = 240$	$T = 480$
Novel portfolios								
Expected Sharpe (approx)	0.713	0.904	1.055	1.180	0.157	0.269	0.428	0.617
Expected Sharpe (exact)	0.707	0.904	1.055	1.180	0.185	0.285	0.442	0.631
Robust Sharpe (approx)	0.717	0.890	1.030	1.159	0.099	0.175	0.291	0.455
Robust Sharpe (exact)	0.708	0.897	1.047	1.177	0.156	0.241	0.392	0.596
Benchmark portfolios								
Expected Utility ($\gamma = 5$)	0.719	0.909	1.061	1.181	0.177	0.357	0.579	0.834
GMV	0.630	0.732	0.777	0.798	0	0	0	0
Tangency	0.546	0.811	1.019	1.168	1	1	1	1
$1/N$	0.485	0.485	0.485	0.485	/	/	/	/
25OPIN ($\theta = 1.134$, $\gamma_{sr} = 5.53$)	$T = 60$	$T = 120$	$T = 240$	$T = 480$	$T = 60$	$T = 120$	$T = 240$	$T = 480$
Novel portfolios								
Expected Sharpe (approx)	0.583	0.727	0.836	0.935	0.142	0.230	0.343	0.513
Expected Sharpe (exact)	0.575	0.725	0.836	0.935	0.167	0.243	0.353	0.525
Robust Sharpe (approx)	0.592	0.723	0.818	0.913	0.089	0.146	0.222	0.354
Robust Sharpe (exact)	0.580	0.724	0.828	0.929	0.138	0.199	0.297	0.473
Benchmark portfolios								
Expected Utility ($\gamma = 5$)	0.589	0.729	0.840	0.939	0.146	0.265	0.395	0.586
GMV	0.549	0.634	0.673	0.691	0	0	0	0
Tangency	0.397	0.601	0.779	0.915	1	1	1	1
$1/N$	0.478	0.478	0.478	0.478	/	/	/	/
30IND ($\theta = 0.867$, $\gamma_{sr} = 4.55$)	$T = 60$	$T = 120$	$T = 240$	$T = 480$	$T = 60$	$T = 120$	$T = 240$	$T = 480$
Novel portfolios								
Expected Sharpe (approx)	0.371	0.489	0.571	0.643	0.108	0.176	0.248	0.364
Expected Sharpe (exact)	0.362	0.487	0.570	0.643	0.133	0.186	0.254	0.371
Robust Sharpe (approx)	0.384	0.492	0.561	0.623	0.068	0.112	0.156	0.236
Robust Sharpe (exact)	0.368	0.489	0.565	0.634	0.110	0.149	0.202	0.308
Benchmark portfolios								
Expected Utility ($\gamma = 5$)	0.381	0.491	0.572	0.645	0.094	0.185	0.263	0.381
GMV	0.368	0.447	0.481	0.498	0	0	0	0
Tangency	0.211	0.355	0.493	0.615	1	1	1	1
$1/N$	0.468	0.468	0.468	0.468	/	/	/	/

Notes. This table reports the annualized expected out-of-sample Sharpe ratio, and the average intensity of the GMV portfolio toward the tangency portfolio ($\hat{\gamma}_{sr}\kappa$), for the simulation analysis of Section 6.3. We use four datasets of monthly excess returns described in Section 6.1. We estimate the covariance matrix with the sample estimator in (1). The eight portfolio strategies are described in Section 6.2 and Table 1. The expected out-of-sample Sharpe ratio is obtained as an average across 10,000 datasets simulated from a multivariate normal distribution, with the population mean and covariance matrix calibrated to the full sample of each dataset. The sample size used to estimate the portfolios is $T = 60$, 120, 240, and 480 months. The parameters θ in (3) and γ_{sr} in (2) reported besides the dataset name are the maximum Sharpe ratio attainable without parameter uncertainty and the risk-aversion coefficient implied by the tangency portfolio, respectively.

Table 4: Results with empirical data and Ledoit-Wolf covariance matrix

	$T = 120$				$T = 240$			
10MOM	SR	Net SR	Turnover	$\hat{\gamma}_{sr\kappa}$	SR	Net SR	Turnover	$\hat{\gamma}_{sr\kappa}$
Novel portfolios								
Expected Sharpe (approx)	0.936**	0.897**	0.830	0.489	0.928	0.902	0.569	0.580
Expected Sharpe (exact)	0.934**	0.894*	0.869	0.509	0.925	0.898	0.590	0.600
Robust Sharpe (approx)	0.912***	0.878**	0.547	0.220	0.966**	0.942*	0.392	0.316
Robust Sharpe (exact)	0.929**	0.891*	0.808	0.386	0.936	0.910	0.559	0.496
Combination Sharpe	0.936**	0.898**	0.752	0.401	0.942	0.916	0.521	0.498
Benchmark portfolios								
Bootstrap Sharpe	0.930	0.891	0.850	0.513	0.929	0.902	0.580	0.594
Expected Utility ($\gamma = 5$)	0.917	0.879	0.925	0.619	0.898	0.873	0.584	0.641
GMV	0.641	0.620	0.264	0	0.697	0.684	0.159	0
Tangency	0.899	0.853	1.294	1	0.872	0.841	0.804	1
MAXSER	0.865	0.814	1.087	/	0.876	0.841	0.757	/
$1/N$	0.449	0.447	0.042	/	0.494	0.491	0.041	/
25SBM	SR	Net SR	Turnover	$\hat{\gamma}_{sr\kappa}$	SR	Net SR	Turnover	$\hat{\gamma}_{sr\kappa}$
Novel portfolios								
Expected Sharpe (approx)	0.968	0.889	1.111	0.381	1.290**	1.230*	0.791	0.503
Expected Sharpe (exact)	0.965	0.885	1.124	0.385	1.288**	1.229*	0.797	0.508
Robust Sharpe (approx)	0.961**	0.893*	0.813	0.195	1.249**	1.200**	0.594	0.306
Robust Sharpe (exact)	0.961	0.887	0.979	0.276	1.262**	1.206*	0.722	0.426
Combination Sharpe	0.971	0.896	0.992	0.309	1.279**	1.222**	0.721	0.436
Benchmark portfolios								
Bootstrap Sharpe	1.002	0.917	1.129	0.373	1.359	1.297	0.791	0.468
Expected Utility ($\gamma = 5$)	1.009	0.918	2.060	1.076	1.069	1.012	1.385	1.197
GMV	0.831	0.778	0.574	0	1.081	1.043	0.396	0
Tangency	0.806	0.701	2.462	1	1.245	1.171	1.407	1
MAXSER	0.935	0.796	2.381	/	1.332	1.237	1.394	/
$1/N$	0.522	0.520	0.044	/	0.499	0.496	0.043	/
25OPIN	SR	Net SR	Turnover	$\hat{\gamma}_{sr\kappa}$	SR	Net SR	Turnover	$\hat{\gamma}_{sr\kappa}$
Novel portfolios								
Expected Sharpe (approx)	0.767	0.715	0.764	0.290	0.968	0.934	0.463	0.350
Expected Sharpe (exact)	0.765	0.711	0.775	0.293	0.966	0.932	0.469	0.353
Robust Sharpe (approx)	0.755	0.712	0.541	0.112	0.985	0.958	0.336	0.162
Robust Sharpe (exact)	0.745	0.697	0.659	0.165	0.984	0.952	0.408	0.234
Combination Sharpe	0.764	0.715	0.673	0.215	0.981	0.949	0.413	0.275
Benchmark portfolios								
Bootstrap Sharpe	0.748	0.693	0.841	0.355	0.956	0.919	0.526	0.436
Expected Utility ($\gamma = 5$)	0.743	0.680	1.057	0.535	0.917	0.879	0.609	0.556
GMV	0.711	0.678	0.380	0	0.931	0.910	0.236	0
Tangency	0.677	0.604	1.857	1	0.763	0.716	1.010	1
MAXSER	0.819	0.724	1.628	/	0.908	0.853	0.870	/
$1/N$	0.522	0.519	0.041	/	0.560	0.557	0.040	/
30IND	SR	Net SR	Turnover	$\hat{\gamma}_{sr\kappa}$	SR	Net SR	Turnover	$\hat{\gamma}_{sr\kappa}$
Novel portfolios								
Expected Sharpe (approx)	0.512	0.473	0.471	0.160	0.578	0.552	0.271	0.144
Expected Sharpe (exact)	0.513	0.473	0.474	0.160	0.578	0.553	0.272	0.144
Robust Sharpe (approx)	0.543	0.512	0.335	0.031	0.635	0.614	0.214	0.024
Robust Sharpe (exact)	0.542	0.508	0.368	0.044	0.625	0.602	0.228	0.034
Combination Sharpe	0.533	0.497	0.398	0.099	0.608	0.585	0.238	0.087
Benchmark portfolios								
Bootstrap Sharpe	0.466	0.423	0.540	0.211	0.535	0.510	0.281	0.167
Expected Utility ($\gamma = 5$)	0.498	0.454	0.567	0.264	0.554	0.527	0.299	0.188
GMV	0.537	0.508	0.302	0	0.653	0.633	0.199	0
Tangency	0.242	0.181	1.838	1	0.253	0.209	0.991	1
MAXSER	0.278	0.203	1.537	/	0.406	0.359	0.693	/
$1/N$	0.498	0.494	0.048	/	0.533	0.530	0.047	/

Notes. This table reports the annualized gross and net-of-cost out-of-sample Sharpe ratio, the average monthly turnover, and the average intensity of the GMV portfolio toward the tangency portfolio ($\hat{\gamma}_{sr\kappa}$), for the empirical analysis of Section 6.4. We use four datasets of monthly excess returns described in Section 6.1. The 11 portfolio strategies are described in Section 6.2 and Table 1. We estimate the covariance matrix with the nonlinear shrinkage estimator of Ledoit and Wolf (2020). The sample size used to estimate the portfolios is $T = 120$ and 240 months. The net-of-cost Sharpe ratio uses proportional transaction costs of 10 basis points. The stars $\star$, $\star\star$, and $\star\star\star$ in superscript of the SR and Net SR of the five novel strategies mean that the Ledoit and Wolf (2008) two-sided p-value for the difference in performance relative to the GMV portfolio, using 1,000 bootstrap resamples and a block size of five, is smaller than 10%, 5%, and 1%, respectively.

Table 5: Results with empirical data and sample covariance matrix

	$T = 120$				$T = 240$			
10MOM	SR	Net SR	Turnover	$\hat{\gamma}_{sr}\kappa$	SR	Net SR	Turnover	$\hat{\gamma}_{sr}\kappa$
Novel portfolios								
Expected Sharpe (approx)	0.914*	0.872*	0.873	0.440	0.918	0.890	0.583	0.551
Expected Sharpe (exact)	0.909**	0.866*	0.948	0.479	0.911	0.883	0.616	0.581
Robust Sharpe (approx)	0.903***	0.866**	0.592	0.253	0.951*	0.926*	0.413	0.348
Robust Sharpe (exact)	0.907*	0.864*	0.889	0.418	0.923	0.895	0.583	0.526
Combination Sharpe	0.914**	0.872	0.821*	0.397	0.928	0.901	0.547	0.502
Benchmark portfolios								
Bootstrap Sharpe	0.910	0.867	0.930	0.479	0.912	0.884	0.604	0.578
Expected Utility ($\gamma = 5$)	0.899	0.859	0.909	0.515	0.895	0.868	0.578	0.586
GMV	0.639	0.615	0.300	0	0.697	0.683	0.168	0
Tangency	0.864	0.813	1.496	1	0.852	0.819	0.864	1
MAXSER	0.865	0.814	1.087	/	0.876	0.841	0.757	/
$1/N$	0.449	0.447	0.042	/	0.494	0.491	0.041	/
25SBM	SR	Net SR	Turnover	$\hat{\gamma}_{sr}\kappa$	SR	Net SR	Turnover	$\hat{\gamma}_{sr}\kappa$
Novel portfolios								
Expected Sharpe (approx)	0.927	0.828	1.409	0.295	1.269*	1.204*	0.873	0.446
Expected Sharpe (exact)	0.917	0.816	1.468	0.310	1.266*	1.200	0.895	0.459
Robust Sharpe (approx)	0.960*	0.870	1.143	0.205	1.257**	1.199**	0.705	0.315
Robust Sharpe (exact)	0.934	0.837	1.352	0.271	1.257*	1.194*	0.829	0.411
Combination Sharpe	0.937	0.841	1.335	0.270	1.265**	1.202*	0.822	0.408
Benchmark portfolios								
Bootstrap Sharpe	0.933	0.823	1.516	0.311	1.329	1.260	0.891	0.428
Expected Utility ($\gamma = 5$)	0.997	0.891	2.087	0.664	1.088	1.025	1.379	0.955
GMV	0.841	0.767	0.842	0	1.087	1.043	0.468	0
Tangency	0.663	0.534	3.743	1	1.190	1.107	1.697	1
MAXSER	0.935	0.796	2.381	/	1.332	1.237	1.394	/
$1/N$	0.522	0.520	0.044	/	0.499	0.496	0.043	/
25OPIN	SR	Net SR	Turnover	$\hat{\gamma}_{sr}\kappa$	SR	Net SR	Turnover	$\hat{\gamma}_{sr}\kappa$
Novel portfolios								
Expected Sharpe (approx)	0.694	0.630	0.995	0.225	0.939	0.901	0.520	0.311
Expected Sharpe (exact)	0.686	0.621	1.038	0.236	0.936	0.897	0.531	0.318
Robust Sharpe (approx)	0.721	0.663	0.786	0.147	0.959	0.927	0.409	0.202
Robust Sharpe (exact)	0.691	0.629	0.950	0.195	0.950	0.913	0.481	0.265
Combination Sharpe	0.700	0.638	0.937	0.201	0.948	0.912	0.482	0.274
Benchmark portfolios								
Bootstrap Sharpe	0.676	0.608	1.114	0.296	0.917	0.875	0.611	0.407
Expected Utility ($\gamma = 5$)	0.692	0.622	1.107	0.330	0.913	0.871	0.621	0.443
GMV	0.678	0.631	0.578	0	0.915	0.891	0.288	0
Tangency	0.544	0.458	2.819	1	0.702	0.648	1.236	1
MAXSER	0.819	0.724	1.628	/	0.908	0.853	0.870	/
$1/N$	0.522	0.519	0.041	/	0.560	0.557	0.040	/
30IND	SR	Net SR	Turnover	$\hat{\gamma}_{sr}\kappa$	SR	Net SR	Turnover	$\hat{\gamma}_{sr}\kappa$
Novel portfolios								
Expected Sharpe (approx)	0.437	0.383	0.658	0.117	0.560	0.530	0.322	0.125
Expected Sharpe (exact)	0.435	0.380	0.676	0.122	0.559	0.529	0.324	0.127
Robust Sharpe (approx)	0.454	0.407	0.553	0.075	0.591	0.564	0.277	0.078
Robust Sharpe (exact)	0.449	0.399	0.597	0.092	0.582	0.554	0.291	0.092
Combination Sharpe	0.444	0.393	0.616	0.101	0.574	0.545	0.301	0.105
Benchmark portfolios								
Bootstrap Sharpe	0.396	0.337	0.790	0.171	0.519	0.488	0.344	0.156
Expected Utility ($\gamma = 5$)	0.447	0.394	0.659	0.146	0.556	0.526	0.330	0.143
GMV	0.454	0.411	0.488	0	0.618	0.594	0.249	0
Tangency	0.170	0.091	3.082	1	0.234	0.183	1.269	1
MAXSER	0.278	0.203	1.537	/	0.406	0.359	0.693	/
$1/N$	0.498	0.494	0.048	/	0.533	0.530	0.047	/

Notes. This table reports the annualized gross and net-of-cost out-of-sample Sharpe ratio, the average monthly turnover, and the average intensity of the GMV portfolio toward the tangency portfolio ($\hat{\gamma}_{sr}\kappa$), for the empirical analysis of Section 6.4. We use four datasets of monthly excess returns described in Section 6.1. The 11 portfolio strategies are described in Section 6.2 and Table 1. We estimate the covariance matrix with the sample estimator in (1). The sample size used to estimate the portfolios is $T = 120$ and 240 months. The net-of-cost Sharpe ratio uses proportional transaction costs of 10 basis points. The stars $\star$, $\star\star$, and $\star\star\star$ in superscript of the SR and Net SR of the five novel strategies mean that the Ledoit and Wolf (2008) two-sided p-value for the difference in performance relative to the GMV portfolio, using 1,000 bootstrap resamples and a block size of five, is smaller than 10%, 5%, and 1%, respectively.

References

Ao, M., Y. Li, and X. Zheng. 2019. Approaching mean-variance efficiency for large portfolios. *Review of Financial Studies* 32:2890–919.

Barroso, P., and K. Saxena. 2021. Lest we forget: Using out-of-sample forecast errors in portfolio optimization. *Review of Financial Studies* 35:1222–78.

Bodnar, T., Y. Okhrin, and N. Parolya. 2023. Optimal shrinkage-based portfolio selection in high dimensions. *Journal of Business & Economic Statistics* 41:140–56.

Branger, N., K. Lučivjanská, and A. Weissensteiner. 2019. Optimal granularity for portfolio choice. *Journal of Empirical Finance* 50:125–46.

Chakrabarti, D. 2021. Parameter-free robust optimization for the maximum-Sharpe portfolio problem. *European Journal of Operational Research* 293:388–99.

DeMiguel, V., L. Garlappi, F. Nogales, and R. Uppal. 2009. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science* 55:798–812.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22:1915–53.

DeMiguel, V., A. Martín-Utrera, and F. Nogales. 2013. Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking and Finance* 37:3018–34.

———. 2015. Parameter uncertainty in multiperiod portfolio optimization with transaction costs. *Journal of Financial and Quantitative Analysis* 50:1443–71.

Ding, Y., Y. Li, and X. Zheng. 2021. High dimensional minimum variance portfolio estimation under statistical factor models. *Journal of Econometrics* 222:502–15.

Engle, R., R. Ferstenberg, and J. Russell. 2012. Measuring and modeling execution cost and risk. *Journal of Portfolio Management* 38:14–28.

Frahm, G., and C. Memmel. 2010. Dominating estimators for minimum-variance portfolios. *Journal of Econometrics* 159:289–302.

Frazzini, A., R. Israel, and T. J. Moskowitz. 2018. Trading costs. *SSRN working paper*.

Garlappi, L., R. Uppal, and T. Wang. 2007. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies* 20:41–81.

Kan, R., and N. Lassance. 2025. Optimal portfolio choice with fat tails and parameter uncertainty. *Journal of Financial and Quantitative Analysis* 60:3753–90.

Kan, R., and X. Wang. 2023. Optimal portfolio choice with unknown benchmark efficiency. *Management Science* 70:6117–38.

Kan, R., X. Wang, and X. Zheng. 2024. In-sample and out-of-sample Sharpe ratios of multi-factor asset pricing models. *Journal of Financial Economics* 155.

Kan, R., X. Wang, and G. Zhou. 2021. Optimal portfolio choice with estimation risk: No risk-free asset case. *Management Science* 68:2047–68.

Kan, R., and G. Zhou. 2007. Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis* 42:621–56.

Kelly, B., S. Malamud, M. Pourmohammadi, and F. Trojani. 2025. Universal portfolio shrinkage. *SSRN working paper*.

Kelly, B., S. Malamud, and K. Zhou. 2024. The virtue of complexity in return prediction. *Journal of Finance* 79:459–503.

Khashanah, K., M. Simaan, and Y. Simaan. 2022. Do we need higher-order comoments to enhance mean-variance portfolios? Evidence from a simplified jump process. *International Review of Financial Analysis* 81:102068–.

Kirby, C., and B. Ostdiek. 2012. It's all in the timing: Simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis* 47:437–67.

Kircher, F., and D. Rosch. 2021. A shrinkage approach for Sharpe ratio optimal portfolios with estimation risks. *Journal of Banking and Finance* 133.

Lassance, N. 2022. Reconciling mean-variance portfolio theory with non-Gaussian returns. *European Journal of Operational Research* 297:729–40.

Lassance, N., V. DeMiguel, and F. Vrins. 2022. Optimal portfolio diversification via independent component analysis. *Operations Research* 70:55–72.

Lassance, N., R. Kan, and X. Wang. 2025. The distribution of out-of-sample returns of estimated optimal portfolios. *SSRN working paper*.

Lassance, N., A. Martín-Utrera, and M. Simaan. 2024. The risk of expected utility under parameter uncertainty. *Management Science* 70:7644–63.

Lassance, N., R. Vanderveken, and F. Vrins. 2024. On the combination of naive and mean-variance portfolio strategies. *Journal of Business & Economic Statistics* 42:875–89.

Ledoit, O., and M. Wolf. 2004. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis* 88:365–411.

———. 2008. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance* 15:850–9.

———. 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. *Annals of Statistics* 48:3043–65.

Maller, R. A., and D. A. Turkington. 2003. New light on the portfolio allocation problem. *Mathematical Methods of Operations Research* 56:501–11.

Nguyen, V. A., F. Zhang, S. Wang, J. Blanchet, E. Delage, and Y. Ye. 2024. Robustifying conditional portfolio decisions via optimal transport. *Operations Research* 73:2801–29.

Olivares-Nadal, A. V., and V. DeMiguel. 2018. A robust perspective on transaction costs in portfolio optimization. *Operations Research* 66:733–9.

Simaan, Y. 2014. The opportunity cost of mean-variance choice under estimation risk. *European Journal of Operational Research* 234:382–91.

Tu, J., and G. Zhou. 2011. Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics* 99:204–15.

Wu, R., Y. Yang, H. L. Shang, and H. Zhu. 2025. Making distributionally robust portfolios feasible in high dimension. *Journal of Econometrics* 252:106118–.

Yuan, M., and G. Zhou. 2024. Why naïve 1/N diversification is not so naïve, and how to beat it? *Journal of Financial and Quantitative Analysis* 59:3601–32.

Zhao, L., D. Chakrabarti, and K. Muthuraman. 2019. Portfolio construction by mitigating error amplification: The bounded-noise portfolio. *Operations Research* 67:965–83.

Electronic Companion to

Maximizing the Out-of-Sample Sharpe Ratio

Nathan Lassance

In Section EC.1, we describe the portfolio strategy of Kircher and Rosch (2021) that maximizes the expected out-of-sample Sharpe ratio using bootstrap. In Section EC.2, we describe how we adapt the MAXSER portfolio strategy of Ao, Li, and Zheng (2019) to the setting without a risk-free asset. In Section EC.3, we explain how we estimate the parameters in the optimal combination coefficients. In Section EC.4, we test the robustness of our empirical results to imposing lower and upper bounds on the portfolio weights. In Section EC.5, we report the proofs of all theoretical results.

EC.1 Bootstrap portfolio strategy of Kircher and Rosch (2021)

In this section, we explain the bootstrap methodology of Kircher and Rosch (2021) to maximize the expected out-of-sample Sharpe ratio. In our notation and under the sample covariance matrix $\hat{\mathbf{\Sigma}}$ in (1), this strategy consists of bootstrapping with replacement $B = 1{,}000$ times T returns from the sample of observations and finding the optimal combination coefficient κ that maximizes the expected out-of-sample Sharpe ratio as

$$\kappa_{boot}^{\hat{\sigma}} = \arg\max_{\kappa\in\mathbb{R}^+} \frac{1}{B}\sum_{b=1}^{B} \frac{\hat{w}_b(\kappa)'\hat{\mu}_a}{\sqrt{\hat{w}_b(\kappa)'\hat{\mathbf{\Sigma}}\hat{w}_b(\kappa)}}. \tag{EC1}$$

In Equation (EC1), $\hat{w}_b(\kappa)$ is the fully invested sample mean-variance portfolio in (35) for the bth bootstrap sample and the adjusted mean-return estimator $\hat{\mu}_a$ is defined as

$$\hat{\mu}_a = \hat{\mu}_{sh} + \max\left(0, \frac{\hat{\gamma}_a - \iota'\hat{\mathbf{\Sigma}}^{-1}\hat{\mu}_{sh}}{\iota'\hat{\mathbf{\Sigma}}^{-1}\iota}\right)\iota \tag{EC2}$$

where

$$\hat{\gamma}_a = \max\left(\gamma_{\min}, \frac{T-N-2}{T}\iota'\hat{\mathbf{\Sigma}}^{-1}\hat{\mu}\right),\ \hat{\mu}_{sh} = (1-\hat{\alpha}^\star)\hat{\mu} + \hat{\alpha}^\star\bar{\mu}\iota,\ \bar{\mu} = \iota'\hat{\mu}/N,\ \hat{\alpha}^\star = 1 - \sqrt{\widetilde{X}/X},$$
$$X = \frac{T(\hat{\mu}-\bar{\mu}\iota)'(\hat{\mu}-\bar{\mu}\iota)}{\mathrm{tr}(\hat{\mathbf{\Sigma}})/N - \iota'\hat{\mathbf{\Sigma}}\iota/N^2},\quad \widetilde{X} = X - N + \frac{2X^{N/2}e^{-X/2}}{\int_0^X t^{N/2-1}e^{-t/2}\mathrm{d}t}. \tag{EC3}$$

Kircher and Rosch (2021) use $\gamma_{\min} = 3$ in their analysis. However, we use $\gamma_{\min} = 5$ to be consistent with the value used in Equation (EC9) and also because it delivers a better empirical performance. Finally, when relying on the nonlinear shrinkage estimator of the covariance matrix of Ledoit and Wolf (2020), denoted $\hat{\mathbf{\Sigma}}_{nl}$, we obtain the combination coefficient κ_{boot}^{σ} by replacing $\hat{\mathbf{\Sigma}}$ by $\hat{\mathbf{\Sigma}}_{nl}$ in (EC1) (we also make the change in $\hat{w}_b(\kappa)$), but keeping $\hat{\mu}_a$ unchanged because the underlying theory for this estimator builds on the sample covariance matrix.

EC.2 Fully invested version of the MAXSER portfolio strategy

In this section, we explain how we adapt the MAXSER methodology of Ao, Li, and Zheng (2019), which is designed for the mean-variance portfolio that targets a given return volatility and invests in the risk-free asset, to the fully invested mean-variance portfolio maximizing the Sharpe ratio, i.e., we aim to estimate the tangency portfolio w_{sr} in (2).

As in Ao, Li, and Zheng (2019), we start by deriving a least-squares regression representation of w_{sr} in the next proposition.

Proposition EC.1. *The tangency portfolio, w_{sr} in (2), admits the following least-squares regression representation:*

$$w_{sr} = \arg\min_{w\in\mathbb{R}^N} \mathbb{E}[(r_c - w'r_{t+1})^2] \quad \textit{subject to} \quad w'\iota = 1, \tag{EC4}$$

where r_{t+1} is the random vector of asset excess returns and

$$r_c = \frac{1+\theta^2}{\gamma_{sr}}, \tag{EC5}$$

with γ_{sr} and θ defined in (2) and (3), respectively.

Given Proposition EC.1, we follow Ao, Li, and Zheng (2019) and obtain the MAXSER estimator of the tangency portfolio w_{sr} given a sample of monthly returns of size T, $(r_{t-T+1}, \ldots, r_t)$, as

$$\hat{w}_{maxser} = \arg\min_{w\in\mathbb{R}^N} \frac{1}{T}\sum_{m=1}^{T}(\hat{r}_c - w'r_{t-T+m})^2 \ \text{ subject to } \ w'\iota = 1, \ ||w||_1 \le \eta, \tag{EC6}$$

where $\hat{r}_c = (1+\hat{\theta}^2)/\hat{\gamma}_{sr}$, $\hat{\theta}^2$ is the adjusted estimator of θ^2 in Kan and Zhou (2007, Equation (51)), and $\hat{\gamma}_{sr}$ is given by (EC9). We find the 1-norm threshold η so as to maximize the out-of-sample Sharpe ratio, using the same 10-fold cross-validation method as in Ao, Li, and Zheng (2019). Specifically, the optimal η_i in each fold, where $i = 1, \ldots, 10$, is obtained such that the portfolio in (EC6) achieves the maximum Sharpe ratio on the testing set, i.e., on the nine remaining folds. The final η is taken as the average of all 10 η_i's. We consider a grid of 50 values of η in the interval between one (i.e., the minimum 1-norm) and the 1-norm of the portfolio that solves (EC6) without the 1-norm constraint.

EC.3 Estimation of the optimal combination coefficients

We now explain how we estimate the parameters μ_g, σ_g^2, and ψ^2 in (5), as well as γ_{sr} in (2), that are needed to find the optimal combination coefficients κ in the strategies Expected Sharpe (approx) ($\tilde{\kappa}_{sr}^{\sigma}$ and $\tilde{\kappa}_{sr}^{\hat{\sigma}}$), Expected Sharpe (exact) (κ_{sr}^{σ} and $\kappa_{sr}^{\hat{\sigma}}$), Robust Sharpe (approx) ($\tilde{\kappa}_{\lambda}^{\sigma}$ and $\tilde{\kappa}_{\lambda}^{\hat{\sigma}}$), Robust Sharpe (exact) ($\kappa_{\lambda}^{\sigma}$ and $\kappa_{\lambda}^{\hat{\sigma}}$), Expected utility ($\kappa_{\gamma}^{\sigma}$ and $\kappa_{\gamma}^{\hat{\sigma}}$), and Tangency ($\kappa = 1/\gamma_{sr}$).

First, we estimate ψ^2 using the adjusted estimator of Kan and Zhou (2007), i.e.,

$$\hat{\psi}^2 = \frac{(T-N-1)\hat{\psi}_p^2 - (N-1)}{T} + \frac{2(\hat{\psi}_p^2)^{\frac{N-1}{2}}(1+\hat{\psi}_p^2)^{-\frac{T-2}{2}}}{T \times B(\hat{\psi}_p^2/(1+\hat{\psi}_p^2); \frac{N-1}{2}, \frac{T-N+1}{2})} \geq 0, \tag{EC7}$$

where $\hat{\psi}_p^2 = \hat{\mu}'\hat{\boldsymbol{B}}\hat{\mu}$ is the plug-in estimator of ψ^2 and $B(x;a,b) = \int_0^x y^{a-1}(1-y)^{b-1}\mathrm{d}y$ is the incomplete beta function. Second, we estimate σ_g^2 by plugging the nonlinear shrinkage covariance matrix estimator of Ledoit and Wolf (2020), denoted by $\hat{\boldsymbol{\Sigma}}_{nl}$. That is,

$$\hat{\sigma}_g^2 = \frac{1}{\iota'\hat{\boldsymbol{\Sigma}}_{nl}^{-1}\iota}. \tag{EC8}$$

Third, we estimate γ_{sr} by plugging $\hat{\boldsymbol{\Sigma}}_{nl}$ and $\hat{\mu}$ and, following Kircher and Rosch (2021), we set a lower bound $\gamma_{\min}$ that γ_{sr} cannot cross. This gives

$$\hat{\gamma}_{sr} = \max\left(\gamma_{\min}, \iota'\hat{\boldsymbol{\Sigma}}_{nl}^{-1}\hat{\mu}\right). \tag{EC9}$$

Imposing a lower bound $\gamma_{\min}$ is useful because the tangency portfolio is obtained for a weight $1/\gamma_{sr}$ on the zero-cost portfolio in (27). Therefore, as noted by Kirby and Ostdiek (2012) and Barroso, Saxena, and Wang (2025), the tangency portfolio is overly exposed to sample mean returns and performs poorly when γ_{sr} is estimated close to zero, which also negatively impacts portfolios that combine the sample GMV and tangency portfolios. We have tested $\gamma_{\min} = 3$ and $\gamma_{\min} = 5$ and choose the latter because it delivers a better performance in empirical data, both for our novel portfolios and the benchmarks. Finally, we estimate $\mu_g = \sigma_g^2\gamma_{sr}$ as $\hat{\mu}_g = \hat{\sigma}_g^2\hat{\gamma}_{sr}$ using (EC8)–(EC9).

EC.4 Lower and upper bounds on the portfolio weights

In the main body of the paper, we find the optimal combination coefficient κ, which determines the optimal location on the sample mean-variance efficient frontier in terms of expected out-of-sample Sharpe ratio, by considering unconstrained portfolios (except for the full-investment constraint

$w'\iota = 1$). This allows us to study our portfolio problem analytically. Nonetheless, constraints on the portfolio weights would typically be imposed in practice. Given that the optimal κ is unknown in the presence of constraints, we follow an ad-hoc approach where we start from the unconstrained mean-variance portfolio $\hat{w}(\kappa)$ using the optimal κ, and then find the portfolio that is as close as possible to $\hat{w}(\kappa)$ while satisfying the desired constraints. Specifically, we consider lower and upper bounds on the portfolio weights and solve

$$\hat{w}_{cst}(\kappa) = \underset{\substack{w\in\mathbb{R}^N \\ w'\iota=1}}{\operatorname{argmin}} ||w - \hat{w}(\kappa)||_2 \quad \text{subject to} \quad l_i \leq w_i \leq u_i \ \forall i.$$

This ad-hoc approach can be readily implemented for all our proposed portfolio strategies as well as the benchmarks. We set $l_i = 1/N - 0.5$ and $u_i = 1/N + 0.5$ for all i.

The empirical out-of-sample performance of our proposed portfolio strategies and the benchmarks under lower and upper-bound constraints is reported in Tables EC.1 and EC.2 for the Ledoit-Wolf and sample covariance matrix, respectively. We observe that the performance differences across portfolio strategies are not as large as in the case with unconstrained portfolios in Tables 4 and 5, as expected. However, our proposed portfolio strategies remain overall on par or superior relative to the benchmarks. For example, under the Ledoit-Wolf covariance matrix and a sample size $T = 120$, the Combination Sharpe portfolio that equally combines our four novel portfolio strategies delivers an out-of-sample Sharpe ratio of 0.725 on average across the four datasets, versus 0.701, 0.629, 0.686, 0.589, 0.670, and 0.498 for the Bootstrap Sharpe, Expected utility, GMV, Tangency, MAXSER, and $1/N$ benchmarks, respectively.

Table EC.1: Results with empirical data and Ledoit-Wolf covariance matrix under lower and upper-bound constraints

	$T = 120$			$T = 240$		
10MOM	SR	Net SR	Turnover	SR	Net SR	Turnover
Novel portfolios						
Expected Sharpe (approx)	0.731	0.707	0.347	0.691	0.672	0.250
Expected Sharpe (exact)	0.732	0.708	0.348	0.691	0.673	0.251
Robust Sharpe (approx)	0.746	0.721	0.333	0.704	0.688	0.216
Robust Sharpe (exact)	0.729	0.705	0.343	0.690	0.672	0.240
Combination Sharpe	0.730	0.705	0.340	0.691	0.673	0.239
Benchmark portfolios						
Bootstrap Sharpe	0.734	0.710	0.345	0.689	0.671	0.250
Expected Utility ($\gamma = 5$)	0.728	0.703	0.365	0.691	0.673	0.255
GMV	0.650	0.630	0.238	0.718	0.705	0.154
Tangency	0.684	0.656	0.420	0.723	0.703	0.292
MAXSER	0.674	0.637	0.510	0.705	0.679	0.362
$1/N$	0.449	0.447	0.042	0.494	0.491	0.041
25SBM	SR	Net SR	Turnover	SR	Net SR	Turnover
Novel portfolios						
Expected Sharpe (approx)	0.817	0.750	0.832	0.955	0.910	0.604
Expected Sharpe (exact)	0.816	0.749	0.836	0.952	0.907	0.607
Robust Sharpe (approx)	0.890	0.831	0.676	1.034	0.993	0.503
Robust Sharpe (exact)	0.863	0.800	0.753	0.979	0.936	0.568
Combination Sharpe	0.853	0.789	0.767	0.982	0.939	0.568
Benchmark portfolios						
Bootstrap Sharpe	0.808	0.735	0.884	0.966	0.919	0.629
Expected Utility ($\gamma = 5$)	0.552	0.490	1.089	0.700	0.657	0.785
GMV	0.841	0.790	0.539	1.041	1.005	0.378
Tangency	0.637	0.568	1.204	0.715	0.667	0.821
MAXSER	0.829	0.721	1.426	0.952	0.875	1.044
$1/N$	0.522	0.520	0.044	0.499	0.496	0.043
25OPIN	SR	Net SR	Turnover	SR	Net SR	Turnover
Novel portfolios						
Expected Sharpe (approx)	0.786	0.736	0.651	0.951	0.918	0.423
Expected Sharpe (exact)	0.784	0.733	0.657	0.950	0.917	0.427
Robust Sharpe (approx)	0.768	0.727	0.504	0.982	0.955	0.325
Robust Sharpe (exact)	0.767	0.721	0.576	0.975	0.944	0.381
Combination Sharpe	0.784	0.737	0.589	0.970	0.940	0.386
Benchmark portfolios						
Bootstrap Sharpe	0.795	0.742	0.713	0.927	0.891	0.481
Expected Utility ($\gamma = 5$)	0.737	0.677	0.837	0.881	0.843	0.528
GMV	0.715	0.682	0.378	0.928	0.908	0.236
Tangency	0.686	0.618	1.104	0.788	0.742	0.729
MAXSER	0.807	0.717	1.203	0.881	0.825	0.743
$1/N$	0.522	0.519	0.041	0.560	0.557	0.040
30IND	SR	Net SR	Turnover	SR	Net SR	Turnover
Novel portfolios						
Expected Sharpe (approx)	0.514	0.474	0.470	0.578	0.552	0.271
Expected Sharpe (exact)	0.514	0.474	0.473	0.578	0.553	0.272
Robust Sharpe (approx)	0.543	0.511	0.335	0.634	0.613	0.214
Robust Sharpe (exact)	0.542	0.508	0.368	0.624	0.601	0.228
Combination Sharpe	0.533	0.498	0.397	0.608	0.585	0.238
Benchmark portfolios						
Bootstrap Sharpe	0.466	0.423	0.557	0.536	0.510	0.295
Expected Utility ($\gamma = 5$)	0.501	0.457	0.564	0.554	0.527	0.299
GMV	0.538	0.510	0.302	0.652	0.632	0.199
Tangency	0.348	0.293	1.324	0.260	0.216	0.846
MAXSER	0.369	0.297	1.291	0.442	0.393	0.682
$1/N$	0.498	0.494	0.048	0.533	0.530	0.047

Notes. This table reports the annualized gross and net-of-cost out-of-sample Sharpe ratio, and the monthly turnover, for the empirical analysis of Section 6.4. We use four datasets of monthly excess returns described in Section 6.1. The 11 portfolio strategies are described in Section 6.2 and Table 1. We impose lower and upper-bound constraints on the portfolio weights as explained in Section EC.4. We estimate the covariance matrix with the nonlinear shrinkage estimator of Ledoit and Wolf (2020). The sample size used to estimate the portfolios is $T = 120$ and 240 months. The net-of-cost Sharpe ratio uses proportional transaction costs of 10 basis points.

Table EC.2: Results with empirical data and sample covariance matrix under lower and upper-bound constraints

	$T = 120$			$T = 240$		
10MOM	SR	Net SR	Turnover	SR	Net SR	Turnover
Novel portfolios						
Expected Sharpe (approx)	0.731	0.707	0.347	0.691	0.672	0.250
Expected Sharpe (exact)	0.732	0.708	0.348	0.691	0.673	0.251
Robust Sharpe (approx)	0.746	0.721	0.333	0.704	0.688	0.216
Robust Sharpe (exact)	0.729	0.705	0.343	0.690	0.672	0.240
Combination Sharpe	0.730	0.705	0.340	0.691	0.673	0.239
Benchmark portfolios						
Bootstrap Sharpe	0.732	0.708	0.348	0.688	0.670	0.251
Expected Utility ($\gamma = 5$)	0.728	0.703	0.365	0.691	0.673	0.255
GMV	0.650	0.630	0.238	0.718	0.705	0.154
Tangency	0.684	0.656	0.420	0.723	0.703	0.292
MAXSER	0.674	0.637	0.510	0.705	0.679	0.362
$1/N$	0.449	0.447	0.042	0.494	0.491	0.041
25SBM	SR	Net SR	Turnover	SR	Net SR	Turnover
Novel portfolios						
Expected Sharpe (approx)	0.817	0.750	0.832	0.955	0.910	0.604
Expected Sharpe (exact)	0.816	0.749	0.836	0.952	0.907	0.607
Robust Sharpe (approx)	0.890	0.831	0.676	1.034	0.993	0.503
Robust Sharpe (exact)	0.863	0.800	0.753	0.979	0.936	0.568
Combination Sharpe	0.853	0.789	0.767	0.982	0.939	0.568
Benchmark portfolios						
Bootstrap Sharpe	0.813	0.740	0.876	0.963	0.916	0.628
Expected Utility ($\gamma = 5$)	0.552	0.490	1.089	0.700	0.657	0.785
GMV	0.841	0.790	0.539	1.041	1.005	0.378
Tangency	0.637	0.568	1.204	0.715	0.667	0.821
MAXSER	0.829	0.721	1.426	0.952	0.875	1.044
$1/N$	0.522	0.520	0.044	0.499	0.496	0.043
25OPIN	SR	Net SR	Turnover	SR	Net SR	Turnover
Novel portfolios						
Expected Sharpe (approx)	0.786	0.736	0.651	0.951	0.918	0.423
Expected Sharpe (exact)	0.784	0.733	0.657	0.950	0.917	0.427
Robust Sharpe (approx)	0.768	0.727	0.504	0.982	0.955	0.325
Robust Sharpe (exact)	0.767	0.721	0.576	0.975	0.944	0.381
Combination Sharpe	0.784	0.737	0.589	0.970	0.940	0.386
Benchmark portfolios						
Bootstrap Sharpe	0.793	0.739	0.713	0.924	0.888	0.479
Expected Utility ($\gamma = 5$)	0.737	0.677	0.837	0.881	0.843	0.528
GMV	0.715	0.682	0.378	0.928	0.908	0.236
Tangency	0.686	0.618	1.104	0.788	0.742	0.729
MAXSER	0.807	0.717	1.203	0.881	0.825	0.743
$1/N$	0.522	0.519	0.041	0.560	0.557	0.040
30IND	SR	Net SR	Turnover	SR	Net SR	Turnover
Novel portfolios						
Expected Sharpe (approx)	0.514	0.474	0.470	0.578	0.552	0.271
Expected Sharpe (exact)	0.514	0.474	0.473	0.578	0.553	0.272
Robust Sharpe (approx)	0.543	0.511	0.335	0.634	0.613	0.214
Robust Sharpe (exact)	0.542	0.508	0.368	0.624	0.601	0.228
Combination Sharpe	0.533	0.498	0.397	0.608	0.585	0.238
Benchmark portfolios						
Bootstrap Sharpe	0.468	0.424	0.556	0.533	0.506	0.294
Expected Utility ($\gamma = 5$)	0.501	0.457	0.564	0.554	0.527	0.299
GMV	0.538	0.510	0.302	0.652	0.632	0.199
Tangency	0.348	0.293	1.324	0.260	0.216	0.846
MAXSER	0.369	0.297	1.291	0.442	0.393	0.682
$1/N$	0.498	0.494	0.048	0.533	0.530	0.047

Notes. This table reports the annualized gross and net-of-cost out-of-sample Sharpe ratio, and the monthly turnover, for the empirical analysis of Section 6.4. We use four datasets of monthly excess returns described in Section 6.1. The 11 portfolio strategies are described in Section 6.2 and Table 1. We impose lower and upper-bound constraints on the portfolio weights as explained in Section EC.4. We estimate the covariance matrix with the sample estimator in (1). The sample size used to estimate the portfolios is $T = 120$ and 240 months. The net-of-cost Sharpe ratio uses proportional transaction costs of 10 basis points.

EC.5 Proofs of theoretical results

Proof of Proposition 1

The proof is contained in Section 3.1 in the text preceding the proposition.

Proof of Proposition 2

Part 1. The approximate expected out-of-sample Sharpe ratio $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ depends on the vector η and the matrix $\boldsymbol{S}$ in (9)–(10) as $\widetilde{\mathcal{S}}(\hat{w}(\kappa)) = \kappa'\eta/\sqrt{\kappa'\boldsymbol{S}\kappa}$. For the case of the three-fund rule in (14) under i.i.d. multivariate normality, η and $\boldsymbol{S}$ are given by (15), which gives the formula for $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ in (18) provided $\kappa_1 > 0$. Setting the derivative of $\widetilde{\mathcal{S}}(\hat{w}(\kappa))$ with respect to ω equal to zero yields the approximate optimal $\tilde{\omega}_{sr}$ in (19).

Part 2. We start by deriving a stochastic representation under i.i.d. multivariate normality for the out-of-sample Sharpe ratio of the three-fund rule $\hat{w}(\kappa)$ in (14), i.e., $\hat{w}(\kappa)'\mu/\sqrt{\hat{w}(\kappa)'\boldsymbol{\Sigma}\hat{w}(\kappa)}$. For that purpose, we define

$$\mu_p = \kappa_1\theta^2 + \kappa_2\mu_g/\sigma_g^2, \tag{EC10}$$

$$\sigma_p^2 = \kappa_1^2\theta^2 + 2\kappa_1\kappa_2\mu_g/\sigma_g^2 + \kappa_2^2/\sigma_g^2, \tag{EC11}$$

$$\lambda = \frac{\sigma_p^2}{\kappa_1^2}\left(1 - \frac{\theta_p^2}{\theta^2}\right) = \frac{\omega^2\psi^2}{\sigma_g^2\theta^2}, \tag{EC12}$$

where $\theta_p = \mu_p/\sigma_p$ and $\omega = \kappa_2/\kappa_1$, $\kappa_1 > 0$. Let $\boldsymbol{P}$ be an $N \times N$ orthonormal matrix whose first two columns are

$$\nu = \frac{\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mu + \omega\iota)}{[(\mu + \omega\iota)'\boldsymbol{\Sigma}^{-1}(\mu + \omega\iota)]^{\frac{1}{2}}}, \tag{EC13}$$

$$\varphi = \frac{(\boldsymbol{I}_N - \nu\nu')\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu}{[\mu'\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{I}_N - \nu\nu')\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu]^{\frac{1}{2}}}. \tag{EC14}$$

Let

$$z = \sqrt{T}\boldsymbol{P}'\boldsymbol{\Sigma}^{-\frac{1}{2}}(\hat{\mu} + \omega\iota) \sim \mathcal{N}(\mu_z, \boldsymbol{I}_N), \tag{EC15}$$

$$\boldsymbol{W} = T\boldsymbol{P}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\hat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P} \sim \mathcal{W}_N(T-1, \boldsymbol{I}_N), \tag{EC16}$$

and they are independent of each other. We can show that

$$\mu_z = \sqrt{T}\boldsymbol{P}'\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mu+\omega\iota) = \sqrt{T}[(\mu+\omega\iota)'\boldsymbol{\Sigma}^{-1}(\mu+\omega\iota)]^{\frac{1}{2}}e_1 = \sqrt{T}\sigma_p e_1/\kappa_1, \tag{EC17}$$

where the $N \times 1$ vector e_i is a vector of zeros expect for the ith entry that is equal to one. Using $\boldsymbol{W}$ and z, we can write $\hat{w}(\kappa)$ as

$$\hat{w}(\kappa) = \kappa_1\sqrt{T}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\boldsymbol{W}^{-1}z. \tag{EC18}$$

The out-of-sample mean and variance of $\hat{w}(\kappa)$ are

$$\tilde{\mu}_p = \hat{w}(\kappa)'\mu = \kappa_1\sqrt{T}z'\boldsymbol{W}^{-1}\boldsymbol{P}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu = \kappa_1\sqrt{T}z'\boldsymbol{W}^{-1}(\nu'\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu e_1 + \varphi'\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu e_2), \tag{EC19}$$

$$\tilde{\sigma}_p^2 = \hat{w}(\kappa)'\boldsymbol{\Sigma}\hat{w}(\kappa) = \kappa_1^2 T z'\boldsymbol{W}^{-2}z. \tag{EC20}$$

Letting

$$\delta = \nu'\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu e_1 + \varphi'\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu e_2, \tag{EC21}$$

we can write

$$\tilde{\mu}_p = \kappa_1\sqrt{T}\delta'\boldsymbol{W}^{-1}z. \tag{EC22}$$

Note that

$$\nu'\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu = \theta_p, \tag{EC23}$$

$$\varphi'\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu = \sqrt{\theta^2 - \theta_p^2}, \tag{EC24}$$

which allows us to write

$$\delta = \theta_p e_1 + \sqrt{\theta^2 - \theta_p^2}e_2 \tag{EC25}$$

and

$$\tilde{\mu}_p = \kappa_1\sqrt{T}\Big(\theta_p e_1'\boldsymbol{W}^{-1}z + \sqrt{\theta^2 - \theta_p^2}e_2'\boldsymbol{W}^{-1}z\Big). \tag{EC26}$$

Next, we define an $N \times N$ orthonormal matrix $\boldsymbol{Q} = [\tilde{z},\ \boldsymbol{Q}_1]$ whose first column is $\tilde{z} \equiv z/(z'z)^{\frac{1}{2}}$. Let

$$\boldsymbol{A} = (\boldsymbol{Q}'\boldsymbol{W}^{-1}\boldsymbol{Q})^{-1} = \begin{bmatrix} \tilde{z}'\boldsymbol{W}^{-1}\tilde{z} & \tilde{z}'\boldsymbol{W}^{-1}\boldsymbol{Q}_1 \\ \boldsymbol{Q}_1'\boldsymbol{W}^{-1}\tilde{z} & \boldsymbol{Q}_1'\boldsymbol{W}^{-1}\boldsymbol{Q}_1 \end{bmatrix}^{-1} \equiv \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \sim \mathcal{W}_N(T-1, \boldsymbol{I}_N), \tag{EC27}$$

where $\boldsymbol{A}_{11}$ is the $(1,1)$ element of $\boldsymbol{A}$. Muirhead (1982, Theorem 3.2.10) implies that

$$u_1 = \boldsymbol{A}_{11\cdot 2} \equiv \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} \sim \chi^2_{T-N}, \tag{EC28}$$

and it is independent of $\boldsymbol{A}_{12}$ and $\boldsymbol{A}_{22}$. In addition, using the result of Dickey (1967), we have

$$-\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} \sim \frac{x}{\sqrt{u_2}}, \tag{EC29}$$

where $x \sim \mathcal{N}(0_{N-1}, \boldsymbol{I}_{N-1})$, $u_2 \sim \chi^2_{T-N+1}$, and they are independent of each other and u_1. Given that the distribution of $\boldsymbol{A}$ is independent of z, we have that x, u_1, and u_2 are also independent of z. Using the formula for the inverse of a partitioned matrix, we can verify that

$$\tilde{z}'\boldsymbol{W}^{-1}\tilde{z} = \boldsymbol{A}_{11\cdot 2}^{-1} = \frac{1}{u_1}, \tag{EC30}$$

$$\boldsymbol{Q}_1'\boldsymbol{W}^{-1}\tilde{z} = -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11\cdot 2}^{-1} = \frac{x}{u_1\sqrt{u_2}}. \tag{EC31}$$

With these identities, we can write

$$z'\boldsymbol{W}^{-2}z = z'\boldsymbol{W}^{-1}(\tilde{z}\tilde{z}' + \boldsymbol{Q}_1\boldsymbol{Q}_1')\boldsymbol{W}^{-1}z = \frac{z'z}{u_1^2}\left(1 + \frac{x'x}{u_2}\right). \tag{EC32}$$

Without loss of generality, let the first column of $\boldsymbol{Q}_1$ be

$$\xi = \frac{(\boldsymbol{I}_N - \tilde{z}\tilde{z}')\delta}{[\delta'(\boldsymbol{I}_N - \tilde{z}\tilde{z}')\delta]^{\frac{1}{2}}}. \tag{EC33}$$

From (EC31), we know that

$$\frac{x_1}{u_1\sqrt{u_2}} = \xi'\boldsymbol{W}^{-1}\tilde{z} = \frac{\delta'\boldsymbol{W}^{-1}\tilde{z} - \frac{\delta'\tilde{z}}{u_1}}{\sqrt{\delta'\delta - \frac{(\delta'z)^2}{z'z}}} = \frac{\tilde{\delta}'\boldsymbol{W}^{-1}z - \frac{\tilde{\delta}'z}{u_1}}{\sqrt{z'z - (\tilde{\delta}'z)^2}}, \tag{EC34}$$

where $\tilde{\delta} = \delta/(\delta'\delta)^{\frac{1}{2}}$ and x_1 is the first element of x. Let

$$u = z'z - (\tilde{\delta}'z)^2 \sim \chi^2_{N-1}(T\sigma_p^2/\kappa_1^2 - (\tilde{\delta}'\mu_z)^2), \tag{EC35}$$

and it is independent of $y \equiv \tilde{\delta}'z \sim \mathcal{N}(\tilde{\delta}'\mu_z, 1)$. Note that

$$\tilde{\delta}'\mu_z = \frac{\sqrt{T}\mu_p}{\kappa_1\theta}, \tag{EC36}$$

and we can write

$$u \sim \chi^2_{N-1}(\lambda T). \tag{EC37}$$

where λ is defined in (EC12). It follows that

$$\tilde{\delta}'W^{-1}z = \frac{y}{u_1} + \frac{x_1\sqrt{u}}{u_1\sqrt{u_2}} = \frac{1}{u_1}\left(y + \frac{x_1\sqrt{u}}{\sqrt{u_2}}\right), \tag{EC38}$$

and therefore,

$$\tilde{\mu}_p = \kappa_1\theta\sqrt{T}\tilde{\delta}'\boldsymbol{W}^{-1}z = \kappa_1\sqrt{T}\theta\frac{1}{u_1}\left(y + \frac{x_1\sqrt{u}}{\sqrt{u_2}}\right). \tag{EC39}$$

Letting $u_3 = x'x - x_1^2 \sim \chi^2_{N-2}$, we can write

$$\tilde{\sigma}_p^2 = \kappa_1^2 T\frac{z'z}{u_1^2}\left(1 + \frac{x'x}{u_2}\right) = \kappa_1^2 T\left(\frac{u+y^2}{u_1^2}\right)\left(1 + \frac{x_1^2+u_3}{u_2}\right). \tag{EC40}$$

Using the stochastic representation for the out-of-sample mean, $\tilde{\mu}_p = \hat{w}(\kappa)'\mu$ in (EC39), and the out-of-sample variance, $\tilde{\sigma}_p^2 = \hat{w}(\kappa)'\boldsymbol{\Sigma}\hat{w}(\kappa)$ in (EC40), we obtain the following stochastic representation for the out-of-sample Sharpe ratio:

$$\tilde{\theta}_p = \frac{\tilde{\mu}_p}{\tilde{\sigma}_p} = \theta\frac{y\sqrt{u_2} + x_1\sqrt{u}}{\sqrt{(y^2+u)(x_1^2+u_2+u_3)}}. \tag{EC41}$$

To simplify this representation, let

$$q = \frac{y}{\sqrt{y^2+u}}, \tag{EC42}$$

$$r = \frac{x_1}{\sqrt{x_1^2+u_2+u_3}} \sim n_1/(n'n)^{\frac{1}{2}} \text{ with } n \sim \mathcal{N}(0_{N-1}, \boldsymbol{I}_{N-1}), \tag{EC43}$$

$$b = \frac{u_2}{u_2 + x_1^2 + u_3} \sim \text{Beta}\Big(\frac{T-N+1}{2}, \frac{N-1}{2}\Big), \tag{EC44}$$

which are independent of each other. Then, the stochastic representation for $\tilde{\theta}_p$ becomes

$$\tilde{\theta}_p = \theta\Big(\sqrt{b}q + \sqrt{(1-b)(1-q^2)}r\Big). \tag{EC45}$$

Given that $\mathbb{E}[r] = 0$, the expected out-of-sample Sharpe ratio is $\mathbb{E}[\tilde{\theta}_p] = \theta\mathbb{E}[\sqrt{b}]\mathbb{E}[q]$. Using Kan, Wang, and Zheng (2024, Equation (A.26)), we have

$$\mathbb{E}[\sqrt{b}] = \frac{\text{B}(\frac{T-N+2}{2}, \frac{N-1}{2})}{\text{B}(\frac{T-N+1}{2}, \frac{N-1}{2})} = \frac{\Gamma(\frac{T}{2})\Gamma(\frac{T-N+2}{2})}{\Gamma(\frac{T+1}{2})\Gamma(\frac{T-N+1}{2})}. \tag{EC46}$$

Moreover, using Kan, Wang, and Zheng (2024, Equation (A.30)), we have

$$\mathbb{E}[q] = \frac{\sqrt{T}\mu_p}{\kappa_1\theta}\mathbb{E}\left[\frac{1}{\sqrt{\chi^2_{N+2}(T\sigma_p^2/\kappa_1^2)}}\right], \tag{EC47}$$

where we use the identity $T\lambda + (\frac{\sqrt{T}\mu_p}{\kappa_1\theta})^2 = T\sigma_p^2/\kappa_1^2$. Then, using Kan, Wang, and Zheng (2024, Equation (A.25)),

$$\mathbb{E}\left[\frac{1}{\sqrt{\chi^2_{N+2}(T\sigma_p^2/\kappa_1^2)}}\right] = \frac{\Gamma(\frac{N+1}{2})}{\sqrt{2}\Gamma(\frac{N+2}{2})}{}_1F_1\Big(\frac{1}{2}; \frac{N+2}{2}; -\frac{T\sigma_p^2}{2\kappa_1^2}\Big). \tag{EC48}$$

Finally, using $\mathbb{E}[\tilde{\theta}_p] = \theta\mathbb{E}[\sqrt{b}]\mathbb{E}[q]$ with Equations (EC46)–(EC48) yields the exact expected out-of-sample Sharpe ratio of the three-fund rule in (20)–(21). This completes the proof.

Proof of Proposition 3

The proof of this proposition is contained in Section 4.1.

Proof of Proposition 4

Part 1. Under the i.i.d. assumption, the sample mean $\hat{\mu}$ has mean μ and covariance matrix $\boldsymbol{\Sigma}/T$. Therefore, we have

$$\mathbb{E}[\hat{w}^\sigma(\kappa)'\mu] = \mu_g + \kappa\mu'\boldsymbol{B}\mu = \mu_g + \kappa\psi^2, \tag{EC49}$$

$$\mathbb{E}[\hat{w}^\sigma(\kappa)'\boldsymbol{\Sigma}\hat{w}^\sigma(\kappa)] = \sigma_g^2 + \kappa^2\mathbb{E}[\hat{\mu}'\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}\hat{\mu}] = \sigma_g^2 + \frac{\kappa^2(T\psi^2 + N - 1)}{T}, \tag{EC50}$$

where we use the properties $\mu'\boldsymbol{B}\mu = \psi^2$, $w_g'\boldsymbol{\Sigma}w_z = 0$, $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B} = \boldsymbol{B}$, $\mathbb{E}[\hat{\mu}'\boldsymbol{B}\hat{\mu}] = \mu'\boldsymbol{B}\mu + \text{tr}(\mathbb{V}[\hat{\mu}]\boldsymbol{B})$, and $\text{tr}(\boldsymbol{\Sigma}\boldsymbol{B}) = N-1$. The approximate expected out-of-sample Sharpe ratio in (31) is then obtained by plugging (EC49) and (EC50) into (7). Taking the derivative of (31) with respect to κ and setting it to zero yields the optimum $\kappa = \tilde{\kappa}_{sr}^{\sigma}$ in (32) given our assumption that $\gamma_{sr} = \iota'\boldsymbol{\Sigma}^{-1}\mu > 0$. Plugging (32) into (31) yields the optimal approximate expected out-of-sample Sharpe ratio in (33).

Part 2. We start by deriving a stochastic representation under i.i.d. multivariate normality for the out-of-sample Sharpe ratio of the fully invested portfolio combination $\hat{w}^{\sigma}(\kappa)$ in (28), i.e., $\hat{w}^{\sigma}(\kappa)'\mu/\sqrt{\hat{w}^{\sigma}(\kappa)'\boldsymbol{\Sigma}\hat{w}^{\sigma}(\kappa)}$. Let $\boldsymbol{P}$ be an $N \times N$ orthonormal matrix whose first two columns are

$$\nu = \sigma_g\boldsymbol{\Sigma}^{-1}\iota, \tag{EC51}$$

$$\varphi = \frac{\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mu - \iota\mu_g)}{\psi}. \tag{EC52}$$

Let

$$z = \sqrt{T}\boldsymbol{P}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\hat{\mu} \sim \mathcal{N}(\mu_z, \boldsymbol{I}_N), \tag{EC53}$$

where $\mu_z = \sqrt{T}(\theta_g e_1 + \psi e_2)$ and the $N \times 1$ vector e_i is a vector of zeros expect for the ith entry that is equal to one. Using z, we can write $\hat{w}^{\sigma}(\kappa) = w_g + \kappa\hat{w}_z^{\sigma}$ as

$$\hat{w}^{\sigma}(\kappa) = \sigma_g\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}e_1 + \frac{\kappa}{\sqrt{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}(\boldsymbol{I}_N - e_1e_1')z. \tag{EC54}$$

Let

$$\tilde{w} = \boldsymbol{P}'\boldsymbol{\Sigma}^{\frac{1}{2}}\hat{w}^{\sigma}(\kappa) = \sigma_g e_1 + \frac{\kappa}{\sqrt{T}}(\boldsymbol{I}_N - e_1e_1')z, \tag{EC55}$$

which satisfies $\tilde{w}_1 = \sigma_g$ and $(\tilde{w}_2, \ldots, \tilde{w}_N)' = \kappa z/\sqrt{T}$. The out-of-sample mean and variance of $\hat{w}(\kappa)$ are then distributed as

$$\tilde{\mu}_p = \hat{w}^{\sigma}(\kappa)'\mu = \hat{w}(\kappa)'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\boldsymbol{P}'\boldsymbol{\Sigma}^{\frac{1}{2}}\mu = \frac{\tilde{w}'\mu_z}{\sqrt{T}} = \mu_g + \frac{\kappa\psi}{\sqrt{T}}z_2, \tag{EC56}$$

$$\tilde{\sigma}_p^2 = \hat{w}^{\sigma}(\kappa)'\boldsymbol{\Sigma}\hat{w}^{\sigma}(\kappa) = \hat{w}(\kappa)'\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{P}\boldsymbol{P}'\boldsymbol{\Sigma}^{\frac{1}{2}}\hat{w}(\kappa) = \tilde{w}'\tilde{w} = \sigma_g^2 + \frac{\kappa^2}{T}(z_2^2 + u), \tag{EC57}$$

where $u = z_3^2 + \cdots + z_N^2 \sim \chi_{N-2}^2$ and it is independent of $z_2 \sim \mathcal{N}(\sqrt{T}\psi, 1)$. It follows that the

out-of-sample Sharpe ratio of $\hat{w}(\kappa)$ is distributed as

$$\tilde{\theta}_p = \frac{\tilde{\mu}_p}{\tilde{\sigma}_p} = \frac{\mu_g + \frac{\kappa\psi}{\sqrt{T}} z_2}{\sqrt{\sigma_g^2 + \frac{\kappa^2}{T}(z_2^2 + u)}}. \tag{EC58}$$

To find the expectation of $\tilde{\theta}_p$ in (EC58), we apply Meng (2005, Lemma 1). Specifically, given $M_{\tilde{\mu}_p,\tilde{\sigma}_p^2}(s,t)$ the joint moment generating function of $\tilde{\mu}_p$ and $\tilde{\sigma}_p^2$, we have

$$\mathcal{S}(\hat{w}^{\sigma}(\kappa)) = \mathbb{E}[\tilde{\theta}_p] = \frac{1}{\sqrt{\pi}} \int_0^{\infty} \left. \frac{\partial M_{\tilde{\mu}_p,\tilde{\sigma}_p^2}(s,-t)}{\partial s} \right|_{s=0} t^{-1/2} \mathrm{d}t. \tag{EC59}$$

We can show that

$$M_{\tilde{\mu}_p,\tilde{\sigma}_p^2}(s,t) = \int_0^{\infty} \int_{-\infty}^{\infty} \exp\left\{ \left(\mu_g + \frac{\kappa\psi}{\sqrt{T}} z \right) s + \left(\sigma_g^2 + \frac{\kappa^2}{T}(z^2 + u) \right) t \right\} f_{z_2}(z) f_u(u) \mathrm{d}z \mathrm{d}u, \tag{EC60}$$

where $f_{z_2}(z)$ and $f_u(u)$ are the pdfs of $z_2 \sim \mathcal{N}(\sqrt{T}\psi, 1)$ and $u \sim \chi^2_{N-2}$, evaluates to

$$M_{\tilde{\mu}_p,\tilde{\sigma}_p^2}(s,t) = \left(1 - \frac{2\kappa^2 t}{T} \right) \exp\left\{ \mu_g s + \sigma_g^2 t - \frac{(\kappa^2 s^2 + 2\kappa s T + 2\kappa^2 t T)\psi^2}{2(2\kappa^2 t - T)} \right\}. \tag{EC61}$$

Therefore,

$$\left. \frac{\partial M_{\tilde{\mu}_p,\tilde{\sigma}_p^2}(s,-t)}{\partial s} \right|_{s=0} = \left(1 + \frac{2\kappa^2 t}{T} \right)^{-\frac{N+1}{2}} \left[\left(1 + \frac{2\kappa^2 t}{T} \right) \mu_g + \kappa\psi^2 \right] \exp\left\{ -t \left(\sigma_g^2 + \frac{\kappa^2 T \psi^2}{T + 2\kappa^2 t} \right) \right\}. \tag{EC62}$$

Finally, plugging (EC62) into (EC59) gives the exact expected out-of-sample Sharpe ratio in (34). This completes the proof.

Proof of Proposition 5

Part 1. From Kan, Wang, and Zhou (2021, Lemma 1), the expected out-of-sample mean and variance of the fully invested portfolio combination $\hat{w}^{\hat{\sigma}}(\kappa)$ in (35) are given by

$$\mathbb{E}[\hat{w}^{\hat{\sigma}}(\kappa)'\mu] = \mu_g + \frac{\kappa T \psi^2}{T - N - 1}, \tag{EC63}$$

$$\mathbb{E}[\hat{w}^{\hat{\sigma}}(\kappa)'\mathbf{\Sigma}\hat{w}^{\hat{\sigma}}(\kappa)] = \frac{T-2}{T-N-1}\sigma_g^2 + \frac{\kappa^2 T(T-2)(T\psi^2 + N - 1)}{(T-N)(T-N-1)(T-N-3)}. \tag{EC64}$$

The approximate expected out-of-sample Sharpe ratio in (37) is then obtained by plugging (EC63)–(EC64) into (7). Taking the derivative of (37) with respect to κ and setting it to zero yields the optimum $\kappa = \tilde{\kappa}_{sr}^{\hat{\sigma}}$ in (38) given our assumption that $\gamma_{sr} = \iota'\mathbf{\Sigma}^{-1}\mu > 0$. It is smaller than the optimal κ under the sample covariance matrix, $\tilde{\kappa}_{sr}^{\sigma}$ in (32), because $\frac{(T-N)(T-N-3)}{T(T-N-1)} < 1$. Plugging (38) into (37) yields the optimal approximate expected out-of-sample Sharpe ratio in (39). To see that it is smaller than that under the sample covariance matrix in (33), we write (33) as

$$\widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\tilde{\kappa}_{sr}^{\sigma})) = \sqrt{\frac{\psi^2\theta^2 + \frac{N-1}{T}\theta_g^2}{\psi^2 + \frac{N-1}{T}}} = \sqrt{\theta_g^2 + \left(\frac{\psi^2}{\psi^2 + \frac{N-1}{T}}\right)\psi^2}. \tag{EC65}$$

Comparing (39) and (EC65), we have $\widetilde{\mathcal{S}}(\hat{w}^{\hat{\sigma}}(\tilde{\kappa}_{sr}^{\hat{\sigma}})) \leq \widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\tilde{\kappa}_{sr}^{\sigma}))$ because both $\frac{T-N-1}{T-2}$ and $\frac{(T-N)(T-N-3)}{(T-2)(T-N-1)}$ are strictly smaller than one.

Part 2. We start by deriving a stochastic representation for the out-of-sample Sharpe ratio of the fully invested portfolio combination $\hat{w}^{\hat{\sigma}}(\kappa)$ in (35) under i.i.d. normality, i.e., $\hat{w}^{\hat{\sigma}}(\kappa)'\mu/\sqrt{\hat{w}^{\hat{\sigma}}(\kappa)'\mathbf{\Sigma}\hat{w}^{\hat{\sigma}}(\kappa)}$. To achieve this, we use the stochastic representation in Lassance, Kan, and Wang (2025, Theorem 1) for the out-of-sample mean and variance of any portfolio of the form $\hat{c}\hat{\mathbf{\Sigma}}^{-1}\hat{\mu} + \hat{d}\hat{\mathbf{\Sigma}}^{-1}\iota$, where $\hat{c}$ and $\hat{d}$ are functions of $(\hat{\mu}_g, \hat{\sigma}_g^2, \hat{\psi}^2)$, i.e., the sample plug-in estimators of $(\mu_g, \sigma_g^2, \psi^2)$. By choosing $\hat{c} = \kappa$ and $\hat{d} = \hat{\sigma}_g^2 - \kappa\hat{\mu}_g$, we obtain the portfolio $\hat{w}^{\hat{\sigma}}(\kappa)$ and the following stochastic representation for its out-of-sample mean and variance. Let $z_1 \sim \mathcal{N}(\sqrt{T}\psi, 1)$, $z_2 \sim \mathcal{N}(0,1)$, $z_3 \sim \mathcal{N}(0,1)$, $u_1 \sim \chi^2_{N-3}$, $u_2 \sim \chi^2_{T-N+1}$, $b \sim \text{Beta}(\frac{T-N+2}{2}, \frac{N-2}{2})$, and they are independent of each other. Let $(z_3, u_1, b) = (0, 0, 1)$ when $N = 2$, and $u_1 = 0$ when $N = 3$. Then, we have

$$\tilde{\mu}_p = \hat{w}^{\hat{\sigma}}(\kappa)'\mu = \mu_g + \frac{\psi}{\hat{\psi}\sqrt{u_2}}\left(z_1 s + z_2\sqrt{\sigma_g^2 + s^2}\sqrt{(1-b)/b}\right), \tag{EC66}$$

$$\tilde{\sigma}_p^2 = \hat{w}^{\hat{\sigma}}(\kappa)'\mathbf{\Sigma}\hat{w}^{\hat{\sigma}}(\kappa) = \frac{\sigma_g^2 + s^2}{b}, \tag{EC67}$$

where

$$\hat{\psi}^2 = \frac{z_1^2 + z_2^2 + u_1}{u_2}, \tag{EC68}$$

$$s = \frac{1}{\sqrt{u_2}}\left(\sigma_g z_3 + \kappa\sqrt{T}\hat{\psi}\right). \tag{EC69}$$

It follows that the out-of-sample Sharpe ratio of $\hat{w}^{\hat{\sigma}}(\kappa)$ is distributed as

$$\tilde{\theta}_p = \frac{\tilde{\mu}_p}{\tilde{\sigma}_p} = X\sqrt{b} + \frac{\psi z_2\sqrt{1-b}}{\sqrt{z_1^2+z_2^2+u_1}}, \tag{EC70}$$

where

$$X = \frac{\mu_g + \frac{\psi z_1 s}{\sqrt{z_1^2+z_2^2+u_1}}}{\sqrt{\sigma_g^2+s^2}}. \tag{EC71}$$

Given that the expectation of the second term in (EC70) is zero by symmetry, the expected out-of-sample Sharpe ratio is $\mathbb{E}[\tilde{\theta}_p] = \mathbb{E}[\sqrt{b}]\mathbb{E}[X]$. Similarly to (EC46), we have

$$\mathbb{E}[\sqrt{b}] = \frac{\Gamma(\frac{T}{2})\Gamma(\frac{T-N+3}{2})}{\Gamma(\frac{T+1}{2})\Gamma(\frac{T-N+2}{2})}. \tag{EC72}$$

Moreover, we can express $\mathbb{E}[X]$ as a double integral. For that purpose, let $u = z_2^2 + u_1 \sim \chi^2_{N-2}$ and $v = u_2 \sim \chi^2_{T-N+1}$, which are independent of each other and of z_1 and z_2. Then, we can write

$$X = \frac{\mu_g + \frac{\psi z_1 s}{\sqrt{z_1^2+u}}}{\sqrt{\sigma_g^2+s^2}}, \quad s = \frac{\sigma_g z_2}{\sqrt{v}} + \frac{\kappa\sqrt{T}\sqrt{z_1^2+u}}{v}. \tag{EC73}$$

Letting $\tilde{\kappa} = \sqrt{T}\kappa$, we first rewrite X as

$$X = \frac{U}{\sqrt{V}}, \tag{EC74}$$

where

$$U = \left(\theta_g\sqrt{v} + \frac{\tilde{\kappa}\psi z_1}{\sigma_g\sqrt{v}}\right) + \frac{\psi z_1 z_2}{\sqrt{z_1^2+u}}, \tag{EC75}$$

$$V = v + \left(z_2 + \frac{\tilde{\kappa}\sqrt{z_1^2+u}}{\sigma_g\sqrt{v}}\right)^2. \tag{EC76}$$

Let $M_{U,V}(s,t)$ be the joint moment generating function of U and V. We use Meng (2005, Lemma 1) to obtain

$$\mathbb{E}[X|z_1,u,v] = \mathbb{E}\left[\frac{U}{\sqrt{V}}\middle| z_1,u,v\right] = \frac{1}{\sqrt{\pi}}\int_0^\infty \frac{\partial M_{U,V}(s,-t)}{\partial s}\bigg|_{s=0} t^{-\frac{1}{2}}\mathrm{d}t. \tag{EC77}$$

Let

$$\tilde{a} = \theta_g \sqrt{v}, \tag{EC78}$$

$$\tilde{b} = \frac{\tilde{\kappa}\psi z_1}{\sigma_g \sqrt{v}}, \tag{EC79}$$

$$\tilde{c} = \frac{\tilde{\kappa}^2 (z_1^2 + u)}{\sigma_g^2 v}. \tag{EC80}$$

We can then show that

$$\mathbb{E}[X|z_1, u, v] = \frac{1}{\sqrt{\pi}} \int_0^\infty \exp\left(-\frac{\tilde{c}t}{1+2t} - vt\right) t^{-\frac{1}{2}} \left(\frac{\tilde{a}}{(1+2t)^{\frac{1}{2}}} + \frac{\tilde{b}}{(1+2t)^{\frac{3}{2}}}\right) \mathrm{d}t. \tag{EC81}$$

We then take the expectation of this expression with respect to u and z_1 to obtain $\mathbb{E}[X|v]$. We can show that for $\alpha > 0$, we have

$$\mathbb{E}[\exp(-\alpha u)] = (1+2\alpha)^{-\frac{N-2}{2}}. \tag{EC82}$$

Setting $\alpha = \tilde{\kappa}^2 t / [\sigma_g^2 v (1+2t)]$, we obtain

$$\mathbb{E}[X|z_1, v] = \frac{1}{\sqrt{\pi}} \int_0^\infty \exp(-\alpha z_1^2 - vt) \frac{t^{-\frac{1}{2}}}{(1+2\alpha)^{\frac{N-2}{2}}} \left(\frac{\tilde{a}}{(1+2t)^{\frac{1}{2}}} + \frac{\tilde{b}}{(1+2t)^{\frac{3}{2}}}\right) \mathrm{d}t. \tag{EC83}$$

In addition, we can show that

$$\mathbb{E}[\exp(-\alpha z_1^2)] = \frac{\exp\left(-\frac{\alpha T \psi^2}{1+2\alpha}\right)}{(1+2\alpha)^{\frac{1}{2}}}, \tag{EC84}$$

$$\mathbb{E}[\exp(-\alpha z_1^2) z_1] = \frac{\sqrt{T}\psi \exp\left(-\frac{\alpha T \psi^2}{1+2\alpha}\right)}{(1+2\alpha)^{\frac{3}{2}}}. \tag{EC85}$$

Using these identities, we obtain

$$\begin{aligned} \mathbb{E}[X|v] = \frac{1}{\sqrt{\pi}} \int_0^\infty \exp\left(-\frac{\alpha T \psi^2}{1+2\alpha} - vt\right) \frac{(tv)^{-\frac{1}{2}}}{(1+2\alpha)^{\frac{N-1}{2}}} \\ \times \left(\frac{\theta_g v}{(1+2t)^{\frac{1}{2}}} + \frac{\tilde{\kappa}\sqrt{T}\psi^2}{\sigma_g (1+2\alpha)(1+2t)^{\frac{3}{2}}}\right) \mathrm{d}t. \end{aligned} \tag{EC86}$$

The unconditional expectation $\mathbb{E}[X]$ is then obtained by integrating out the density of v,

$$\mathbb{E}[X] = \int_0^\infty \mathbb{E}[X|v] \mathrm{d}v. \tag{EC87}$$

Finally, plugging $\mathbb{E}[\sqrt{b}]$ in (EC72) and $\mathbb{E}[X]$ in (EC86)–(EC87) into $\mathbb{E}[\tilde{\theta}_p] = \mathbb{E}[\sqrt{b}]\mathbb{E}[X]$ gives the expression for the exact expected out-of-sample Sharpe ratio in (40). This completes the proof.

Proof of Proposition 6

Part 1. Following the proof of part 1 of Proposition 4, and applying Meng (2005, Lemma 1), the second moment of the out-of-sample Sharpe ratio of $\hat{w}^{\sigma}(\kappa)$ in (28), $\tilde{\theta}_p$ distributed as (EC58), is

$$\mathbb{E}[\tilde{\theta}_p^2] = \int_0^{\infty} \left. \frac{\partial^2 M_{\tilde{\mu}_p, \tilde{\sigma}_p^2}(s, -t)}{\partial s^2} \right|_{s=0} \mathrm{d}t. \tag{EC88}$$

Given the joint moment generating function $M_{\tilde{\mu}_p, \tilde{\sigma}_p^2}(s,t)$ in (EC61), we can show that

$$\begin{aligned} \left. \frac{\partial^2 M_{\tilde{\mu}_p, \tilde{\sigma}_p^2}(s, -t)}{\partial s^2} \right|_{s=0} = & \left(1 + \frac{2\kappa^2 t}{T}\right)^{-\frac{N+3}{2}} \left\{ \left[\left(1 + \frac{2\kappa^2 t}{T}\right)\mu_g + \kappa\psi^2 \right]^2 + \frac{\kappa^2\psi^2(T + 2\kappa^2 t)}{T^2} \right\} \\ & \times \exp\left\{ -t\left(\sigma_g^2 + \frac{\kappa^2\psi^2 T}{T + 2\kappa^2 t}\right) \right\}. \end{aligned} \tag{EC89}$$

Plugging (EC89) into (EC88) gives the second moment of the out-of-sample Sharpe ratio in (44).

Part 2. Following the proof of part 2 of Proposition 5, the second moment of the out-of-sample Sharpe ratio of $\hat{w}^{\hat{\sigma}}(\kappa)$ in (35), $\tilde{\theta}_p$ distributed as (EC70), is

$$\mathbb{E}[\tilde{\theta}_p^2] = \mathbb{E}[bX^2] + 2\psi\mathbb{E}\left[\frac{z_2 X\sqrt{b(1-b)}}{\sqrt{z_1^2 + z_2^2 + u_1}}\right] + \psi^2\mathbb{E}\left[\frac{z_2^2(1-b)}{z_1^2 + z_2^2 + u_1}\right], \tag{EC90}$$

where $z_1 \sim \mathcal{N}(\sqrt{T}\psi, 1)$, $z_2 \sim \mathcal{N}(0,1)$, $u_1 \sim \chi^2_{N-3}$, $b \sim \text{Beta}(\frac{T-N+2}{2}, \frac{N-2}{2})$, and they are independent of each other. Moreover, X is distributed as (EC71). By symmetry, the second expectation in (EC90) is zero. Moreover, because b and X are independent, the first expectation in (EC90) is

$$\mathbb{E}[bX^2] = \mathbb{E}[b]\mathbb{E}[X^2] = \frac{T - N + 2}{T}\mathbb{E}[X^2]. \tag{EC91}$$

Concerning the third expectation in (EC90), we have

$$\mathbb{E}\left[\frac{z_2^2(1-b)}{z_1^2 + z_2^2 + u_1}\right] = \mathbb{E}[1-b]\mathbb{E}\left[\frac{z_2^2}{z_1^2 + z_2^2 + u_1}\right] = \frac{N-2}{T}\mathbb{E}\left[\frac{z'\boldsymbol{A}z}{z'z}\right], \tag{EC92}$$

where $\boldsymbol{A} = \text{diag}([1, 0'_{N-2}])$, $z = (z_1, z_2, \ldots z_{N-1})'$, $z_1 \sim \mathcal{N}(\sqrt{T}\psi, 1)$, $z_k \sim \mathcal{N}(0,1)$ for $k \geq 2$, and all entries of z are independent of each other. Using Hillier, Kan, and Wang (2014, Theorem 4), we

have

$$\mathbb{E}\left[\frac{z'\boldsymbol{A}z}{z'z}\right] = \frac{1}{N-1}{}_1F_1\left(1; \frac{N+1}{2}; -\frac{T\psi^2}{2}\right). \tag{EC93}$$

Putting everything together, $\mathbb{E}[\tilde{\theta}_p^2]$ in (EC90) becomes

$$\mathbb{E}[\tilde{\theta}_p^2] = \frac{T-N+2}{T}\mathbb{E}[X^2] + \frac{(N-2)\psi^2}{(N-1)T}{}_1F_1\left(1; \frac{N+1}{2}; -\frac{T\psi^2}{2}\right). \tag{EC94}$$

We now express $\mathbb{E}[X^2]$ as a double integral. Writing $X = U/\sqrt{V}$ with U and V defined in (EC75)–(EC76), we use Meng (2005, Lemma 1) to obtain

$$\mathbb{E}[X^2|z_1,u,v] = \mathbb{E}\left[\frac{U^2}{V}\middle| z_1,u,v\right] = \int_0^\infty \frac{\partial^2 M_{U,V}(s,-t)}{\partial s^2}\bigg|_{s=0} \mathrm{d}t, \tag{EC95}$$

where $M_{U,V}(s,t)$ is the joint moment generating function of U and V. Define $\tilde{a}$, $\tilde{b}$, and $\tilde{c}$ as in (EC78)–(EC80), as well as

$$\tilde{d} = \frac{2\tilde{\kappa}\psi\theta_g z_1}{\sigma_g}, \tag{EC96}$$

$$\tilde{e} = \frac{\tilde{\kappa}^2(z_1^2+u)}{\sigma_g^2 v}, \tag{EC97}$$

where $\tilde{\kappa} = \sqrt{T}\kappa$. We can show that

$$\mathbb{E}[X^2|z_1,u,v] = \int_0^\infty \exp\left(-\frac{\tilde{e}t}{1+2t} - vt\right)\left(\frac{\tilde{a}^2}{(1+2t)^{\frac{1}{2}}} + \frac{\tilde{c}+\tilde{d}}{(1+2t)^{\frac{3}{2}}} + \frac{\tilde{b}^2}{(1+2t)^{\frac{5}{2}}}\right)\mathrm{d}t. \tag{EC98}$$

We then take the expectation of this expression with respect to u and z_1 to obtain $\mathbb{E}[X^2|v]$. To do so, we need the two expectations in (EC84)–(EC85). In addition, given $\alpha > 0$, we need

$$\mathbb{E}[\exp(-\alpha z_1^2)z_1^2] = \exp\left(-\frac{\alpha T\psi^2}{1+2\alpha}\right)\left(\frac{1}{(1+2\alpha)^{\frac{3}{2}}} + \frac{T\psi^2}{(1+2\alpha)^{\frac{5}{2}}}\right) \tag{EC99}$$

as well as an expression for

$$\mathbb{E}\left[\frac{\exp(-\alpha(z_1^2+u))z_1^2}{z_1^2+u}\right]. \tag{EC100}$$

For the latter, using the density of $z_1 \sim \mathcal{N}(\sqrt{T}\psi, 1)$ and $u \sim \chi^2_{N-2}$, we have

$$\mathbb{E}\left[\frac{\exp(-\alpha(z_1^2+u))z_1^2}{z_1^2+u}\right] = \int_0^\infty \int_{-\infty}^\infty \frac{z_1^2}{z_1^2+u} e^{-\alpha(z_1^2+u)} \frac{e^{-\frac{1}{2}(z_1-\sqrt{T}\psi)^2}}{\sqrt{2\pi}} \frac{\left(\frac{u}{2}\right)^{\frac{N-4}{2}} e^{-\frac{u}{2}}}{2\Gamma\left(\frac{N-2}{2}\right)} \mathrm{d}z_1 \mathrm{d}u. \tag{EC101}$$

Making a change of variables $\tilde{z}_2 = \sqrt{1+2\alpha} z_1$ and $\tilde{u} = (1+2\alpha)u$, we can then write

$$\begin{aligned}\mathbb{E}\left[\frac{\exp(-\alpha(z_1^2+u))z_1^2}{z_1^2+u}\right] &= \frac{\exp\left(-\frac{T\psi^2\alpha}{1+2\alpha}\right)}{(1+2\alpha)^{\frac{N-1}{2}}} \int_0^\infty \int_{-\infty}^\infty \frac{\tilde{z}_2^2}{\tilde{z}_2^2+\tilde{u}} \frac{e^{-\frac{1}{2}\left(\tilde{z}_2-\frac{\sqrt{T}\psi}{\sqrt{1+2\alpha}}\right)^2}}{\sqrt{2\pi}} \frac{\left(\frac{\tilde{u}}{2}\right)^{\frac{N-4}{2}} e^{-\frac{\tilde{u}}{2}}}{2\Gamma\left(\frac{N-2}{2}\right)} \mathrm{d}\tilde{z}_2 \mathrm{d}\tilde{u} \\ &= \frac{\exp\left(-\frac{T\psi^2\alpha}{1+2\alpha}\right)}{(1+2\alpha)^{\frac{N-1}{2}}} \mathbb{E}\left[\frac{\tilde{z}_2^2}{\tilde{z}_2^2+\tilde{u}}\right],\end{aligned} \tag{EC102}$$

where $\tilde{z}_2 \sim N(\sqrt{T}\psi/\sqrt{1+2\alpha}, 1)$ and $\tilde{u} \sim \chi^2_{N-2}$, and they are independent of each other. Using Hillier, Kan, and Wang (2014, Theorem 4), we can show that

$$\mathbb{E}\left[\frac{\tilde{z}_2^2}{\tilde{z}_2^2+\tilde{u}}\right] = 1 - \mathbb{E}\left[\frac{\tilde{u}}{\tilde{z}_2^2+\tilde{u}}\right] = 1 - \frac{N-2}{N-1} {}_1F_1\left(1; \frac{N+1}{2}; -\frac{T\psi^2}{2(1+\alpha)}\right). \tag{EC103}$$

Therefore, it holds that

$$\mathbb{E}\left[\frac{\exp(-\alpha(z_1^2+u))z_1^2}{z_1^2+u}\right] = \frac{\exp\left(-\frac{\psi^2 T\alpha}{1+2\alpha}\right)}{(1+2\alpha)^{\frac{N-1}{2}}} \left[1 - \frac{N-2}{N-1} {}_1F_1\left(1; \frac{N+1}{2}; -\frac{\psi^2 T}{2(1+2\alpha)}\right)\right]. \tag{EC104}$$

Combining (EC84), (EC85), (EC99), and (EC104), we obtain

$$\begin{aligned}\mathbb{E}[X^2|v] = \int_0^\infty (1+2\alpha)^{-\frac{N-1}{2}} \exp\left(-\frac{T\psi^2\alpha}{1+2\alpha} - vt\right) \\ \times \left\{\frac{\theta_g^2 v}{(1+2t)^{\frac{1}{2}}} + \frac{2\kappa T\theta_g\psi^2}{\sigma_g(1+2\alpha)(1+2t)^{\frac{3}{2}}} + \frac{\kappa^2 T\psi^2(1+2\alpha+T\psi^2)}{\sigma_g^2 v(1+2\alpha)^2(1+2t)^{\frac{5}{2}}}\right. \\ \left. + \frac{\psi^2}{(1+2t)^{\frac{3}{2}}}\left[1 - \frac{N-2}{N-1} {}_1F_1\left(1; \frac{N+1}{2}; -\frac{T\psi^2}{2(1+2\alpha)}\right)\right]\right\} f_v(v)\mathrm{d}t,\end{aligned} \tag{EC105}$$

where $\alpha = \kappa^2 Tt/[\sigma_g^2 v(1+2t)]$. The unconditional expectation $\mathbb{E}[X^2]$ is then obtained by integrating out the density of v,

$$\mathbb{E}[X^2] = \int_0^\infty \mathbb{E}[X^2|v]\mathrm{d}v. \tag{EC106}$$

Finally, plugging $\mathbb{E}[X^2]$ in (EC105)–(EC106) into $\mathbb{E}[\tilde{\theta}_p^2]$ in (EC94) gives the expression for the

second moment of the out-of-sample Sharpe ratio in (45). This completes the proof.

Proof of Proposition 7

Part 1. Under the i.i.d. assumption, the sample mean $\hat{\mu}$ has expectation μ and covariance matrix $\boldsymbol{\Sigma}/T$. Using the property $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B} = \boldsymbol{B}$, this means that the covariance matrix of $\hat{w}^{\sigma}(\kappa)$ in (28) is

$$\mathbb{V}[\hat{w}^{\sigma}(\kappa)] = \frac{\kappa^2}{T}\boldsymbol{B}. \tag{EC107}$$

Therefore, the numerator of $\widetilde{\mathcal{V}}(\hat{w}^{\sigma}(\kappa))$ is $\mathbb{V}[\hat{w}^{\sigma}(\kappa)'\mu] = \mu'\mathbb{V}[\hat{w}^{\sigma}(\kappa)]\mu = \kappa^2\psi^2/T$, where we use $\mu'\boldsymbol{B}\mu = \psi^2$. Moreover, the denominator of $\widetilde{\mathcal{V}}(\hat{w}^{\sigma}(\kappa))$ is given by (EC50), which yields (50). Then, using Equations (31) and (50), the approximate robust $\tilde{\kappa}^{\sigma}_{\lambda}$ in (48) is found as

$$\tilde{\kappa}^{\sigma}_{\lambda} = \arg\max_{\kappa\in\mathbb{R}} \frac{\mu_g + \kappa\psi^2 - \lambda|\kappa|\psi/\sqrt{T}}{\sqrt{\sigma_g^2 + \kappa^2(\psi^2 + \frac{N-1}{T})}}. \tag{EC108}$$

Observe that the optimal κ is necessarily positive. This is because $\widetilde{\mathcal{V}}(\hat{w}^{\sigma}(\kappa))$ in (50) is an even function of κ but $\widetilde{\mathcal{S}}(\hat{w}^{\sigma}(\kappa))$ in (31) is larger for $\kappa = 0$ than for any $\kappa < 0$. Therefore, we can get rid of the absolute value in (EC108). Then, taking the derivative of (EC108) with respect to κ and setting it to zero yields the solution

$$\kappa = \tilde{\kappa}^{\sigma}_{sr}\left(1 - \frac{\lambda}{\sqrt{T}\psi}\right), \tag{EC109}$$

where $\tilde{\kappa}^{\sigma}_{sr}$ is given by (32). If $\lambda \leq \sqrt{T}\psi$, the optimal κ in (EC109) is positive and optimal. Otherwise, the optimal $\kappa = 0$ as discussed above. Therefore, we obtain the solution for $\tilde{\kappa}^{\sigma}_{\lambda}$ in (51).

Part 2. Using (47), the numerator of $\widetilde{\mathcal{V}}(\hat{w}^{\hat{\sigma}}(\kappa))$ is given by $\mathbb{V}[\hat{w}^{\hat{\sigma}}(\kappa)'\mu]$. As shown by Lassance, Martín-Utrera, and Simaan (2024, Equation (IA.44)), this variance is equal to

$$\mathbb{V}[\hat{w}^{\hat{\sigma}}(\kappa)'\mu] = \frac{1}{T-N-1}\left(\sigma_g^2\psi^2 + \kappa^2\psi^2\frac{T(T-2)(T-N-1) + 2T^2(T-N)\psi^2}{(T-N)(T-N-1)(T-N-3)}\right). \tag{EC110}$$

Moreover, the denominator of $\widetilde{\mathcal{V}}(\hat{w}^{\hat{\sigma}}(\kappa))$ is given by (EC64). Then, to prove the properties satisfied by the robust $\tilde{\kappa}^{\hat{\sigma}}_{\lambda}$ in (49), we first show that the approximate variance of the out-of-sample Sharpe ratio, $\widetilde{\mathcal{V}}(\hat{w}^{\hat{\sigma}}(\kappa))$ in (52), is minimized at $\kappa = 0$. This is the case because, after taking the derivative

of $\widetilde{\mathcal{V}}(\hat{w}^{\hat{\sigma}}(\kappa))$ with respect to κ and setting it equal to zero, we obtain the first-order condition

$$\kappa(T\psi^2 + T - N - 1) = 0. \tag{EC111}$$

This condition holds if and only if $\kappa = 0$, and thus, $\widetilde{\mathcal{V}}(\hat{w}^{\hat{\sigma}}(\kappa))$ is minimized at $\kappa = 0$. This result means that, as $\lambda \to \infty$, $\tilde{\kappa}_\lambda^{\hat{\sigma}}$ in (49) is equal to zero. Moreover, when $\lambda = 0$, $\tilde{\kappa}_\lambda^{\hat{\sigma}} = \tilde{\kappa}_{sr}^{\hat{\sigma}}$ in (38) by definition. Finally, when $\lambda \in (0, \infty)$, we necessarily have $0 \leq \tilde{\kappa}_\lambda^{\hat{\sigma}} \leq \tilde{\kappa}_{sr}^{\hat{\sigma}}$. This is because $\widetilde{\mathcal{V}}(\hat{w}^{\hat{\sigma}}(\kappa))$ is increasing in κ for $\kappa \geq 0$, and $\widetilde{\mathcal{S}}(\hat{w}^{\hat{\sigma}}(\kappa))$ is increasing in κ for $\kappa \in [0, \tilde{\kappa}_{sr}^{\hat{\sigma}}]$ and decreasing in κ for $\kappa \geq \tilde{\kappa}_{sr}^{\hat{\sigma}}$, so that any $\kappa > \tilde{\kappa}_{sr}^{\hat{\sigma}}$ necessarily delivers a smaller value of the objective function in (49) than choosing $\kappa \leq \tilde{\kappa}_{sr}^{\hat{\sigma}}$. Finally, the inequalities are strict, $0 < \tilde{\kappa}_\lambda^{\hat{\sigma}} < \tilde{\kappa}_{sr}^{\hat{\sigma}}$, if $\psi^2 > 0$. Indeed, the derivative of the objective function evaluated at $\kappa = 0$ is given by

$$\frac{\partial}{\partial \kappa}\left[\widetilde{\mathcal{S}}(\hat{w}^{\hat{\sigma}}(\kappa)) - \lambda\sqrt{\widetilde{\mathcal{V}}(\hat{w}^{\hat{\sigma}}(\kappa))}\right]\Bigg|_{\kappa=0} = \frac{T\psi^2}{\sqrt{(T-2)(T-N-1)\sigma_g^2}}, \tag{EC112}$$

which is strictly positive if $\psi^2 > 0$. This completes the proof.

Proof of Proposition EC.1

For a given r_c, consider the portfolio

$$\begin{aligned} w_{r_c} &= \arg\min_{w \in \mathbb{R}^N} \mathbb{E}[(r_c - w' r_{t+1})^2] \;\text{ subject to }\; w'\iota = 1, \\ &= \arg\min_{w \in \mathbb{R}^N} (r_c - w'\mu)^2 + w'\mathbf{\Sigma} w \;\text{ subject to }\; w'\iota = 1. \end{aligned} \tag{EC113}$$

Lassance and Vrins (2023, Equation (IA49)) show that the solution to w_{r_c} is

$$w_{r_c} = \frac{(1+\theta^2)\sigma_g^2 - r_c\mu_g}{(1+\theta^2)\sigma_g^2 - \mu_g^2} w_g + \frac{\mu_g(\mu_g - r_c)}{\mu_g^2 - (1+\theta^2)\sigma_g^2} w_{sr}. \tag{EC114}$$

Therefore, we find that $w_{r_c} = w_{sr}$ when $r_c = (1+\theta^2)\sigma_g^2/\mu_g = (1+\theta^2)/\gamma_{sr}$. This completes the proof.

References in Electronic Companion

Ao, M., Y. Li, and X. Zheng. 2019. Approaching mean-variance efficiency for large portfolios. *Review of Financial Studies* 32:2890–919.

Barroso, P., K. Saxena, and H. Wang. 2025. Calm your portfolio: The importance of disciplining intelligent but fickle forecasts in portfolio optimization. *SSRN working paper*.

Dickey, J. 1967. Matricvariate generalizations of the multivariate t distribution and the inverted multivariate t distribution. *Annals of Mathematical Statistics* 38:511–8.

Hillier, G., R. Kan, and X. Wang. 2014. Generating functions and short recursions, with applications to the moments of quadratic forms in noncentral normal vectors. *Econometric Theory* 30:436–73.

Kan, R., X. Wang, and X. Zheng. 2024. In-sample and out-of-sample Sharpe ratios of multi-factor asset pricing models. *Journal of Financial Economics* 155.

Kan, R., X. Wang, and G. Zhou. 2021. Optimal portfolio choice with estimation risk: No risk-free asset case. *Management Science* 68:2047–68.

Kan, R., and G. Zhou. 2007. Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis* 42:621–56.

Kirby, C., and B. Ostdiek. 2012. It's all in the timing: Simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis* 47:437–67.

Kircher, F., and D. Rosch. 2021. A shrinkage approach for Sharpe ratio optimal portfolios with estimation risks. *Journal of Banking and Finance* 133.

Lassance, N., R. Kan, and X. Wang. 2025. The distribution of out-of-sample returns of estimated optimal portfolios. *SSRN working paper*.

Lassance, N., A. Martín-Utrera, and M. Simaan. 2024. The risk of expected utility under parameter uncertainty. *Management Science* 70:7644–63.

Lassance, N., and F. Vrins. 2023. Portfolio selection: A target-distribution approach. *European Journal of Operational Research* 310:302–14.

Ledoit, O., and M. Wolf. 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. *Annals of Statistics* 48:3043–65.

Meng, X.-L. 2005. From unit root to Stein's estimator to fisher's K statistics: If you have a moment, I can tell you more. *Statistical Science* 20:141–62.

Muirhead, R. 1982. *Aspects of Multivariate Statistical Theory.* New Jersey: John Wiley & Sons.