THE RISK OF EXPECTED UTILITY UNDER PARAMETER UNCERTAINTY

Nathan Lassance, Alberto Martín-Utrera, Majeed Simaan

REPRINT • 2023 / 11

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.html

The Risk of Expected Utility under Parameter Uncertainty*

Nathan Lassance Alberto Martín-Utrera Majeed Simaan

August 18, 2023

Abstract

We derive analytical expressions for the risk of an investor's expected utility under parameter uncertainty. In particular, our analysis focuses on characterizing the out-of-sample utility variance of three portfolios: the classic mean-variance portfolio, the minimum-variance portfolio, and a shrinkage portfolio that combines both. We then use our analytical expressions to study a robustness measure that balances out-of-sample utility mean and volatility. We show that neither the sample mean-variance portfolio nor the sample minimum-variance portfolio exhibits maximal robustness individually, and one needs to combine both to optimize portfolio robustness. Accordingly, we introduce a robust shrinkage portfolio that delivers an optimal tradeoff between out-of-sample utility mean and volatility and is more resilient to estimation errors. Our results highlight the importance of considering out-of-sample performance risk in designing and evaluating investment strategies and stochastic discount factor models.

Keywords: parameter uncertainty, mean-variance portfolio, shrinkage.
JEL Classification: G11, G12

*Lassance, LFIN/LIDAM, UCLouvain, corresponding author, e-mail: nathan.lassance@uclouvain.be; Martín-Utrera, Iowa State University, e-mail: amutrera@iastate.edu; Simaan, Stevens Institute of Technology, e-mail: msimaan@stevens.edu. Previous versions of this manuscript were circulated under the titles *"A Robust Approach to Optimal Portfolio Choice with Parameter Uncertainty"* and *"The Risk of Out-of-Sample Portfolio Performance"*. We thank Kay Giesecke (the Editor), an anonymous Associate Editor, as well as three anonymous referees for their valuable feedback. We also thank Vitali Alekseev, Stephen Brown, German Creamer, Victor DeMiguel, Kristoffer Glover, Raymond Kan, Ralph Koijen, Petter Kolm, Allen Li, Yan Liu, Yueliang Lu, Jean Pauphilet, Harry Scheule, Frédéric Vrins, Xiaolu Wang, Alex Weissensteiner, Guofu Zhou, as well as seminar and conference participants at London Business School, NYU Courant, Stevens Institute of Technology, UCLouvain, University of Technology Sydney, 15th International Conference on Computational and Financial Econometrics, Financial Management Association Annual Meeting, INFORMS Annual Meeting, Northern Finance Association Annual Meeting, and Silicon Prairie Finance Conference for their comments. Lassance gratefully acknowledges financial support by the Fonds de la Recherche Scientifique [grant number J.0115.22].

I think that when we know that we actually do live in uncertainty, then we ought to admit it.

—Richard P. Feynman, *The Pleasure of Finding Things Out*

1 Introduction

Financial institutions utilize quantitative strategies to create investment funds that give investors access to optimally diversified portfolios. The fund's performance is typically assessed based on standard metrics such as the fund's average return, beta, or Sharpe ratio. The construction of these well-diversified portfolios and their performance measures often relies on a single realization of historical return data. However, this approach presents at least two challenges for optimizing and evaluating the performance of investment strategies. First, constructing optimal portfolios solely relying on historical data and ignoring sampling volatility delivers an abysmal *out-of-sample* performance (DeMiguel, Garlappi, and Uppal, 2009; Tu and Zhou, 2011). Second, out-of-sample portfolio performance varies with each different draw of the data-generating process, and therefore investment strategies should not be evaluated based on a single observation of realized returns (Lo, 2002; Harvey and Liu, 2014).

This study addresses these issues by characterizing and optimizing the out-of-sample performance risk of quantitative strategies, where we measure out-of-sample performance as the investor's expected utility under parameter uncertainty. In particular, we present analytical expressions for the out-of-sample utility (OOSU) variance of three types of sample portfolios: the mean-variance portfolio (SMV), the global-minimum-variance portfolio (SGMV), and any linear combination of the two. Our closed-form expression for the out-of-sample utility variance of these sample portfolios provides a comprehensive understanding of the performance uncertainties faced by investors who utilize quantitative strategies in the presence of estimation risk.[1] Importantly, our work extends

[1]For example, the OOSU 95% interval of the SMV portfolio calibrated from a dataset of 25 portfolios of stocks sorted on size and book-to-market with 120 monthly return observations is $[-12.7\%, 0.66\%]$. This large variation

the pioneering research of Kan and Zhou (2007), which focused on the out-of-sample utility mean of sample portfolios.

Our manuscript makes four contributions to the existing literature on parameter uncertainty and portfolio selection. First, we analytically characterize the OOSU variance of the SMV portfolio, the SGMV portfolio, and any convex combination of these two portfolios. Using our analytical results, we document that the SMV portfolio's OOSU volatility is substantially larger than that of the SGMV portfolio. For instance, calibrating the model with a dataset of 25 portfolios of stocks sorted on size and book to market, we show that the OOSU standard deviation of the SMV portfolio is 29 times larger than that of the SGMV portfolio when both portfolios are estimated using 120 monthly observations. In this particular setting, an unrealistically larger sample size of over 13,000 monthly observations —more than 1,000 years of data— is required for the SMV portfolio to deliver an out-of-sample performance as stable as that of the SGMV portfolio. Interestingly, even though the SGMV portfolio has a much lower OOSU risk, combining the SMV and SGMV portfolios is still optimal for minimizing OOSU risk.

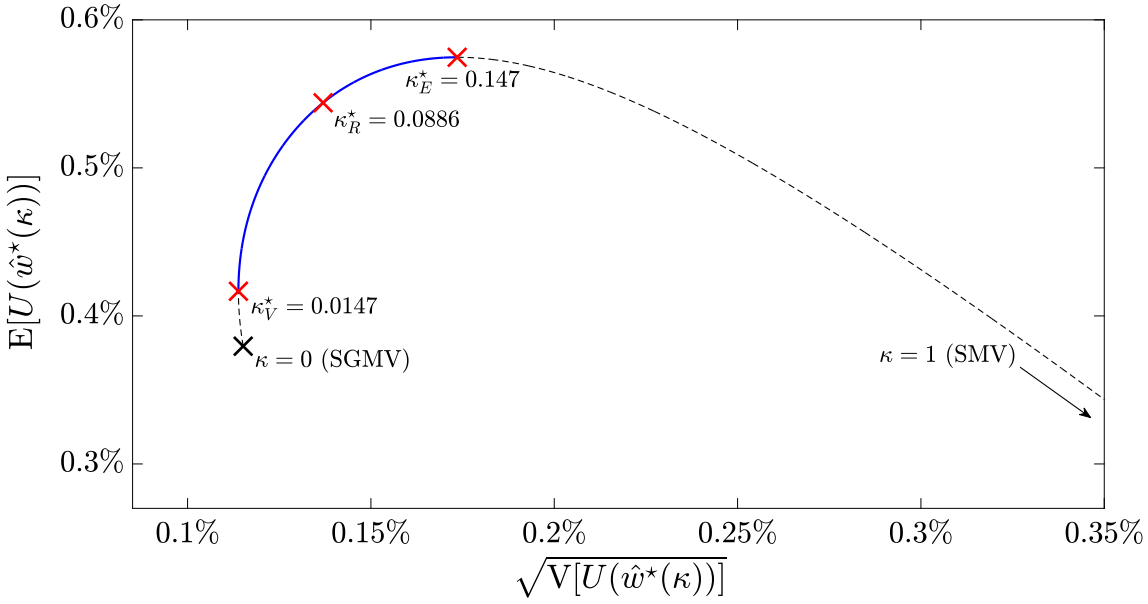
Our second contribution is to propose a novel measure of portfolio robustness defined as the difference between OOSU mean and a multiple of OOSU standard deviation. We assess the robustness of shrinkage portfolios that combine the SMV and SGMV portfolios as in Garlappi, Uppal, and Wang (2007) and Kan, Wang, and Zhou (2022b).[2] We show theoretically that neither the SMV portfolio nor the SGMV portfolio offers maximal robust performance individually, and one must *optimally combine* both to achieve a better tradeoff between OOSU mean and volatility. We term this strategy as the *robust shrinkage portfolio*.

Figure 1 illustrates the main idea behind our robust shrinkage portfolio. The vertical axis de-

in the monthly OOSU of quantitative portfolio strategies represents a paramount concern for investors who deem performance uncertainty a critical factor affecting their investment decisions.

[2]Our portfolio robustness metric can accommodate a broader range of portfolio combinations. Indeed, in Section IA.9.5 of the Internet Appendix, we extend our analysis to a shrinkage portfolio that combines the SMV portfolio, the SGMV portfolio, and the equally weighted portfolio as in Tu and Zhou (2011).

Figure 1: Out-of-sample utility efficient frontier



Notes. This figure depicts the out-of-sample utility mean and standard deviation of shrinkage portfolios $\hat{w}^\star(\kappa)$ that combine the sample global-minimum-variance (SGMV) and sample mean-variance (SMV) portfolios for different values of κ. The shrinkage intensity $\kappa = 0$ corresponds to the SGMV portfolio and $\kappa = 1$ to the SMV portfolio. The population vector of means and covariance matrix of stock excess returns are calibrated from the monthly return data of the 25 portfolios of stocks sorted on size and book-to-market. The shrinkage portfolios are estimated using a sample size of $T = 120$ months and a risk-aversion coefficient of $\gamma = 3$. The solid blue line corresponds to the efficient tradeoff between out-of-sample utility mean and standard deviation provided by the shrinkage portfolios whose shrinkage intensity κ is in the interval $[\kappa_V^\star, \kappa_E^\star]$. The shrinkage intensity $\kappa_R^\star$ maximizes the portfolio robustness measure in Section 5 with $\lambda = 2$.

picts the portfolio OOSU mean, and the horizontal axis depicts its OOSU standard deviation. The parameter κ is the shrinkage intensity defining the combination between the SMV and SGMV portfolios. The shrinkage portfolio exploiting $\kappa_E^\star$ maximizes OOSU mean, and the shrinkage portfolio exploiting $\kappa_R^\star$ maximizes our proposed measure of portfolio robustness. Figure 1 shows that the robust shrinkage portfolio exploiting $\kappa_R^\star$ decreases OOSU standard deviation by 21% at the expense of only a 5% reduction in OOSU mean relative to the shrinkage portfolio exploiting $\kappa_E^\star$. That is, our robustness criterion allows us to construct investment strategies that achieve a substantially more stable out-of-sample performance while sacrificing only a small out-of-sample average performance.[3]

[3]The robustness criterion we propose to combine portfolios also resembles the diversification idea behind the mean-variance efficient frontier of Markowitz (1952). Instead of finding a combination of stocks that achieves an optimal tradeoff between the mean and variance of portfolio returns, we obtain the combination of estimated portfolios that delivers the optimal tradeoff between OOSU mean and volatility.

Our third contribution is to provide a theoretical explanation, from the perspective of parameter uncertainty, of why in-sample near-arbitrage opportunities do not appear to reliably persist out of sample, as noted by Kozak, Nagel, and Santosh (2018). To do this, we use our theoretical results to study how the presence of in-sample near-arbitrage opportunities, which Kozak et al. (2018) define as low-variance principal components that deliver extremely high Sharpe ratios, affects OOSU risk. We demonstrate that in-sample near-arbitrage opportunities and OOSU risk are positively correlated, indicating that portfolios that deliver high in-sample Sharpe ratios are tied to a more uncertain out-of-sample utility. In addition, we show that using a shrinkage covariance matrix in the construction of portfolio strategies alleviates the impact that low-variance principal components have on the performance of sample portfolios, and thus, it helps reduce OOSU risk.

Our fourth contribution is to evaluate the out-of-sample performance of our proposed robust shrinkage portfolio. Our simulations show that the robust shrinkage portfolio delivers a better tradeoff between OOSU mean and standard deviation than the portfolio maximizing only OOSU mean. An appealing feature of our method is that the robust shrinkage portfolio also delivers a larger OOSU mean than the portfolios specifically designed to maximize OOSU mean in cases where returns are not Gaussian and shrinkage intensities are estimated from the data. This demonstrates that our proposed portfolio framework is resilient to estimation errors and model misspecification.

We then study the performance of the robust shrinkage portfolio on six empirical datasets of monthly returns. We document that the proposed shrinkage portfolio outperforms in terms of certainty-equivalent return and Sharpe ratio. Specifically, for an estimation window of 120 monthly observations, the median improvement in terms of certainty-equivalent return (Sharpe ratio) net of transaction costs across the six datasets is 79% (29%) relative to the shrinkage portfolio only maximizing OOSU mean, 151% (30%) relative to the SMV portfolio, 50% (21%) relative to the SGMV portfolio, 100% (51%) relative to the timing strategy of Kirby and Ostdiek (2012), and 179%

(74%) relative to the equally weighted portfolio. In addition, we use three-year non-overlapping windows to evaluate the variability of the shrinkage portfolios' out-of-sample performance. We find that the out-of-sample certainty-equivalent return of our robust portfolio is larger on average and more stable over time than that of the shrinkage portfolio that only maximizes OOSU mean.

Our work has implications for the evaluation of quantitative strategies. Our theory shows that sample strategies exploiting low-variance principal components to achieve large Sharpe ratios are inherently riskier due to their larger OOSU variance. This insight is distinct from Harvey and Liu (2014) who point out that high-Sharpe-ratio trading strategies may exist by chance and not because they truly provide investors with a skill that will continue to outperform out of sample. In contrast, our theory does not rule out the existence of high-Sharpe-ratio quantitative strategies, but it highlights the important caveat that OOSU risk increases for strategies that exploit *in-sample* near-arbitrage opportunities.

Our theory also has implications for the evaluation of asset pricing models that define the stochastic discount factor (SDF) as a linear combination of test assets. In Section IA.2 of the Internet Appendix, we show that there is a link between the OOSU of shrinkage portfolios and the out-of-sample R-squared of a particular robust SDF model. Therefore, our analytical characterization of OOSU risk can be applied to assess the uncertainty of the out-of-sample R-squared of SDF models. Our analysis suggests that SDF models estimated in high-dimensional settings and that capture near-arbitrage opportunities will deliver, in general, unreliable results due to their large out-of-sample R-squared risk.

2 Literature review

We build on the literature pioneered by Kan and Zhou (2007), who study the *average* out-of-sample utility losses of mean-variance portfolios. Kan and Zhou use their analytical characterization of

OOSU mean to build shrinkage portfolios that mitigate the impact of parameter uncertainty on performance.[4] Inspired by the seminal work of Kan and Zhou, an extensive literature studies the average out-of-sample performance of different sample portfolios to propose new strategies that work well in the presence of parameter uncertainty.[5] In contrast to this literature, our manuscript theoretically characterizes the OOSU *variance* of sample portfolios and highlights the relevance of accounting for out-of-sample performance risk in both evaluating and constructing quantitative investment strategies and asset pricing models.

Our work is also related to several papers that study the *distribution* of out-of-sample portfolio performance measures. For instance, Kan and Smith (2008) and Kan et al. (2022b) derive the distribution of the out-of-sample mean return and variance of efficient portfolios, and Kan, Wang, and Zheng (2022a) derive the distribution of the out-of-sample Sharpe ratio of the tangency portfolio. We complement these papers by characterizing the OOSU variance of any convex combination of the SMV, SGMV, and equally weighted portfolios. In addition, we also derive in Appendix IA.5 the asymptotic OOSU distribution of any linear combination between the SMV and SGMV portfolios. Moreover, unlike these papers, we use the OOSU mean and volatility to measure the robustness of sample portfolios and derive an optimal robust shrinkage portfolio. To the best of our knowledge, our work is the first to exploit out-of-sample performance volatility to construct quantitative investment strategies.

The approach we adopt in this manuscript is closely related to the literature on portfolios that exploit *shrinkage estimators* to mitigate the impact of parameter uncertainty. Such estimators are traditionally applied to alleviate the statistical errors affecting the inputs like the mean (Jorion,

[4]Earlier studies consider OOSU mean as a portfolio choice criterion under parameter uncertainty using a Bayesian framework, such as Brown (1976) and Frost and Savarino (1986). However, Kan and Zhou (2007) are the first to characterize the average out-of-sample utility losses from parameter uncertainty analytically.

[5]See Zhou (2008), DeMiguel et al. (2009), Frahm and Memmel (2010), Tu and Zhou (2011), DeMiguel, Martín-Utrera, and Nogales (2013, 2015), Branger, Lučivjanská, and Weissensteiner (2019), Kircher and Rosch (2021), Füss, Koeppel, and Miebs (2021), Kan et al. (2022b), Bodnar, Okhrin, and Parolya (2023), Kan and Wang (2023), and Lassance, Vanderveken, and Vrins (2023).

1986; Barroso and Saxena, 2021) and the covariance matrix (Ledoit and Wolf, 2003, 2004, 2017, 2020). Unlike these papers, we focus on combining *portfolios* to attain an optimal tradeoff between OOSU mean and volatility.

The shrinkage portfolios we consider in this manuscript share fundamental elements with regularization techniques, one of the most common machine learning approaches adopted in the recent asset pricing literature (Giglio, Kelly, and Xiu, 2022; Bryzgalova, Pelger, and Zhu, 2021). Like regularization, the robust shrinkage portfolio proposed in this manuscript helps mitigate the impact of sampling volatility on the estimated portfolio weights. We show that our robust approach to shrinkage portfolios refines and improves the out-of-sample performance of investment strategies over existing methods.

The portfolio approach we propose in this study not only provides a practical method for constructing portfolios that deliver robust performance, but it is also economically sound. Specifically, we show in Section IA.3 of the Internet Appendix that our robust shrinkage portfolio is equivalent to the optimal mean-variance portfolio of an ambiguity-averse investor who does not know the true vector of means, as in Garlappi et al. (2007). This connection provides a microfoundation for the methodology introduced in this manuscript.

The proposed robust shrinkage portfolio also shares elements with the robust portfolio optimization literature. Goldfarb and Iyengar (2003) show that constructing the portfolio that is optimal under the worst-case scenario is a powerful technique "to combat the sensitivity of the optimal portfolio to statistical errors." Relatedly, Appendix IA.4 shows that constructing portfolio combinations that maximize our proposed robustness measure is equivalent to solving a robust optimization problem where the investor maximizes the worst-case scenario of the unknown OOSU mean. Thus, our proposed robust shrinkage portfolio implicitly accounts for statistical errors affecting the OOSU mean.

Finally, our work also speaks to the debate on the replicability of finance research (Harvey, Liu, and Zhu, 2016). In this debate, academics try to find out-of-sample evidence that confirms the validity of cross-sectional anomalies (Jensen, Kelly, and Pedersen, 2021). Similar to this line of research, our work is motivated by the need to create evaluation measures that assess the out-of-sample robustness of quantitative strategies. Our work addresses this issue by providing a theory to measure the out-of-sample performance risk of quantitative strategies that suffer from parameter uncertainty.

3 Mean-variance portfolios and out-of-sample utility

In this section, we review some general concepts in portfolio selection. Section 3.1 defines the classic mean-variance framework, Section 3.2 lays out the assumptions of our theoretical analysis, and Section 3.3 provides the definition of out-of-sample performance.

3.1 Mean-variance portfolios

We first review the portfolio framework introduced by Markowitz (1952) where the investor has mean-variance preferences and knows the true distributional properties of stock returns. There are N stocks whose returns in excess of the risk-free rate have a vector of means μ and a positive-definite covariance matrix $\mathbf{\Sigma}$. In addition, we impose the standard constraint that the investor's wealth is fully allocated to the N risky assets, i.e., $w^\top e = 1$, where e is the N-dimensional vector of ones and w is a vector of portfolio weights. Then, the optimal mean-variance portfolio is the solution to the following quadratic program

$$\max_{w:w^\top e=1} \quad U(w) = w^\top \mu - \frac{\gamma}{2} w^\top \mathbf{\Sigma} w, \tag{1}$$

where $U(w)$ is the *expected utility* of portfolio w and $\gamma > 0$ is the investor's risk-aversion coefficient. The solution to problem (1) is

$$w^\star = w_g + \frac{1}{\gamma} w_z, \tag{2}$$

where w_g is the global-minimum-variance (GMV) portfolio,

$$w_g = \mathbf{\Sigma}^{-1} e (e^\top \mathbf{\Sigma}^{-1} e)^{-1}, \tag{3}$$

and w_z is a zero-cost portfolio (i.e., $w_z^\top e = 0$) defined as

$$w_z = \mathbf{B}\mu, \quad \mathbf{B} = \mathbf{\Sigma}^{-1}(\mathbf{I} - e w_g^\top). \tag{4}$$

For notational simplicity, we define the mean return and variance of the GMV portfolio w_g, and the return variance of the zero-cost portfolio w_z, as

$$\mu_g = w_g^\top \mu = \mu^\top \mathbf{\Sigma}^{-1} e (e^\top \mathbf{\Sigma}^{-1} e)^{-1}, \tag{5}$$

$$\sigma_g^2 = w_g^\top \mathbf{\Sigma} w_g = (e^\top \mathbf{\Sigma}^{-1} e)^{-1}, \tag{6}$$

$$\psi^2 = w_z^\top \mathbf{\Sigma} w_z = \mu^\top \mathbf{\Sigma}^{-1} \mu - \mu_g^2 / \sigma_g^2, \tag{7}$$

respectively. Note that the return variance of the zero-cost portfolio, $\psi^2 \geq 0$, is equal to the difference of squared Sharpe ratios of the tangency portfolio and the GMV portfolio. In Section 6, we show that ψ^2 determines the contribution of low-variance principal components to the maximum attainable Sharpe ratio, and therefore, it indicates whether near-arbitrage opportunities are available to investors.

It is straightforward to show that the expected utility of the mean-variance portfolio $w^\star$ is

$$U(w^\star) = U(w_g) + \frac{\psi^2}{2\gamma}. \tag{8}$$

Because $\psi^2/(2\gamma)$ is always positive, the optimal mean-variance portfolio always delivers a higher in-sample utility than the GMV portfolio. However, the mean-variance portfolio's in-sample optimality does not hold out of sample because of estimation errors affecting the inputs of the portfolio problem. In particular, the impact of estimation errors in the vector of means on portfolio performance can be severe (Merton, 1980; Chopra and Ziemba, 1993). Therefore, it is essential to account for parameter uncertainty in the construction of investment strategies, which we address next.

3.2 Theoretical assumptions

Let us consider the time $T+1$ portfolio return $w^\top r_{T+1}$, where r_{T+1} is the N-dimensional vector of stock returns in excess of the risk-free rate with mean μ and covariance matrix $\mathbf{\Sigma}$. Using historical return data over the past T months $(r_1, \ldots, r_T)$, the investor estimates μ and $\mathbf{\Sigma}$ with their sample counterparts:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_t, \quad \widehat{\mathbf{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T}(r_t - \hat{\mu})(r_t - \hat{\mu})^\top. \tag{9}$$

Consistent with prior literature, we make the following two assumptions.

Assumption 1. *There are at least two stocks, $N \geq 2$, and the sample size is $T > N + 7$.*

Assumption 2. *The vector of stock returns at time t, r_t, follows a multivariate Gaussian distribution with vector of means μ and covariance matrix $\mathbf{\Sigma}$, and all return observations are independent and identically distributed (iid) over time.*

The condition $T > N + 7$ in Assumption 1 is needed to ensure that the out-of-sample utility variance derived in Section 4 exists. Assumption 2 is a standard assumption in the literature used for analytical tractability (Kan and Zhou, 2007; Ao, Li, and Zheng, 2019).

While it is unlikely that stock returns follow a Gaussian distribution, there are two reasons why Assumption 2 does not compromise our analysis. First, the economic losses of optimal portfolios that ignore fat tails in the distribution of stock returns are small, as demonstrated by Tu and Zhou (2004). Second, our empirical results show that the shrinkage portfolios calibrated under the assumption of Gaussian returns deliver good out-of-sample performance even for datasets with empirical return data where returns are not Gaussian.

3.3 Sample portfolios and out-of-sample utility

In practice, investors do not know the true vector of means and covariance matrix of stock returns, and instead, they use the sample estimates in Equation (9). Accordingly, the sample estimate of the mean-variance portfolio in (2), hereafter the SMV portfolio, is

$$\hat{w}^{\star} = \hat{w}_g + \frac{1}{\gamma}\hat{w}_z, \tag{10}$$

where $\hat{w}_g$ is the sample GMV portfolio, hereafter the SGMV portfolio, which is a function of $\widehat{\mathbf{\Sigma}}$ alone, and $\hat{w}_z$ is the sample zero-cost portfolio, which is a function of both $\hat{\mu}$ and $\widehat{\mathbf{\Sigma}}$.

It is well known that the estimation risk affecting the SMV portfolio leads to suboptimal out-of-sample performance (DeMiguel et al., 2009). To combat the impact of parameter uncertainty, we consider shrinkage techniques that help mitigate the impact of statistical errors on the performance of mean-variance portfolios. For instance, Kan et al. (2022b) show that one can improve the *average* out-of-sample performance by *combining* the SMV portfolio $\hat{w}^{\star}$ with the SGMV portfolio $\hat{w}_g$. Following Kan et al. (2022b), we offer a novel methodology to combine the SMV and

SGMV portfolios.

There are at least two reasons why this is a reasonable set of sample portfolios to combine. First, they are both portfolios on the in-sample efficient frontier, and therefore, a combination of both is also an in-sample efficient portfolio. Second, Appendix IA.3 shows that this combination is also economically sound because it has a direct connection with the ambiguity-averse portfolios of Garlappi et al. (2007). Nonetheless, Appendix IA.9.5 extends our theory to construct shrinkage portfolios that combine the SMV and SGMV portfolios with the equally weighted portfolio as in Tu and Zhou (2011).

Like Kan et al. (2022b), we consider a linear combination between the SMV and SGMV portfolios determined by the shrinkage intensity $\kappa \in [0, 1]$:

$$\hat{w}^\star(\kappa) = (1 - \kappa)\hat{w}_g + \kappa\hat{w}^\star \quad \text{with} \quad \kappa \in [0, 1]. \tag{11}$$

Note that the shrinkage portfolio (11) contains as special cases the SMV and SGMV portfolios for $\kappa = 1$ and $\kappa = 0$, respectively. For this reason, the theory we develop in this manuscript for the shrinkage portfolio (11) is applicable to the SMV and SGMV portfolios.

To evaluate the performance of $\hat{w}^\star(\kappa)$ while accounting for estimation risk, we follow Kan and Zhou (2007) and define the *out-of-sample utility* (OOSU) of an estimated portfolio $\hat{w}^\star(\kappa)$ as

$$U(\hat{w}^\star(\kappa)) = \hat{w}^\star(\kappa)^\top \mu - \frac{\gamma}{2}\hat{w}^\star(\kappa)^\top \mathbf{\Sigma}\hat{w}^\star(\kappa). \tag{12}$$

Note that because $\hat{w}^\star(\kappa)$ is estimated from a random sample, the OOSU is on its own a random variable. In Appendix IA.5, we characterize the asymptotic distribution of OOSU when the sample size T and the number of assets N go to infinity and the ratio N/T converges to a constant in $[0, 1)$. We show that the asymptotic distribution is Gaussian and is fully characterized by its mean

and variance. In the next section, we study the finite-sample OOSU variance of sample portfolios.

4 Out-of-sample utility risk

The out-of-sample performance of estimated portfolios is uncertain due to estimation errors. This section sheds new light on the stochastic nature of portfolio performance by characterizing the OOSU variance of portfolios estimated from finite samples. In Section 4.1, we derive the closed-form analytical expression of the OOSU variance for the shrinkage portfolio defined in Equation (11), which contains as particular cases the OOSU variance of the SMV and SGMV portfolios,[6] and in Section 4.2 we discuss the monotonicity properties of the OOSU variance of shrinkage portfolios.

4.1 Out-of-sample utility variance

In the following proposition, we derive the closed-form analytical expression of the OOSU variance of the shrinkage portfolio defined in Equation (11).[7] All proofs are available in the Internet Appendix.

Proposition 1. *Let Assumptions 1 and 2 hold. Then, the out-of-sample utility variance of the shrinkage portfolio* $\hat{w}^\star(\kappa)$ *is*

$$\mathbb{V}[U(\hat{w}^\star(\kappa))] = \mathbb{V}[U(\hat{w}_g)] + \Delta(\kappa), \tag{13}$$

where

$$\mathbb{V}[U(\hat{w}_g)] = \frac{\sigma_g^2\psi^2}{T-N-1} + \frac{\gamma^2\sigma_g^4(N-1)(T-2)}{2(T-N-1)^2(T-N-3)} \tag{14}$$

is the out-of-sample utility variance of the sample GMV portfolio and $\Delta(\kappa)$ *is a fourth-degree polynomial in* κ*,*

$$\Delta(\kappa) = a_1\kappa^4 + a_2\kappa^3 + a_3\kappa^2 + a_4\kappa, \tag{15}$$

[6] Appendix IA.1 reviews the results of Kan et al. (2022b) who characterize the OOSU mean of the shrinkage portfolio $\hat{w}^\star(\kappa)$ in Equation (11).

[7] We are thankful to Raymond Kan for his helpful feedback, which helped us in deriving the result in Proposition 1.

with the coefficients (a_1, a_2, a_3, a_4) *being functions of* γ, T, N, σ_g^2, *and* ψ^2*:*

$$a_1 = \frac{1}{2\gamma^2}\frac{T^2(T-2)C(T,N,\psi^2)}{(T-N)^2(T-N-1)^2(T-N-2)(T-N-3)^2(T-N-5)(T-N-7)}, \tag{16}$$

$$a_2 = -\frac{2\psi^2}{\gamma^2}\frac{T^2(T-2)(T+N-3+2T\psi^2)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}, \tag{17}$$

$$\begin{aligned} a_3 = {} & \frac{\psi^2}{\gamma^2}\frac{2T(N+1)+T^2(T-N-3+2(T-N)\psi^2)}{(T-N)(T-N-1)^2(T-N-3)} \\ & + \sigma_g^2\frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}, \end{aligned} \tag{18}$$

$$a_4 = -2\sigma_g^2\psi^2\frac{T(T-2)}{(T-N-1)^2(T-N-3)}, \tag{19}$$

and

$$\begin{aligned} C(T,N,\psi^2) = {} & (2T\psi^2+N-1)(N^4+N^3T-3N^3-4N^2T^2+22N^2T-31N^2+NT^3 \\ & -7NT^2+13NT-5N+T^4-12T^3+53T^2-100T+70)+T^2\psi^4(N^3 \\ & +2N^2T-6N^2-7NT^2+40NT-53N+4T^3-34T^2+88T-70). \end{aligned}$$

Proposition 1 shows that the OOSU variance of the shrinkage portfolio only depends on six parameters: the shrinkage intensity κ, the investor's risk-aversion coefficient γ, the number of stocks N, the sample size T, the return variance of the GMV portfolio σ_g^2, and the return variance of the zero-cost portfolio ψ^2. We now use the analytical expression of OOSU variance in Proposition 1 to obtain the following corollary.

Corollary 1. *Provided that* ψ^2 *is strictly positive, there is a nonzero shrinkage intensity* $\kappa \in (0,1)$ *for which the shrinkage portfolio* $\hat{w}^\star(\kappa)$ *delivers a lower out-of-sample utility variance than that of the SMV and SGMV portfolios.*

Corollary 1 demonstrates that neither the SMV portfolio nor the SGMV portfolio delivers,

individually, the lowest OOSU variance, and it is optimal to combine them to minimize OOSU variance. We illustrate this point in Figure 1, where we depict the OOSU standard deviation of different shrinkage portfolios in the horizontal axis using the closed-form expression obtained in Proposition 1. For this specific case, the shrinkage intensity that minimizes OOSU variance is $\kappa_V^\star = 0.0147 > 0$.

4.2 Monotonicity properties of out-of-sample utility variance

We now study the monotonicity properties of the OOSU variance of the shrinkage portfolio in (11), which we highlight in the following proposition.

Proposition 2. *Let Assumptions 1 and 2 hold. Then, the out-of-sample utility variance of the shrinkage portfolio* $\hat{w}^\star(\kappa)$

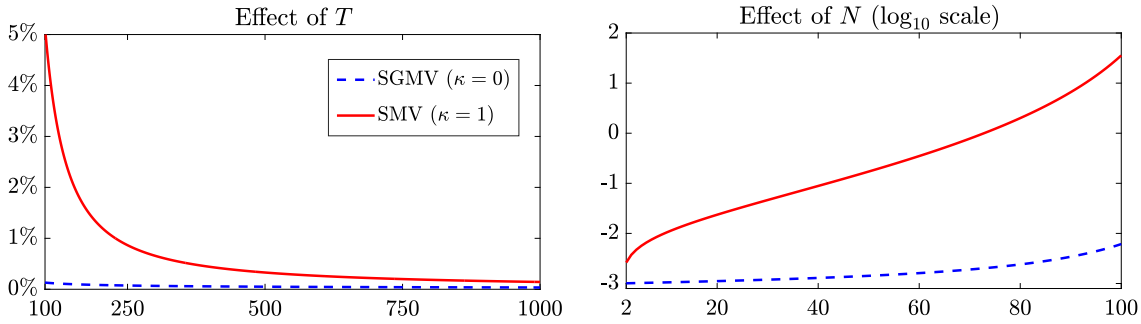
1. *decreases with the sample size* T *and converges to zero as* $T \to \infty$,

2. *increases with the number of stocks* N*, the return variance of the GMV portfolio* σ_g^2*, the return variance of the zero-cost portfolio* ψ^2*, and the shrinkage intensity* κ *if* $\kappa \geq \kappa_E^\star$*, where* $\kappa_E^\star$*, defined in Equation* (IA3) *of the Internet Appendix, is the shrinkage intensity that maximizes the out-of-sample utility mean of the shrinkage portfolio.*

Proposition 2 provides the intuitive result that the OOSU variance of sample portfolios is substantial in high-dimensional settings where the number of assets N is large relative to the sample size T.[8] Moreover, OOSU volatility increases with parameters σ_g^2 and ψ^2, which are the return variances of the GMV portfolio w_g and the zero-cost portfolio w_z, respectively. Finally, Proposition 2 shows that for a shrinkage intensity $\kappa \geq \kappa_E^\star$, the substantial exposure to the SMV portfolio leads to an increasing OOSU volatility as we increase κ and the shrinkage portfolio gets

[8] We use the terms OOSU risk and OOSU volatility to refer to either the OOSU variance or its square root, the OOSU standard deviation, of shrinkage portfolios.

Figure 2: Effect of sample size and number of stocks on out-of-sample utility volatility



Notes. This figure depicts the out-of-sample utility standard deviation of the SGMV portfolio ($\kappa = 0$, dashed blue line) and the SMV portfolio ($\kappa = 1$, solid red line). The population vector of means and covariance matrix of stock excess returns are calibrated from the monthly return data of the 25 portfolios of stocks sorted on size and book-to-market. This gives a value of $\sigma_g = 0.0437$ and a value of $\psi^2 = 0.0625$. We set a risk-aversion coefficient of $\gamma = 3$. We vary the sample size T in the left panel while keeping a fixed $N = 25$, and we vary the number of stocks N in the right panel while keeping a fixed $T = 120$. The values in the right panel are in log-scale for visibility.

closer to the SMV portfolio. Therefore, while κ should be greater than zero to minimize OOSU risk as demonstrated in Corollary 1, it must be smaller than $\kappa_E^\star$.

Figure 2 illustrates the results in Proposition 2. For conciseness, we only show the results for the sample size T and the number of stocks N. We calibrate the distributional parameters using the sample moments of the 25SBTM dataset described in Section 7.1. In addition, we set a risk-aversion coefficient of $\gamma = 3$. In the left panel, we vary T while keeping a fixed $N = 25$. In the right panel, we vary N while keeping a fixed $T = 120$. We focus on the OOSU volatility of the SMV and SGMV portfolios, which correspond to the special cases with $\kappa = 1$ and $\kappa = 0$ in the shrinkage portfolio (11), respectively.

The left panel in Figure 2 shows that the OOSU standard deviation of the SMV portfolio is substantially larger than that of the SGMV portfolio. Specifically, for a realistic sample size of $T = 120$ monthly observations, the monthly OOSU standard deviation of the SMV portfolio is approximately 3.34%, which is 29 times larger than that of the SGMV portfolio. Additionally, consistent with Proposition 2, we observe that the OOSU standard deviation decreases with the sample size. However, it is worth noting that the SMV portfolio requires an unrealistically large

sample size of $T = 13{,}340$ monthly observations to have a smaller OOSU volatility than the SGMV portfolio.

The right panel in Figure 2 shows that OOSU standard deviation increases with the number of stocks N as demonstrated in Proposition 2. The effect of the number of stocks N is particularly severe for the SMV portfolio, for which we see that OOSU volatility increases much more rapidly than for the SGMV portfolio.

5 A new portfolio robustness measure

We now use the results in Section 4 to propose a novel measure of portfolio robustness defined as the difference between OOSU mean and a multiple of OOSU standard deviation. Section 5.1 introduces the robustness measure. Section 5.2 studies the shrinkage portfolio that optimizes the proposed robustness measure. Section 5.3 explains how we estimate the shrinkage intensities. Section 5.4 describes the monotonicity properties of the robustness measure.

5.1 A new robustness measure

In our view, a robust portfolio should deliver good average performance *and* a stable performance. In line with this view, the robustness measure of a given estimated portfolio $\hat{w}$ is equal to the difference between OOSU mean and a multiple of OOSU standard deviation:

$$R(\hat{w}) = \mathbb{E}[U(\hat{w})] - \lambda\sqrt{\mathbb{V}[U(\hat{w})]}, \tag{20}$$

where $\lambda \geq 0$ determines the weight that OOSU risk has on our robustness measure. Note that for $\lambda = 0$, we recover the OOSU mean criterion proposed by Kan and Zhou (2007). We dub our robustness measure the *mean-risk OOSU*.

In addition, this measure of portfolio robustness is related to the OOSU Value-at-Risk, as we highlight in the following proposition.

Proposition 3. *Let $N, T \to \infty$, $N/T \to \rho \in [0, 1)$, and Φ be the standard normal cumulative distribution function. Then, the mean-risk out-of-sample utility of the shrinkage portfolio $\hat{w}^\star(\kappa)$ converges to the $\Phi(-\lambda)$-Value-at-Risk of its out-of-sample utility.*[9]

Proposition 3 shows that the mean-risk OOSU in Equation (20) converges asymptotically to the OOSU Value-at-Risk at a specific confidence level. This result demonstrates that maximizing our proposed robustness measure minimizes asymptotically the left-tail risk in the out-of-sample utility of estimated portfolios.

5.2 The robust shrinkage portfolio

We now define our robust shrinkage portfolio that combines the SMV and SGMV portfolios to maximize the mean-risk OOSU measure defined in Section 5.1. Formally, the intensity of the robust shrinkage portfolio is the solution to the following problem:

$$\kappa_R^\star = \arg \max_{\kappa \in [0,1]} R(\hat{w}^\star(\kappa)). \tag{21}$$

Problem (21) can be easily solved numerically using the analytical expressions for the OOSU mean in Equation (IA2) of the Internet Appendix and for the OOSU variance in Proposition 1. Note that we recover $\kappa_R^\star = \kappa_E^\star$ when $\lambda = 0$ and $\kappa_R^\star = \kappa_V^\star$ when $\lambda \to \infty$, where $\kappa_E^\star$ is the shrinkage intensity maximizing OOSU mean and $\kappa_V^\star$ is the shrinkage intensity minimizing OOSU variance.

The spirit of our robust portfolio approach resembles the fundamental risk-diversification tenet of Markowitz (1952). In our case, instead of obtaining the combination of stocks that achieves the

[9] We define the α-Value-at-Risk (e.g., $\alpha = 5\%$) of OOSU such that $\mathbb{P}(U(\hat{w}^\star(\kappa)) \leq \text{VaR}_\alpha) = \alpha$.

optimal tradeoff between the mean and variance of portfolio returns, we obtain the combination of estimated portfolios that achieves the optimal tradeoff between OOSU mean and variance. Figure 1 depicts the OOSU efficient frontier for the 25SBTM dataset, $T = 120$, and $\gamma = 3$. First, note that the shrinkage intensity $\kappa_E^\star$ proposed by Kan et al. (2022b) delivers the portfolio on the OOSU efficient frontier that maximizes OOSU mean. Second, the shrinkage portfolio that maximizes the mean-risk OOSU metric with $\lambda = 2$ delivers an OOSU standard deviation 21% lower than that of the shrinkage portfolio maximizing OOSU mean. To achieve this substantial reduction in OOSU risk, the shrinkage portfolio that exploits $\kappa_R^\star$ only sacrifices a small 5.3% reduction in OOSU mean relative to the shrinkage portfolio that exploits $\kappa_E^\star$.

In the following proposition, we formally prove two important properties of the shrinkage intensity $\kappa_R^\star$.

Proposition 4. *Let Assumptions 1 and 2 hold. Then, the shrinkage intensity $\kappa_R^\star$ solving (21) has the following properties:*

1. *$\kappa_R^\star \to 1$ as $T \to \infty$. Thus, $\hat{w}^\star(\kappa_R^\star)$ is a consistent estimator of $w^\star$.*

2. *$\kappa_V^\star \leq \kappa_R^\star \leq \kappa_E^\star$, where $\kappa_V^\star$ minimizes the out-of-sample utility variance and $\kappa_E^\star$ maximizes the out-of-sample utility mean of the shrinkage portfolio, respectively.*

The first result of Proposition 4 shows that our proposed robust shrinkage portfolio is asymptotically optimal. The second result of Proposition 4 demonstrates that while an investor facing parameter uncertainty can increase her average OOSU by shrinking the SMV portfolio toward the SGMV portfolio, a more substantial shrinkage toward the SGMV portfolio is needed to reduce OOSU variance further and enhance portfolio robustness. Using the insight in Corollary 1 that $\kappa_V^\star > 0$ if $\psi^2 > 0$, the second result of Proposition 4 also implies that neither the SMV portfolio nor the SGMV portfolio optimizes our proposed measure of portfolio robustness. Thus, combining

them using intensity $\kappa_R^\star$ is necessary to achieve the maximal robust performance.

5.3 Estimation of optimal shrinkage intensities

The optimal shrinkage intensity $\kappa_R^\star$ maximizing the mean-risk OOSU of the shrinkage portfolio depends on the true moments of stock returns via σ_g^2 and ψ^2. Similarly, the shrinkage intensity $\kappa_E^\star$ maximizing OOSU mean in Equation (IA3) in the Internet Appendix depends on parameter ψ^2. Since these parameters are unknown, Appendix IA.6 provides details for the consistent estimators of these parameters, which we use in the construction of the shrinkage intensities. Specifically, we estimate ψ^2 with the estimator of Kan and Zhou (2007), and we estimate σ_g^2 via the estimator of Frahm and Memmel (2010). In the rest of the manuscript, we denote the estimated intensities that exploit the consistent estimators of ψ^2 and σ_g^2 as $\hat{\kappa}_R^\star$ and $\hat{\kappa}_E^\star$.[10]

5.4 Monotonicity properties of the robustness measure

In the following proposition, we provide some monotonicity properties for the mean-risk OOSU measure of the shrinkage portfolio (11).

Proposition 5. *The mean-risk out-of-sample utility of the shrinkage portfolio $\hat{w}^\star(\kappa)$ satisfies the following monotonicity properties:*

1. *It increases with the sample size T and the mean return of the GMV portfolio μ_g.*
2. *It decreases with the number of stocks N, the return variance of the GMV portfolio σ_g^2, and the shrinkage intensity κ if $\kappa \geq \kappa_E^\star$.*
3. *Provided that $\kappa \leq \frac{2(T-N)(T-N-3)}{T(T-2)}$, it increases with ψ^2 if $\lambda < \lambda_0$ and decreases with ψ^2 if*

[10] The result in Part 2 of Proposition 4 holds for any value of σ_g^2 and ψ^2. Hence, the estimated shrinkage intensities also obey the inequality $\hat{\kappa}_R^\star \leq \hat{\kappa}_E^\star$. Moreover, they remain asymptotically optimal.

$\lambda > \lambda_0$, *where*

$$\lambda_0 = \frac{\frac{\partial}{\partial \psi^2}\mathbb{E}[U(\hat{w}^\star(\kappa))]}{\frac{\partial}{\partial \psi^2}\sqrt{\mathbb{V}[U(\hat{w}^\star(\kappa))]}} \geq 0. \tag{22}$$

Proposition 5 demonstrates that the mean-risk OOSU measure increases with T and decreases with N. This is a desirable property of our proposed robustness metric because increasing T and decreasing N reduces the statistical errors affecting the estimated moments of stock returns and their impact on sample portfolios. Moreover, the mean-risk OOSU measure increases with μ_g because the OOSU mean increases with μ_g and the OOSU standard deviation is independent of μ_g. On the contrary, the mean-risk OOSU decreases with σ_g^2 because the OOSU mean is decreasing in σ_g^2, and the OOSU standard deviation is increasing in σ_g^2 as shown in Proposition 2. Proposition 5 also demonstrates that allocating more weight to the SMV portfolio than $\kappa_E^\star$ necessarily deteriorates the mean-risk OOSU measure, which is why $\kappa_R^\star \leq \kappa_E^\star$ as highlighted in Proposition 4.

Finally, Proposition 5 also shows that when λ is higher than a certain threshold, a higher ψ^2 leads to a decrease in our robustness measure. This is because while ψ^2 has a positive impact on OOSU mean, it also positively impacts OOSU volatility.

6 Relation with near-arbitrage opportunities

In this section, we use the results introduced in Section 4 to illustrate the connection between the OOSU risk of shrinkage portfolios and near-arbitrage opportunities. We first introduce the following assumption for tractability, which is also used by Kozak et al. (2018).

Assumption 3. *Consider the eigenvalue decomposition of the covariance matrix of stock returns,* $\mathbf{\Sigma} = \mathbf{V}\mathbf{D}\mathbf{V}^\top$, *where* $\mathbf{D} = \text{diag}(d_1, \ldots, d_N)$ *is the diagonal matrix of eigenvalues, and* $\mathbf{V} = [v_1, \ldots, v_N]$ *is the matrix of eigenvectors. We assume that the first eigenvector is proportional to the equally weighted portfolio, that is,* $v_1 = e/\sqrt{N}$.

Assumption 3 is not necessary for the analytical expressions of out-of-sample utility risk presented in Section 4. We only introduce this assumption for tractability and ease of exposition for the analysis in this section.[11] The following proposition employs Assumption 3 only to decompose the return variance of the zero-cost portfolio, ψ^2 in (7), as the sum of the squared Sharpe ratios of low-variance principal components.

Proposition 6. *Let Assumption 3 hold. Then, the return variance of the zero-cost portfolio is*

$$\psi^2 = \sum_{i>1}^{N} SR_{PC_i}^2, \tag{23}$$

where $SR_{PC_i} = v_i^\top \mu / \sqrt{d_i}$ *is the Sharpe ratio of the* i*th principal component of stock returns.*

Proposition 6 reveals that ψ^2 is determined by the squared Sharpe ratios of low-variance principal components, and thus, a large value of ψ^2 indicates the presence of *near-arbitrage opportunities*. Kozak et al. (2018) document that low-variance principal components with substantial contribution to the overall performance of the tangency portfolio are hard to exploit out of sample. Kozak et al. (2018) also argue that "[it is] implausible that a principal component with low eigenvalue could contribute substantially to the volatility of the SDF and hence to the overall maximum squared Sharpe ratio." Our theoretical result in Proposition 2 that the OOSU variance of shrinkage portfolios increases with ψ^2 complements and substantiates these claims. Specifically, even though our results do not suggest that it is implausible to find near-arbitrage opportunities, we demonstrate that they are risky in nature because they command higher out-of-sample performance volatility.

Finally, from an investment perspective, it is interesting to understand whether quantitative strategies exploiting *in-sample* near-arbitrage opportunities are subject to higher OOSU risk. The following proposition addresses this concern.

[11] Section IA.7 in the Internet Appendix formally characterizes the correlation between the first principal component of stock returns and the equally weighted portfolio.

Proposition 7. *Let Assumptions 1 and 2 hold. Then, provided that $T > N + 9$, the covariance between $\hat{\psi}^2 = \hat{\mu}^\top \widehat{\mathbf{\Sigma}}^{-1} \hat{\mu} - (\hat{\mu}_g / \hat{\sigma}_g)^2$ and $(U(\hat{w}(\kappa)) - \mathbb{E}[U(\hat{w}(\kappa))])^2$ exists and is always positive.*

Proposition 7 demonstrates that sample portfolios exploiting in-sample near-arbitrage opportunities face larger OOSU risk because the in-sample measure of near-arbitrage opportunities, $\hat{\psi}^2$, covaries positively with the squared difference between the OOSU of a shrinkage portfolio and its mean.

6.1 Mitigating the impact of in-sample arbitrage opportunities

We have shown that exploiting in-sample arbitrage opportunities, defined as in-sample low-variance principal components that deliver high Sharpe ratios, is associated with higher OOSU risk. We now discuss a technique that mitigates the impact of low-variance principal components, which helps reduce OOSU risk. The considered method relies on shrinking the sample covariance matrix toward the identity as in Ledoit and Wolf (2004), which is

$$\widehat{\mathbf{\Sigma}}_{\text{sh}} = (1 - \delta)\widehat{\mathbf{\Sigma}} + \delta \bar{\sigma}^2 \mathbf{I}, \tag{24}$$

where $\bar{\sigma}^2$ is the cross-sectional average of sample variances and δ is the shrinkage intensity. The eigenvalue decomposition of this shrinkage covariance matrix is

$$\widehat{\mathbf{\Sigma}}_{\text{sh}} = \widehat{\boldsymbol{V}} \Big((1 - \delta)\widehat{\boldsymbol{D}} + \delta \bar{\sigma}^2 \mathbf{I} \Big) \widehat{\boldsymbol{V}}^\top, \tag{25}$$

where $\widehat{\boldsymbol{V}}$ and $\widehat{\boldsymbol{D}}$ are the matrix of eigenvectors and the diagonal matrix of eigenvalues, respectively, obtained from $\widehat{\mathbf{\Sigma}}$. Taking the derivative of the shrinkage covariance matrix eigenvalue $d_{\text{sh},i}$ with

respect to the shrinkage intensity δ gives

$$\frac{\partial d_{\text{sh},i}}{\partial \delta} = \frac{\partial\Big((1-\delta)\hat{d}_i + \delta\bar{\sigma}^2\Big)}{\partial \delta} = \bar{\sigma}^2 - \hat{d}_i, \tag{26}$$

where $\hat{d}_i$ is the ith eigenvalue of matrix $\widehat{\boldsymbol{\Sigma}}$. Expression (26) indicates that shrinking the covariance matrix toward the identity will increase the eigenvalues of the sample covariance matrix, provided that $\bar{\sigma}^2 - \hat{d}_i > 0$, and this effect will be larger among the lowest eigenvalues. Therefore, by shrinking the sample covariance matrix, we inflate the low eigenvalues of the estimated covariance matrix, which allows us to attenuate the impact of low-variance principal components on the out-of-sample performance of sample portfolios. The analysis in Section 7 assesses the impact that shrinking the covariance matrix has on the performance of sample portfolios with simulated and empirical data.[12]

7 Performance analysis

In this section, we illustrate the economic gains of our proposed robust shrinkage portfolio. In Section 7.1, we describe the data used in the performance analysis. In Section 7.2, we evaluate the performance of our strategy using simulated data and empirically assess the relation between OOSU risk and near-arbitrage opportunities. In Section 7.3, we evaluate the performance of our strategy using empirical data.

7.1 Data

We use the monthly excess returns of six datasets. The first four datasets are downloaded from Kenneth French's website: (i) 10 momentum portfolios (*10MOM*) from January 1927 through December

[12]Section IA.8 in the Internet Appendix develops theoretically the OOSU risk of a shrinkage portfolio that exploits a linear shrinkage covariance matrix. In Section IA.9.7 of the Internet Appendix, we assess the performance of the shrinkage portfolios that exploit the linear shrinkage covariance matrix of Ledoit and Wolf (2004). In addition, in Section IA.9.8 of the Internet Appendix, we assess the empirical performance of the shrinkage portfolios that exploit the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020).

2019, (ii) 25 portfolios formed on size and book-to-market (*25SBTM*) from January 1927 through December 2019, (iii) 25 portfolios formed on operating profitability and investment (*25OPINV*) from July 1963 through December 2019, and (iv) 49 industry portfolios (*49IND*) from July 1969 through December 2019. The last two datasets come from the 23 anomalies considered by Novy-Marx and Velikov (2016) and are downloaded from Robert Novy-Marx's website: (v) the long and short legs of eight low-turnover anomalies (*16LTANOM*) from July 1963 through December 2013 and (vi) the long and short legs of all the 23 anomalies (*46ANOM*) from July 1973 through December 2013.[13]

7.2 Simulated returns

We use two different methods to simulate monthly return data. In the first method, we draw observations from a multivariate Gaussian distribution. In the second method, our return data is not Gaussian, and instead, we simulate data using the bootstrap method of Efron (1979).

We now explain how we construct the Gaussian simulated returns. For each of the six empirical datasets in Section 7.1, we compute the sample vector of means $\hat{\mu}$ and sample covariance matrix $\widehat{\mathbf{\Sigma}}$ employing the entire sample and use these sample estimates as the population parameters of a multivariate Gaussian distribution $\mathcal{N}(\hat{\mu}, \widehat{\mathbf{\Sigma}})$ from which we draw T observations with $T \in (120, 180, 240)$. We construct $M = 100{,}000$ simulated datasets of T observations using this method and compute the estimated shrinkage portfolio $\hat{w}_m(\hat{\kappa}_m)$ for each of the M simulated datasets, where $m = 1, \ldots, M$. Then, the OOSU mean, OOSU variance, and mean-risk OOSU of the estimated shrinkage portfolio

[13] In addition to these datasets, in Section IA.9.9 of the Internet Appendix, we assess the performance of the different portfolio strategies in a high-dimensional dataset of $N = 155$ assets, where we merge all six datasets for the time period for which we have available return data across the six datasets. The results in this section confirm that our main insight about the importance of accounting for OOSU risk in the construction of portfolio strategies is robust in high-dimensional settings where the number of assets is large relative to the number of observations. We thank Kenneth French and Robert Novy-Marx for making their data publicly available.

$\hat{w}^\star(\hat{\kappa})$ are approximated as

$$\mathbb{E}[U(\hat{w}^\star(\hat{\kappa}))] \approx \frac{1}{M}\sum_{m=1}^{M} U(\hat{w}_m^\star(\hat{\kappa}_m)), \tag{27}$$

$$\mathbb{V}[U(\hat{w}^\star(\hat{\kappa}))] \approx \frac{1}{M}\sum_{m=1}^{M} (U(\hat{w}_m^\star(\hat{\kappa}_m)) - \mathbb{E}[U(\hat{w}^\star(\kappa))])^2, \tag{28}$$

$$R(\hat{w}^\star(\hat{\kappa})) \approx \mathbb{E}[U(\hat{w}^\star(\hat{\kappa}))] - \lambda\sqrt{\mathbb{V}[U(\hat{w}^\star(\hat{\kappa}))]}, \tag{29}$$

where $U(\hat{w}_m^\star(\hat{\kappa}_m))$ is the investor's out-of-sample utility, defined as in Equation (12), of the estimated shrinkage portfolio $\hat{w}_m^\star(\hat{\kappa}_m)$ obtained from the mth simulated dataset. We set the risk-aversion coefficient to $\gamma = 3$ as in Kan and Zhou (2007) and Kan et al. (2022b), and coefficient $\lambda = 2$. Therefore, the robustness measure we consider in the performance analysis is asymptotically equivalent to the 2.5%-Value-at-Risk of OOSU as shown in Proposition 3.[14]

The simulation with Gaussian data is interesting because the theoretical results rely on the assumption that stock returns are iid multivariate Gaussian; see Assumption 2. However, this assumption does not hold in practice, and therefore, the second type of simulated data is obtained by bootstrapping return data from the original sample. In particular we create 1,000 bootstrap samples of $2T$ return observations, where $T \in (120, 180, 240)$. For each bootstrap sample of $2T$ observations, we use the first half to estimate the shrinkage portfolio and evaluate its performance in the second half of the sample. We compute the OOSU mean, variance, and the mean-risk OOSU as in Equations (27)–(29) from the 1,000 OOSU observations obtained with bootstrap.

Panel A of Table 1 reports the OOSU mean, standard deviation, and the mean-risk OOSU for the simulated Gaussian data. We consider the shrinkage portfolio with the estimated intensity $\hat{\kappa}_E^\star$ that maximizes OOSU mean and the shrinkage portfolio with the estimated intensity $\hat{\kappa}_R^\star$ that maximizes our proposed robustness measure with $\lambda = 2$. We observe that for a sample size of

[14] In Section IA.9.6 of the Internet Appendix, we show that our insights are robust to cross-validating λ.

$T = 120$ observations, the shrinkage portfolio that exploits $\hat{\kappa}_R^\star$ delivers a larger OOSU mean than that of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ for four datasets. This suggests that $\hat{\kappa}_R^\star$ emerges as a robust shrinkage intensity that is subject to lower estimation risk than $\hat{\kappa}_E^\star$, which allows our robust shrinkage portfolio to outperform in terms of OOSU mean the shrinkage portfolio designed to optimize it. Indeed, in unreported results, we find that the estimated $\hat{\kappa}_E^\star$ has a larger sampling variability than $\hat{\kappa}_R^\star$, and thus, the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ is more sensitive to estimation errors.[15] In addition, we observe that the shrinkage portfolio that exploits $\hat{\kappa}_E^\star$ delivers an OOSU that is more volatile than that of the shrinkage portfolio that exploits $\hat{\kappa}_R^\star$. The difference is particularly large for the case with a sample size of $T = 120$ observations. Accordingly, the shrinkage portfolio that exploits the intensity $\hat{\kappa}_R^\star$ delivers a better mean-risk OOSU than that obtained by the shrinkage portfolio that exploits $\hat{\kappa}_E^\star$.[16]

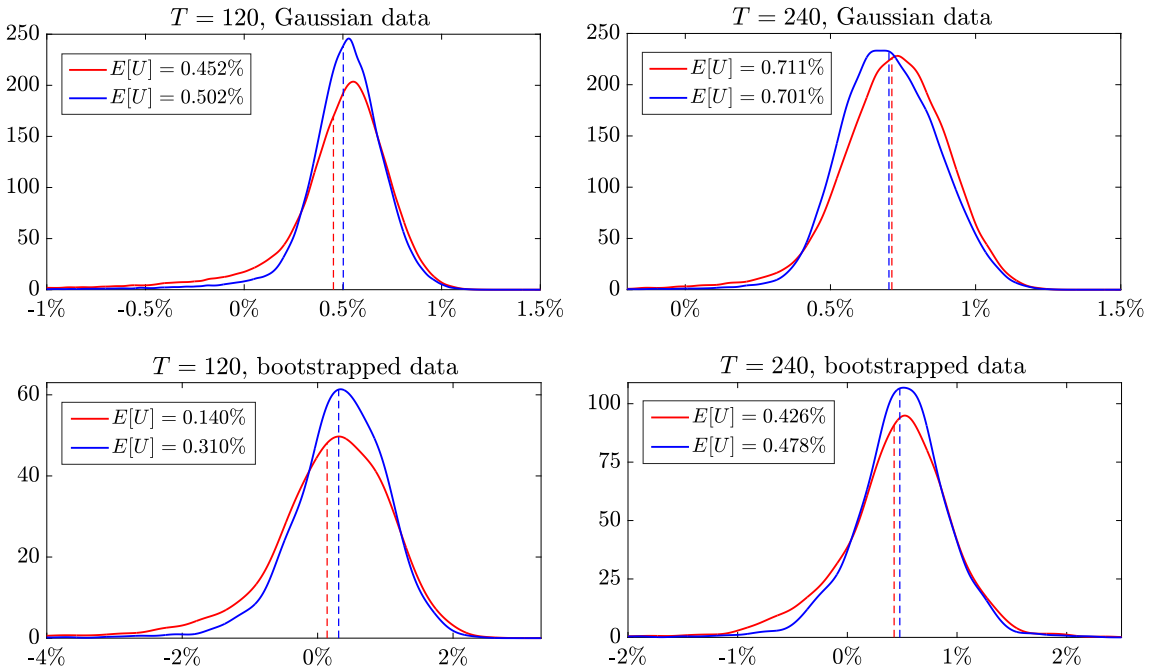
Panel B of Table 1 reports the performance of the two shrinkage portfolios in the bootstrap experiment. In terms of OOSU mean, we see that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ performs remarkably better. Specifically, it outperforms the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ in five out of six datasets across all sample sizes. In addition, the shrinkage intensity $\hat{\kappa}_E^\star$ yields an OOSU volatility that is, on average across all datasets, 32%, 24%, and 21% larger than that of $\hat{\kappa}_R^\star$ for $T = 120$, 180, and 240 months, respectively. The results in this panel suggest that the performance of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ deteriorates relative to that of the robust shrinkage portfolio when the data is not Gaussian.

Figure 3 depicts the OOSU density of the estimated shrinkage portfolios for the simulated return data that uses the *25SBTM* dataset. The figure shows that, when the sample size is $T = 120$ and

[15] In Section IA.9.10 of the Internet Appendix, we assess the performance of the oracle shrinkage portfolios where the shrinkage intensities are computed using the population vector of means and covariance matrix. This allows us to study the impact that the estimation errors affecting $\kappa_E^\star$ and $\kappa_R^\star$ have on the performance of the shrinkage portfolios. We find that the performance loss caused by estimation errors in shrinkage intensities is larger for the shrinkage portfolio exploiting $\kappa_E^\star$ than $\kappa_R^\star$.

[16] In unreported results, we show that the shrinkage portfolio constructed with $\hat{\kappa}_R^\star$ outperforms the SMV and SGMV portfolios in terms of OOSU mean and mean-risk OOSU.

Figure 3: Density of monthly out-of-sample utility in simulated data

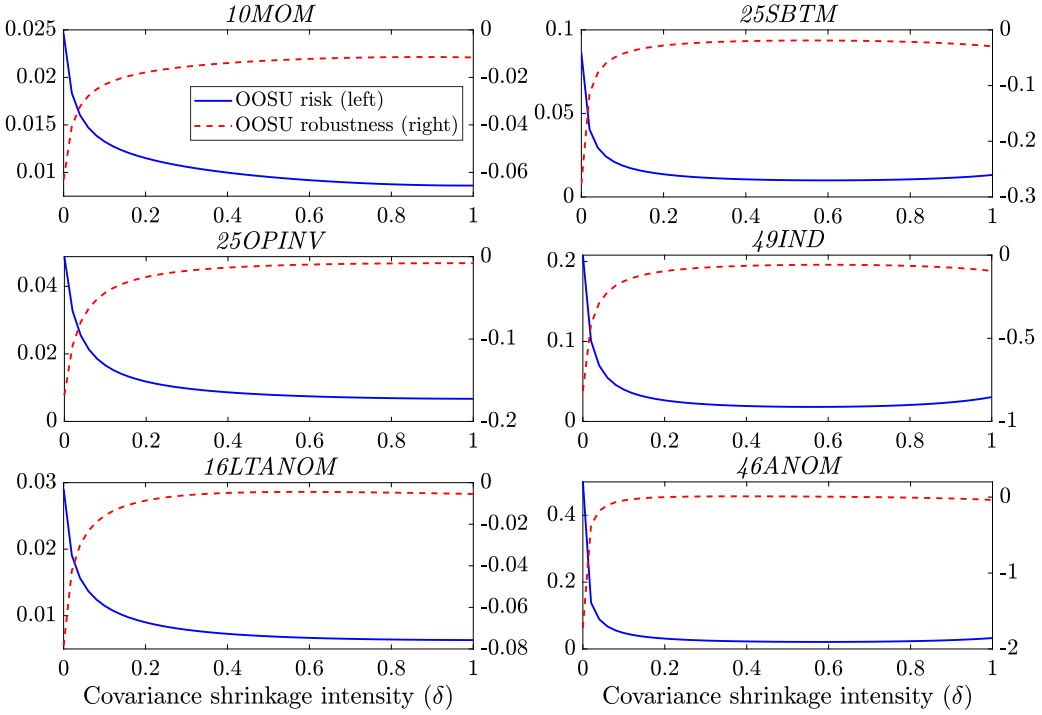


Notes. This figure depicts the monthly out-of-sample utility densities of the estimated shrinkage portfolios maximizing out-of-sample utility mean ($\hat{\kappa}_E^\star$, in red) and the portfolio robustness measure in Section 5 ($\hat{\kappa}_R^\star$, in blue). The top figures depict the density function of the out-of-sample utilities of the shrinkage portfolios from the 100,000 simulated samples of T observations drawn from a multivariate Gaussian distribution whose moments are calibrated from the dataset of monthly excess returns of 25 portfolios of stocks sorted on size and book-to-market. The bottom figures are obtained by bootstrapping (with replacement) 1,000 samples of $2T$ observations from the dataset of 25 portfolios of stocks sorted on size and book-to-market, where the first half of the bootstrap sample is used to estimate the two shrinkage portfolios and the second half is used to evaluate the out-of-sample utility of the shrinkage portfolios estimated in the first half of the sample. We consider a sample size of $T = 120$ and 240 monthly observations, a risk-aversion coefficient of $\gamma = 3$, and a coefficient $\lambda = 2$ for the portfolio robustness measure.

when the data is not Gaussian, the shrinkage portfolio that exploits $\hat{\kappa}_R^\star$ has a larger OOSU mean and a smaller OOSU volatility than the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$. In addition, the OOSU density of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ has a heavier left tail than that of the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$. This result is consistent with the idea that our robust portfolio minimizes the tail risk of OOSU as we highlight in Section 5.1. The lower downside risk offered by our robust shrinkage portfolio represents an additional advantage of our proposed methodology.

Finally, as shown in Section 6.1, shrinking the sample covariance matrix toward the identity attenuates the impact that low-variance principal components have on the performance of sample

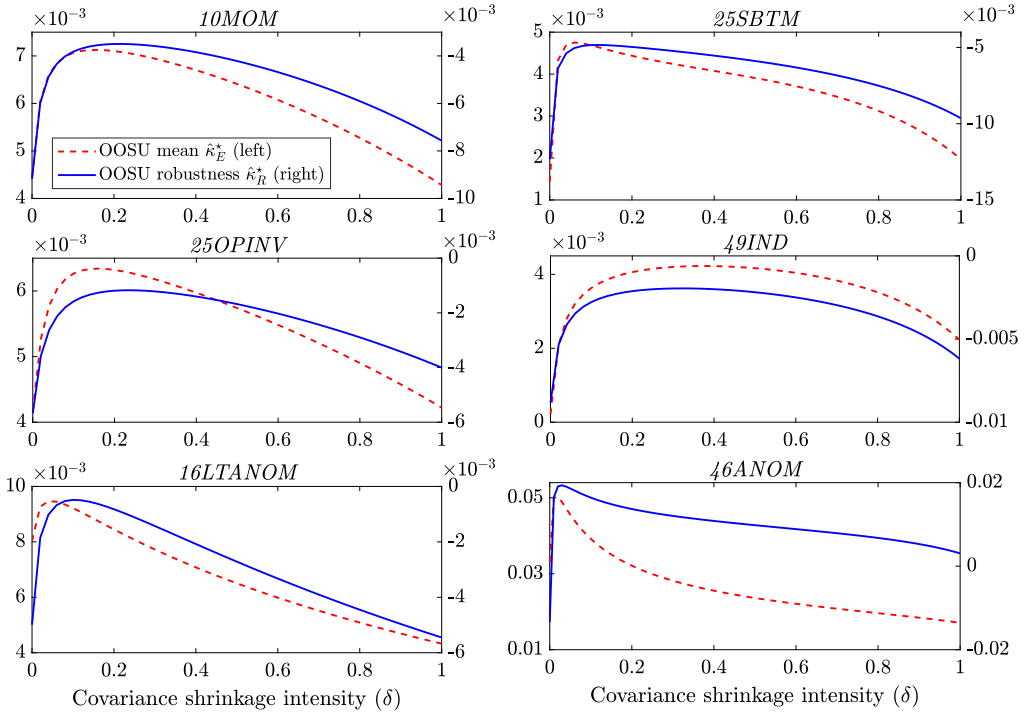
Figure 4: Shrinkage covariance matrix and out-of-sample utility of SMV portfolio



Notes. This figure depicts the out-of-sample utility standard deviation (left axis) and the portfolio robustness measure in Section 5 (right axis) for the sample mean-variance (SMV) portfolio as a function of the intensity with which the sample covariance matrix is shrunk toward the identity. More precisely, we use the covariance matrix $\widehat{\mathbf{\Sigma}}_{\text{sh}} = (1-\delta)\widehat{\mathbf{\Sigma}} + \delta\bar{\sigma}^2\mathbf{I}$ where $\bar{\sigma}^2$ is the cross-sectional average of asset-return variances and δ is the shrinkage intensity. For each dataset, we draw 10,000 bootstrap samples of $2T$ observations (with replacement), where $T = 120$. We estimate the SMV portfolio using the first T observations and the shrinkage covariance matrix $\widehat{\mathbf{\Sigma}}_{\text{sh}}$. Then, we evaluate the OOSU of this portfolio in the second half of the sample. Finally, we compute the OOSU risk and robustness of the SMV portfolio across the 10,000 bootstrap samples. We run this experiment for a shrinkage intensity δ between zero and one. We consider a risk-aversion coefficient of $\gamma = 3$ and a coefficient $\lambda = 2$ for the portfolio robustness measure.

portfolios and helps to reduce OOSU risk. Accordingly, Figure 4 demonstrates how shrinking the covariance matrix toward the identity as in (24) attenuates the impact that low-variance principal components have on the OOSU risk and robustness of the SMV portfolio. We draw 10,000 bootstrap samples of $2T$ observations for each dataset, where $T = 120$, and estimate the SMV portfolio with $\gamma = 3$ exploiting the shrinkage covariance matrix $\widehat{\mathbf{\Sigma}}_{\text{sh}}$ using the first T observations. Then, we evaluate the OOSU of this portfolio in the second half of the sample. The left vertical axis plots OOSU risk, which is the standard deviation of the OOSU of the SMV portfolio across all 10,000 bootstrap samples. Similarly, the right vertical axis plots the mean-risk OOSU with $\lambda = 2$.

Figure 5: Shrinkage covariance matrix and out-of-sample utility of shrinkage portfolios



Notes. This figure depicts the out-of-sample utility mean of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ (left axis) and the portfolio robustness measure in Section 5 of the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ (right axis) as a function of the intensity with which the sample covariance matrix is shrunk toward the identity. More precisely, we use the covariance matrix $\widehat{\mathbf{\Sigma}}_{\text{sh}} = (1-\delta)\widehat{\mathbf{\Sigma}} + \delta\bar{\sigma}^2\mathbf{I}$ where $\bar{\sigma}^2$ is the cross-sectional average of asset-return variances and δ is the shrinkage intensity. For each dataset, we draw 10,000 bootstrap samples of $2T$ observations (with replacement), where $T = 120$. We estimate the two shrinkage portfolios using the first T observations and the shrinkage covariance matrix $\widehat{\mathbf{\Sigma}}_{\text{sh}}$. Then, we evaluate the OOSU of the two portfolios in the second half of the sample. Finally, we compute the OOSU mean and robustness across the 10,000 bootstrap samples. We run this experiment for a shrinkage intensity δ between zero and one. We consider a risk-aversion coefficient of $\gamma = 3$ and a coefficient $\lambda = 2$ for the portfolio robustness measure.

Even though low-variance principal components help improve *in-sample* performance (Kozak et al., 2018), the empirical exercise in Figure 4 corroborates the results in Propositions 2 and 7 that exploiting low-variance principal components contributes to having larger OOSU risk. To see this, note in Figure 4 that as we increase the shrinkage intensity δ, the impact of low-variance principal components on the performance of the SMV portfolio is attenuated, and therefore, the OOSU risk decreases while the mean-risk OOSU increases.

In addition to the previous analysis, we are also interested in whether shrinking the sample covariance matrix can further improve the performance of shrinkage portfolios that optimally combine

the SMV and SGMV portfolios. Figure 5 illustrates the relation between the shrinkage intensity of the covariance matrix and the out-of-sample performance of the shrinkage portfolios that optimally combine the SMV and SGMV portfolios. For the shrinkage portfolio that maximizes OOSU mean (i.e., using intensity $\hat{\kappa}_E^\star$), we illustrate the relation between the shrinkage intensity of the covariance matrix and the OOSU mean. For the shrinkage portfolio that maximizes the mean-risk OOSU with $\lambda = 2$ (i.e., using intensity $\hat{\kappa}_R^\star$), we depict the relation between the shrinkage intensity of the covariance matrix and the mean-risk OOSU. There are three important insights we can obtain from this figure. First, we see that across all datasets, shrinking the sample covariance matrix can further improve the shrinkage portfolios' out-of-sample performance. Second, it is not optimal to fully shrink the sample covariance matrix (i.e., $\delta < 1$). Third, the optimal shrinkage intensity of the covariance matrix δ is different depending on the criterion one wants to optimize (i.e., OOSU mean versus mean-risk OOSU). In the next section, we capitalize on these insights and construct the shrinkage portfolios using the shrinkage covariance matrix in (24) where the shrinkage intensity δ is calibrated using the same criterion used to obtain the optimal combination between the SMV and SGMV portfolios.

7.3 Empirical returns

In this section, we evaluate the out-of-sample performance of the robust shrinkage portfolio and several benchmark strategies using empirical return data. Consistent with the analysis with simulated data, the results in this section confirm that our proposed shrinkage portfolio is a robust strategy that delivers favorable average out-of-sample performance *and* a stable out-of-sample performance.

We use the six datasets of characteristic and industry-sorted portfolios described in Section 7.1. We study the performance of six portfolio strategies. First, the equally weighted (EW) portfolio. Second, the reward-to-risk (RTR) timing strategy of Kirby and Ostdiek (2012). Third, the sample

global-minimum-variance (SGMV) portfolio. Fourth, the sample mean-variance (SMV) portfolio. Fifth, the shrinkage portfolio that exploits the intensity $\hat{\kappa}_E^\star$ maximizing OOSU mean as in Kan et al. (2022b). The last portfolio strategy is our proposed robust shrinkage portfolio that exploits the intensity $\hat{\kappa}_R^\star$ maximizing the mean-risk OOSU criterion that we introduce in Section 5, with a fixed value of $\lambda = 2$ as in the simulated return data experiment of Section 7.2. We set the risk-aversion coefficient to $\gamma = 3$ as in Kan and Zhou (2007) and Kan et al. (2022b).[17]

Our performance analysis exploits a covariance matrix that shrinks the sample covariance matrix toward the identity. The SMV and SGMV portfolios are constructed using the shrinkage covariance matrix of Ledoit and Wolf (2004), and the shrinkage portfolios use a covariance matrix whose shrinkage intensity is calibrated to optimize the same criterion used to combine the SGMV and SMV portfolios. In particular, for the shrinkage portfolio that maximizes OOSU mean, we numerically find the shrinkage intensity of the covariance matrix that maximizes the portfolio's OOSU mean. For the shrinkage portfolio that maximizes the mean-risk OOSU with $\lambda = 2$, we numerically find the shrinkage intensity of the covariance matrix that maximizes the portfolio mean-risk OOSU.[18] While this empirical setting imposes a higher hurdle for our proposed robust portfolio, we still find that our methodology consistently outperforms.

Similar to DeMiguel et al. (2009), we use a rolling-window approach to evaluate the out-of-sample performance of the different portfolio strategies. In particular, let τ be the total number of

[17] In Appendix IA.9.3, we also consider risk aversions of $\gamma = 1$ and 5. In addition, in Appendix IA.9.5, we consider a shrinkage portfolio that combines the SMV, SGMV, and EW portfolios as in Tu and Zhou (2011).

[18] We first obtain the optimal shrinkage intensity κ without applying shrinkage to the sample covariance matrix. With that optimal κ, we identify the optimal shrinkage intensity δ for the covariance matrix using a nonparametric approach. In particular, we consider a grid of different values for the shrinkage intensity of the covariance matrix in $[0, 1]$ and for each value, we draw 1,000 bootstrap samples of T observations (with replacement) from the estimation window. We then split the bootstrap sample into five folds of size $T/5$, estimate the shrinkage portfolio using the shrinkage covariance matrix with four out of the five folds, and evaluate the estimated portfolio in the remaining fold. We do this for every four-fold combination, which gives T out-of-sample returns that we use to approximate the OOSU of the shrinkage portfolio in that particular bootstrap sample. We repeat this procedure for every bootstrap sample to obtain a distribution of the OOSU of the shrinkage portfolio that exploits a shrinkage covariance matrix. Finally, we compute the OOSU mean and mean-risk OOSU of the estimated shrinkage portfolio across the 1,000 bootstrap samples using (27) and (29). We then find the optimal value of the shrinkage intensity δ for the shrinkage covariance matrix that maximizes the OOSU mean of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ and that maximizes the mean-risk OOSU of the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$.

monthly returns in the dataset and T the sample size used to estimate the portfolios. We estimate portfolio w using an estimation window containing the first T monthly returns of our sample, and compute its first out-of-sample return in month $T+1$ as $\tilde{p}_{T+1} = w^\top r_{T+1}$, where r_{T+1} is the vector of stock returns in month $T+1$. We then move the estimation window one month ahead and construct the out-of-sample return in month $T+2$ similarly. We repeat this process until the end of the sample, which gives a time series of $\tau - T$ out-of-sample returns, i.e., $\tilde{p}_t$, $t = T+1, \ldots, \tau$. Our experiments consider estimation window sizes of $T = 120$ and 240 monthly observations.

We compute the portfolio turnover at time t as

$$\text{Turnover}_t = \sum_{i=1}^{N} |w_{i,t} - w_{i,(t-1)^+}|, \quad t = T+1, \ldots, \tau, \tag{30}$$

where $w_{i,t}$ is the weight of stock i in month t and $w_{i,(t-1)^+}$ is the prior-month weight before rebalancing in month t that takes into account portfolio growth. We use the portfolio turnover at time t to compute out-of-sample portfolio returns net of proportional transaction costs as

$$p_{T+1} = \tilde{p}_{T+1} \quad \text{and} \quad p_t = (1+\tilde{p}_t)(1 - c \times \text{Turnover}_{t-1}) - 1, \quad t = T+2, \ldots, \tau, \tag{31}$$

where c is the proportional cost required to rebalance the portfolio. We report the results for the case without transaction costs (i.e., $c = 0$) and for the case where $c = 20$ basis points, which is the same level of transaction costs as in Kan et al. (2022b).[19]

We then compute the out-of-sample mean and variance of portfolio returns net of costs as

$$\mu_p = \frac{1}{\tau - T} \sum_{t=T+1}^{\tau} p_t \quad \text{and} \quad \sigma_p^2 = \frac{1}{\tau - T} \sum_{t=T+1}^{\tau} (p_t - \mu_p)^2.$$

[19] In Appendix IA.9.4, we show that our conclusions are robust to considering a larger c of 30 basis points.

We compare the six portfolio strategies in terms of their annualized out-of-sample certainty-equivalent return (CER) and Sharpe ratio (SR):

$$\text{OOS CER} = 12 \times \left(\mu_p - \frac{\gamma}{2}\sigma_p^2\right), \tag{32}$$

$$\text{OOS SR} = \sqrt{12} \times \mu_p/\sigma_p. \tag{33}$$

The OOS CER corresponds to the empirical out-of-sample utility of estimated portfolios. We also test the null hypothesis that the OOS CER or SR delivered by the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ are equal to those delivered by the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ and the EW portfolio, against the alternative hypothesis that $\hat{\kappa}_R^\star$ yields larger OOS CER or SR. To compute the test p-values, we generate 1,000 bootstrap samples using the stationary block bootstrap approach of Politis and Romano (1994) with an average block size of five and use the methodology of Ledoit and Wolf (2008, Remark 3.2.) to produce the resulting p-values.

Finally, we also evaluate the out-of-sample performance risk of shrinkage portfolios by dividing the out-of-sample portfolio returns into non-overlapping three-year windows.[20] We then compute the OOS CER in each window, and the mean and standard deviation of the OOS CER across all windows. This additional experimental setting allows us to evaluate the average out-of-sample performance delivered by the shrinkage portfolios and their out-of-sample performance risk, in line with the robustness measure we introduce in this manuscript.

Table 2 reports the out-of-sample results for the portfolios constructed with the standard estimation window of $T = 120$ monthly observations. We document that the proposed robust shrinkage portfolio outperforms in terms of certainty-equivalent return and Sharpe ratio. In particular, the median improvement in terms of certainty-equivalent return (Sharpe ratio) net of transaction costs

[20] We use three-year windows to obtain a reasonable tradeoff between performance evaluation frequency and the number of observations in each window. The conclusions are robust to using other window lengths.

across the six datasets is 79% (29%) relative to the shrinkage portfolio only maximizing OOSU mean, 151% (30%) relative to the SMV portfolio, 50% (21%) relative to the SGMV portfolio, 100% (51%) relative to the timing portfolio, and 179% (74%) relative to the EW portfolio. The outperformance of the robust shrinkage portfolio relative to the benchmark portfolios is similar in magnitude in the absence of transaction costs.

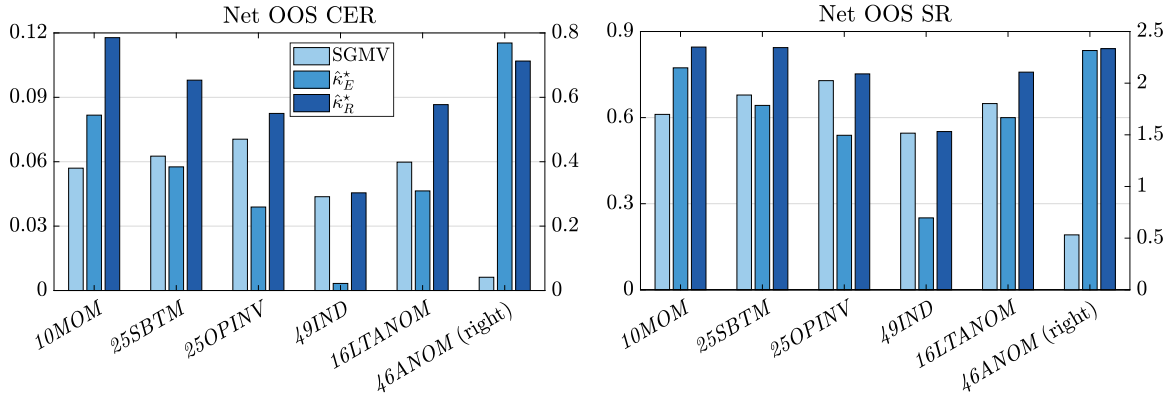
Table 3 reports the out-of-sample results for the portfolios constructed with a larger estimation window of $T = 240$ monthly observations. The results in this new experimental setting are consistent with those presented in Table 2. In particular, the median improvement in terms of certainty-equivalent return (Sharpe ratio) net of transaction costs across the six datasets is 66% (16%) relative to the shrinkage portfolio only maximizing OOSU mean, 75% (17%) relative to the SMV portfolio, 62% (13%) relative to the SGMV portfolio, 124% (50%) relative to the timing portfolio and 203% (80%) relative to the EW portfolio. Like in the case with $T = 120$, the robust shrinkage portfolio also outperforms the benchmark portfolios in the absence of transaction costs for $T = 240$.[21]

Figure 6 presents an overview of the results for the case with an estimation window of $T = 120$ monthly observations. In particular, this figure depicts the certainty-equivalent return (OOS CER) and Sharpe ratio (OOS SR) net of transaction costs. From this figure, it is easy to recognize the favorable performance exhibited by our robust shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ relative to the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ and the SGMV portfolio.

Finally, in addition to outperforming the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ in terms of average

[21] Note that for the *46ANOM* dataset with $T = 240$, the performance of the mean-variance portfolio and the shrinkage portfolios is substantially worse than for the case with $T = 120$ observations. This is because when using $T = 240$, the out-of-sample period is different from that when using $T = 120$ observations. To see this, note that for a dataset with a total of τ observations, our first estimation window corresponds with the first T observations, and hence the out-of-sample period contains $\tau - T$ observations. While this implies that the out-of-sample period is different for the cases with $T = 120$ and $T = 240$, this approach allows us to assess the performance of each portfolio with the longest out-of-sample period we have available to gain statistical power. In unreported results, we show that the performance of the shrinkage portfolios for the case with $T = 120$ substantially deteriorates when we assess these portfolios over the same out-of-sample period as that for the case with $T = 240$. That is, we assess the portfolios over the last $\tau - 240$ observations of the dataset. Indeed, we see that in this case, the performance of the shrinkage portfolios for $T = 120$ and $T = 240$ is similar both in terms of certainty-equivalent return and Sharpe ratio.

Figure 6: Out-of-sample certainty-equivalent return and Sharpe ratio in empirical data



Notes. This figure depicts the annualized out-of-sample certainty-equivalent return and Sharpe ratio of the SGMV portfolio, the shrinkage portfolio maximizing out-of-sample utility mean ($\hat{\kappa}_E^\star$), and the shrinkage portfolio maximizing the portfolio robustness measure in Section 5 ($\hat{\kappa}_R^\star$) for the six datasets of monthly excess returns described in Section 7.1. We use a sample size of $T = 120$ monthly observations and a risk-aversion coefficient of $\gamma = 3$. The certainty-equivalent return and Sharpe ratio are net of proportional transaction costs of 20 basis points. The right axis reports the performance for the *46ANOM* dataset for visibility.

OOS CER, the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ also delivers a substantially more stable out-of-sample performance. To see this, we report in Table 4 the mean and standard deviation of out-of-sample CER across three-year non-overlapping windows. We find that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ systematically delivers an OOS CER that is both larger on average and more stable over time.[22]

8 Conclusion

We characterize the out-of-sample utility (OOSU) risk of the sample mean-variance (SMV) portfolio, the sample global-minimum-variance (SGMV) portfolio, and any linear combination of the two portfolios. We show that SMV portfolios need unrealistically large sample sizes —in some cases over 1,000 years of monthly return data— to deliver an out-of-sample performance as stable as that of SGMV portfolios. In general, the OOSU risk of sample portfolios is substantial in high-dimensional settings where the number of assets is large relative to the sample size, and when these portfolios

[22] Appendix IA.9 confirms that our main insight about the benefits of accounting for OOSU risk in the construction of shrinkage portfolios is robust to a number of different empirical specifications.

exploit in-sample near-arbitrage opportunities. We then use our characterization of OOSU risk to propose a novel measure of portfolio robustness that strikes a balance between OOSU mean and OOSU volatility. We show that neither the SMV portfolio nor the SGMV portfolio delivers maximal robustness individually. In particular, one needs to *combine* both portfolios to obtain an optimal tradeoff between OOSU mean and volatility. We find that shrinkage portfolios that optimize our measure of portfolio robustness deliver higher certainty-equivalent returns and Sharpe ratios than several prominent benchmark portfolios.

Our robust portfolio framework admits a broader range of settings than the one considered in the main body of the manuscript. For instance, we also use our framework in the Internet Appendix to construct shrinkage portfolios that combine the SMV portfolio, the SGMV portfolio, and the equally weighted portfolio as in Tu and Zhou (2011).

Our manuscript highlights the value of accounting for out-of-sample performance risk when evaluating and constructing quantitative investment strategies. On the evaluation front, financial institutions can use our theory to provide investors with valuable information about performance uncertainty, in line with the idea that investment strategies should not be judged solely based on a single past performance realization. On the construction side, we show that sample portfolios that exploit near-arbitrage opportunities in high-dimensional settings are doomed to experience large OOSU risk. In contrast, portfolios that account for OOSU risk are more resilient to estimation errors and deliver robust performance.

Tables

Table 1: Out-of-sample performance of shrinkage portfolios in simulated data

		OOSU mean			OOSU risk			Mean-risk OOSU		
	T	120	180	240	120	180	240	120	180	240
Panel A: Gaussian data (in %)										
10MOM	$\hat{\kappa}_E^\star$	0.63	0.72	0.77	0.27	0.18	0.15	0.09	0.35	0.47
	$\hat{\kappa}_R^\star$	0.66	0.72	0.76	0.20	0.15	0.13	0.27	0.43	0.50
25SBTM	$\hat{\kappa}_E^\star$	0.45	0.61	0.71	0.37	0.24	0.20	-0.28	0.13	0.31
	$\hat{\kappa}_R^\star$	0.50	0.62	0.70	0.25	0.19	0.17	0.01	0.24	0.36
25OPINV	$\hat{\kappa}_E^\star$	0.67	0.85	0.98	0.38	0.27	0.23	-0.10	0.32	0.52
	$\hat{\kappa}_R^\star$	0.70	0.84	0.95	0.26	0.22	0.21	0.17	0.40	0.52
49IND	$\hat{\kappa}_E^\star$	0.33	0.56	0.70	0.44	0.29	0.23	-0.56	-0.02	0.24
	$\hat{\kappa}_R^\star$	0.40	0.57	0.69	0.26	0.20	0.19	-0.11	0.17	0.31
16LTANOM	$\hat{\kappa}_E^\star$	1.11	1.37	1.54	0.42	0.34	0.29	0.27	0.70	0.97
	$\hat{\kappa}_R^\star$	1.06	1.32	1.50	0.37	0.34	0.31	0.32	0.63	0.89
46ANOM	$\hat{\kappa}_E^\star$	5.84	8.04	9.31	1.48	1.11	0.91	2.89	5.81	7.48
	$\hat{\kappa}_R^\star$	5.70	7.96	9.27	1.27	1.03	0.87	3.17	5.90	7.54
Panel B: Boostrapped data (in %)										
10MOM	$\hat{\kappa}_E^\star$	0.45	0.54	0.61	0.93	0.63	0.49	-1.41	-0.72	-0.36
	$\hat{\kappa}_R^\star$	0.53	0.58	0.63	0.77	0.54	0.43	-1.01	-0.50	-0.22
25SBTM	$\hat{\kappa}_E^\star$	0.14	0.29	0.43	1.05	0.72	0.57	-1.96	-1.16	-0.71
	$\hat{\kappa}_R^\star$	0.31	0.38	0.48	0.77	0.57	0.46	-1.23	-0.76	-0.45
25OPINV	$\hat{\kappa}_E^\star$	0.37	0.50	0.59	0.82	0.56	0.41	-1.27	-0.61	-0.24
	$\hat{\kappa}_R^\star$	0.47	0.55	0.61	0.63	0.44	0.34	-0.80	-0.33	-0.06
49IND	$\hat{\kappa}_E^\star$	0.07	0.18	0.23	0.69	0.44	0.35	-1.31	-0.69	-0.47
	$\hat{\kappa}_R^\star$	0.21	0.27	0.30	0.50	0.33	0.27	-0.78	-0.39	-0.24
16LTANOM	$\hat{\kappa}_E^\star$	0.75	0.99	1.16	0.85	0.62	0.48	-0.95	-0.24	0.21
	$\hat{\kappa}_R^\star$	0.75	0.96	1.12	0.67	0.52	0.42	-0.60	-0.08	0.29
46ANOM	$\hat{\kappa}_E^\star$	3.53	4.95	5.93	3.59	2.71	2.05	-3.65	-0.46	1.84
	$\hat{\kappa}_R^\star$	3.90	5.24	6.13	2.61	2.19	1.71	-1.32	0.85	2.70

Notes. This table reports the out-of-sample performance of estimated shrinkage portfolios across six different datasets of simulated data. The first two blocks of three columns report the mean and standard deviation of the monthly out-of-sample utility (in percentage) of the estimated shrinkage portfolios. The third block of three columns reports the mean-risk out-of-sample utility, which is the proposed robustness metric in Section 5. We report the results for the estimated shrinkage portfolio maximizing out-of-sample utility mean ($\hat{\kappa}_E^\star$) and for the estimated shrinkage portfolio maximizing the proposed robustness measure ($\hat{\kappa}_R^\star$). In Panel A, we define the population parameters of a multivariate Gaussian distribution with the sample moments of each of the six datasets of monthly excess returns described in Section 7.1, and draw 100,000 samples of size T. For each simulated sample, we construct the two shrinkage portfolios and their corresponding out-of-sample utilities using Equation (12). Finally, we use the out-of-sample utilities of the 100,000 simulated samples to construct our performance metrics as in (27)-(29). We proceed similarly in Panel B by generating 1,000 bootstrap samples of size $2T$ monthly observations, constructing the two shrinkage portfolios using the first T observations and computing the out-of-sample utility in the remaining T observations, and finally evaluating the performance metrics across the 1,000 simulated bootstrap samples. We consider sample sizes of $T = 120$, 180, and 240 monthly observations, a risk-aversion coefficient of $\gamma = 3$, and a coefficient $\lambda = 2$ for the portfolio robustness measure.

Table 2: Out-of-sample performance in empirical data ($T = 120$)

		EW	RTR	SGMV	SMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$
10MOM	Gross OOS CER	0.034	0.051	0.061	0.082	0.115	$\mathbf{0.134}_{\circ\circ\circ}$
	Net OOS CER	0.033	0.049	0.057	0.047	0.082	$\mathbf{0.118}^{\star\star}_{\circ\circ\circ}$
	Gross OOS SR	0.455	0.558	0.641	0.877	0.865	$\mathbf{0.909}_{\circ\circ\circ}$
	Net OOS SR	0.449	0.547	0.611	0.798	0.773	$\mathbf{0.846}^{\star\star}_{\circ\circ\circ}$
	Average $\hat{\kappa}$	/	/	0	1	0.395	0.296
25SBTM	Gross OOS CER	0.045	0.052	0.072	-0.012	**0.123**	$\mathbf{0.123}_{\circ\circ\circ}$
	Net OOS CER	0.044	0.051	0.063	-0.107	0.058	$\mathbf{0.098}^{\star\star}_{\circ\circ\circ}$
	Gross OOS SR	0.522	0.561	0.748	0.813	0.861	$\mathbf{0.991}^{\star\star}_{\circ\circ\circ}$
	Net OOS SR	0.516	0.554	0.679	0.631	0.643	$\mathbf{0.844}^{\star\star\star}_{\circ\circ\circ}$
	Average $\hat{\kappa}$	/	/	0	1	0.214	0.145
25OPINV	Gross OOS CER	0.044	0.055	0.079	-0.269	0.080	$\mathbf{0.102}_{\circ\circ\circ}$
	Net OOS CER	0.043	0.053	0.071	-0.364	0.039	$\mathbf{0.083}^{\star\star\star}_{\circ\circ}$
	Gross OOS SR	0.515	0.586	0.794	0.614	0.699	$\mathbf{0.872}^{\star\star\star}_{\circ\circ\circ}$
	Net OOS SR	0.509	0.575	0.729	0.461	0.538	$\mathbf{0.752}^{\star\star\star}_{\circ}$
	Average $\hat{\kappa}$	/	/	0	1	0.191	0.120
49IND	Gross OOS CER	0.050	0.052	0.053	-1.456	0.039	$\mathbf{0.057}^{\star\star}$
	Net OOS CER	0.049	**0.050**	0.044	-1.565	0.008	$0.046^{\star\star\star}$
	Gross OOS SR	0.552	0.565	0.624	0.264	0.486	$\mathbf{0.645}^{\star\star\star}$
	Net OOS SR	0.544	**0.554**	0.546	0.125	0.279	$0.551^{\star\star\star}$
	Average $\hat{\kappa}$	/	/	0	1	0.043	0.024
16LTANOM	Gross OOS CER	0.027	0.044	0.066	0.016	0.098	$\mathbf{0.109}_{\circ\circ\circ}$
	Net OOS CER	0.026	0.042	0.060	-0.040	0.046	$\mathbf{0.087}^{\star\star}_{\circ\circ}$
	Gross OOS SR	0.419	0.516	0.695	0.734	0.775	$\mathbf{0.887}^{\star}_{\circ\circ\circ}$
	Net OOS SR	0.413	0.505	0.649	0.609	0.600	$\mathbf{0.758}^{\star\star}_{\circ\circ}$
	Average $\hat{\kappa}$	/	/	0	1	0.309	0.218
46ANOM	Gross OOS CER	0.016	0.052	0.054	-0.532	**0.831**	$0.765_{\circ\circ\circ}$
	Net OOS CER	0.015	0.051	0.041	-0.609	**0.769**	$0.713_{\circ\circ\circ}$
	Gross OOS SR	0.358	0.566	0.643	2.272	2.431	$\mathbf{2.442}_{\circ\circ\circ}$
	Net OOS SR	0.353	0.555	0.532	2.134	2.316	$\mathbf{2.334}_{\circ\circ\circ}$
	Average $\hat{\kappa}$	/	/	0	1	0.250	0.207

Notes. This table reports the out-of-sample performance of the six portfolio strategies described in Section 7.3 for the six datasets of monthly excess returns described in Section 7.1. Each estimated portfolio is constructed using a sample size of $T = 120$ monthly observations. The mean-variance portfolios consider a risk-aversion coefficient of $\gamma = 3$. The table reports the annualized out-of-sample certainty-equivalent return (OOS CER) and the annualized out-of-sample Sharpe ratio (OOS SR), and for both criteria we report the gross performance and the performance net of proportional transaction costs of 20 basis points. We also report the average estimated shrinkage intensity $\hat{\kappa}$ over time, except for the EW and RTR portfolios that do not combine the SMV and SGMV portfolios. The stars $\star, \star\star, \star\star\star$ establish that the OOS CER and SR of the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ is larger than that of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ at a confidence level of 10%, 5%, and 1%, respectively. The circles $\circ, \circ\circ, \circ\circ\circ$ convey the same information relative to the EW portfolio. The numbers in bold font identify the best portfolio in terms of OOS CER and SR.

Table 3: Out-of-sample performance in empirical data ($T = 240$)

		EW	RTR	SGMV	SMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$
10MOM	Gross OOS CER	0.039	0.057	0.069	0.106	0.134	$\mathbf{0.152}_{\circ\circ\circ}$
	Net OOS CER	0.038	0.056	0.065	0.083	0.114	$\mathbf{0.138}^{\star\star}_{\circ\circ\circ}$
	Gross OOS SR	0.484	0.607	0.706	0.917	0.917	$\mathbf{0.959}^{\star}_{\circ\circ\circ}$
	Net OOS SR	0.478	0.600	0.682	0.864	0.862	$\mathbf{0.912}^{\star\star}_{\circ\circ\circ}$
	Average $\hat{\kappa}$	/	/	0	1	0.546	0.471
25SBTM	Gross OOS CER	0.053	0.060	0.081	0.092	0.139	$\mathbf{0.144}_{\circ\circ\circ}$
	Net OOS CER	0.052	0.059	0.074	0.031	0.095	$\mathbf{0.111}_{\circ\circ\circ}$
	Gross OOS SR	0.567	0.612	0.833	0.917	0.914	$\mathbf{1.012}^{\star\star}_{\circ\circ\circ}$
	Net OOS SR	0.561	0.606	0.778	0.788	0.776	$\mathbf{0.851}^{\star}_{\circ\circ}$
	Average $\hat{\kappa}$	/	/	0	1	0.364	0.289
25OPINV	Gross OOS CER	0.050	0.060	0.102	-0.103	0.124	$\mathbf{0.135}_{\circ\circ\circ}$
	Net OOS CER	0.049	0.059	0.097	-0.154	0.099	$\mathbf{0.119}^{\star}_{\circ\circ}$
	Gross OOS SR	0.556	0.633	0.976	0.660	0.872	$\mathbf{1.017}^{\star\star\star}_{\circ\circ\circ}$
	Net OOS SR	0.550	0.627	**0.935**	0.565	0.772	$0.928^{\star\star\star}_{\circ\circ}$
	Average $\hat{\kappa}$	/	/	0	1	0.320	0.232
49IND	Gross OOS CER	0.052	**0.054**	0.044	-0.784	0.024	$0.046^{\star\star}$
	Net OOS CER	0.051	**0.053**	0.038	-0.846	0.008	$0.036^{\star\star\star}$
	Gross OOS SR	0.569	**0.592**	0.557	0.111	0.391	$0.557^{\star\star}$
	Net OOS SR	0.562	**0.585**	0.501	0.014	0.293	$0.475^{\star\star}$
	Average $\hat{\kappa}$	/	/	0	1	0.087	0.054
16LTANOM	Gross OOS CER	0.028	0.045	0.095	0.123	0.172	$\mathbf{0.175}_{\circ\circ\circ}$
	Net OOS CER	0.027	0.044	0.091	0.086	0.136	$\mathbf{0.158}_{\circ\circ\circ}$
	Gross OOS SR	0.422	0.520	0.942	0.947	1.016	$\mathbf{1.132}^{\star}_{\circ\circ\circ}$
	Net OOS SR	0.416	0.513	0.904	0.863	0.910	$\mathbf{1.051}^{\star\star}_{\circ\circ\circ}$
	Average $\hat{\kappa}$	/	/	0	1	0.522	0.449
46ANOM	Gross OOS CER	0.022	0.057	0.098	-2.160	0.138	$\mathbf{0.347}^{\star\star\star}_{\circ\circ}$
	Net OOS CER	0.021	0.056	0.088	-2.159	0.022	$\mathbf{0.295}^{\star\star\star}_{\circ}$
	Gross OOS SR	0.399	0.600	1.037	1.522	1.536	$\mathbf{1.648}^{\star\star}_{\circ\circ\circ}$
	Net OOS SR	0.393	0.593	0.945	1.427	1.371	$\mathbf{1.573}^{\star\star\star}_{\circ\circ\circ}$
	Average $\hat{\kappa}$	/	/	0	1	0.511	0.471

Notes. This table reports the out-of-sample performance of the six portfolio strategies described in Section 7.3 for the six datasets of monthly excess returns described in Section 7.1. Each estimated portfolio is constructed using a sample size of $T = 240$ monthly observations. The mean-variance portfolios consider a risk-aversion coefficient of $\gamma = 3$. The table reports the annualized out-of-sample certainty-equivalent return (OOS CER) and the annualized out-of-sample Sharpe ratio (OOS SR), and for both criteria we report the gross performance and the performance net of proportional transaction costs of 20 basis points. We also report the average estimated shrinkage intensity $\hat{\kappa}$ over time, except for the EW and RTR portfolios that do not combine the SMV and SGMV portfolios. The stars $\star, \star\star, \star\star\star$ establish that the OOS CER and SR of the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ is larger than that of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ at a confidence level of 10%, 5%, and 1%, respectively. The circles $\circ, \circ\circ, \circ\circ\circ$ convey the same information relative to the EW portfolio. The numbers in bold font identify the best portfolio in terms of OOS CER and SR.

Table 4: Out-of-sample certainty-equivalent returns of shrinkage portfolios in empirical data

		Mean OOS CER		Std dev OOS CER	
		$T = 120$	$T = 240$	$T = 120$	$T = 240$
Panel A: Without transaction costs (in %)					
10MOM	$\hat{\kappa}_E^\star$	1.027	1.205	2.444	2.631
	$\hat{\kappa}_R^\star$	1.138	1.312	1.576	2.031
25SBTM	$\hat{\kappa}_E^\star$	1.074	1.230	1.555	1.739
	$\hat{\kappa}_R^\star$	1.028	1.255	1.071	1.051
25OPINV	$\hat{\kappa}_E^\star$	0.613	0.945	1.315	1.198
	$\hat{\kappa}_R^\star$	0.831	1.118	0.997	0.888
49IND	$\hat{\kappa}_E^\star$	0.361	0.252	1.178	0.927
	$\hat{\kappa}_R^\star$	0.522	0.425	0.841	0.786
16LTANOM	$\hat{\kappa}_E^\star$	0.813	1.448	1.497	2.458
	$\hat{\kappa}_R^\star$	0.925	1.443	1.12	1.722
46ANOM	$\hat{\kappa}_E^\star$	7.182	2.709	6.423	4.821
	$\hat{\kappa}_R^\star$	6.509	4.287	5.646	5.534
Panel B: Net of transaction costs (in %)					
10MOM	$\hat{\kappa}_E^\star$	0.743	1.037	2.422	2.628
	$\hat{\kappa}_R^\star$	1.002	1.201	1.565	2.040
25SBTM	$\hat{\kappa}_E^\star$	0.515	0.861	1.362	1.769
	$\hat{\kappa}_R^\star$	0.819	0.977	1.012	1.002
25OPINV	$\hat{\kappa}_E^\star$	0.260	0.737	1.339	1.188
	$\hat{\kappa}_R^\star$	0.666	0.986	1.001	0.878
49IND	$\hat{\kappa}_E^\star$	0.098	0.113	1.170	0.943
	$\hat{\kappa}_R^\star$	0.423	0.344	0.834	0.807
16LTANOM	$\hat{\kappa}_E^\star$	0.390	1.162	1.478	2.438
	$\hat{\kappa}_R^\star$	0.740	1.303	1.109	1.730
46ANOM	$\hat{\kappa}_E^\star$	6.645	1.597	6.207	4.996
	$\hat{\kappa}_R^\star$	6.064	3.762	5.437	5.364

Notes. This table reports the out-of-sample certainty-equivalent return (OOS CER) mean and standard deviation. We consider the estimated shrinkage portfolio that maximizes out-of-sample utility mean ($\hat{\kappa}_E^\star$), and the estimated shrinkage portfolio that maximizes the portfolio robustness measure in Section 5 ($\hat{\kappa}_R^\star$). We report the performance results of the shrinkage portfolios for the six datasets of monthly excess returns described in Section 7.1. We estimate the shrinkage portfolios with sample sizes of $T = 120$ and $T = 240$ monthly observations, a risk-aversion coefficient of $\gamma = 3$, and a coefficient $\lambda = 2$ for the shrinkage portfolio that maximizes the proposed robustness measure. We obtain the performance measures applying Equations (27)-(29) to the OOS CER's of the shrinkage portfolios obtained by dividing the out-of-sample portfolio returns into non-overlapping three-year windows and computing for each three-year window the OOS CER of the shrinkage portfolios. Panel A considers the case without transaction costs, and Panel B considers the case with proportional transaction costs of 20 basis points.

References

Ao, M., Li, Y., Zheng, X., 2019. Approaching mean-variance efficiency for large portfolios. The Review of Financial Studies 32, 2890–2919.

Barroso, P., Saxena, K., 2021. Lest we forget: Using out-of-sample forecast errors in portfolio optimization. The Review of Financial Studies 35, 1222–1278.

Bodnar, T., Okhrin, Y., Parolya, N., 2023. Optimal shrinkage-based portfolio selection in high dimensions. Journal of Business & Economic Statistics 41, 140–156.

Branger, N., Lučivjanská, K., Weissensteiner, A., 2019. Optimal granularity for portfolio choice. Journal of Empirical Finance 50, 125–146.

Brown, S., 1976. Optimal Portfolio Choice under Uncertainty. Ph.D. Thesis, University of Chicago.

Bryzgalova, S., Pelger, M., Zhu, J., 2021. Forest through the trees: Building cross-sections of stock returns. SSRN working paper.

Chopra, V. K., Ziemba, W. T., 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. The Journal of Portfolio Management 19, 6–11.

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? The Review of Financial Studies 22, 1915–1953.

DeMiguel, V., Martín-Utrera, A., Nogales, F., 2013. Size matters: Optimal calibration of shrinkage estimators for portfolio selection. Journal of Banking and Finance 37, 3018–3034.

DeMiguel, V., Martín-Utrera, A., Nogales, F., 2015. Parameter uncertainty in multiperiod portfolio optimization with transaction costs. Journal of Financial and Quantitative Analysis 50, 1443–1471.

Efron, B., 1979. Bootstrap methods: Another look at the jackknife. The Annals of Statistics 7, 1–26.

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

Frost, P., Savarino, J., 1986. An empirical bayes approach to efficient portfolio selection. Journal of Financial and Quantitative Analysis 21, 293–305.

Füss, R., Koeppel, C., Miebs, F., 2021. Diversifying estimation errors: An efficient averaging rule for portfolio optimization. University of St.Gallen, School of Finance Research Paper 2021/05.

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

Giglio, S., Kelly, B., Xiu, D., 2022. Factor models, machine learning, and asset pricing. Annual Review of Financial Economics 14, 337–368.

Goldfarb, D., Iyengar, G., 2003. Robust portfolio selection problems. Mathematics of Operations Research 28, 1–38.

Harvey, C., Liu, Y., 2014. Evaluating trading strategies. The Journal of Portfolio Management 40, 108–118.

Harvey, C., Liu, Y., Zhu, H., 2016. ... and the cross-section of expected returns. The Review of Financial Studies 29, 5–68.

Jensen, T. I., Kelly, B. T., Pedersen, L. H., 2021. Is there a replication crisis in finance? Tech. rep., National Bureau of Economic Research.

Jorion, P., 1986. Bayes-stein estimation for portfolio analysis. Journal of Financial and Quantitative Analysis 21, 279–292.

Kan, R., Smith, D., 2008. The distribution of the sample minimum-variance frontier. Management Science 54, 1364–1380.

Kan, R., Wang, X., 2023. Optimal portfolio choice with unknown benchmark efficiency. Management Science. *Forthcoming.*

Kan, R., Wang, X., Zheng, X., 2022a. In-sample and out-of-sample Sharpe ratios of multi-factor asset pricing models. SSRN working paper.

Kan, R., Wang, X., Zhou, G., 2022b. Optimal portfolio choice with estimation risk: No risk-free asset case. Management Science 68, 2047–2068.

Kan, R., Zhou, G., 2007. Optimal portfolio choice with parameter uncertainty. Journal of Financial and Quantitative Analysis 42, 621–656.

Kirby, C., Ostdiek, B., 2012. It's all in the timing: Simple active portfolio strategies that outperform naive diversification. Journal of Financial and Quantitative Analysis 47, 437–467.

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

Kozak, S., Nagel, S., Santosh, S., 2018. Interpreting factor models. The Journal of Finance 73, 1183–1223.

Lassance, N., Vanderveken, R., Vrins, F., 2023. On the combination of naive and mean-variance portfolio strategies. SSRN working paper.

Ledoit, O., Wolf, M., 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. Journal of Empirical Finance 10, 603–621.

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

Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the sharpe ratio. Journal of Empirical Finance 15, 850–859.

Ledoit, O., Wolf, M., 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. The Review of Financial Studies 30, 4349–4388.

Ledoit, O., Wolf, M., 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. The Annals of Statistics 48, 3043–3065.

Lo, A. W., 2002. The statistics of Sharpe ratios. Financial Analysts Journal 58, 36–52.

Markowitz, H., 1952. Portfolio selection. The Journal of Finance 7, 77–91.

Merton, R. C., 1980. On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8, 323–361.

Novy-Marx, R., Velikov, M., 2016. A taxonomy of anomalies and their trading costs. The Review of Financial Studies 29, 104–147.

Politis, D., Romano, J., 1994. The stationary bootstrap. Journal of the American Statistical Association 89, 1303–1313.

Tu, J., Zhou, G., 2004. Data-generating process uncertainty: What difference does it make in portfolio decisions? Journal of Financial Economics 72, 385–421.

Tu, J., Zhou, G., 2011. Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. Journal of Financial Economics 99, 204–215.

Zhou, G., 2008. On the fundamental law of active portfolio management: What happens if our estimates are wrong? The Journal of Portfolio Management 34, 26–33.

Internet Appendix to

The Risk of Expected Utility under Parameter Uncertainty*

Nathan Lassance Alberto Martín-Utrera Majeed Simaan

August 18, 2023

*Lassance, LFIN/LIDAM, UCLouvain, corresponding author, e-mail: nathan.lassance@uclouvain.be; Martín-Utrera, Iowa State University, e-mail: amutrera@iastate.edu; Simaan, Stevens Institute of Technology, e-mail: msimaan@stevens.edu. Previous versions of this manuscript were circulated under the titles *"A Robust Approach to Optimal Portfolio Choice with Parameter Uncertainty"* and *"The Risk of Out-of-Sample Portfolio Performance"*. We thank Kay Giesecke (the Editor), an anonymous Associate Editor, as well as three anonymous referees for their valuable feedback. We also thank Vitali Alekseev, Stephen Brown, German Creamer, Victor DeMiguel, Kristoffer Glover, Raymond Kan, Ralph Koijen, Petter Kolm, Allen Li, Yan Liu, Yueliang Lu, Jean Pauphilet, Harry Scheule, Frédéric Vrins, Xiaolu Wang, Alex Weissensteiner, Guofu Zhou, as well as seminar and conference participants at London Business School, NYU Courant, Stevens Institute of Technology, UCLouvain, University of Technology Sydney, 15th International Conference on Computational and Financial Econometrics, Financial Management Association Annual Meeting, INFORMS Annual Meeting, Northern Finance Association Annual Meeting, and Silicon Prairie Finance Conference for their comments. Lassance gratefully acknowledges financial support by the Fonds de la Recherche Scientifique [grant number J.0115.22].

In this Internet Appendix, we first review in Section IA.1 the analytical results of Kan et al. (2022b), who characterize the OOSU mean of sample portfolios. We then provide several extensions of our analysis and the proofs for our theoretical results. In Section IA.2, we provide the link between our theory and cross-sectional asset pricing. In Section IA.3, we show the theoretical relation between the shrinkage portfolio approach and ambiguity-averse portfolios. In Section IA.4, we relate our robust shrinkage portfolio with robust optimization. In Section IA.5, we derive the asymptotic distribution of out-of-sample utility of shrinkage portfolios. In Section IA.6, we give details for the feasible estimators we use in the empirical analysis. In Section IA.7, we formally show the relation between the first principal component and the equally weighted portfolio. In Section IA.8, we derive the out-of-sample utility risk of shrinkage portfolios that utilize a linear shrinkage covariance matrix. In Section IA.9, we check the robustness of our main insights to several variations of the experimental setting considered in the main body of the manuscript. Finally, In Section IA.10, we report the proofs for all the theoretical results in the manuscript and this appendix.

IA.1 Out-of-sample utility mean

In the following proposition, we review some of the main results of Kan et al. (2022b) concerning the OOSU mean of the shrinkage portfolio $\hat{w}^\star(\kappa)$ in Equation (11).

Proposition IA.1 (Kan et al. (2022b))**.** *Let Assumptions 1 and 2 hold. Then,*

1. *The out-of-sample utility mean of the sample GMV portfolio is*

$$\mathbb{E}[U(\hat{w}_g)] = \mu_g - \frac{\gamma}{2}\frac{T-2}{T-N-1}\sigma_g^2. \tag{IA1}$$

2. *The out-of-sample utility mean of the shrinkage portfolio $\hat{w}^\star(\kappa)$ is*

$$\begin{aligned}\mathbb{E}[U(\hat{w}^\star(\kappa))] &= \mathbb{E}[U(\hat{w}_g)] \\ &+ \frac{1}{\gamma}\frac{T}{T-N-1}\left(\kappa\psi^2 - \kappa^2\left(\psi^2 + \frac{N-1}{T}\right)\frac{T(T-2)}{2(T-N)(T-N-3)}\right).\end{aligned} \tag{IA2}$$

3. *The shrinkage intensity $\kappa_E^\star$ maximizing out-of-sample utility mean in (IA2) is*

$$\kappa_E^\star = \frac{(T-N)(T-N-3)}{T(T-2)}\frac{\psi^2}{\psi^2 + \frac{N-1}{T}} \in [0,1]. \tag{IA3}$$

Finally, $\kappa_E^\star \to 1$ *as* $T \to \infty$*. Thus,* $\hat{w}^\star(\kappa_E^\star)$ *is a consistent estimator of* $w^\star$*.*

Note that the optimal shrinkage intensity $\kappa_E^\star$ in Proposition IA.1 is an oracle estimator that depends on the unknown parameter ψ^2. Kan et al. (2022b) rely on a feasible estimator of $\kappa_E^\star$ using the estimator of ψ^2 proposed by Kan and Zhou (2007) and find that the resulting portfolio delivers better out-of-sample performance than several benchmark portfolios.

IA.2 Relation with cross-sectional asset pricing

This section shows that our characterization of out-of-sample utility (OOSU) risk can be applied to assess the uncertainty of the out-of-sample fit of a particular robust stochastic discount factor (SDF). Specifically, we exploit the well-known link between the SDF and the returns of a mean-variance portfolio (Cochrane, 2005; Kozak, Nagel, and Santosh, 2020), where the estimated mean-variance portfolio $\hat{w}$ corresponds to the SDF loadings. Accordingly, we show that our measure of out-of-sample performance defined in Equation (12) is equivalent to the out-of-sample fit of an SDF model that prices the cross-section of stock returns. We define the OOS R-squared of an SDF model as

$$\hat{R}^2_{OOS} = 1 - \frac{(\mu - \boldsymbol{\Sigma}\hat{w})^\top \boldsymbol{\Sigma}^{-1}(\mu - \boldsymbol{\Sigma}\hat{w})}{\theta^2}, \tag{IA4}$$

where $\theta^2 = \mu^\top \boldsymbol{\Sigma}^{-1}\mu$. Like Kozak et al. (2020), our measure of OOS fit determines how well the SDF model, defined by the SDF loadings $\hat{w}$, explains the vector of out-of-sample mean returns, μ. Now, consider the following scaled OOS R-squared

$$\frac{\theta^2}{2}\hat{R}^2_{OOS} = \hat{w}^\top \mu - \frac{1}{2}\hat{w}^\top \boldsymbol{\Sigma}\hat{w}, \tag{IA5}$$

which corresponds to the OOSU of the estimated portfolio $\hat{w}$ for an investor with risk-aversion coefficient $\gamma = 1$. Like the OOSU, the scaled $\hat{R}^2_{OOS}$ is a random variable because it depends on the estimated SDF loadings $\hat{w}$.

The shrinkage mean-variance portfolio we consider in the main body of the manuscript embeds as a particular case the robust SDF loadings that solve the following constrained quadratic program:

$$\hat{w} = \underset{w:\, w^\top e = 1}{\operatorname{argmin}} \quad (\mu - \widehat{\boldsymbol{\Sigma}} w)^\top \widehat{\boldsymbol{\Sigma}}^{-1}(\mu - \widehat{\boldsymbol{\Sigma}} w) \tag{IA6}$$

$$\text{subject to} \quad \min_{\mu} \quad (\hat{\mu}-\mu)^{\top}\widehat{\boldsymbol{\Sigma}}^{-1}(\hat{\mu}-\mu) \leq \varepsilon^{2}. \tag{IA7}$$

The solution to this mathematical program delivers the ambiguity-averse mean-variance portfolio problem in Section IA.3 for an investor with risk-aversion coefficient $\gamma = 1$. Due to the link between constraint (IA7) and the shrinkage intensity κ, which we highlight in Proposition IA.2, the shrinkage portfolios considered in the main body of the manuscript can be interpreted as the robust SDF loadings that minimize the Hansen-Jagannathan distance (Hansen and Jagannathan, 1997) under the economically motivated constraint faced by ambiguity-averse investors.

Given the link between the economically motivated SDF loadings in problem (IA6) and shrinkage portfolios, we can apply our measure of OOSU risk to evaluate the risk of the OOS fit of robust SDF models. Our theoretical results in Section 4 demonstrate that the OOS R-squared risk is large in high-dimensional settings and when the estimated SDF loadings of the model exploit near-arbitrage opportunities. Accordingly, SDF models constructed from many test assets with short time series and that capture near-arbitrage opportunities are, in general, unreliable.

IA.3 Microeconomic foundation of robust portfolios

In this section, we provide a microeconomic foundation of the robust shrinkage portfolio introduced in Section 5.2. In particular, we show that there is an explicit relation between the robust shrinkage portfolio and the optimal portfolio of an ambiguity-averse investor. The insights provided in this section build on the work of Garlappi et al. (2007), who account for ambiguity by considering a joint uncertainty set for the vector of means. This uncertainty set serves as a constraint in the portfolio problem of a mean-variance investor who solves the following mathematical program:

$$\max_{w:w^{\top}e=1} \min_{\mu} \; w^{\top}\mu - \frac{\gamma}{2}w^{\top}\widehat{\boldsymbol{\Sigma}}w \quad \text{subject to } (\hat{\mu}-\mu)^{\top}\widehat{\boldsymbol{\Sigma}}^{-1}(\hat{\mu}-\mu) \leq \varepsilon^{2}, \tag{IA8}$$

where $(\hat{\mu}-\mu)^{\top}\widehat{\boldsymbol{\Sigma}}^{-1}(\hat{\mu}-\mu)$ measures the distance between the sample vector of means $\hat{\mu}$ and the true vector of means μ. Intuitively, a larger degree of ambiguity aversion is equivalent to having a larger value of ε in the constraint of portfolio problem (IA8). Garlappi et al. (2007) show that the closed-form solution to this problem is

$$\hat{w}^{\star}(\varepsilon) = \frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\left(\frac{1}{1+\varepsilon/(\gamma\sigma_{P}^{\star})}\right)\left(\hat{\mu} - \frac{B-\gamma(1+\varepsilon/(\gamma\sigma_{P}^{\star}))}{A}e\right), \tag{IA9}$$

where $A = e^\top \widehat{\boldsymbol{\Sigma}}^{-1} e$, $B = \hat{\mu}^\top \widehat{\boldsymbol{\Sigma}}^{-1} e$, and $\sigma_P^\star$ is the unique positive real root to a specific fourth-degree polynomial and it is monotonically decreasing in ε. The following proposition provides the link between any shrinkage portfolio that combines the SMV and SGMV portfolios and the optimal portfolio of the ambiguity-averse investor solving problem (IA8).

Proposition IA.2. *The shrinkage portfolio $\hat{w}^\star(\kappa)$ is equal to the ambiguity-averse portfolio $\hat{w}^\star(\varepsilon)$ in Equation (IA9) when*

$$\kappa = \left(1 + \frac{\varepsilon}{\gamma \sigma_P^\star}\right)^{-1}, \tag{IA10}$$

which is monotonically decreasing in ε.

Proposition IA.2 uncovers the explicit relation between the shrinkage portfolio considered in the main body of the manuscript and the ambiguity-averse portfolio of Garlappi et al. (2007). In particular, Equation (IA10) shows that a high degree of ambiguity in mean returns (i.e., a higher ε) results in an ambiguity-averse portfolio whose weights lean more strongly toward those of the SGMV portfolio, corresponding to a smaller value of κ. As shown in Proposition 4, our robustness criterion establishes a larger tilt toward the SGMV portfolio than the traditional shrinkage criterion that only maximizes OOSU mean. Therefore, the result in Proposition IA.2 indicates that our robust shrinkage portfolio is equivalent to an ambiguity-averse portfolio for an investor with a more considerable degree of ambiguity in mean returns than that implied by an investor that maximizes only OOSU mean.[1]

IA.4 Relation with robust optimization

In this section, we interpret our proposed mean-risk OOSU criterion through the lenses of robust optimization.[2] Under this approach, the mean-variance investor is averse to the *ambiguity* around the true, *but unknown*, OOSU mean and maximizes the worst-case scenario assuming that the true OOSU mean lies within a bounded region. In particular, we assume that the true OOSU mean belongs to the following uncertainty set:

$$\mathcal{S}(\lambda, \kappa) = \left\{ x \in \left[\widehat{\mathbb{E}}[U(\hat{w}^\star(\kappa))] - \lambda \sqrt{\widehat{\mathbb{V}}[U(\hat{w}^\star(\kappa))]}, \widehat{\mathbb{E}}[U(\hat{w}^\star(\kappa))] + \lambda \sqrt{\widehat{\mathbb{V}}[U(\hat{w}^\star(\kappa))]} \right] \right\}, \tag{IA11}$$

[1] For example, for the 25SBTM dataset considered in Figure 1, we find that the shrinkage intensity $\kappa_E^\star = 0.147$ maximizing OOSU mean corresponds to $\varepsilon = 0.77$, whereas the shrinkage intensity $\kappa_R^\star = 0.0886$ maximizing our proposed robustness measure for $\lambda = 2$ corresponds to $\varepsilon = 1.36$, which is larger.

[2] There is extensive literature on robust optimization and portfolio theory. One of the most prominent papers in this literature is Goldfarb and Iyengar (2003).

where $\widehat{\mathbb{E}}[U(\hat{w}^\star(\kappa))]$ and $\widehat{\mathbb{V}}[U(\hat{w}^\star(\kappa))]$ are the estimated OOSU mean and variance of the shrinkage portfolio $\hat{w}^\star(\kappa)$. Note that in this case parameter $\lambda \geq 0$ determines the level of uncertainty around the OOSU mean. The uncertainty set in (IA11) can be interpreted as a confidence interval similar to Garlappi et al. (2007). Accordingly, an ambiguity-averse investor who wants to maximize OOSU mean solves the robust optimization problem

$$\max_{\kappa\in[0,1]} \min_{\mathcal{S}(\lambda,\kappa)} \mathbb{E}[U(\hat{w}^\star(\kappa))] = \max_{\kappa\in[0,1]} \widehat{\mathbb{E}}[U(\hat{w}^\star(\kappa))] - \lambda\sqrt{\widehat{\mathbb{V}}[U(\hat{w}^\star(\kappa))]}. \tag{IA12}$$

Problem (IA12) delivers our estimated robust shrinkage intensity, i.e., $\hat{\kappa}_R^\star$. Therefore, our methodology implicitly accounts for the estimation errors in OOSU mean, which can contaminate the estimated shrinkage intensity that maximizes OOSU mean and deteriorate portfolio performance (Kan and Wang, 2023). We confirm this finding in the simulation analysis of Section 7.2 in the main body of the manuscript, where we find that our robust shrinkage portfolio often delivers a larger OOSU mean than that of the shrinkage portfolio that is designed to maximize OOSU mean.

IA.5 Asymptotic distribution of out-of-sample utility

In this section, we characterize the asymptotic distribution of the OOSU for the shrinkage portfolio defined in Equation (11). The asymptotic distribution is derived as both the sample size T and the number of assets N go to infinity but their ratio converges to a constant, similar to the analysis in Ledoit and Wolf (2017), Ao et al. (2019), and Kan et al. (2022a). This result contains, as a particular case, the asymptotic distribution of the OOSU for the SMV and the SGMV portfolios when the shrinkage intensity is one and zero, respectively.

Proposition IA.3. *Let $N, T \to \infty$, $N/T \to \rho \in [0, 1)$, and Assumption 2 holds. Then, the out-of-sample utility of the shrinkage portfolio $\hat{w}^\star(\kappa)$ is asymptotically Gaussian,*

$$\sqrt{T}(U(\hat{w}^\star(\kappa)) - u(\kappa, \rho)) \xrightarrow{d} \mathcal{N}(0, v(\kappa, \rho)), \tag{IA13}$$

where the asymptotic mean is

$$u(\kappa, \rho) = \mu_g - \frac{\gamma}{2}\frac{\sigma_g^2}{1-\rho} + \frac{1}{\gamma}\frac{1}{1-\rho}\left(\kappa\psi^2 - \frac{\kappa^2}{2}\frac{\psi^2+\rho}{(1-\rho)^2}\right),$$

and the asymptotic variance is

$$v(\kappa,\rho) = v(0,\rho) + a_1(\rho)\kappa^4 + a_2(\rho)\kappa^3 + a_3(\rho)\kappa^2 + a_4(\rho)\kappa$$

with

$$\begin{aligned}
v(0,\rho) &= \frac{\sigma_g^2\psi^2}{1-\rho} + \frac{\gamma^2\sigma_g^4\rho}{2(1-\rho)^3},\\
a_1(\rho) &= \frac{(\psi^2+\rho/2)(\rho^2+3\rho+1)+\psi^4(2+\rho/2)}{\gamma^2(1-\rho)^7},\\
a_2(\rho) &= -\frac{2\psi^2(2\psi^2+\rho+1)}{\gamma^2(1-\rho)^5},\\
a_3(\rho) &= \frac{\psi^2(4\psi^2+1)}{\gamma^2(1-\rho)^3} + \frac{\sigma_g^2(1+\rho)(\psi^2+\rho)}{(1-\rho)^5},\\
a_4(\rho) &= -\frac{2\sigma_g^2\psi^2}{(1-\rho)^3}.
\end{aligned}$$

Proposition IA.3 shows that the OOSU of sample portfolios follows asymptotically a Gaussian distribution. One of the important implications of this result is that the entire distribution can be defined with the first two moments. Consequently, as N and T increase with $N/T = \rho \in [0,1)$, the α-Value-at-Risk (e.g., $\alpha = 5\%$) of OOSU can be defined as

$$u(\kappa,\rho) - \lambda_{1-\alpha}\sqrt{v(\kappa,\rho)/T}, \tag{IA14}$$

where $\lambda_{1-\alpha}$ is the $(1-\alpha)$th percentile of a standard normal distribution. This corresponds to the portfolio robustness measure we propose in Section 5.

IA.6 Feasible estimators of shrinkage intensities

The shrinkage intensities $\kappa_E^\star$ and $\kappa_R^\star$ in Equations (IA3) and (21) are unfeasible because they depend on the unknown distributional parameters of stock returns. In this section, we provide the details of the feasible estimators we use in the performance analysis in Section 7.

For the return variance of the zero-cost portfolio, which is defined in (7) as ψ^2, we use the adjusted estimator proposed by Kan and Zhou (2007). Let $\hat{\psi}^2 = \hat{\mu}^\top\widehat{\mathbf{B}}\hat{\mu}$ be the plug-in estimator,

then we estimate ψ^2 as

$$\hat{\psi}_{kz}^2 = \frac{(T-N-1)\hat{\psi}^2 - (N-1)}{T} + \frac{2(\hat{\psi}^2)^{\frac{N-1}{2}}(1+\hat{\psi}^2)^{-\frac{T-2}{2}}}{T \times B\left(\frac{\hat{\psi}^2}{1+\hat{\psi}^2}; \frac{N-1}{2}, \frac{T-N+1}{2}\right)}, \tag{IA15}$$

where $B(x;a,b) = \int_0^x y^{a-1}(1-y)^{b-1}\mathrm{d}y$ is the incomplete beta function.

For the return variance of the GMV portfolio, which is defined as σ_g^2 in (6), we rely on the shrinkage portfolio estimator proposed by Frahm and Memmel (2010, Theorem 2), which gives a smaller mean out-of-sample variance than the SGMV portfolio. Specifically, we estimate σ_g^2 as

$$\hat{\sigma}_g^2 = \hat{w}_{fm}^\top \widehat{\mathbf{\Sigma}} \hat{w}_{fm}, \tag{IA16}$$

where $\hat{w}_{fm}$ combines the equally weighted portfolio and the SGMV portfolio as

$$\hat{w}_{fm} = \hat{\pi}_{fm} w_{ew} + (1 - \hat{\pi}_{fm})\hat{w}_g,$$

with a shrinkage intensity

$$\hat{\pi}_{fm} = \min\left(1, \frac{N-3}{T-N+2} \frac{\hat{w}_g^\top \widehat{\mathbf{\Sigma}} \hat{w}_g}{w_{ew}^\top \widehat{\mathbf{\Sigma}} w_{ew} - \hat{w}_g^\top \widehat{\mathbf{\Sigma}} \hat{w}_g}\right).$$

IA.7 Relation between first PC and equally weighted portfolio

In Assumption 3 in the main body of the manuscript, we assume that the first eigenvector is proportional to the equally weighted portfolio. In this section, we formally characterize the correlation between the first principal component of stock returns and the equally weighted portfolio return.

Proposition IA.4. *Let the vector of excess returns at time t, r_t, follow a one-factor model of the form*

$$r_{it} = \beta_i f_t + \varepsilon_{it}, \ \ i = 1, \ldots, N, \tag{IA17}$$

where $(\varepsilon_{it})_{1 \le i \le N}$ are i.i.d. and ε_{it} is uncorrelated with f_t for all i. Then, the correlation between

the first principal component and the equally weighted portfolio return is

$$\mathbb{C}\text{orr}(v_1^\top r_t, w_{ew}^\top r_t) = \frac{\sigma_f^2 \bar{\beta}_1 \bar{\beta}_2 + \sigma_\varepsilon^2 \bar{\beta}_1 / N}{\sqrt{\left(\sigma_f^2 \bar{\beta}_2^2 + \sigma_\varepsilon^2 \bar{\beta}_2 / N\right)\left(\sigma_f^2 \bar{\beta}_1^2 + \sigma_\varepsilon^2 / N\right)}}, \tag{IA18}$$

where $(\sigma_f^2, \sigma_\varepsilon^2)$ *are the variances of* (f_t, ε_{it}), $\bar{\beta}_1 = \frac{1}{N}\sum_{i=1}^N \beta_i$, *and* $\bar{\beta}_2 = \frac{1}{N}\sum_{i=1}^N \beta_i^2$. *Moreover, this correlation converges to one when i)* $\sigma_\varepsilon^2 \to 0$ *while* $\bar{\beta}_1 > 0$, *or ii)* $N \to \infty$ *while* $0 < \lim_{N\to\infty} \bar{\beta}_1 < \infty$ *and* $0 < \lim_{N\to\infty} \bar{\beta}_2 < \infty$.

IA.8 Out-of-sample utility risk with shrinkage covariance matrix

In this section, we develop analytically the OOSU risk of a shrinkage portfolio that exploits a linear shrinkage covariance matrix like the one in Equation (24), under the assumption that only the vector of mean returns is affected by parameter uncertainty, which has been identified as the main source of estimation risk in portfolio selection (Jagannathan and Ma, 2003; Kozak et al., 2020).[3] We then use this novel result to theoretically study the impact of shrinking the covariance matrix on the OOSU risk of shrinkage portfolios.

Even though we assume a known covariance matrix in the analysis presented in this section, we recognize that it is still useful to shrink the covariance matrix toward the identity matrix to improve portfolio performance. This is because by doing so, one can attenuate the impact that estimation errors in the vector of means have on the portfolio's out-of-sample performance. Accordingly, we define the shrinkage portfolio strategy as

$$\hat{w}^\star(\delta, \kappa) = (1-\kappa) w_{sg} + \kappa \hat{w}^\star = w_{sg} + \frac{\kappa}{\gamma} \hat{w}_{sz}, \tag{IA19}$$

where w_{sg} and $\hat{w}_{sz}$ are the shrinkage GMV and zero-cost portfolios,

$$w_{sg} = \frac{\mathbf{\Sigma}_{sh}^{-1} e}{e^\top \mathbf{\Sigma}_{sh}^{-1} e}, \tag{IA20}$$

[3] Our derivation of the OOSU risk in Proposition 1 relies on the stochastic representation of the out-of-sample mean return and variance derived by Kan et al. (2022b). This stochastic representation relies on the Wishart distribution of the sample covariance matrix. However, the linear shrinkage covariance matrix we use in this paper no longer follows a Wishart distribution. Thus, analytical expressions for the OOSU variance of sample portfolios that exploit a shrinkage covariance matrix are unavailable. To overcome this challenge, we assume that the covariance matrix is known, which is not a strong assumption given that 1) the critical source of parameter uncertainty in portfolio selection stems from the vector of mean returns, as stated in the main text, and 2) the shrinkage covariance matrix substantially reduces the estimator's mean-squared error.

$$\hat{w}_{sz} = \mathbf{B}_{sh}\hat{\mu}, \tag{IA21}$$

with

$$\mathbf{\Sigma}_{sh} = (1-\delta)\mathbf{\Sigma} + \delta\bar{\sigma}^2\mathbf{I}, \tag{IA22}$$

$$\mathbf{B}_{sh} = \mathbf{\Sigma}_{sh}^{-1}(\mathbf{I} - ew_{sg}^{\top}), \tag{IA23}$$

and $\bar{\sigma}^2 = \text{Tr}(\mathbf{\Sigma})/N$ corresponds to the cross-sectional average of asset return variances. In the next proposition, we derive the OOSU variance of $\hat{w}^{\star}(\delta,\kappa)$. For that purpose, we define

$$\mathbf{S} = \mathbf{B}_{sh}\mathbf{\Sigma}. \tag{IA24}$$

Proposition IA.5. *Assume that the covariance matrix* $\mathbf{\Sigma}$ *is known,* $T > N$*, and Assumption 2 in the main body of the manuscript holds. Then, the out-of-sample utility variance of* $\hat{w}^{\star}(\delta,\kappa)$ *in* (IA19) *is*

$$\mathbb{V}[U(\hat{w}^{\star}(\delta,\kappa))] = \frac{\kappa^2}{T}\Big(a_1(\delta)\kappa^2 + a_2(\delta)\kappa + a_3(\delta)\Big), \tag{IA25}$$

where

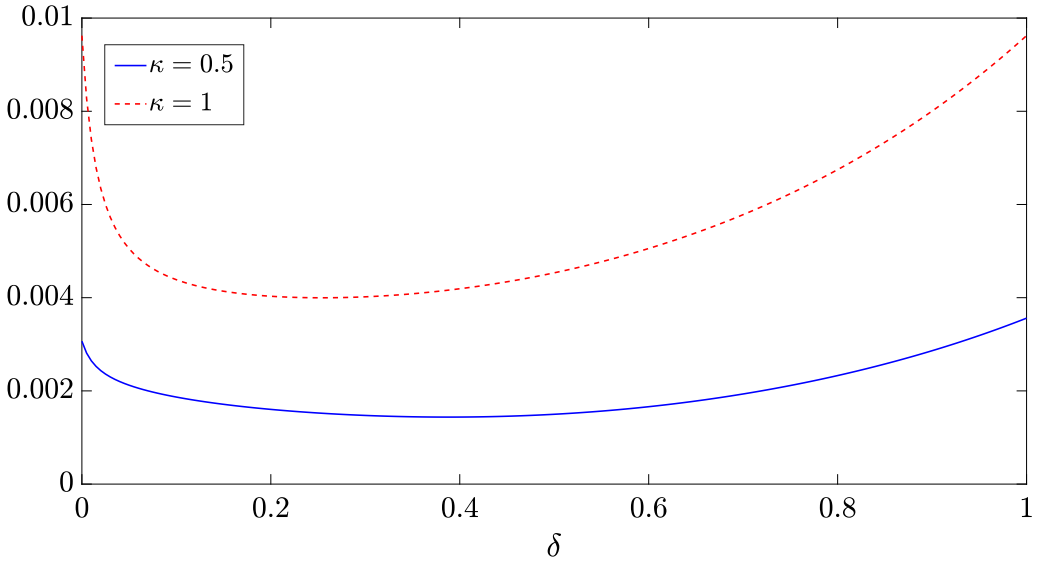
$$a_1(\delta) = \frac{1}{2\gamma^2}\left(\frac{1}{T}\text{Tr}(\mathbf{S}^4) + 2w_{sz}^{\top}\mathbf{\Sigma}\mathbf{S}^2w_{sz}\right), \tag{IA26}$$

$$a_2(\delta) = \frac{2}{\gamma^2}\Big(\gamma w_{sg}^{\top}\mathbf{\Sigma}\mathbf{S}^2w_{sz} - \mu^{\top}\mathbf{S}^2w_{sz}\Big), \tag{IA27}$$

$$a_3(\delta) = \frac{1}{\gamma^2}\mu^{\top}\mathbf{S}w_{sz} - \frac{2}{\gamma}\mu^{\top}\mathbf{S}^2w_{sg} + w_{sg}^{\top}\mathbf{\Sigma}\mathbf{S}^2w_{sg}. \tag{IA28}$$

Finally, we illustrate in Figure IA.1 how shrinking the covariance matrix affects the OOSU risk of the shrinkage portfolio. We define the population vector of means μ and covariance matrix $\mathbf{\Sigma}$ with the sample moments of the 25SBTM dataset, and we set a sample size of $T = 120$ observations, with $N = 25$ assets, a risk-aversion coefficient of $\gamma = 3$, and shrinkage intensities for the portfolio weights of $\kappa = 0.5$ and 1. We depict the OOSU standard deviation of $\hat{w}^{\star}(\delta,\kappa)$, defined in Proposition IA.5, for different values of the shrinkage intensity δ. The figure shows that shrinking the covariance matrix helps decrease OOSU risk, particularly when κ is large, which is consistent with the simulation results in Figures 4 and 5 in the main body of the manuscript.

Figure IA.1: Impact of shrinking the covariance matrix on out-of-sample utility risk



Notes. This figure depicts the out-of-sample utility standard deviation of the shrinkage portfolio $w^\star(\delta, \kappa)$ defined in Proposition IA.5. We define the population vector of means μ and covariance matrix $\mathbf{\Sigma}$ with the sample moments of the monthly return data of the 25 portfolios of stocks sorted on size and book-to-market, and we set a sample size of $T = 120$ observations, with $N = 25$ assets, a risk-aversion coefficient of $\gamma = 3$, and $\kappa = 0.5$ and 1. We depict the out-of-sample utility standard deviation as a function of the shrinkage intensity δ applied to the covariance matrix.

IA.9 Robustness checks of empirical analysis

We now assess the robustness of our main insight that OOSU risk is an important element to consider in the construction of robust investment strategies. In particular, we produce a number of additional empirical and theoretical results that expand the main ideas explored in the main body of the manuscript. Specifically, we check the robustness of our results to considering: (i) the tail risk of the portfolios, (ii) the cumulative wealth derived from the portfolios, (iii) different risk-aversion coefficients, (iv) a higher level of transaction costs, (v) a different combination of portfolios that exploits the SMV portfolio, the SGMV portfolio, and the equally weighted portfolio as in Tu and Zhou (2011), (vi) the robust shrinkage portfolio with a cross-validated λ, (vii) shrinkage portfolios that use the linear shrinkage covariance matrix of Ledoit and Wolf (2004), (viii) shrinkage portfolios that use the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020), (ix) a high-dimensional set of assets, and (x) the impact of estimation error in the shrinkage intensities.

IA.9.1 Value-at-Risk

Table IA.1 shows that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ delivers lower Value-at-Risk than that of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$, which is an additional empirical feature offered by our proposed robust shrinkage portfolio.

IA.9.2 Cumulative wealth

To gauge the economic magnitude of the outperformance delivered by the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$, Table IA.2 reports the cumulative wealth obtained by investing in the shrinkage portfolios using $\hat{\kappa}_E^\star$ and $\hat{\kappa}_R^\star$.[4] We only consider the overlapping out-of-sample period across the six datasets for comparison purposes. The table shows that the outperformance delivered by the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ is economically significant, translating into a large increase of cumulative wealth relative to the shrinkage portfolio using $\hat{\kappa}_E^\star$. In particular, the cumulative wealth increases by 255% when $T = 120$, and by 75.8% when $T = 240$, on average across the six datasets.

IA.9.3 Different risk-aversion coefficients

Tables IA.3 and IA.4 replicate the results in Tables 2 and 3 for the case where the risk-aversion coefficient is $\gamma = 1$ and 5, respectively. For conciseness, we do not report the performance of the EW portfolio because the RTR portfolio always outperforms it, and we do not report the abysmal performance of the SMV portfolio either.

For the case with $\gamma = 1$ in Table IA.3, we observe that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ delivers a better OOS CER than the RTR portfolio, the SGMV portfolio, and the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$, both before and after transaction costs. We also see that the outperformance net of transaction costs delivered by $\hat{\kappa}_R^\star$ relative to $\hat{\kappa}_E^\star$ is larger than in the case with $\gamma = 3$. For the case with $\gamma = 5$ in Table IA.4, we observe that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ delivers a better OOS CER and OOS Sharpe ratio than the RTR portfolio, the SGMV portfolio, and the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$, both before and after transaction costs. Overall, the results in this section confirm that the insights shown in the main body of the manuscript are robust to considering different degrees of risk aversion.

[4] We standardize their out-of-sample returns to have the same volatility for comparison purposes. We take as target volatility that of the market portfolio over the same out-of-sample period, which we download from Kenneth French's website.

IA.9.4 Higher level of transaction costs

In Table IA.5, we replicate the results in Tables 2 and 3 using a higher level of proportional transaction cost of 30 basis points in Equation (31). For conciseness, we do not report the performance of the EW portfolio because the RTR portfolio always outperforms it, and we do not report the abysmal performance of the SMV portfolio either.

Our main insights in the main body of the manuscript is robust to considering a higher level of proportional transaction costs. Because the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ is on average more exposed to the SMV portfolio than the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$, the impact of a higher level of transaction costs on the performance of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ is more severe than that of the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$. Comparing the SGMV portfolio and the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$, we find that $\hat{\kappa}_R^\star$ continues to deliver a better performance net of transaction costs in nearly all cases. Finally, the shrinkage portfolio exploiting intensity $\hat{\kappa}_R^\star$ maintains a better net performance than that of the RTR portfolio, except for the *49IND* dataset.

IA.9.5 Combining with the equally weighted portfolio

Motivated by the finding of DeMiguel et al. (2009) that the equally weighted (EW) portfolio often outperforms mean-variance portfolios out of sample, Tu and Zhou (2011) extend the framework introduced by Kan and Zhou (2007) and combine several estimates of the mean-variance portfolio with the EW portfolio. This section follows a similar approach and applies our methodology to the shrinkage portfolio that combines the SMV, SGMV, and EW portfolios.

Let $w_{ew} = e/N$ denote the EW portfolio. Then, the three-fund shrinkage portfolio that combines the SMV, SGMV, and EW portfolios is

$$\hat{w}^\star(\pi, \kappa) = (1-\pi)w_{ew} + \pi\hat{w}^\star(\kappa) = (1-\pi)w_{ew} + \pi((1-\kappa)\hat{w}_g + \kappa\hat{w}^\star), \tag{IA29}$$

where π and $\kappa \in [0, 1]$. In the next proposition, we derive closed-form expressions for the OOSU mean and variance of the shrinkage portfolio $\hat{w}^\star(\pi, \kappa)$. This result allows us to compute the mean-risk OOSU and find the corresponding optimal shrinkage intensities $(\pi_R^\star, \kappa_R^\star)$. For notational simplicity, we introduce the following terms

$$\mu_{ew} = w_{ew}^\top \mu \quad \text{and} \quad \sigma_{ew}^2 = w_{ew}^\top \mathbf{\Sigma} w_{ew}, \tag{IA30}$$

for the mean return and variance of the EW portfolio.

Proposition IA.6. *Let Assumptions 1 and 2 hold. Then,*

1. *The out-of-sample utility mean of the three-fund shrinkage portfolio $\hat{w}^\star(\pi,\kappa)$ is*

$$\begin{aligned}\mathbb{E}[U(\hat{w}^\star(\pi,\kappa))] &= (1-\pi)\mu_{ew} + \pi\bigg(\mu_g + \frac{\kappa}{\gamma}\frac{T}{T-N-1}\psi^2\bigg) - \frac{\gamma}{2}\bigg((1-\pi)^2\sigma_{ew}^2 \\ &+ \pi^2\bigg(\frac{T-2}{T-N-1}\sigma_g^2 + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)(T\psi^2+N-1)}{(T-N)(T-N-1)(T-N-3)}\bigg) \\ &+ 2\pi(1-\pi)\bigg(\sigma_g^2 + \frac{\kappa}{\gamma}\frac{T}{T-N-1}(\mu_{ew}-\mu_g)\bigg)\bigg). \end{aligned} \tag{IA31}$$

2. *The out-of-sample utility variance of the three-fund shrinkage portfolio $\hat{w}^\star(\pi,\kappa)$ is*

$$\begin{aligned}\mathbb{V}[U(\hat{w}^\star(\pi,\kappa))] &= \mathbb{V}\Big[\hat{w}^\star(\pi,\kappa)^\top\mu\Big] + \frac{\gamma^2}{4}\mathbb{V}\Big[\hat{w}^\star(\pi,\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\pi,\kappa)\Big] \\ &\quad - \gamma\,\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\pi,\kappa)^\top\mu, \hat{w}^\star(\pi,\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\pi,\kappa)\Big], \end{aligned} \tag{IA32}$$

where the variance of the out-of-sample mean return is

$$\mathbb{V}\Big[\hat{w}^\star(\pi,\kappa)^\top\mu\Big] = \pi^2\bigg(\frac{\sigma_g^2\psi^2}{T-N-1} + \frac{\kappa^2\psi^2}{\gamma^2}\frac{2T(N+1)+T^2(T-N-3+2(T-N)\psi^2)}{(T-N)(T-N-1)^2(T-N-3)}\bigg), \tag{IA33}$$

and the variance of the out-of-sample return variance is

$$\begin{aligned}\mathbb{V}\Big[\hat{w}^\star(\pi,\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\pi,\kappa)\Big] &= \pi^4\bigg(\frac{2\sigma_g^4(N-1)(T-2)}{(T-N-1)^2(T-N-3)} \\ &+ \frac{4\kappa^2\sigma_g^2}{\gamma^2}\frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} \\ &+ \frac{2\kappa^4}{\gamma^4}\frac{T^2(T-2)C(T,N,\psi^2)}{(T-N)^2(T-N-1)^2(T-N-2)(T-N-3)^2(T-N-5)(T-N-7)}\bigg) \\ &+ 4\pi^2(1-\pi)^2\bigg((\sigma_{ew}^2-\sigma_g^2)\bigg(\frac{\sigma_g^2}{T-N-1} + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)+T^2\psi^2}{(T-N)(T-N-1)(T-N-3)}\bigg) \\ &+ \frac{\kappa^2}{\gamma^2}\frac{T^2(T-N+1)}{(T-N)(T-N-1)^2(T-N-3)}(\mu_{ew}-\mu_g)^2\bigg) + 8\pi^3(1-\pi)\frac{\kappa}{\gamma}(\mu_{ew}-\mu_g) \\ &\bigg(\sigma_g^2\frac{T(T-2)}{(T-N-1)^2(T-N-3)} + \frac{\kappa^2}{\gamma^2}\frac{T^2(T-2)(T+N-3+2T\psi^2)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}\bigg), \end{aligned} \tag{IA34}$$

where $C(T, N, \psi^2)$ is defined in Proposition 1, and the covariance between the out-of-sample mean return and variance is

$$\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\pi,\kappa)^\top\mu, \hat{w}^\star(\pi,\kappa)^\top\boldsymbol{\Sigma}\hat{w}^\star(\pi,\kappa)\Big] = 2\pi^3\frac{\kappa}{\gamma}\Bigg(\sigma_g^2\psi^2\frac{T(T-2)}{(T-N-1)^2(T-N-3)} + \frac{\kappa^2\psi^2}{\gamma^2}\frac{T^2(T-2)(T+N-3+2T\psi^2)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}\Bigg) + 2\pi^2(1-\pi)(\mu_{ew}-\mu_g)\Bigg(\frac{\sigma_g^2}{T-N-1} + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)(T-N-1)+2T^2(T-N)\psi^2}{(T-N)(T-N-1)^2(T-N-3)}\Bigg). \tag{IA35}$$

Using the result in Proposition IA.6, we can find numerically the shrinkage intensities $(\pi_E^\star, \kappa_E^\star)$ maximizing OOSU mean and the shrinkage intensities $(\pi_R^\star, \kappa_R^\star)$ maximizing our proposed mean-risk OOSU robustness metric in (20). Those shrinkage intensities depend on the following distributional parameters of stock returns: μ_g, σ_g^2, ψ^2, μ_{ew}, and σ_{ew}^2. We estimate σ_g^2 and ψ^2 as in Appendix IA.6. We estimate μ_{ew} and σ_{ew}^2 via the plug-in estimators

$$\hat{\mu}_{ew} = w_{ew}^\top\hat{\mu} \quad \text{and} \quad \hat{\sigma}_{ew}^2 = w_{ew}^\top\hat{\boldsymbol{\Sigma}}w_{ew}, \tag{IA36}$$

as in Tu and Zhou (2011). Finally, we estimate μ_g as the mean return of the shrinkage portfolio that combines the EW and SGMV portfolios instead of using the plug-in estimator of μ_g, which is highly contaminated by estimation errors. We select the shrinkage intensity ν of this shrinkage portfolio as the parameter ν that minimizes the mean squared error of the out-of-sample mean return of the portfolio.

Proposition IA.7. *Let Assumptions 1 and 2 hold. Then, the shrinkage portfolio $\hat{w}(\nu) = \nu w_{ew} + (1-\nu)\hat{w}_g$ that minimizes the mean squared error $\mathbb{E}[(\hat{w}(\nu)'\mu - \mu_g)^2]$ is obtained for*

$$\nu = \frac{\sigma_g^2\psi^2}{\sigma_g^2\psi^2 + (\mu_{ew}-\mu_g)(T-N-1)}. \tag{IA37}$$

Using Proposition IA.7, we estimate μ_g as

$$\hat{\mu}_g = \hat{w}(\hat{\nu})^\top\hat{\mu} \quad \text{with} \quad \hat{\nu} = \frac{\hat{\sigma}_g^2\hat{\psi}_{kz}^2}{\hat{\sigma}_g^2\hat{\psi}_{kz}^2 + (\hat{\mu}_{ew}-\hat{w}_g^\top\hat{\mu})(T-N-1)}, \tag{IA38}$$

where $\hat{\psi}_{kz}^2$ is defined in (IA15), $\hat{\sigma}_g^2$ in (IA16), and $\hat{\mu}_{ew}$ in (IA36). Finally, using those estimators of the distributional parameters of stock returns, we can obtain the estimated shrinkage intensities $(\hat{\pi}_E^\star, \hat{\kappa}_E^\star)$ and $(\hat{\pi}_R^\star, \hat{\kappa}_R^\star)$ numerically using the results in Proposition IA.6.

In Tables IA.6, we report the out-of-sample performance of the shrinkage portfolio $\hat{w}^\star(\pi, \kappa)$ in (IA29) that combines the SMV, SGMV, and EW portfolios using the intensities $(\hat{\pi}_E^\star, \hat{\kappa}_E^\star)$ and $(\hat{\pi}_R^\star, \hat{\kappa}_R^\star)$. As in Tables 2 and 3, we estimate the portfolios with a covariance matrix that shrinks the sample covariance matrix toward the identity. The SGMV portfolio uses the shrinkage approach of Ledoit and Wolf (2004) and the two shrinkage portfolios use a shrinkage intensity calibrated via bootstrap (see Footnote 18) to optimize the same criterion used to combine the SGMV and SMV portfolios.

We observe that the results in the manuscript are robust to adding the EW portfolio in the shrinkage portfolios. Specifically, the robust shrinkage portfolio exploiting $(\hat{\pi}_R^\star, \hat{\kappa}_R^\star)$ delivers a greater OOS CER than the shrinkage portfolio maximizing OOSU mean in all cases except one before transaction costs and all cases after transaction costs. It also systematically delivers a greater Sharpe ratio. Similarly, the robust shrinkage portfolio outperforms the SGMV portfolio in most cases, before and after transaction costs.

Overall, these results confirm that our proposed robustness measure can be applied in the construction of other investment strategies and outperform combinations of portfolios that only focus on maximizing OOSU mean as well as the individual portfolios being combined.

IA.9.6 Performance of robust portfolios with cross-validated λ

We now test the robustness of our results to considering a data-driven approach to calibrate λ. Our method relies on cross-validation. In particular, we consider a grid of values for λ and divide each estimation window into five folds. Then, for a value of λ inside the grid, we construct the robust portfolio using four folds and evaluate the out-of-sample performance of the portfolio in the fifth fold that is not used for the estimation of the portfolio. We repeat this procedure for all the combinations of four folds with which we can construct the robust portfolio. We finally choose the λ that maximizes the out-of-sample certainty-equivalent return.

Table IA.7 reports the out-of-sample performance for the robust portfolio when λ is calibrated via cross-validation. Our results confirm that an optimally calibrated λ still outperforms. In particular, we observe that the average improvement relative to the shrinkage portfolio designed only to optimize OOSU mean in terms of certainty-equivalent return (Sharpe ratio) net of transaction costs across the six datasets and sample sizes is about 182% (24%).

IA.9.7 Performance of portfolios with a linear shrinkage covariance matrix

In Section 7.2, we illustrate the superior performance of the robust shrinkage portfolio that exploits the sample covariance matrix with simulated data. However, in Section 7.3, *all* our shrinkage portfolios exploit a linear shrinkage covariance matrix even though the optimal shrinkage intensities are obtained under the assumption that one uses the sample covariance matrix. We use a linear shrinkage covariance matrix because, in Section 6.1, we demonstrate that this allows one to further alleviate the impact of parameter uncertainty on OOSU risk.

We now study the performance of the shrinkage portfolios when the shrinkage intensities are obtained under the assumption that one uses a linear shrinkage covariance matrix. In particular, we now assess the performance of the shrinkage portfolios that use the linear shrinkage covariance matrix of Ledoit and Wolf (2004). In this context, to find the optimal shrinkage intensities maximizing OOSU mean and mean-risk OOSU, we assume that the vector of mean returns is the only source of parameter uncertainty, which has been identified as the main source of estimation risk in portfolio selection (Jagannathan and Ma, 2003; Kozak et al., 2020). To obtain the optimal κ, we need to define the OOSU mean and the OOSU variance of the shrinkage portfolios constructed with a linear shrinkage covariance matrix. Proposition IA.5 derives the OOSU variance for this class of shrinkage portfolios. The proposition below gives the analytical expression for the OOSU mean of the shrinkage portfolio that exploits a linear shrinkage covariance matrix and the vector of means is the only source of parameter uncertainty.

Proposition IA.8. *Assume that the covariance matrix $\mathbf{\Sigma}$ is known, $T > N$, and Assumption 2 in the main body of the manuscript holds. Then, the out-of-sample utility mean of $\hat{w}^\star(\delta, \kappa)$ in (IA19) is*

$$\mathbb{E}[U(\hat{w}^\star(\delta,\kappa))] = w_{sg}^\top \mu + \frac{\kappa}{\gamma} w_{sz}^\top \mu - \frac{\gamma}{2} w_{sg}^\top \mathbf{\Sigma} w_{sg} - \kappa w_{sg}^\top \mathbf{\Sigma} w_{sz} - \frac{\kappa^2}{2\gamma}\left(w_{sz}^\top \mathbf{\Sigma} w_{sz} + \frac{1}{T}\mathrm{Tr}(\mathbf{S}^2)\right). \quad \text{(IA39)}$$

We now use the analytical expressions for the OOSU mean and variance derived in Propositions IA.8 and IA.5, respectively, to estimate the optimal shrinkage intensities $\kappa_E^\star$ and $\kappa_R^\star$. We test the out-of-sample performance of the resulting shrinkage portfolios and report the results in Table IA.8. We consider two robust shrinkage portfolios. One where the optimal shrinkage intensity is obtained using a fixed $\lambda = 2$ in Equation (20). In addition, we consider a robust shrinkage portfolio where λ is calibrated via cross-validation in each estimation window. The results in Table IA.8 show that the robust shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ outperforms, in general, the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$. For instance, for the cross-validated λ we see that the average improvement in terms

of certainty-equivalent return (Sharpe ratio) net of transaction costs across the six datasets is 76.6% (46.1%) relative to the shrinkage portfolio only maximizing OOSU mean.

IA.9.8 Performance of portfolios with a nonlinear shrinkage covariance matrix

An interesting extension of our work is to study the performance of shrinkage portfolios that use a *nonlinear* shrinkage estimator of the covariance matrix instead of the sample covariance matrix or a linear shrinkage estimator of the covariance matrix.

The nonlinear shrinkage estimator of the covariance matrix we use is that of Ledoit and Wolf (2020), which is optimal for the minimum-variance loss function in Ledoit and Wolf (2017), and it is computationally efficient due to the closed-form solutions given by Ledoit and Wolf (2020). Analytical expressions for the OOSU mean and variance of the portfolios exploiting a nonlinear shrinkage covariance matrix are not available, and thus, the optimal shrinkage intensities maximizing OOSU mean and mean-risk OOSU are unknown. As a remedy, Kan et al. (2022b) propose to employ the optimal κ derived under the *sample* covariance matrix. However, this approach overstates the impact of estimation errors affecting the covariance matrix. Instead of using the shrinkage intensities that take μ and $\mathbf{\Sigma}$ as unknown, we construct shrinkage portfolios assuming that only the vector of means μ is unknown, which has been identified as the main source of estimation risk in portfolio selection (Jagannathan and Ma, 2003; Kozak et al., 2020).[5] Thus, we first provide a new theoretical result where we analytically characterize the mean-risk OOSU of a shrinkage portfolio with a known covariance matrix. Second, we exploit this new theoretical result to combine the mean-variance and minimum-variance portfolios constructed with a nonlinear shrinkage covariance matrix and study the performance of this class of shrinkage portfolios with simulated and empirical data.

In the next proposition, we derive a closed-form expression for the mean-risk OOSU of the shrinkage portfolio when only the vector of means μ is unknown.

Proposition IA.9. *Let $N \geq 2$, $T > N$, and Assumption 2 hold. When the covariance matrix $\mathbf{\Sigma}$*

[5] Our derivation of the OOSU risk in Proposition 1 builds on the stochastic representation of the out-of-sample mean return and variance derived by Kan et al. (2022b). This stochastic representation relies on the Wishart distribution of the sample covariance matrix. However, the nonlinear shrinkage covariance matrix we use in this section no longer follows a Wishart distribution. Thus, analytical expressions for the OOSU variance of sample portfolios that exploit a shrinkage covariance matrix are unavailable. To overcome this challenge, we assume that the covariance matrix is known, which is not a strong assumption given that 1) the critical source of parameter uncertainty in portfolio selection stems from the vector of mean returns, as stated in the main text, and 2) the nonlinear shrinkage covariance matrix substantially reduces the estimator's mean-squared error.

is known, the mean-risk out-of-sample utility of the shrinkage portfolio $\hat{w}^\star(\kappa)$ *is*

$$R(\hat{w}^\star(\kappa)) = \mu_g - \frac{\gamma}{2}\sigma_g^2 + \frac{1}{\gamma}\left(\kappa\psi^2 - \frac{\kappa^2}{2}\left(\psi^2 + \frac{N-1}{T}\right) - \lambda\frac{\kappa\psi}{\sqrt{T}}\sqrt{(1-\kappa)^2 + \kappa^2\frac{N-1}{2T\psi^2}}\right). \tag{IA40}$$

Moreover, the optimal shrinkage intensity $\kappa_R^\star$ *is bounded such that* $\kappa_R^\star \leq \kappa_E^\star$*, where* $\kappa_E^\star = \frac{\psi^2}{\psi^2+(N-1)/T}$ *maximizes the out-of-sample utility mean of the shrinkage portfolio.*

Note that the optimal shrinkage intensity maximizing the mean-risk OOSU, $\kappa_R^\star$, does not have an explicit closed-form expression as $\kappa_E^\star$ but can be easily obtained numerically with the analytical expression provided in Proposition IA.9.

We now assess the performance of the shrinkage portfolios that use the optimal shrinkage intensities characterized in Proposition IA.9 constructed with the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020). Table IA.9 reports the performance of the shrinkage portfolios with simulated data. Panel A of Table IA.9 reports the OOSU mean, standard deviation, and the mean-risk OOSU for simulated Gaussian data. We consider the shrinkage portfolio with the estimated intensity $\hat{\kappa}_E^\star$ that maximizes OOSU mean and the shrinkage portfolio with the estimated intensity $\hat{\kappa}_R^\star$ that maximizes our proposed robustness measure with $\lambda = 2$. We observe that for a sample size of $T = 120$ observations, the shrinkage portfolio that exploits $\hat{\kappa}_R^\star$ delivers a larger OOSU mean than that of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ in four datasets. This result is consistent with our baseline findings and suggests that $\hat{\kappa}_R^\star$ emerges as a robust shrinkage intensity that is subject to lower estimation risk than $\hat{\kappa}_E^\star$, which allows our robust shrinkage portfolio to outperform in terms of OOSU mean the shrinkage portfolio designed to optimize it. In addition, we observe that the shrinkage portfolio that exploits $\hat{\kappa}_E^\star$ delivers an OOSU that is more volatile than that of the shrinkage portfolio that exploits $\hat{\kappa}_R^\star$. Similar to the main results with simulated data, the difference is particularly large for the case with a sample size of $T = 120$ observations. Thus, the shrinkage portfolio that exploits the intensity $\hat{\kappa}_R^\star$ delivers a better mean-risk OOSU than that of the shrinkage portfolio that exploits $\hat{\kappa}_E^\star$.

Panel B of Table IA.9 reports the performance of the two shrinkage portfolios with bootstrapped data. In terms of OOSU mean, we see that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ performs remarkably better. Specifically, it outperforms the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ in all six datasets across all sample sizes. In addition, the shrinkage intensity $\hat{\kappa}_E^\star$ yields an OOSU volatility that is, on average across all datasets, 178%, 183%, and 151% larger than that of $\hat{\kappa}_R^\star$ for $T = 120$, 180, and 240 months, respectively. The results in this panel confirm our main insight in the main body of the

manuscript that accounting for OOSU risk in the construction of shrinkage portfolios is particularly important when the data is not Gaussian.

Finally, we test the out-of-sample performance with empirical return data of the shrinkage portfolios that exploit a nonlinear shrinkage covariance matrix and report the results in Table IA.10. We consider two robust shrinkage portfolios. One where the optimal shrinkage intensity is obtained using a fixed $\lambda = 2$ in Equation (20) in the main body of the manuscript. In addition, we consider a robust shrinkage portfolio where λ is calibrated via cross-validation in each estimation window. The results in Table IA.10 show that the robust shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ outperforms, in general, the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$. For instance, for the cross-validated λ, we see that the average improvement in terms of certainty-equivalent return (Sharpe ratio) net of transaction costs across the six datasets and sample sizes is 121% (7%) relative to the shrinkage portfolio only maximizing OOSU mean.

IA.9.9 Performance of shrinkage portfolios with many assets

The case where N is large is an important case to consider for portfolio selection in the presence of parameter uncertainty. We address this issue in two ways. First, we note that from Propositions 2 and 5, the OOSU variance increases with the number of assets N, and the robustness measure introduced in Section 5 decreases with N, which gives theoretical support for accounting for out-of-sample utility risk in the construction of a high-dimensional optimal portfolio. Second, we study the performance of the shrinkage portfolios in a high-dimensional setting with $N = 155$ assets. We build this dataset of 155 assets by merging all six datasets in the empirical analysis for the time period for which we have available return data across the six datasets.

To construct the shrinkage portfolios in the empirical application, we rely on the insights of Ledoit and Wolf (2017), who show that nonlinear shrinkage covariance matrices can outperform linear shrinkage covariance matrices in high-dimensional settings where the number of assets is large. Accordingly, we follow Ledoit and Wolf (2017) and construct the portfolios in this high-dimensional setting of $N = 155$ assets using the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020) and the optimal shrinkage intensities derived in Proposition IA.9.

Table IA.11 reports the performance of the shrinkage portfolios with simulated data. Panel A of Table IA.11 reports the OOSU mean, standard deviation, and the mean-risk OOSU for the simulated Gaussian data. We consider the shrinkage portfolio with the estimated intensity $\hat{\kappa}_E^\star$ that maximizes OOSU mean and the shrinkage portfolio with the estimated intensity $\hat{\kappa}_R^\star$ that

maximizes our proposed robustness measure with $\lambda = 2$. We observe that for all sample sizes T, the shrinkage portfolio that exploits $\hat{\kappa}_R^\star$ delivers a larger OOSU mean than that of the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$. Thus, $\hat{\kappa}_R^\star$ is a robust shrinkage intensity that is subject to lower estimation risk than $\hat{\kappa}_E^\star$, which allows our robust shrinkage portfolio to outperform in terms of OOSU mean the shrinkage portfolio designed to optimize it. In addition, we observe that the shrinkage portfolio that exploits $\hat{\kappa}_E^\star$ delivers an OOSU that is more volatile than that of the shrinkage portfolio that exploits $\hat{\kappa}_R^\star$. Thus, the shrinkage portfolio that exploits the intensity $\hat{\kappa}_R^\star$ delivers a better mean-risk OOSU than that of the shrinkage portfolio that exploits $\hat{\kappa}_E^\star$.

Panel B of Table IA.11 reports the performance of the two shrinkage portfolios with bootstrapped data. In terms of OOSU mean, we see that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ performs remarkably better. Specifically, it outperforms the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ across all sample sizes. In addition, the shrinkage intensity $\hat{\kappa}_E^\star$ yields an OOSU volatility that is 138% and 275% larger than that of $\hat{\kappa}_R^\star$ for $T = 180$ and 240 months, respectively. The results in this panel confirm one of the key insights in the main body of the manuscript that accounting for OOSU risk in the construction of shrinkage portfolios is particularly important when the data is not Gaussian.

Finally, we test the out-of-sample performance with empirical return data of the shrinkage portfolios and report the results in Table IA.12. We consider two robust shrinkage portfolios. One where the optimal shrinkage intensity is obtained using a fixed $\lambda = 2$ in Equation (20) and a second robust shrinkage portfolio where λ is calibrated via cross-validation. The results in Table IA.12 show that the robust shrinkage portfolio exploiting $\kappa_R^\star$ outperforms, in general, the shrinkage portfolio exploiting $\kappa_E^\star$. For instance, for the cross-validated λ we see that the average improvement across sample sizes in terms of certainty-equivalent return (Sharpe ratio) net of transaction costs is about 101% (10%) relative to the shrinkage portfolio only maximizing OOSU mean.

IA.9.10 Impact of estimation errors in shrinkage intensities

It is interesting to look at the performance of the oracle shrinkage portfolios in the simulations because this allows us to study the impact that the estimation errors affecting $\kappa_E^\star$ and $\kappa_R^\star$ have on the performance of the shrinkage portfolios. Table IA.13 reports the performance of the shrinkage portfolios with simulated data when the shrinkage intensities $\kappa_E^\star$ and $\kappa_R^\star$ are estimated and suffer from estimation errors (Panel A), and when the shrinkage intensities are constructed from the population moments and do not suffer from estimation errors (Panel B). We make several observations from these results. First, as one would expect, the shrinkage portfolio with the oracle $\kappa_E^\star$ delivers a

higher OOSU mean across all datasets and for all estimation windows than the shrinkage portfolio with the oracle $\kappa_R^\star$. On the contrary, the shrinkage portfolio with the oracle $\kappa_R^\star$ delivers a higher mean-risk OOSU across all datasets and for all estimation windows than the shrinkage portfolio with the oracle $\kappa_E^\star$. Second, we see that the performance deterioration driven by the estimation errors affecting the shrinkage intensities is larger for the shrinkage portfolio optimizing OOSU mean. For instance, for an estimation window of $T = 120$ observations, the performance deterioration in terms of OOSU mean when the shrinkage intensity is estimated is about 14.5% for $\kappa_E^\star$ and only 7.0% for $\kappa_R^\star$ on average across all datasets. In terms of mean-risk OOSU, the performance deterioration is about 103.4% for $\kappa_E^\star$ and 56.7% for $\kappa_R^\star$ on average across all datasets. Thus, the results in Table IA.13 confirm our insight in Section 7.2 that our robust portfolio is more resilient to estimation errors affecting the shrinkage intensities.

Table IA.1: Value-at-Risk of shrinkage portfolios

		OOS 1% VaR		OOS 5% VaR	
		$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$
Panel A: Without transaction costs					
$T = 120$	*10MOM*	0.249	0.184	0.119	0.081
	25SBTM	0.203	0.102	0.099	0.058
	25OPINV	0.207	0.125	0.094	0.062
	49IND	0.129	0.084	0.067	0.056
	16LTANOM	0.197	0.122	0.107	0.076
	46ANOM	0.186	0.157	0.117	0.091
$T = 240$	*10MOM*	0.287	0.218	0.132	0.102
	25SBTM	0.215	0.155	0.099	0.064
	25OPINV	0.203	0.136	0.103	0.071
	49IND	0.142	0.092	0.070	0.060
	16LTANOM	0.183	0.116	0.117	0.079
	46ANOM	0.612	0.408	0.247	0.212
Panel B: Net of transaction costs					
$T = 120$	*10MOM*	0.252	0.187	0.121	0.083
	25SBTM	0.212	0.105	0.110	0.060
	25OPINV	0.211	0.127	0.097	0.063
	49IND	0.131	0.084	0.074	0.056
	16LTANOM	0.200	0.122	0.114	0.076
	46ANOM	0.189	0.160	0.121	0.095
$T = 240$	*10MOM*	0.291	0.221	0.133	0.104
	25SBTM	0.218	0.158	0.102	0.068
	25OPINV	0.204	0.136	0.106	0.071
	49IND	0.143	0.093	0.071	0.060
	16LTANOM	0.186	0.117	0.120	0.080
	46ANOM	0.616	0.412	0.259	0.218

Notes. This table reports the out-of-sample 1% and 5% Value-at-Risk of the estimated shrinkage portfolio that maximizes out-of-sample utility mean ($\hat{\kappa}_E^\star$) and the estimated shrinkage portfolio that maximizes the portfolio robustness measure in Section 5 ($\hat{\kappa}_R^\star$) for the six datasets of monthly excess returns described in Section 7.1. We estimate the shrinkage portfolios using sample sizes of $T = 120$ and $T = 240$ monthly observations, a risk-aversion coefficient of $\gamma = 3$, and a coefficient $\lambda = 2$ for the shrinkage portfolio that maximizes the proposed robustness measure. The out-of-sample Value-at-Risk is computed as the negative 1st and 5th percentiles of the out-of-sample monthly excess returns. Panel A considers the case without transaction costs, and panel B considers the case with proportional transaction costs of 20 basis points.

Table IA.2: Cumulative wealth net of transaction costs of shrinkage portfolios

	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$	% increase
Panel A: $T = 120$ (July 1983 – December 2013)			
10MOM	88.5	140	59%
25SBTM	149	740	396%
25OPINV	34.6	166	381%
49IND	5.44	24.2	346%
16LTANOM	27.5	119	332%
46ANOM	49,686	56,731	14%
Panel B: $T = 240$ (July 1993 – December 2013)			
10MOM	5.86	7.70	31%
25SBTM	18.3	23.4	28%
25OPINV	9.83	17.2	75%
49IND	2.18	4.70	116%
16LTANOM	8.84	19.0	115%
46ANOM	64.8	123	90%

Notes. This table reports the cumulative wealth net of proportional transaction costs of 20 basis points of the estimated shrinkage portfolios. We consider the estimated shrinkage portfolio that maximizes out-of-sample utility mean ($\hat{\kappa}_E^\star$), and the estimated shrinkage portfolio that maximizes the portfolio robustness measure in Section 5 ($\hat{\kappa}_R^\star$). We report the cumulative wealth of the shrinkage portfolios for the six datasets of monthly excess returns described in Section 7.1. The out-of-sample returns include the risk-free rate and are standardized to have the same volatility as that of the market factor during the same time period. We only consider the overlapping out-of-sample period across the six datasets, which spans July 1983 through December 2013 when the portfolios are estimated with a sample size of $T = 120$ (panel A) and July 1993 through December 2013 when the portfolios are estimated with a sample size of $T = 240$ (panel B). We use a risk-aversion coefficient of $\gamma = 3$ and a coefficient $\lambda = 2$ for the shrinkage portfolio that maximizes the proposed robustness measure.

Table IA.3: Out-of-sample performance with lower risk aversion

		$T = 120$				$T = 240$			
		RTR	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$	RTR	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$
10MOM	Gross OOS CER	0.077	0.082	0.238	**0.286**	0.079	0.088	0.274	**0.323**
	Net OOS CER	0.075	0.078	0.150	**0.243**	0.078	0.084	0.218	**0.289**
	Gross OOS SR	0.558	0.641	0.745	**0.770**	0.607	0.706	0.784	**0.804**
	Net OOS SR	0.547	0.611	0.654	**0.702**	0.600	0.682	0.727	**0.761**
	Average $\hat{\kappa}$	/	0	0.395	0.273	/	0	0.546	0.454
25SBTM	Gross OOS CER	0.087	0.088	**0.231**	**0.231**	0.086	0.096	0.253	**0.271**
	Net OOS CER	0.086	0.079	0.052	**0.177**	0.085	0.089	0.133	**0.199**
	Gross OOS SR	0.561	0.748	0.693	**0.866**	0.612	0.833	0.730	**0.802**
	Net OOS SR	0.554	0.679	0.464	**0.708**	0.606	**0.778**	0.589	0.653
	Average $\hat{\kappa}$	/	0	0.214	0.131	/	0	0.364	0.280
25OPINV	Gross OOS CER	0.079	0.097	0.107	**0.154**	0.081	0.119	0.191	**0.209**
	Net OOS CER	0.077	0.088	-0.008	**0.107**	0.080	0.114	0.119	**0.167**
	Gross OOS SR	0.586	**0.794**	0.495	0.656	0.633	**0.976**	0.619	0.745
	Net OOS SR	0.575	**0.729**	0.318	0.504	0.627	**0.935**	0.505	0.635
	Average $\hat{\kappa}$	/	0	0.191	0.104	/	0	0.320	0.221
49IND	Gross OOS CER	**0.075**	0.067	0.068	0.070	**0.073**	0.058	0.020	0.056
	Net OOS CER	**0.074**	0.058	0.011	0.051	**0.072**	0.051	-0.017	0.031
	Gross OOS SR	0.565	**0.624**	0.395	0.550	**0.592**	0.557	0.234	0.391
	Net OOS SR	**0.554**	0.546	0.172	0.423	**0.585**	0.501	0.131	0.256
	Average $\hat{\kappa}$	/	0	0.043	0.016	/	0	0.087	0.041
16LTANOM	Gross OOS CER	0.077	0.084	0.191	**0.206**	0.071	0.111	0.341	**0.351**
	Net OOS CER	0.075	0.078	0.046	**0.151**	0.070	0.107	0.235	**0.295**
	Gross OOS SR	0.516	0.695	0.640	**0.766**	0.520	**0.942**	0.836	0.907
	Net OOS SR	0.505	**0.649**	0.452	0.611	0.513	**0.904**	0.724	0.808
	Average $\hat{\kappa}$	/	0	0.309	0.201	/	0	0.522	0.443
46ANOM	Gross OOS CER	0.079	0.067	**2.378**	2.179	0.083	0.111	0.240	**0.850**
	Net OOS CER	0.077	0.054	**2.177**	2.011	0.082	0.101	-0.179	**0.736**
	Gross OOS SR	0.566	0.643	2.364	**2.372**	0.600	1.037	1.474	**1.565**
	Net OOS SR	0.555	0.532	2.249	**2.265**	0.593	0.945	1.228	**1.491**
	Average $\hat{\kappa}$	/	0	0.250	0.205	/	0	0.511	0.471

Notes. This table reports the out-of-sample performance of four portfolio strategies described in Section 7.3 for the six datasets of monthly excess returns described in Section 7.1. Each estimated portfolio is constructed using a sample size of $T = 120$ and 240 monthly observations. The mean-variance portfolios consider a risk-aversion coefficient of $\gamma = 1$. The table reports the annualized out-of-sample certainty-equivalent return (OOS CER) and the annualized out-of-sample Sharpe ratio (OOS SR), and for both criteria we report the gross performance and the performance net of proportional transaction costs of 20 basis points. We also report the average estimated shrinkage intensity $\hat{\kappa}$ over time, except for the RTR portfolio that does not combine the SMV and SGMV portfolios.

Table IA.4: Out-of-sample performance with higher risk aversion

		$T = 120$				$T = 240$			
		RTR	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$	RTR	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$
10MOM	Gross OOS CER	0.025	0.040	0.072	**0.084**	0.035	0.050	0.091	**0.101**
	Net OOS CER	0.023	0.036	0.051	**0.072**	0.034	0.047	0.078	**0.090**
	Gross OOS SR	0.558	0.641	0.915	**0.916**	0.607	0.706	0.985	**1.005**
	Net OOS SR	0.547	0.611	0.829	**0.851**	0.600	0.682	0.931	**0.952**
	Average $\hat{\kappa}$	/	0	0.395	0.312	/	0	0.546	0.483
25SBTM	Gross OOS CER	0.018	0.055	**0.089**	0.086	0.034	0.065	0.102	**0.109**
	Net OOS CER	0.016	0.046	0.046	**0.067**	0.033	0.058	0.074	**0.083**
	Gross OOS SR	0.561	0.748	0.949	**0.958**	0.612	0.833	1.012	**1.077**
	Net OOS SR	0.554	0.679	0.745	**0.828**	0.606	0.778	0.881	**0.920**
	Average $\hat{\kappa}$	/	0	0.214	0.155	/	0	0.364	0.298
25OPINV	Gross OOS CER	0.031	0.062	0.059	**0.075**	0.039	0.085	0.097	**0.105**
	Net OOS CER	0.029	0.053	0.031	**0.060**	0.038	0.080	0.081	**0.093**
	Gross OOS SR	0.586	0.794	0.787	**0.882**	0.633	0.976	0.985	**1.070**
	Net OOS SR	0.575	0.729	0.638	**0.777**	0.627	0.935	0.899	**0.992**
	Average $\hat{\kappa}$	/	0	0.191	0.132	/	0	0.320	0.244
49IND	Gross OOS CER	0.028	0.039	0.029	**0.044**	0.034	0.031	0.016	**0.036**
	Net OOS CER	0.026	0.030	0.009	**0.034**	**0.033**	0.025	0.003	0.029
	Gross OOS SR	0.565	0.624	0.553	**0.668**	0.592	0.557	0.459	**0.598**
	Net OOS SR	0.554	0.546	0.401	**0.582**	**0.585**	0.501	0.359	0.540
	Average $\hat{\kappa}$	/	0	0.043	0.029	/	0	0.087	0.061
16LTANOM	Gross OOS CER	0.011	0.048	0.064	**0.076**	0.018	0.079	0.125	**0.127**
	Net OOS CER	0.009	0.042	0.029	**0.059**	0.017	0.075	0.103	**0.114**
	Gross OOS SR	0.516	0.695	0.829	**0.878**	0.520	0.942	1.118	**1.183**
	Net OOS SR	0.505	0.649	0.658	**0.766**	0.513	0.904	1.021	**1.104**
	Average $\hat{\kappa}$	/	0	0.309	0.232	/	0	0.522	0.455
46ANOM	Gross OOS CER	0.026	0.041	**0.513**	0.472	0.032	0.085	0.109	**0.237**
	Net OOS CER	0.024	0.028	**0.474**	0.439	0.031	0.075	0.002	**0.203**
	Gross OOS SR	0.566	0.643	**2.445**	2.435	0.600	1.037	1.596	**1.717**
	Net OOS SR	0.555	0.532	**2.327**	2.324	0.593	0.945	1.377	**1.641**
	Average $\hat{\kappa}$	/	0	0.250	0.208	/	0	0.511	0.472

Notes. This table reports the out-of-sample performance of four portfolio strategies described in Section 7.3 for the six datasets of monthly excess returns described in Section 7.1. Each estimated portfolio is constructed using a sample size of $T = 120$ and 240 monthly observations. The mean-variance portfolios consider a risk-aversion coefficient of $\gamma = 5$. The table reports the annualized out-of-sample certainty-equivalent return (OOS CER) and the annualized out-of-sample Sharpe ratio (OOS SR), and for both criteria we report the gross performance and the performance net of proportional transaction costs of 20 basis points. We also report the average estimated shrinkage intensity $\hat{\kappa}$ over time, except for the RTR portfolio that does not combine the SMV and SGMV portfolios.

Table IA.5: Out-of-sample performance with higher transaction costs

		$T = 120$				$T = 240$			
		RTR	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$	RTR	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$
10MOM	Gross OOS CER	0.051	0.061	0.114	**0.131**	0.057	0.069	0.135	**0.152**
	Net OOS CER	0.049	0.055	0.065	**0.105**	0.056	0.064	0.104	**0.131**
	Gross OOS SR	0.558	0.641	0.864	**0.898**	0.607	0.706	0.917	**0.958**
	Net OOS SR	0.542	0.597	0.727	**0.793**	0.597	0.670	0.834	**0.888**
	Average $\hat{\kappa}$	/	0	0.395	0.296	/	0	0.546	0.471
25SBTM	Gross OOS CER	0.052	0.072	0.122	**0.123**	0.060	0.081	0.139	**0.146**
	Net OOS CER	0.050	0.058	0.022	**0.086**	0.058	0.070	0.073	**0.096**
	Gross OOS SR	0.561	0.748	0.858	**0.995**	0.612	0.833	0.914	**1.014**
	Net OOS SR	0.550	0.644	0.523	**0.768**	0.603	0.750	0.705	**0.773**
	Average $\hat{\kappa}$	/	0	0.214	0.145	/	0	0.364	0.289
25OPINV	Gross OOS CER	0.055	0.079	0.080	**0.101**	0.060	0.102	0.124	**0.133**
	Net OOS CER	0.052	0.066	0.018	**0.073**	0.059	0.094	0.087	**0.109**
	Gross OOS SR	0.586	0.794	0.699	**0.866**	0.633	0.976	0.872	**1.004**
	Net OOS SR	0.569	**0.696**	0.458	**0.696**	0.623	**0.915**	0.722	0.872
	Average $\hat{\kappa}$	/	0	0.191	0.120	/	0	0.320	0.232
49IND	Gross OOS CER	0.052	0.053	0.043	**0.059**	**0.054**	0.044	0.024	0.046
	Net OOS CER	**0.049**	0.039	-0.005	0.041	**0.052**	0.035	0.001	0.031
	Gross OOS SR	0.565	0.624	0.510	**0.661**	**0.592**	0.557	0.392	0.558
	Net OOS SR	**0.549**	0.507	0.194	0.519	**0.581**	0.473	0.250	0.437
	Average $\hat{\kappa}$	/	0	0.043	0.024	/	0	0.087	0.054
16LTANOM	Gross OOS CER	0.044	0.066	0.099	**0.110**	0.045	0.095	0.172	**0.175**
	Net OOS CER	0.041	0.057	0.020	**0.077**	0.043	0.088	0.118	**0.149**
	Gross OOS SR	0.516	0.695	0.779	**0.893**	0.520	0.942	1.016	**1.130**
	Net OOS SR	0.499	0.626	0.508	**0.703**	0.509	0.886	0.856	**1.011**
	Average $\hat{\kappa}$	/	0	0.309	0.218	/	0	0.522	0.449
46ANOM	Gross OOS CER	0.052	0.054	**0.832**	0.765	0.057	0.098	0.147	**0.347**
	Net OOS CER	0.050	0.035	**0.738**	0.687	0.056	0.083	0.003	**0.268**
	Gross OOS SR	0.566	0.643	2.433	**2.442**	0.600	1.037	1.543	**1.648**
	Net OOS SR	0.550	0.476	2.259	**2.278**	0.589	0.900	1.347	**1.534**
	Average $\hat{\kappa}$	/	0	0.250	0.207	/	0	0.511	0.471

Notes. This table reports the out-of-sample performance of four portfolio strategies described in Section 7.3 for the six datasets of monthly excess returns described in Section 7.1. Each estimated portfolio is constructed using a sample size of $T = 120$ and 240 monthly observations. The mean-variance portfolios consider a risk-aversion coefficient of $\gamma = 3$. The table reports the annualized out-of-sample certainty-equivalent return (OOS CER) and the annualized out-of-sample Sharpe ratio (OOS SR), and for both criteria we report the gross performance and the performance net of proportional transaction costs of 30 basis points. We also report the average estimated shrinkage intensity $\hat{\kappa}$ over time, except for the RTR portfolio that does not combine the SMV and SGMV portfolios.

Table IA.6: Out-of-sample performance exploiting the equally weighted portfolio

		$T = 120$				$T = 240$			
		EW	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$	EW	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$
10MOM	Gross OOS CER	0.034	0.061	0.110	**0.120**	0.039	0.069	0.132	**0.147**
	Net OOS CER	0.033	0.057	0.079	**0.102**	0.038	0.065	0.111	**0.133**
	Gross OOS SR	0.455	0.641	0.851	**0.854**	0.484	0.706	0.909	**0.942**
	Net OOS SR	0.449	0.611	0.764	**0.784**	0.478	0.682	0.851	**0.894**
	Average $\hat{\kappa}$	/	0	0.692	0.674	/	0	0.784	0.781
	Average $\hat{\pi}$	0	/	0.637	0.473	0	/	0.724	0.629
25SBTM	Gross OOS CER	0.045	0.072	**0.122**	0.120	0.053	0.081	0.142	**0.151**
	Net OOS CER	0.044	0.063	0.059	**0.090**	0.052	0.074	0.099	**0.103**
	Gross OOS SR	0.522	0.748	0.858	**0.945**	0.567	0.833	0.923	**1.026**
	Net OOS SR	0.516	0.679	0.649	**0.775**	0.561	0.778	0.788	**0.803**
	Average $\hat{\kappa}$	/	0	0.648	0.628	/	0	0.663	0.658
	Average $\hat{\pi}$	0	/	0.427	0.312	0	/	0.600	0.471
25OPINV	Gross OOS CER	0.044	0.079	0.074	**0.090**	0.050	0.102	0.119	**0.126**
	Net OOS CER	0.043	**0.071**	0.032	0.069	0.049	0.097	0.094	**0.111**
	Gross OOS SR	0.515	0.794	0.671	**0.802**	0.556	**0.976**	0.851	0.969
	Net OOS SR	0.509	**0.729**	0.505	0.673	0.550	**0.935**	0.749	0.883
	Average $\hat{\kappa}$	/	0	0.820	0.786	/	0	0.867	0.866
	Average $\hat{\pi}$	0	/	0.272	0.170	0	/	0.381	0.265
49IND	Gross OOS CER	0.050	**0.053**	0.038	0.049	**0.052**	0.044	0.037	0.047
	Net OOS CER	**0.049**	0.044	0.013	0.039	**0.051**	0.038	0.024	0.038
	Gross OOS SR	0.552	**0.624**	0.481	0.560	**0.569**	0.557	0.471	0.547
	Net OOS SR	0.544	**0.546**	0.336	0.491	**0.562**	0.501	0.393	0.485
	Average $\hat{\kappa}$	/	0	0.431	0.365	/	0	0.418	0.393
	Average $\hat{\pi}$	0	/	0.232	0.131	0	/	0.249	0.133
16LTANOM	Gross OOS CER	0.027	0.066	0.094	**0.109**	0.028	0.095	0.165	**0.167**
	Net OOS CER	0.026	0.060	0.043	**0.086**	0.027	0.091	0.128	**0.150**
	Gross OOS SR	0.419	0.695	0.762	**0.883**	0.422	0.942	0.994	**1.103**
	Net OOS SR	0.413	0.649	0.588	**0.750**	0.416	0.904	0.886	**1.020**
	Average $\hat{\kappa}$	/	0	0.783	0.784	/	0	0.839	0.859
	Average $\hat{\pi}$	0	/	0.443	0.301	0	/	0.672	0.573
46ANOM	Gross OOS CER	0.016	0.054	0.302	**0.765**	0.022	0.098	-1.007	**0.374**
	Net OOS CER	0.015	0.041	0.126	**0.713**	0.021	0.088	-1.067	**0.320**
	Gross OOS SR	0.358	0.643	1.999	**2.436**	0.399	1.037	1.415	**1.665**
	Net OOS SR	0.353	0.532	1.746	**2.328**	0.393	0.945	1.279	**1.587**
	Average $\hat{\kappa}$	/	0	0.727	0.713	/	0	0.799	0.801
	Average $\hat{\pi}$	0	/	0.494	0.454	0	/	0.698	0.651

Notes. This table reports the out-of-sample performance of the shrinkage portfolios that combine the SMV, SGMV, and equally weighted portfolios introduced in Appendix IA.9.5 for the six datasets of monthly excess returns described in Section 7.1. Each estimated portfolio is constructed using a sample size of $T = 120$ and 240 monthly observations. The mean-variance portfolios consider a risk-aversion coefficient of $\gamma = 3$. The table reports the annualized out-of-sample certainty-equivalent return (OOS CER) and the annualized out-of-sample Sharpe ratio (OOS SR), and for both criteria we report the gross performance and the performance net of proportional transaction costs of 20 basis points. We also report the average estimated shrinkage intensities $\hat{\pi}$ and $\hat{\kappa}$ over time, which determine the combination of the three portfolios in Equation (IA29).

Table IA.7: Out-of-sample performance in empirical data with cross-validation of λ

		$T = 120$				$T = 240$			
		SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$		SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$	
				$\lambda = 2$	$\lambda = \hat{\lambda}_{cv}$			$\lambda = 2$	$\lambda = \hat{\lambda}_{cv}$
10MOM	Gross OOS CER	0.061	0.115	0.134	**0.139**	0.069	0.134	**0.152**	0.150
	Net OOS CER	0.057	0.082	**0.118**	**0.118**	0.065	0.114	**0.138**	0.131
	Gross OOS SR	0.641	0.865	0.909	**0.923**	0.706	0.917	**0.959**	0.950
	Net OOS SR	0.611	0.773	**0.846**	0.844	0.682	0.862	**0.912**	0.888
	Average $\hat{\kappa}$	0	0.395	0.296	0.350	0	0.546	0.471	0.493
25SBTM	Gross OOS CER	0.072	0.123	0.123	0.132	0.081	0.139	0.144	**0.145**
	Net OOS CER	0.063	0.058	0.098	0.102	0.074	0.095	0.111	**0.112**
	Gross OOS SR	0.748	0.861	0.991	**0.997**	0.833	0.914	**1.012**	1.000
	Net OOS SR	0.679	0.643	**0.844**	0.831	0.778	0.776	**0.851**	0.845
	Average $\hat{\kappa}$	0	0.214	0.145	0.194	0	0.364	0.289	0.338
25OPINV	Gross OOS CER	0.079	0.080	**0.102**	0.094	0.102	0.124	**0.135**	0.131
	Net OOS CER	0.071	0.039	**0.083**	0.066	0.097	0.099	**0.119**	0.108
	Gross OOS SR	0.794	0.699	**0.872**	0.809	0.976	0.872	**1.017**	1.002
	Net OOS SR	0.729	0.538	**0.752**	0.641	**0.935**	0.772	0.928	0.874
	Average $\hat{\kappa}$	0	0.191	0.120	0.136	0	0.320	0.232	0.220
49IND	Gross OOS CER	0.053	0.039	0.057	**0.064**	0.044	0.024	0.046	**0.051**
	Net OOS CER	0.044	0.008	0.046	**0.048**	0.038	0.008	0.036	**0.041**
	Gross OOS SR	0.624	0.486	0.645	**0.687**	0.557	0.391	0.557	**0.617**
	Net OOS SR	0.546	0.279	0.551	**0.563**	0.501	0.293	0.475	**0.529**
	Average $\hat{\kappa}$	0	0.043	0.024	0.023	0	0.087	0.054	0.020
16LTANOM	Gross OOS CER	0.066	0.098	**0.109**	0.099	0.095	0.172	**0.175**	0.171
	Net OOS CER	0.060	0.046	**0.087**	0.067	0.091	0.136	**0.158**	0.146
	Gross OOS SR	0.695	0.775	**0.887**	0.814	0.942	1.016	**1.132**	1.101
	Net OOS SR	0.649	0.600	**0.758**	0.639	0.904	0.910	**1.051**	0.987
	Average $\hat{\kappa}$	0	0.309	0.218	0.247	0	0.522	0.449	0.476
46ANOM	Gross OOS CER	0.054	**0.831**	0.765	**0.831**	0.098	0.138	**0.347**	0.296
	Net OOS CER	0.041	**0.769**	0.713	0.768	0.088	0.022	**0.295**	0.241
	Gross OOS SR	0.643	2.431	**2.442**	2.433	1.037	1.536	**1.648**	1.621
	Net OOS SR	0.532	2.316	**2.334**	2.317	0.945	1.371	**1.573**	1.543
	Average $\hat{\kappa}$	0	0.250	0.207	0.245	0	0.511	0.471	0.483

Notes. This table reports the out-of-sample performance of four portfolio strategies described in Section 7.3 for the six datasets of monthly excess returns described in Section 7.1. Each estimated portfolio is constructed using a sample size of $T = 120$ or $T = 240$ monthly observations. The mean-variance portfolios consider a risk-aversion coefficient of $\gamma = 3$. The proposed robust shrinkage portfolio that maximizes the mean-risk out-of-sample utility is estimated with a coefficient λ equal to two or calibrated via five-fold cross-validation. The table reports the annualized out-of-sample certainty-equivalent return (OOS CER) and the annualized out-of-sample Sharpe ratio (OOS SR), and for both criteria we report the gross performance and the performance net of proportional transaction costs of 20 basis points. We also report the average estimated shrinkage intensity $\hat{\kappa}$ over time. The numbers in bold font identify the best portfolio in terms of OOS CER and SR.

Table IA.8: Out-of-sample performance with Ledoit and Wolf (2004) covariance matrix

		$T = 120$				$T = 240$			
		SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$		SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$	
				$\lambda = 2$	$\lambda = \hat{\lambda}_{cv}$			$\lambda = 2$	$\lambda = \hat{\lambda}_{cv}$
10MOM	Gross OOS CER	0.061	0.140	**0.143**	**0.143**	0.069	0.144	0.146	**0.156**
	Net OOS CER	0.057	0.116	**0.122**	0.119	0.065	0.126	0.129	**0.138**
	Gross OOS SR	0.641	0.921	**0.927**	0.926	0.706	0.944	0.945	**0.970**
	Net OOS SR	0.611	0.851	**0.859**	0.846	0.682	0.895	0.896	**0.916**
	Average $\hat{\kappa}$	0	0.552	0.436	0.407	0	0.668	0.586	0.535
25SBTM	Gross OOS CER	0.072	0.128	**0.133**	0.122	0.081	0.157	**0.158**	0.147
	Net OOS CER	0.063	0.071	**0.082**	0.069	0.074	0.117	**0.121**	0.111
	Gross OOS SR	0.748	0.883	**0.893**	0.871	0.833	0.971	**0.975**	0.955
	Net OOS SR	0.679	0.710	**0.722**	0.646	0.778	0.848	**0.854**	0.815
	Average $\hat{\kappa}$	0	0.514	0.434	0.291	0	0.602	0.545	0.401
25OPINV	Gross OOS CER	0.079	0.033	0.055	**0.088**	0.102	0.081	0.095	**0.129**
	Net OOS CER	0.071	-0.027	0.001	**0.048**	0.097	0.047	0.064	**0.105**
	Gross OOS SR	**0.794**	0.689	0.703	0.728	**0.976**	0.767	0.788	0.938
	Net OOS SR	**0.729**	0.540	0.555	0.547	**0.935**	0.673	0.694	0.819
	Average $\hat{\kappa}$	0	0.557	0.488	0.192	0	0.618	0.565	0.243
49IND	Gross OOS CER	0.053	-0.278	-0.198	**0.054**	**0.044**	-0.179	-0.136	0.042
	Net OOS CER	**0.044**	-0.346	-0.260	0.033	**0.038**	-0.215	-0.169	0.031
	Gross OOS SR	**0.624**	0.327	0.342	0.586	**0.557**	0.176	0.190	0.530
	Net OOS SR	**0.546**	0.193	0.209	0.445	**0.501**	0.084	0.100	0.439
	Average $\hat{\kappa}$	0	0.482	0.420	0.034	0	0.486	0.425	0.016
16LTANOM	Gross OOS CER	0.066	0.114	**0.119**	0.100	0.095	0.183	**0.187**	0.169
	Net OOS CER	0.060	0.080	**0.089**	0.066	0.091	0.156	**0.162**	0.140
	Gross OOS SR	0.695	0.826	**0.848**	0.787	0.942	1.051	**1.072**	1.039
	Net OOS SR	0.649	0.707	**0.731**	0.630	0.904	0.968	**0.989**	0.928
	Average $\hat{\kappa}$	0	0.552	0.461	0.308	0	0.681	0.633	0.520
46ANOM	Gross OOS CER	0.054	0.312	0.350	**0.882**	**0.098**	-1.264	-1.231	-0.287
	Net OOS CER	0.041	0.225	0.266	**0.772**	**0.088**	-1.297	-1.264	-0.341
	Gross OOS SR	0.643	2.312	2.316	**2.378**	1.037	1.556	**1.558**	**1.558**
	Net OOS SR	0.532	2.191	2.197	**2.247**	0.945	1.463	**1.466**	1.465
	Average $\hat{\kappa}$	0	0.747	0.727	0.482	0	0.832	0.822	0.569

Notes. This table reports the out-of-sample performance of the shrinkage portfolios and the SGMV portfolio. Each estimated portfolio is constructed using a sample size of $T = 120$ or 240 monthly observations. The covariance matrix is estimated using the linear shrinkage estimator of Ledoit and Wolf (2004) and the corresponding optimal shrinkage intensities $\hat{\kappa}_E^\star$ and $\hat{\kappa}_R^\star$ are estimated using the formulas for out-of-sample utility mean and variance developed in Sections IA.8 and IA.9.7. The mean-variance portfolios consider a risk-aversion coefficient of $\gamma = 3$. The proposed robust shrinkage portfolio that maximizes the mean-risk out-of-sample utility is estimated with a coefficient $\lambda = 2$ or calibrated via cross-validation. The table reports the annualized out-of-sample certainty-equivalent return (OOS CER) and the annualized out-of-sample Sharpe ratio (OOS SR), and for both criteria we report the gross performance and the performance net of proportional transaction costs of 20 basis points. We also report the average estimated shrinkage intensity $\hat{\kappa}$ over time. The numbers in bold font identify the best portfolio in terms of OOS CER and SR.

Table IA.9: Performance with simulated data and nonlinear shrinkage covariance matrix

		OOSU mean			OOSU risk			Mean-risk OOSU		
	T	120	180	240	120	180	240	120	180	240
Panel A: Gaussian data (in %)										
10MOM	$\hat{\kappa}_E^\star$	0.63	0.72	0.77	0.30	0.20	0.16	0.02	0.32	0.46
	$\hat{\kappa}_R^\star$	0.63	0.69	0.73	0.23	0.17	0.15	0.16	0.34	0.43
25SBTM	$\hat{\kappa}_E^\star$	0.44	0.62	0.72	0.49	0.30	0.23	-0.55	0.01	0.26
	$\hat{\kappa}_R^\star$	0.48	0.61	0.70	0.37	0.24	0.21	-0.25	0.12	0.28
25OPINV	$\hat{\kappa}_E^\star$	0.70	0.89	1.01	0.45	0.31	0.25	-0.20	0.28	0.52
	$\hat{\kappa}_R^\star$	0.71	0.86	0.98	0.35	0.27	0.25	0.01	0.31	0.49
49IND	$\hat{\kappa}_E^\star$	0.26	0.55	0.71	0.78	0.43	0.31	-1.30	-0.31	0.09
	$\hat{\kappa}_R^\star$	0.41	0.60	0.72	0.58	0.32	0.25	-0.75	-0.05	0.22
16LTANOM	$\hat{\kappa}_E^\star$	1.11	1.38	1.55	0.52	0.37	0.30	0.07	0.65	0.96
	$\hat{\kappa}_R^\star$	1.05	1.33	1.52	0.48	0.38	0.32	0.09	0.56	0.88
46ANOM	$\hat{\kappa}_E^\star$	6.57	8.46	9.59	2.81	1.76	1.28	0.95	4.94	7.02
	$\hat{\kappa}_R^\star$	6.78	8.57	9.64	2.66	1.70	1.25	1.47	5.16	7.13
Panel B: Bootstrapped data (in %)										
10MOM	$\hat{\kappa}_E^\star$	0.42	0.53	0.60	1.10	0.67	0.56	-1.78	-0.80	-0.53
	$\hat{\kappa}_R^\star$	0.55	0.66	0.72	0.44	0.21	0.20	-0.33	0.24	0.33
25SBTM	$\hat{\kappa}_E^\star$	0.07	0.23	0.42	1.20	0.99	0.62	-2.34	-1.76	-0.81
	$\hat{\kappa}_R^\star$	0.46	0.58	0.68	0.35	0.28	0.21	-0.23	0.03	0.26
25OPINV	$\hat{\kappa}_E^\star$	0.37	0.55	0.59	0.94	0.54	0.47	-1.51	-0.53	-0.36
	$\hat{\kappa}_R^\star$	0.72	0.86	0.93	0.34	0.25	0.26	0.04	0.36	0.40
49IND	$\hat{\kappa}_E^\star$	-0.06	0.16	0.25	1.00	0.55	0.37	-2.05	-0.94	-0.49
	$\hat{\kappa}_R^\star$	0.52	0.61	0.65	0.29	0.17	0.19	-0.06	0.27	0.27
16LTANOM	$\hat{\kappa}_E^\star$	0.83	1.00	1.15	0.94	0.66	0.55	-1.04	-0.32	0.06
	$\hat{\kappa}_R^\star$	1.10	1.41	1.62	0.40	0.36	0.32	0.29	0.68	0.98
46ANOM	$\hat{\kappa}_E^\star$	1.60	4.11	5.57	8.69	5.30	3.77	-15.79	-6.49	-1.97
	$\hat{\kappa}_R^\star$	6.40	9.22	10.81	3.95	1.73	1.01	-1.50	5.75	8.79

Notes. This table reports the out-of-sample performance of estimated shrinkage portfolios across six different datasets of simulated data. The first two blocks of three columns report the mean and standard deviation of the monthly out-of-sample utility (in percentage) of the estimated shrinkage portfolios. The third block of three columns reports the mean-risk out-of-sample utility, which is the proposed robustness metric in Section 5. We report the results for the estimated shrinkage portfolio maximizing out-of-sample utility mean ($\hat{\kappa}_E^\star$) and for the estimated shrinkage portfolio maximizing the proposed robustness measure ($\hat{\kappa}_R^\star$). The two portfolios are constructed with the nonlinear shrinkage estimator of the covariance matrix of Ledoit and Wolf (2020). In Panel A, we define the population parameters of a multivariate Gaussian distribution with the sample moments of each of the six datasets of monthly excess returns described in Section 7.1, and draw 100,000 samples of size T. For each simulated sample, we construct the two shrinkage portfolios and their corresponding out-of-sample utilities using Equation (12). Finally, we use the out-of-sample utilities of the 100,000 simulated samples to construct our performance metrics as in (27)-(29). We proceed similarly in Panel B by generating 1,000 bootstrap samples of size $2T$ monthly observations, constructing the two shrinkage portfolios using the first T observations and computing the out-of-sample utility in the remaining T observations, and finally evaluating the performance metrics across the 1,000 simulated bootstrap samples. We consider sample sizes of $T = 120$, 180, and 240 monthly observations, a risk-aversion coefficient of $\gamma = 3$, and a coefficient $\lambda = 2$ for the portfolio robustness measure.

Table IA.10: Performance with empirical data and nonlinear shrinkage covariance matrix

		$T = 120$				$T = 240$			
		SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$		SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$	
				$\lambda = 2$	$\lambda = \hat{\lambda}_{cv}$			$\lambda = 2$	$\lambda = \hat{\lambda}_{cv}$
10MOM	Gross OOS CER	0.062	0.111	0.112	**0.134**	0.070	0.130	0.132	**0.159**
	Net OOS CER	0.055	0.079	0.084	**0.103**	0.065	0.108	0.112	**0.138**
	Gross OOS SR	0.638	0.882	0.863	**0.907**	0.708	0.922	0.919	**0.976**
	Net OOS SR	0.595	0.802	0.786	**0.813**	0.678	0.867	0.866	**0.912**
	Average $\hat{\kappa}$	0	0.475	0.331	0.340	0	0.597	0.494	0.449
25SBTM	Gross OOS CER	0.073	0.122	0.127	**0.128**	0.086	0.129	0.129	**0.153**
	Net OOS CER	0.060	0.057	**0.073**	0.066	0.077	0.085	0.090	**0.114**
	Gross OOS SR	0.752	0.881	0.875	**0.884**	0.875	0.898	0.884	**0.981**
	Net OOS SR	0.654	0.697	**0.698**	0.641	0.804	0.774	0.764	**0.828**
	Average $\hat{\kappa}$	0	0.346	0.230	0.208	0	0.456	0.362	0.303
25OPINV	Gross OOS CER	0.079	0.085	0.090	**0.093**	0.105	0.125	**0.133**	**0.133**
	Net OOS CER	**0.071**	0.043	0.058	0.055	0.099	0.100	**0.112**	0.111
	Gross OOS SR	**0.792**	0.731	0.736	0.754	**0.994**	0.866	0.928	0.972
	Net OOS SR	**0.726**	0.589	0.597	0.577	**0.953**	0.775	0.835	0.858
	Average $\hat{\kappa}$	0	0.310	0.183	0.163	0	0.401	0.290	0.216
49IND	Gross OOS CER	0.057	0.034	0.047	**0.058**	**0.047**	0.018	0.038	0.043
	Net OOS CER	**0.048**	0.009	0.030	0.038	**0.041**	0.004	0.027	0.033
	Gross OOS SR	**0.659**	0.475	0.536	0.610	**0.581**	0.370	0.480	0.545
	Net OOS SR	**0.583**	0.355	0.426	0.476	**0.528**	0.292	0.405	0.460
	Average $\hat{\kappa}$	0	0.125	0.043	0.038	0	0.139	0.058	0.018
16LTANOM	Gross OOS CER	0.064	0.096	**0.099**	0.087	0.094	0.173	**0.179**	0.172
	Net OOS CER	0.054	0.045	**0.056**	0.042	0.088	0.137	**0.145**	0.137
	Gross OOS SR	0.671	**0.784**	0.774	0.731	0.924	1.020	1.038	**1.053**
	Net OOS SR	0.599	**0.625**	0.618	0.515	0.877	0.919	**0.936**	0.919
	Average $\hat{\kappa}$	0	0.417	0.284	0.197	0	0.603	0.527	0.418
46ANOM	Gross OOS CER	0.045	-1.081	-1.004	**0.827**	**0.095**	-2.406	-2.349	-0.181
	Net OOS CER	0.027	-1.008	-0.926	**0.689**	**0.082**	-2.351	-2.294	-0.242
	Gross OOS SR	0.564	**2.299**	2.296	2.295	0.980	1.515	**1.517**	1.478
	Net OOS SR	0.408	2.154	**2.156**	2.123	0.869	1.411	**1.414**	1.372
	Average $\hat{\kappa}$	0	0.673	0.640	0.326	0	0.788	0.771	0.429

Notes. This table reports the out-of-sample performance of four portfolio strategies described in Section 7.3 for the six datasets of monthly excess returns described in Section 7.1. Each estimated portfolio is constructed using a sample size of $T = 120$ or 240 monthly observations. The covariance matrix is estimated using the nonlinear shrinkage estimator of Ledoit and Wolf (2020). The mean-variance portfolios consider a risk-aversion coefficient of $\gamma = 3$. The proposed robust shrinkage portfolio that maximizes the mean-risk out-of-sample utility is estimated with a coefficient λ equal to two or calibrated via five-fold cross-validation. The table reports the annualized out-of-sample certainty-equivalent return (OOS CER) and the annualized out-of-sample Sharpe ratio (OOS SR), and for both criteria we report the gross performance and the performance net of proportional transaction costs of 20 basis points. We also report the average estimated shrinkage intensity $\hat{\kappa}$ over time. The numbers in bold font identify the best portfolio in terms of OOS CER and SR.

Table IA.11: Performance of shrinkage portfolios with simulated data in high-dimensional dataset

	OOSU mean		OOSU risk		Mean-risk OOSU	
	$T = 180$	$T = 240$	$T = 180$	$T = 240$	$T = 180$	$T = 240$
Panel A) Gaussian data (in %)						
$\hat{\kappa}_E^\star$	-0.10	6.86	9.69	4.68	-19.47	-2.51
$\hat{\kappa}_R^\star$	0.62	7.22	9.49	4.59	-18.36	-1.96
Panel B) Bootstrapped data (in %)						
$\hat{\kappa}_E^\star$	-4.62	-0.88	11.95	6.98	-28.53	-14.84
$\hat{\kappa}_R^\star$	6.96	12.28	5.03	1.86	-3.09	8.55

Notes. This table reports the out-of-sample performance of estimated shrinkage portfolios across simulated data from all six datasets merged together, which corresponds to $N = 155$ assets. The first two blocks of three columns report the mean and standard deviation of the monthly out-of-sample utility (in percentage) of the estimated shrinkage portfolios. The third block of three columns reports the mean-risk out-of-sample utility, which is the proposed robustness metric in Section 5. We report the results for the estimated shrinkage portfolio maximizing out-of-sample utility mean ($\hat{\kappa}_E^\star$) and for the estimated shrinkage portfolio maximizing the proposed robustness measure ($\hat{\kappa}_R^\star$). The two portfolios are constructed with the nonlinear shrinkage estimator of the covariance matrix of Ledoit and Wolf (2020). In Panel A, we define the population parameters of a multivariate Gaussian distribution with the sample moments of the high-dimensional dataset of monthly excess returns, and draw 100,000 samples of size T. For each simulated sample, we construct the two shrinkage portfolios and their corresponding out-of-sample utilities using Equation (12). Finally, we use the out-of-sample utilities of the 100,000 simulated samples to construct our performance metrics as in (27)-(29). We proceed similarly in Panel B by generating 1,000 bootstrap samples of size $2T$ monthly observations, constructing the two shrinkage portfolios using the first T observations and computing the out-of-sample utility in the remaining T observations, and finally evaluating the performance metrics across the 1,000 simulated bootstrap samples. We consider sample sizes of $T = 180$ and 240 monthly observations, a risk-aversion coefficient of $\gamma = 3$, and a coefficient $\lambda = 2$ for the portfolio robustness measure.

Table IA.12: Performance of shrinkage portfolios with empirical data in high-dimensional dataset

		EW	RTR	SGMV	SMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$	
							$\lambda = 2$	$\lambda = \hat{\lambda}_{cv}$
$T = 180$	Gross OOS CER	0.044	0.058	0.051	-13.96	-4.760	-4.432	**0.383**
	Net OOS CER	0.042	0.056	0.005	-11.23	-4.247	-3.970	**0.100**
	Gross OOS SR	0.512	0.610	0.583	1.528	1.596	1.600	**1.725**
	Net OOS SR	0.505	0.602	0.238	1.080	1.183	1.188	**1.326**
	Average $\hat{\kappa}$	/	/	0	1	0.577	0.553	0.194
$T = 240$	Gross OOS CER	0.044	0.058	0.074	-12.85	-5.312	-5.044	**0.178**
	Net OOS CER	0.043	**0.057**	0.047	-11.14	-4.879	-4.640	0.005
	Gross OOS SR	0.513	0.606	0.865	1.415	1.491	1.497	**1.583**
	Net OOS SR	0.506	0.600	0.609	1.166	1.259	1.267	**1.345**
	Average $\hat{\kappa}$	/	/	0	1	0.630	0.609	0.233

Notes. This table reports the out-of-sample performance of six portfolio strategies described in Section 7.3 for a dataset that merges all six datasets of monthly excess returns described in Section 7.1, which corresponds to $N = 155$ assets. Each estimated portfolio is constructed using a sample size of $T = 180$ or 240 monthly observations. The covariance matrix is estimated using the nonlinear shrinkage estimator of Ledoit and Wolf (2020). The mean-variance portfolios consider a risk-aversion coefficient of $\gamma = 3$. The proposed robust shrinkage portfolio that maximizes the mean-risk out-of-sample utility is estimated with a coefficient λ equal to two or calibrated via five-fold cross-validation. The table reports the annualized out-of-sample certainty-equivalent return (OOS CER) and the annualized out-of-sample Sharpe ratio (OOS SR), and for both criteria we report the gross performance and the performance net of proportional transaction costs of 20 basis points. We also report the average estimated shrinkage intensity $\hat{\kappa}$ over time, except for the EW and RTR portfolios that do not combine the SMV and SGMV portfolios. The numbers in bold font identify the best portfolio in terms of OOS CER and SR.

Table IA.13: Effect of estimation errors in shrinkage intensities on out-of-sample performance

		OOSU mean			OOSU risk			Mean-risk OOSU		
	T	120	180	240	120	180	240	120	180	240
Panel A: Estimated shrinkage intensities (in %)										
10MOM	$\hat{\kappa}_E^\star$	0.63	0.72	0.77	0.27	0.18	0.15	0.09	0.35	0.47
	$\hat{\kappa}_R^\star$	0.66	0.72	0.76	0.20	0.15	0.13	0.27	0.43	0.50
25SBTM	$\hat{\kappa}_E^\star$	0.45	0.61	0.71	0.37	0.24	0.20	-0.28	0.13	0.31
	$\hat{\kappa}_R^\star$	0.50	0.62	0.70	0.25	0.19	0.17	0.01	0.24	0.36
25OPINV	$\hat{\kappa}_E^\star$	0.67	0.85	0.98	0.38	0.27	0.23	-0.10	0.32	0.52
	$\hat{\kappa}_R^\star$	0.70	0.84	0.95	0.26	0.22	0.21	0.17	0.40	0.52
49IND	$\hat{\kappa}_E^\star$	0.33	0.56	0.70	0.44	0.29	0.23	-0.56	-0.02	0.24
	$\hat{\kappa}_R^\star$	0.40	0.57	0.69	0.26	0.20	0.19	-0.11	0.17	0.31
16LTANOM	$\hat{\kappa}_E^\star$	1.11	1.37	1.54	0.42	0.34	0.29	0.27	0.70	0.97
	$\hat{\kappa}_R^\star$	1.06	1.32	1.50	0.37	0.34	0.31	0.32	0.63	0.89
46ANOM	$\hat{\kappa}_E^\star$	5.84	8.04	9.31	1.48	1.11	0.91	2.89	5.81	7.48
	$\hat{\kappa}_R^\star$	5.70	7.96	9.27	1.27	1.03	0.87	3.17	5.90	7.54
Panel B: Known shrinkage intensities (in %)										
10MOM	$\kappa_E^\star$	0.73	0.78	0.82	0.13	0.12	0.11	0.46	0.53	0.59
	$\kappa_R^\star$	0.69	0.75	0.79	0.09	0.09	0.09	0.50	0.57	0.62
25SBTM	$\kappa_E^\star$	0.57	0.69	0.78	0.17	0.16	0.16	0.23	0.36	0.46
	$\kappa_R^\star$	0.54	0.66	0.75	0.14	0.13	0.13	0.27	0.40	0.49
25OPINV	$\kappa_E^\star$	0.80	0.95	1.06	0.21	0.20	0.19	0.38	0.55	0.68
	$\kappa_R^\star$	0.74	0.90	1.03	0.15	0.16	0.16	0.44	0.59	0.71
49IND	$\kappa_E^\star$	0.48	0.65	0.78	0.17	0.17	0.17	0.13	0.31	0.44
	$\kappa_R^\star$	0.45	0.62	0.74	0.14	0.13	0.14	0.17	0.35	0.47
16LTANOM	$\kappa_E^\star$	1.25	1.47	1.62	0.33	0.28	0.25	0.60	0.90	1.12
	$\kappa_R^\star$	1.19	1.44	1.60	0.26	0.25	0.23	0.67	0.94	1.14
46ANOM	$\kappa_E^\star$	6.10	8.18	9.40	1.27	1.05	0.87	3.56	6.09	7.65
	$\kappa_R^\star$	5.94	8.10	9.35	1.06	0.95	0.82	3.82	6.21	7.72

Notes. This table reports the out-of-sample performance of shrinkage portfolios across six different datasets of simulated data. The first two blocks of three columns report the mean and standard deviation of the monthly out-of-sample utility (in percentage) of the shrinkage portfolios. The third block of three columns reports the mean-risk out-of-sample utility, which is the proposed robustness metric in Section 5. We define the population parameters of a multivariate Gaussian distribution with the sample moments of each of the six datasets of monthly excess returns described in Section 7.1, and draw 100,000 samples of size T. For each simulated sample, we construct the two shrinkage portfolios and their corresponding out-of-sample utilities using Equation (12). Finally, we use the out-of-sample utilities of the 100,000 simulated samples to construct our performance metrics as in (27)-(29). In Panel A, we report the results for the shrinkage portfolio maximizing out-of-sample utility mean ($\hat{\kappa}_E^\star$) and for the shrinkage portfolio maximizing the proposed robustness measure ($\hat{\kappa}_R^\star$) when shrinkage intensities are estimated. In Panel B, we report results when the shrinkage intensities are known without estimation error. We consider sample sizes of $T = 120$, 180, and 240 monthly observations, a risk-aversion coefficient of $\gamma = 3$, and a coefficient $\lambda = 2$ for the portfolio robustness measure.

IA.10 Proofs of all theoretical results

Throughout the proofs presented in this section, we use the fact that the shrinkage portfolio $\hat{w}^\star(\kappa)$ in (11) can be rewritten as

$$\hat{w}^\star(\kappa) = \hat{w}_g + \frac{\kappa}{\gamma}\hat{w}_z. \tag{IA41}$$

Proof of Proposition 1

The out-of-sample utility variance of a random vector of portfolio weights $\hat{w}$ is

$$\mathbb{V}[U(\hat{w})] = \mathbb{V}\Big[\hat{w}^\top\mu\Big] + \frac{\gamma^2}{4}\mathbb{V}\Big[\hat{w}^\top\mathbf{\Sigma}\hat{w}\Big] - \gamma\,\mathbb{C}\text{ov}\Big[\hat{w}^\top\mu, \hat{w}^\top\mathbf{\Sigma}\hat{w}\Big]. \tag{IA42}$$

We prove this proposition by characterizing each of the three terms in (IA42) for the shrinkage portfolio $\hat{w}^\star(\kappa)$ that combines the SMV and SGMV portfolios.

Expression for $\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top\mu\Big]$. Okhrin and Schmid (2006, Theorem 1) show that the covariance matrix of $\hat{w}^\star(\kappa)$ is

$$\begin{aligned}\mathbb{V}[\hat{w}^\star(\kappa)] = &\left(\frac{\sigma_g^2}{T-N-1} + \frac{\kappa^2}{\gamma^2}\frac{T(T-2+T\psi^2)}{(T-N)(T-N-1)(T-N-3)}\right)\mathbf{B}\\ &+ \frac{\kappa^2}{\gamma^2}\frac{T^2(T-N+1)}{(T-N)(T-N-1)^2(T-N-3)}\mathbf{B}\mu\mu^\top\mathbf{B}.\end{aligned} \tag{IA43}$$

Therefore, given that $\mu^\top\mathbf{B}\mu = \psi^2$, the variance of out-of-sample mean return is

$$\begin{aligned}\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top\mu\Big] &= \mu^\top\mathbb{V}[\hat{w}^\star(\kappa)]\mu\\ &= \frac{\sigma_g^2\psi^2}{T-N-1} + \frac{\kappa^2\psi^2}{\gamma^2}\frac{2T(N+1)+T^2(T-N-3+2(T-N)\psi^2)}{(T-N)(T-N-1)^2(T-N-3)}.\end{aligned} \tag{IA44}$$

For the next two terms in (IA42) we use Kan et al. (2022b, Proposition 1) who derive a stochastic representation for the out-of-sample mean return and variance of the shrinkage portfolio $\hat{w}^\star(\kappa)$, i.e., for the two random variables $\hat{w}^\star(\kappa)^\top\mu$ and $\hat{w}^\star(\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\kappa)$. In the next proposition, we review their result using our notation.

Kan et al. (2022b, Proposition 1). *Let* $z_2 \sim \mathcal{N}(\sqrt{T}\psi, 1)$, $u_0 \sim \chi^2_{N-2}$, $v_2 \sim \chi^2_{T-N+1}$, $w_1 \sim \chi^2_{T-N+3}$, $w_2 \sim \chi^2_{T-N+2}$, $s_1 \sim \chi^2_{N-4}$, $s_2 \sim \chi^2_{N-3}$, $x_{11} \sim \mathcal{N}(0,1)$, $x_{21} \sim \mathcal{N}(0,1)$, $a \sim \mathcal{N}(0,1)$,

$b \sim \mathcal{N}(0,1)$, $c \sim \mathcal{N}(0,1)$, *and they are mutually independent.*[6] *Define*

$$y_1 = \frac{x_{11}}{\sqrt{w_1}} + \frac{b x_{21}}{\sqrt{w_1 w_2}} + \frac{a x_{21}}{\sqrt{v_2 w_2}}, \tag{IA45}$$

$$y_2 = \frac{c}{\sqrt{w_1}} + \frac{b\sqrt{s_2}}{\sqrt{w_1 w_2}} + \frac{a\sqrt{s_2}}{\sqrt{v_2 w_2}}. \tag{IA46}$$

Then, the out-of-sample mean return and variance of the shrinkage portfolio $\hat{w}^\star(\kappa)$ *are distributed as*

$$\hat{w}^\star(\kappa)^\top \mu = \mu_g + \frac{\sigma_g \psi \sqrt{v_2}}{\sqrt{z_2^2 + u_0}} \left(\frac{\sqrt{u_0} y_1}{\sqrt{v_2}} + \frac{a z_2}{v_2} \right) + \frac{\kappa \sqrt{T} \psi}{\gamma v_2} \left(\frac{x_{21}\sqrt{u_0}}{\sqrt{w_2}} + z_2 \right), \tag{IA47}$$

$$\begin{aligned} \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa) = \sigma_g^2 \left(1 + y_1^2 + y_2^2 + \frac{s_1}{w_1} + \frac{a^2}{v_2} \right) + \frac{\kappa^2 T (z_2^2 + u_0)}{\gamma^2 v_2^2} \left(1 + \frac{x_{21}^2 + s_2}{w_2} \right) \\ + \frac{2\kappa\sqrt{T}\sigma_g \sqrt{z_2^2 + u_0}}{\gamma v_2} \left(\frac{a}{\sqrt{v_2}} + \frac{x_{21} y_1}{\sqrt{w_2}} + \frac{\sqrt{s_2} y_2}{\sqrt{w_2}} \right). \end{aligned} \tag{IA48}$$

Using this result, we now find analytical expressions for the variance of $\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)$ and the covariance between $\hat{w}^\star(\kappa)^\top \mu$ and $\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)$.

Expression for $\mathbb{V}\left[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\right]$**.** From Kan et al. (2022b), we know that the expectation of $\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)$ is

$$\mathbb{E}\left[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\right] = \sigma_g^2 \frac{T-2}{T-N-1} + \frac{\kappa^2}{\gamma^2} \frac{T(T-2)(N-1+T\psi^2)}{(T-N)(T-N-1)(T-N-3)}. \tag{IA49}$$

Therefore, we are left with deriving an expression for $\mathbb{E}\left[\left(\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\right)^2\right]$. Using the stochastic representation in (IA48) and dropping expectations that are equal to zero when expanding the square, it corresponds to

$$\begin{aligned} \mathbb{E}\left[\left(\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\right)^2\right] = \sigma_g^4 \mathbb{E}\left[\left(1 + y_1^2 + y_2^2 + \frac{s_1}{w_1} + \frac{a^2}{v_2} \right)^2 \right] \\ + \frac{4\kappa^2}{\gamma^2} T \sigma_g^2 \mathbb{E}\left[z_2^2 + u_0\right] \mathbb{E}\left[\frac{1}{v_2^2} \left(\frac{a}{\sqrt{v_2}} + \frac{x_{21} y_1}{\sqrt{w_2}} + \frac{\sqrt{s_2} y_2}{\sqrt{w_2}} \right)^2 \right] \\ + \frac{2\kappa^2}{\gamma^2} T \sigma_g^2 \mathbb{E}\left[z_2^2 + u_0\right] \mathbb{E}\left[\frac{1}{v_2^2} \left(1 + \frac{x_{21}^2 + s_2}{w_2} \right) \right. \end{aligned}$$

[6] When $N = 3$, set s_1, s_2, and c equal to zero. When $N = 2$, set u_0, x_{11}, x_{21}, s_1, s_2, and c equal to zero.

$$\times \left(1 + y_1^2 + y_2^2 + \frac{s_1}{w_1} + \frac{a^2}{v_2}\right)\Bigg] + \frac{\kappa^4}{\gamma^4} T^2 \mathbb{E}\Big[(z_2^2 + u_0)^2\Big] \mathbb{E}\left[\frac{1}{v_2^4}\left(1 + \frac{x_{21}^2 + s_2}{w_2}\right)^2\right]. \tag{IA50}$$

After some developments, the different expectations in (IA50) are equal to

$$\mathbb{E}\Big[z_2^2 + u_0\Big] = N - 1 + T\psi^2, \tag{IA51}$$

$$\mathbb{E}\Big[(z_2^2 + u_0)^2\Big] = (N+1)(N-1) + 2T(N+1)\psi^2 + T^2\psi^4, \tag{IA52}$$

$$\mathbb{E}\left[\left(1 + y_1^2 + y_2^2 + \frac{s_1}{w_1} + \frac{a^2}{v_2}\right)^2\right] = \frac{(T-2)(T-4)}{(T-N-1)(T-N-3)}, \tag{IA53}$$

$$\mathbb{E}\left[\frac{1}{v_2^2}\left(\frac{a}{\sqrt{v_2}} + \frac{x_{21}y_1}{\sqrt{w_2}} + \frac{\sqrt{s_2}y_2}{\sqrt{w_2}}\right)^2\right] = \frac{(T-2)(T+N-3)}{(T-N+1)(T-N)(T-N-1)(T-N-3)(T-N-5)}, \tag{IA54}$$

$$\mathbb{E}\left[\frac{1}{v_2^2}\left(1 + y_1^2 + y_2^2 + \frac{s_1}{w_1} + \frac{a^2}{v_2}\right)\right] = \frac{(T-2)(T-N-4)}{(T-N)(T-N-1)(T-N-3)(T-N-5)}, \tag{IA55}$$

$$\mathbb{E}\left[\frac{1}{v_2^2 w_2}\left(x_{21}^2 + s_2\right)\left(1 + y_1^2 + y_2^2 + \frac{s_1}{w_1} + \frac{a^2}{v_2}\right)\right] = \frac{(T-2)(N-2)}{(T-N+1)(T-N)(T-N-1)(T-N-5)}, \tag{IA56}$$

$$\mathbb{E}\left[\frac{1}{v_2^4}\left(1 + \frac{x_{21}^2 + s_2}{w_2}\right)^2\right] = \frac{(T-2)(T-4)}{(T-N)(T-N-1)(T-N-2)(T-N-3)(T-N-5)(T-N-7)}. \tag{IA57}$$

Plugging (IA51)–(IA57) into (IA50), we obtain

$$\mathbb{E}\Big[\Big(\hat{w}^\star(\kappa)^\top \boldsymbol{\Sigma} \hat{w}^\star(\kappa)\Big)^2\Big] = \frac{(T-2)(T-4)}{(T-N)(T-N-2)}\Bigg(\sigma_g^4 \frac{(T-N)(T-N-2)}{(T-N-1)(T-N-3)} + \frac{2\kappa^2\sigma_g^2}{\gamma^2}\frac{T(T-N-2)(N-1+T\psi^2)}{(T-N-1)(T-N-3)(T-N-5)} + \frac{\kappa^4}{\gamma^4}\frac{T^2(T^2\psi^4 + 2(N+1)T\psi^2 + (N+1)(N-1))}{(T-N-1)(T-N-3)(T-N-5)(T-N-7)}\Bigg). \tag{IA58}$$

Finally, combining (IA49) and (IA58), the variance of out-of-sample return variance is

$$\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \boldsymbol{\Sigma} \hat{w}^\star(\kappa)\Big] = 2\sigma_g^4 \frac{(N-1)(T-2)}{(T-N-1)^2(T-N-3)}$$

$$+\frac{4\kappa^2\sigma_g^2}{\gamma^2}\frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}$$
$$+\frac{2\kappa^4}{\gamma^4}\frac{T^2(T-2)C(T,N,\psi^2)}{(T-N)^2(T-N-1)^2(T-N-2)(T-N-3)^2(T-N-5)(T-N-7)}, \tag{IA59}$$

where $C(T,N,\psi^2)$ is defined in Proposition 1.

Expression for $\mathbb{C}\text{ov}\Big[\hat{w}^\star(\kappa)^\top\mu, \hat{w}^\star(\kappa)^\top\boldsymbol{\Sigma}\hat{w}^\star(\kappa)\Big]$. We know that the expectation of $\hat{w}^\star(\kappa)^\top\boldsymbol{\Sigma}\hat{w}^\star(\kappa)$ is given by (IA49) and, from Kan et al. (2022b), the expectation of $\hat{w}^\star(\kappa)^\top\mu$ is

$$\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top\mu\Big] = \mu_g + \frac{\kappa}{\gamma}\frac{T\psi^2}{T-N-1}. \tag{IA60}$$

Therefore, we are left with deriving an expression for $\mathbb{E}\Big[\Big(\hat{w}^\star(\kappa)^\top\mu\Big)\Big(\hat{w}^\star(\kappa)^\top\boldsymbol{\Sigma}\hat{w}^\star(\kappa)\Big)\Big]$. Using the stochastic representations in (IA47)–(IA48) and dropping expectations that are equal to zero when expanding the product, it corresponds to

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\hat{w}^\star(\kappa)^\top\mu\Big)\Big(\hat{w}^\star(\kappa)^\top\boldsymbol{\Sigma}\hat{w}^\star(\kappa)\Big)\Big] \\
&= \mu_g\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top\boldsymbol{\Sigma}\hat{w}^\star(\kappa)\Big] + \frac{2\kappa}{\gamma}\sqrt{T}\sigma_g^2\psi\mathbb{E}[z_2]\mathbb{E}\left[\frac{a}{v_2^{3/2}}\left(\frac{a}{\sqrt{v_2}} + \frac{x_{21}y_1}{\sqrt{w_2}} + \frac{\sqrt{s_2}y_2}{\sqrt{w_2}}\right)\right] \\
&\quad + \frac{\kappa}{\gamma}\sqrt{T}\sigma_g^2\psi\mathbb{E}[z_2]\mathbb{E}\left[\frac{1}{v_2}\left(1 + y_1^2 + y_2^2 + \frac{s_1}{w_1} + \frac{a^2}{v_2}\right)\right] \\
&\quad + \frac{\kappa^3}{\gamma^3}T^{3/2}\psi\mathbb{E}\Big[z_2(z_2^2+u_0)\Big]\mathbb{E}\left[\frac{1}{v_2^3}\left(1 + \frac{x_{21}^2+s_2}{w_2}\right)\right].
\end{aligned} \tag{IA61}$$

In addition to $\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top\boldsymbol{\Sigma}\hat{w}^\star(\kappa)\Big]$ that is available in (IA49), we find after some developments that the different elements in (IA42) are equal to

$$\mathbb{E}[z_2] = \sqrt{T}\psi, \tag{IA62}$$
$$\mathbb{E}[z_2(z_2^2+u_0)] = \sqrt{T}\psi(N+1+T\psi^2), \tag{IA63}$$
$$\mathbb{E}\left[\frac{a}{v_2^{3/2}}\left(\frac{a}{\sqrt{v_2}} + \frac{x_{21}y_1}{\sqrt{w_2}} + \frac{\sqrt{s_2}y_2}{\sqrt{w_2}}\right)\right] = \frac{T-2}{(T-N)(T-N-1)(T-N-3)}, \tag{IA64}$$
$$\mathbb{E}\left[\frac{1}{v_2}\left(1 + y_1^2 + y_2^2 + \frac{s_1}{w_1} + \frac{a^2}{v_2}\right)\right] = \frac{(T-2)(T-N-2)}{(T-N)(T-N-1)(T-N-3)}, \tag{IA65}$$
$$\mathbb{E}\left[\frac{1}{v_2^3}\left(1 + \frac{x_{21}^2+s_2}{w_2}\right)\right] = \frac{T-2}{(T-N)(T-N-1)(T-N-3)(T-N-5)}. \tag{IA66}$$

Plugging (IA62)–(IA66) into (IA61), we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\hat{w}^\star(\kappa)^\top \mu\Big)\Big(\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big)\Big] \\
&= \mu_g \sigma_g^2 \frac{T-2}{T-N-1} + \frac{\kappa^2 \mu_g}{\gamma^2} \frac{T(T-2)(N-1+T\psi^2)}{(T-N)(T-N-1)(T-N-3)} + \frac{\kappa \sigma_g^2 \psi^2}{\gamma} \frac{T(T-2)}{(T-N-1)(T-N-3)} \\
&\quad + \frac{\kappa^3 \psi^2}{\gamma^3} \frac{T^2(T-2)(N+1+T\psi^2)}{(T-N)(T-N-1)(T-N-3)(T-N-5)}.
\end{aligned} \tag{IA67}$$

Finally, combining (IA49), (IA60), and (IA67), the covariance between the out-of-sample mean return and variance is

$$\begin{aligned}
&\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big] = \frac{2\kappa \sigma_g^2 \psi^2}{\gamma} \frac{T(T-2)}{(T-N-1)^2(T-N-3)} \\
&+ \frac{2\kappa^3 \psi^2}{\gamma^3} \frac{T^2(T-2)(T+N-3+2T\psi^2)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}.
\end{aligned} \tag{IA68}$$

We find the final expression for the OOSU variance in Proposition 1 by plugging (IA44), (IA59), and (IA68) into (IA42), which concludes the proof.

Proof of Corollary 1

First, we prove that the shrinkage intensity that minimizes OOSU variance, $\kappa_V^\star$, is strictly positive. The derivative of the OOSU variance in (13) with respect to κ, evaluated at $\kappa = 0$, is

$$\frac{\partial \mathbb{V}[U(\hat{w}^\star(\kappa))]}{\partial \kappa}\bigg|_{\kappa=0} = a_4. \tag{IA69}$$

Now, observe that a_4 in Equation (19) is strictly negative when $\psi^2 > 0$ because $\sigma_g^2 > 0$. Therefore, provided that $\psi^2 > 0$, it is always optimal to choose a shrinkage intensity $\kappa > 0$ to minimize OOSU variance.

Second, we prove that $\kappa_V^\star < 1$, which follows from Proposition 4 where we show that $\kappa_V^\star \leq \kappa_E^\star$. Indeed, because $\kappa_E^\star < 1$ as long as the sample size T is finite, it follows that $\kappa_V^\star < 1$.

Proof of Proposition 2

Parameters T and N. From the closed-form expression of $\mathbb{V}[U(\hat{w}^\star(\kappa))]$ in Proposition 1, it is straightforward to see that it is decreasing in T and increasing in N. In particular, it is easy to check that $\mathbb{V}[U(\hat{w}^\star(\kappa))] \to 0$ as $T \to \infty$ for any shrinkage intensity κ.

Parameter σ_g^2. The derivative of the OOSU variance with respect to σ_g^2 is

$$\frac{\partial \mathbb{V}[U(\hat{w}^\star(\kappa))]}{\partial \sigma_g^2} = \frac{\psi^2}{T-N-1} + \sigma_g^2\gamma^2 \frac{(N-1)(T-2)}{(T-N-1)^2(T-N-3)} + \kappa^2 \frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} - 2\kappa\psi^2 \frac{T(T-2)}{(T-N-1)^2(T-N-3)}. \quad \text{(IA70)}$$

The objective is to show that the derivative in (IA70) is always positive. Notice that it increases with σ_g^2, and thus it suffices to show that it is always positive for the case $\sigma_g^2 = 0$. That is, we need to show that

$$\frac{\psi^2}{T-N-1} + \kappa^2 \frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} - 2\kappa\psi^2 \frac{T(T-2)}{(T-N-1)^2(T-N-3)} \geq 0. \quad \text{(IA71)}$$

Notice that the left-hand side of (IA71) is a second-degree polynomial in κ. Because the coefficient in front of κ^2 is positive, we can prove that inequality (IA71) holds by showing that the polynomial discriminant is always negative. That is, after some simplifications,

$$\psi^2 \frac{T(T-2)}{(T-N-1)(T-N-3)} \leq \frac{(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-5)}. \quad \text{(IA72)}$$

Notice that the right-hand side of inequality (IA72) is of the form $a + b\psi^2$ with $a > 0$. Therefore, we can prove the inequality by showing that the coefficient multiplying ψ^2 on the right-hand side is larger than that in front of ψ^2 on the left-hand side. This is equivalent to showing that

$$\frac{T(T+N-3)}{(T-N)(T-N-5)} \geq \frac{T(T-2)}{(T-N-1)(T-N-3)}, \quad \text{(IA73)}$$

which holds under Assumption 1.

Parameter ψ^2. The derivative of the OOSU variance with respect to ψ^2 is

$$\frac{\partial \mathbb{V}[U(\hat{w}^\star(\kappa))]}{\partial \psi^2} = \frac{\sigma_g^2}{T-N-1} + \frac{\kappa^4}{2\gamma^2} \frac{T^2(T-2)\frac{\partial C(T,N,\psi^2)}{\partial \psi^2}}{(T-N)^2(T-N-1)^2(T-N-2)(T-N-3)^2(T-N-5)(T-N-7)} - \frac{2\kappa^3}{\gamma^2} \frac{4\psi^2T^3(T-2)+T^2(T-2)(T+N-3)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}$$

$$
\begin{aligned}
&+ \frac{\kappa^2}{\gamma^2}\frac{2T(N+1)+T^2(T-N-3)+4T^2(T-N)\psi^2}{(T-N)(T-N-1)^2(T-N-3)} \\
&+ \kappa^2\sigma_g^2\frac{T^2(T-2)(T+N-3)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} \\
&- 2\kappa\sigma_g^2\frac{T(T-2)}{(T-N-1)^2(T-N-3)}.
\end{aligned} \tag{IA74}
$$

First, we show that the derivative (IA74) is increasing in σ_g^2 and thus that it suffices to show that it is positive for $\sigma_g^2 = 0$. We have

$$
\begin{aligned}
\frac{\partial}{\partial\sigma_g^2}\left(\frac{\partial\mathbb{V}[U(\hat{w}^\star(\kappa))]}{\partial\psi^2}\right) = \frac{1}{T-N-1} &+ \kappa^2\frac{T^2(T-2)(T+N-3)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} \\
&- 2\kappa\frac{T(T-2)}{(T-N-1)^2(T-N-3)}.
\end{aligned} \tag{IA75}
$$

Following a similar strategy to the case with σ_g^2 as a parameter, the derivative (IA75) is always positive if the polynomial discriminant is negative. This amounts to showing, after some simplifications, that

$$
\frac{(T+N-3)(T-N-1)(T-N-3)}{(T-2)(T-N)(T-N-5)} \geq 1,
$$

which holds under Assumption 1. Therefore, we can now prove the result of the proposition by showing that the derivative in (IA74) is positive for $\sigma_g^2 = 0$. That is, after some simplifications,

$$
\begin{aligned}
&\frac{\kappa^2}{2}\frac{T^2(T-2)\frac{\partial C(T,N,\psi^2)}{\partial\psi^2}}{(T-N)(T-N-2)(T-N-3)(T-N-7)} \\
&- 2\kappa\Big(4T^3(T-2)\psi^2 + T^2(T-2)(T+N-3)\Big) \\
&+ (T-N-5)\Big(2T(N+1)+T^2(T-N-3)+4T^2(T-N)\psi^2\Big) \geq 0.
\end{aligned} \tag{IA76}
$$

We find that the derivative of (IA76) with respect to ψ^2 is positive if

$$
\frac{\partial^2 C(T,N,\psi^2)}{\partial(\psi^2)^2} \geq \frac{8T^2(T-2)(T-N-2)(T-N-3)(T-N-7)}{(T-N-5)}, \tag{IA77}
$$

where $\frac{\partial^2 C(T,N,\psi^2)}{\partial(\psi^2)^2} = 2T^2(N^3+2N^2T-6N^2-7NT^2+40NT-53N+4T^3-34T^2+88T-70)$, and inequality (IA77) holds under Assumption 1. Therefore, we can now prove the result of the

proposition by showing that the derivative in (IA76) is positive for $\psi^2 = 0$. That is,

$$\frac{\kappa^2}{2}\frac{T^2(T-2)\frac{\partial C(T,N,\psi^2)}{\partial \psi^2}\Big|_{\psi^2=0}}{(T-N)(T-N-2)(T-N-3)(T-N-7)} - 2\kappa T^2(T-2)(T+N-3)$$
$$+ (T-N-5)(2T(N+1)+T^2(T-N-3)) \geq 0. \tag{IA78}$$

As usual, we prove inequality (IA78) by showing that the polynomial discriminant is negative, which amounts to showing, after some simplifications, that

$$\frac{\partial C(T,N,\psi^2)}{\partial \psi^2}\Big|_{\psi^2=0} \geq \frac{2T(T-2)(T-N)(T-N-2)(T-N-3)(T+N-3)^2}{(T-N-5)(2(N+1)+T(T-N-3))}, \tag{IA79}$$

which holds under Assumption 1, thus concluding the proof for parameter ψ^2.

Parameter κ. To prove the result, we need to show that the derivative of the OOSU variance with respect to κ is positive if $\kappa \geq \kappa_E^\star$. That is,

$$4a_1\kappa^3 + 3a_2\kappa^2 + 2a_3\kappa + a_4 \geq 0 \tag{IA80}$$

if $\kappa \geq \kappa_E^\star$. As we show below, the derivative in (IA80) decreases with γ when $\kappa \geq \kappa_E^\star$. Therefore, we can derive a sufficient condition on the value of κ for which inequality (IA80) holds by considering the case $\gamma \to \infty$. In that case, the condition in (IA80) becomes

$$2\kappa\sigma_g^2\frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} - 2\sigma_g^2\psi^2\frac{T(T-2)}{(T-N-1)^2(T-N-3)} \geq 0,$$

which after isolating κ reduces to the sufficient condition

$$\kappa \geq \kappa_E^\star\frac{(T-2)(T-N-5)}{(T+N-3)(T-N-3)}, \tag{IA81}$$

which is satisfied when $\kappa \geq \kappa_E^\star$ because the right-hand side of (IA81) is smaller than $\kappa_E^\star$ under Assumption 1.

The only step missing now is to show that the left-hand side of (IA80) is decreasing in γ when $\kappa \geq \kappa_E^\star$. To prove this result, it is useful to introduce the notation

$$\overline{a}_1 = a_1\gamma^2,$$
$$\overline{a}_2 = a_2\gamma^2,$$

$$\overline{a}_{3,1} = \psi^2 \frac{2T(N+1) + T^2(T-N-3+2(T-N)\psi^2)}{(T-N)(T-N-1)^2(T-N-3)},$$

which are all independent of γ. Now, it is straightforward to show that the left-hand side of (IA80) is decreasing in γ when $\kappa \geq \kappa_E^\star$ if

$$4\overline{a}_1\kappa^2 + 3\overline{a}_2\kappa + 2\overline{a}_{3,1} \geq 0 \tag{IA82}$$

when $\kappa \geq \kappa_E^\star$. Since $\overline{a}_1 \geq 0$, inequality (IA82) holds for all κ if the polynomial discriminant is negative. Otherwise, if the discriminant is positive, we need to show that the maximum of the two real roots to the polynomial in (IA82) is smaller than $\kappa_E^\star$. That is,

$$\frac{-3\overline{a}_2 + \sqrt{9\overline{a}_2^2 - 32\overline{a}_1\overline{a}_{3,1}}}{8\overline{a}_1} \leq \kappa_E^\star. \tag{IA83}$$

After some simplifications, proving inequality (IA83) is equivalent to showing that

$$4\overline{a}_1(\kappa_E^\star)^2 + 3\overline{a}_2\kappa_E^\star + 2\overline{a}_{3,1} \geq 0. \tag{IA84}$$

We can reformulate inequality (IA84) as

$$\begin{aligned}
&\frac{\psi^2 C(T,N,\psi^2)}{(T-2)(T-N-2)(T-N-5)(T-N-7)} \\
&- \frac{3\psi^2(T\psi^2+N-1)(T+N-3+2T\psi^2)}{T-N-5} \\
&+ \frac{2T(N+1)+T^2(T-N-3+2(T-N)\psi^2)}{(T-N)(T-N-3)}\left(\psi^2 + \frac{N-1}{T}\right)^2 \geq 0.
\end{aligned} \tag{IA85}$$

Notice that inequality (IA85) holds when $\psi^2 = 0$. Therefore, we can prove (IA85) by showing that the derivative of the left-hand side with respect to ψ^2 is positive. That is,

$$\begin{aligned}
&\frac{1}{(T-2)(T-N-2)(T-N-5)(T-N-7)}\Big[(4T\psi^2+N-1)(N^4+N^3T-3N^3 \\
&-4N^2T^2+22N^2T-31N^2+NT^3-7NT^2+13NT-5N+T^4-12T^3+53T^2-100T \\
&+70)+3T^2\psi^4(N^3+2N^2T-6N^2-7NT^2+40NT-53N+4T^3-34T^2+88T-70)\Big] \\
&-\frac{3T}{T-N-5}\left[(T+N-3)\left(2\psi^2+\frac{N-1}{T}\right)+2T\left(3\psi^4+2\psi^2\frac{N-1}{T}\right)\right] \\
&+\frac{2T}{(T-N)(T-N-3)}\left[(2(N+1)+T(T-N-3))\left(\psi^2+\frac{N-1}{T}\right)+T(T-N)\right.
\end{aligned}$$

$$\left(3\psi^4 + 4\psi^2\frac{N-1}{T} + \left(\frac{N-1}{T}\right)^2\right)\Bigg] \geq 0. \tag{IA86}$$

One can check that inequality (IA86) holds when $\psi^2 = 0$ under Assumption 1. Therefore, we can prove (IA86) by showing as before that the derivative of the left-hand side with respect to ψ^2 is positive. That is,

$$\begin{aligned}
&\frac{2T}{(T-2)(T-N-2)(T-N-5)(T-N-7)}\Big[2(N^4 + N^3T - 3N^3 - 4N^2T^2 + 22N^2T \\
&- 31N^2 + NT^3 - 7NT^2 + 13NT - 5N + T^4 - 12T^3 + 53T^2 - 100T + 70) + 3T\psi^2(N^3 \\
&+ 2N^2T - 6N^2 - 7NT^2 + 40NT - 53N + 4T^3 - 34T^2 + 88T - 70)\Big] - \frac{6T}{T-N-5} \\
&\left[T + N - 3 + 2T\left(3\psi^2 + \frac{N-1}{T}\right)\right] + \frac{2T}{(T-N)(T-N-3)}\Big[2(N+1) + T(T-N-3) \\
&+ 2T(T-N)\left(3\psi^2 + 2\frac{N-1}{T}\right)\Big] \geq 0.
\end{aligned} \tag{IA87}$$

Again, one can check that inequality (IA87) holds when $\psi^2 = 0$ under Assumption 1. Therefore, we prove as usual that the derivative of the left-hand side of (IA87) with respect to ψ^2 is positive. That is,

$$\begin{aligned}
&\frac{N^3 + 2N^2T - 6N^2 - 7NT^2 + 40NT - 53N + 4T^3 - 34T^2 + 88T - 70}{(T-2)(T-N-2)(T-N-5)(T-N-7)} \\
&- \frac{6}{T-N-5} + \frac{2}{T-N-3} \geq 0.
\end{aligned} \tag{IA88}$$

This last inequality holds under Assumption 1, which concludes the demonstration of inequality (IA82) for $\kappa \geq \kappa_E^\star$.

Proof of Proposition 3

This proposition is a direct consequence of Proposition IA.3, which shows that the OOSU distribution of the shrinkage portfolio $\hat{w}^\star(\kappa)$ is asymptotically Gaussian as $N, T \to \infty$ with $N/T \to \rho \in [0, 1)$.

Proof of Proposition 4

Part 1. The proof is direct because, as shown in Proposition 1, the OOSU variance $\mathbb{V}[U(\hat{w}^\star(\kappa))] \to 0$ as $T \to \infty$ and, thus, the shrinkage intensity $\kappa_R^\star$ converges to $\kappa_E^\star$ as $T \to \infty$, and as shown in

Proposition IA.1 this $\kappa_E^\star$ is asymptotically optimal.

Part 2. First, $\kappa_R^\star \geq \kappa_V^\star$ because $\kappa_V^\star$ minimizes OOSU variance by definition and the OOSU mean in (IA2) is increasing in κ for $\kappa \leq \kappa_E^\star$. Since $\kappa_V^\star \leq \kappa_E^\star$ as we will prove next, this means that $\kappa_V^\star$ has a larger OOSU mean and smaller OOSU variance than any $\kappa \leq \kappa_V^\star$. Therefore, $\kappa_R^\star$ maximizing the mean-risk OOSU metric in (20) is necessarily larger than $\kappa_V^\star$.

Second, we prove the inequality $\kappa_R^\star \leq \kappa_E^\star$, which also implies $\kappa_V^\star \leq \kappa_E^\star$. To prove this inequality, note from part 2 of Proposition 2 that OOSU variance increases with κ if $\kappa \geq \kappa_E^\star$. Moreover, OOSU mean in (IA2) is decreasing in κ for $\kappa \geq \kappa_E^\star$. As a result, any $\kappa \geq \kappa_E^\star$ delivers a smaller mean-risk OOSU than $\kappa_E^\star$ and thus $\kappa_R^\star$ is necessarily smaller than $\kappa_E^\star$.

Proof of Proposition 5

Parts 1 and 2. The proof is direct because, on the one hand, the OOSU mean in (IA2) increases with T and μ_g and decreases with N, σ_g^2, and κ if $\kappa \geq \kappa_E^\star$. On the other hand, we show in Proposition 2 that OOSU standard deviation decreases with T and increases with κ if $\kappa \geq \kappa_E^\star$, increases with N and σ_g^2, and also is independent of μ_g.

Part 3. From the OOSU mean formula in Proposition IA.1, it is easy to check that OOSU mean increases with ψ^2 if $\kappa \leq \frac{2(T-N)(T-N-3)}{T(T-2)}$. Moreover, as we show in Proposition 2, OOSU variance always increases with ψ^2. Therefore, there must exist a value of λ in the mean-risk OOSU (20) below which mean-risk OOSU increases with ψ^2 and above which it decreases with ψ^2. This threshold, λ_0, is obtained as the value of λ for which $\frac{\partial}{\partial \psi^2} R(\hat{w}^\star(\kappa)) = 0$, which simplifies to (22).

Proof of Proposition 6

By definition of ψ^2 in (7), it can be written as

$$\psi^2 = \mu^\top \mathbf{\Sigma}^{-1} \mu - \frac{(e^\top \mathbf{\Sigma}^{-1} \mu)^2}{e^\top \mathbf{\Sigma}^{-1} e}. \tag{IA89}$$

Using the eigenvalue decomposition $\mathbf{\Sigma}^{-1} = \mathbf{V}\mathbf{D}^{-1}\mathbf{V}^\top$ and exploiting Assumption 3, the three quantities appearing in (IA89) simplify to

$$\mu^\top \mathbf{\Sigma}^{-1} \mu = \mu^\top \mathbf{V}\mathbf{D}^{-1}\mathbf{V}^\top \mu = \sum_{i=1}^{N} \frac{(v_i^\top \mu)^2}{d_i},$$
$$e^\top \mathbf{\Sigma}^{-1} \mu = \sqrt{N} v_1^\top \mathbf{V}\mathbf{D}^{-1}\mathbf{V}^\top \mu = \sqrt{N} \frac{v_1^\top \mu}{d_1},$$

$$e^\top \mathbf{\Sigma}^{-1} e = N v_1^\top \mathbf{V}\mathbf{D}^{-1}\mathbf{V}^\top v_1 = \frac{N}{d_1}.$$

Using these expressions, it is straightforward to check that (IA89) is equal to $\sum_{i>1}^{N} SR_{PC_i}^2$, which proves the proposition.

Proof of Proposition 7

For conciseness, we denote $U(\kappa) = U(\hat{w}(\kappa))$. The covariance between $\hat{\psi}^2 = \hat{\mu}^\top \widehat{\mathbf{\Sigma}}^{-1}\hat{\mu} - (\hat{\mu}_g/\hat{\sigma}_g)^2$ and $(U(\kappa) - \mathbb{E}[U(\kappa)])^2$ is

$$\mathbb{C}\text{ov}\Big[\hat{\psi}^2, (U(\kappa) - \mathbb{E}[U(\kappa)])^2\Big] = \mathbb{E}\Big[\hat{\psi}^2 (U(\kappa) - \mathbb{E}[U(\kappa)])^2\Big] - \mathbb{E}[\hat{\psi}^2]\mathbb{V}[U(\kappa)]. \tag{IA90}$$

In Equation (IA90), the OOSU variance $\mathbb{V}[U(\kappa)]$ is given in Proposition 1. Moreover, from Kan and Zhou (2007), we have

$$\mathbb{E}[\hat{\psi}^2] = \frac{\psi^2 T + N - 1}{T - N - 1}. \tag{IA91}$$

The last expectation, $\mathbb{E}\Big[\hat{\psi}^2 (U(\kappa) - \mathbb{E}[U(\kappa)])^2\Big]$, decomposes as

$$\mathbb{E}\Big[\hat{\psi}^2 (U(\kappa) - \mathbb{E}[U(\kappa)])^2\Big] = \mathbb{E}\Big[\hat{\psi}^2 U(\kappa)^2\Big] - 2\mathbb{E}[U(\kappa)]\mathbb{E}\Big[\hat{\psi}^2 U(\kappa)\Big] + \mathbb{E}[U(\kappa)]^2 \mathbb{E}[\hat{\psi}^2]. \tag{IA92}$$

In Equation (IA92), the OOSU mean $\mathbb{E}[U(\kappa)]$ is given in Proposition IA.1. Therefore, it remains to evaluate two expectations: $\mathbb{E}\Big[\hat{\psi}^2 U(\kappa)\Big]$ and $\mathbb{E}\Big[\hat{\psi}^2 U(\kappa)^2\Big]$. Using the stochastic representation for $\hat{\psi}^2$ and $U(\kappa)$ available in Kan et al. (2022b, Proposition 1), we find the following analytical expressions for the two expectations:

$$\begin{aligned}
\mathbb{E}\Big[\hat{\psi}^2 U(\kappa)\Big] &= \frac{\psi^2 T + N - 1}{T - N - 1}\left(\mu_g - \frac{\gamma \sigma_g^2 (T-2)(T-N-2)}{2(T-N)(T-N-3)}\right) \\
&\quad + \frac{\kappa \psi^2 T (\psi^2 T + N + 1)}{\gamma (T-N-1)(T-N-3)} \\
&\quad - \frac{\kappa^2 T (T-2)(\psi^4 T^2 + 2\psi^2 T(N+1) + (N+1)(N-1))}{2\gamma (T-N)(T-N-1)(T-N-3)(T-N-5)}, \\
\mathbb{E}\Big[\hat{\psi}^2 U(\kappa)^2\Big] &= \frac{\mu_g^2 (\psi^2 T + N - 1)}{T - N - 1} + \frac{\sigma_g^2 \psi^2 (\psi^2 T (T-N) + N(T-N-1) + 4 - T)}{(T-N)(T-N-1)(T-N-3)} \\
&\quad - \frac{\gamma \mu_g \sigma_g^2 (T-2)(T-N-2)(\psi^2 T + N - 1)}{(T-N)(T-N-1)(T-N-3)} \\
&\quad + \frac{\gamma^2 \sigma_g^4 (T-2)(T-4)(\psi^2 T + N - 1)((T-N-1)(T-N-5) + 3)}{4(T-N)(T-N-1)(T-N-2)(T-N-3)(T-N-5)}
\end{aligned} \tag{IA93}$$

$$
\begin{aligned}
&+\frac{\kappa\psi^2T(\psi^2T+N+1)\big(\frac{2\mu_g}{\gamma}(T-N)(T-N-5)-\sigma_g^2(T-2)(T-N-2)\big)}{(T-N)(T-N-1)(T-N-3)(T-N-5)}\\
&+\frac{\kappa^2\psi^2T\big(\psi^4T^2(T-N)+\psi^2T((T-N)(N+4)+N-2)+(N-1)(T-2)\big)}{\gamma^2(T-N)(T-N-1)(T-N-3)(T-N-5)}\\
&-\frac{\kappa^2\mu_gT(T-2)\big(\psi^4T^2+2\psi^2(N+1)+(N+1)(N-1)\big)}{\gamma(T-N)(T-N-1)(T-N-3)(T-N-5)}\\
&+\frac{\kappa^2\sigma_g^2T(T-2)(T-4)(T-N-4)\big(\psi^4T^2+2\psi^2T(N+1)+(N+1)(N-1)\big)}{2(T-N)(T-N-1)(T-N-2)(T-N-3)(T-N-5)(T-N-7)}\\
&-\frac{\kappa^3\psi^2T^2(T-2)\big(\psi^4T^2+2\psi^2T(N+3)+(N+3)(N+1)\big)}{\gamma^2(T-N)(T-N-1)(T-N-3)(T-N-5)(T-N-7)}\\
&+\frac{\kappa^4T^2(T-2)(T-4)/(T-N-9)}{4\gamma^2(T-N)(T-N-1)(T-N-2)(T-N-3)(T-N-5)(T-N-7)}\\
&\times\big(\psi^6T^3+3\psi^4T^2(N+3)+3\psi^2T(N+3)(N+1)+(N+3)(N+1)(N-1)\big),
\end{aligned}
\tag{IA94}
$$

where $\mathbb{E}\Big[\hat{\psi}^2U(\kappa)^2\Big]$ exists under the assumption that $T>N+9$.

To prove that the covariance is always positive, note first that it is independent of μ_g because $U(\kappa)-\mathbb{E}[U(\kappa)]$ is independent of μ_g. Therefore, without loss of generality, we set $\mu_g=0$. We begin by showing that the covariance increases with σ_g^2, and thus, that it suffices to show that the covariance is positive for $\sigma_g^2=0$. To show that the covariance increases with σ_g^2, we take the derivative with respect to σ_g^2 and observe that it is positive when $\sigma_g^2=0$, and moreover the derivative itself increases with σ_g^2. Indeed, we have

$$
\begin{aligned}
&\frac{\partial^2}{\partial\sigma_g^2\partial\sigma_g^2}\mathbb{C}\mathrm{ov}\Big[\hat{\psi}^2,(U(\kappa)-\mathbb{E}[U(\kappa)])^2\Big]=\\
&\frac{\gamma^2(T-2)(\psi^2T+N-1)}{2(T-N)(T-N-1)^3(T-N-2)(T-N-3)(T-N-5)}\\
&\times\big((T-4)(T-N-1)^2((T-N-1)(T-N-5)+3)-2(T-2)(T-N-2)(T-N-5)\big),
\end{aligned}
\tag{IA95}
$$

which is positive under the assumption that $T>N+9$. Therefore, what remains is to show that the covariance is positive for $\mu_g=0$ and $\sigma_g^2=0$. We proceed in a similar way to the proof of Proposition 2 by showing that the covariance is positive for $\kappa=0$, taking derivatives with respect to κ, and showing each time that these derivatives are positive for $\kappa=0$. At the end of this process,

what remains to show is that the following quantity is positive:

$$\begin{aligned}
&\frac{(T-4)(\psi^6T^3+3\psi^4T^2(N+3)+3\psi^2T(N+3)(N+1)+(N+3)(N+1)(N-1))}{(T-N-2)(T-N-5)(T-N-7)(T-N-9)}\\
&-\frac{2(T-2)(\psi^2T+N-1)(\psi^4T^2+2\psi^2T(N+1)+(N+1)(N-1))}{(T-N)(T-N-1)(T-N-3)(T-N-5)}\\
&+\frac{(T-2)(\psi^2T+N-1)^3}{(T-N)(T-N-1)^2(T-N-3)}\\
&-\frac{2C(T,N,\psi^2)(\psi^2T+N-1)}{(T-N)(T-N-1)^2(T-N-2)(T-N-3)(T-N-5)(T-N-7)}\geq 0,
\end{aligned}\tag{IA96}$$

where $C(T,N,\psi^2)$ is defined in Proposition 1. To show that inequality (IA96) holds, observe that it holds for $\psi^2=0$, and we can proceed as before by taking iterative derivatives with respect to ψ^2 and showing that they are positive for $\psi^2=0$. At the end of this process, what remains to show is the following inequality:

$$\begin{aligned}
&\frac{T-4}{(T-N-2)(T-N-5)(T-N-7)(T-N-9)}\\
&-\frac{2(T-2)}{(T-N)(T-N-1)(T-N-3)(T-N-5)}\\
&+\frac{T-2}{(T-N)(T-N-1)^2(T-N-3)}\\
&-\frac{2(N^3+2N^2T-6N^2-7NT^2+40NT-53N+4T^3-34T^2+88T-70)}{(T-N)(T-N-1)^2(T-N-2)(T-N-3)(T-N-5)(T-N-7)}\geq 0.
\end{aligned}\tag{IA97}$$

This inequality holds under the assumption $T>N+9$, which concludes the proof.

Proof of Proposition IA.1

The proof of this proposition is in Kan et al. (2022b).

Proof of Proposition IA.2

Denote $\hat{\mu}_g=\hat{w}_g^\top\hat{\mu}$ and $\hat{\sigma}_g^2=\hat{w}_g^\top\widehat{\mathbf{\Sigma}}\hat{w}_g$. Then, the coefficients A and B in (IA9) correspond to $A=1/\hat{\sigma}_g^2$ and $B=\hat{\mu}_g/\hat{\sigma}_g^2$. Denote also $f(\varepsilon)=1+\varepsilon/(\gamma\sigma_P^\star)$. Then, the ambiguity-averse portfolio can be rewritten as

$$\hat{w}^\star(\varepsilon)=\frac{1}{\gamma f(\varepsilon)}\widehat{\mathbf{\Sigma}}^{-1}\Big(\hat{\mu}-\hat{\mu}_g e+\gamma f(\varepsilon)\hat{\sigma}_g^2 e\Big)=\hat{w}_g+\frac{1}{\gamma f(\varepsilon)}\widehat{\mathbf{\Sigma}}^{-1}(\hat{\mu}-\hat{\mu}_g e).$$

The result follows by noticing that $\widehat{\mathbf{\Sigma}}^{-1}(\hat{\mu} - \hat{\mu}_g e) = \widehat{\mathbf{B}}\hat{\mu} = \hat{w}_z$, which is the estimated zero-cost portfolio, and therefore $\hat{w}^\star(\varepsilon)$ corresponds to the shrinkage portfolio $\hat{w}^\star(\kappa)$ in (11) if $\kappa = 1/f(\varepsilon) = (1 + \frac{\varepsilon}{\gamma\sigma_P^\star})^{-1}$.

Finally, $\sigma_P^\star$ is monotonically decreasing in ε because Garlappi et al. (2007) show that a higher ε implies a higher exposure to the SGMV portfolio and, thus, a smaller portfolio-return volatility. Consequently, κ in (IA10) is monotonically decreasing in ε.

Proof of Proposition IA.3

Kan et al. (2022b, Proposition 1) show that the finite-sample distribution of out-of-sample mean return and variance, and thus of OOSU, of the shrinkage portfolio $\hat{w}^\star(\kappa)$ is a function of 12 univariate independent random variables. Six of those random variables are normally distributed, and the remaining six are the following chi-square distributions: $u_0 \sim \chi^2_{N-2}$, $s_1 \sim \chi^2_{N-4}$, $s_2 \sim \chi^2_{N-3}$, $v_2 \sim \chi^2_{T-N+1}$, $w_1 \sim \chi^2_{T-N+3}$, and $w_2 \sim \chi^2_{T-N+2}$. In the high-dimensional asymptotic setting where $N, T \to \infty$ and $N/T \to \rho \in [0, 1)$, these six chi-square random variables are normally distributed. Specifically, from the proof of Kan et al. (2022a, Proposition 4), we have that

$$\begin{aligned}
\sqrt{T}(u_0/T - \rho) &\xrightarrow{d} \mathcal{N}(0, 2\rho),\\
\sqrt{T}(s_1/T - \rho) &\xrightarrow{d} \mathcal{N}(0, 2\rho),\\
\sqrt{T}(s_2/T - \rho) &\xrightarrow{d} \mathcal{N}(0, 2\rho),\\
\sqrt{T}(v_2/T - (1-\rho)) &\xrightarrow{d} \mathcal{N}(0, 2(1-\rho)),\\
\sqrt{T}(w_1/T - (1-\rho)) &\xrightarrow{d} \mathcal{N}(0, 2(1-\rho)),\\
\sqrt{T}(w_2/T - (1-\rho)) &\xrightarrow{d} \mathcal{N}(0, 2(1-\rho)).
\end{aligned}$$

Therefore, $\sqrt{T}U(\hat{w}^\star(\kappa))$ is asymptotically a function of 12 independent univariate Gaussian random variables. Using the delta method, it then follows that the asymptotic distribution of $\sqrt{T}U(\hat{w}^\star(\kappa))$ is Gaussian as well.

To find the asymptotic mean and variance, $u(\kappa, \rho)$ and $v(\kappa, \rho)$, we proceed as follows. First, we take the finite-sample expressions for OOSU mean and variance, $\mathbb{E}[U(\hat{w}^\star(\kappa))]$ and $\mathbb{V}[U(\hat{w}^\star(\kappa))]$, which are available in Propositions IA.1 and 1, respectively. Second, we only keep the dominant terms in N and T. Third, in the resulting expressions, we set $N = \rho T$. Following these three steps, it is straightforward to show that $\mathbb{E}[U(\hat{w}^\star(\kappa))] \to u(\kappa, \rho)$ and $T\mathbb{V}[U(\hat{w}^\star(\kappa))] \to v(\kappa, \rho)$ as $N, T \to \infty$

with $N/T \to \rho \in [0,1)$, which proves the proposition.

Proof of Proposition IA.4

Under the one-factor model in (IA17), the covariance matrix of r_t is

$$\mathbf{\Sigma} = \sigma_f^2 \beta\beta^\top + \sigma_\varepsilon^2 \mathbf{I}. \tag{IA98}$$

Given expression (IA98) for the covariance matrix, it is straightforward to see that the first eigenvector v_1 of $\mathbf{\Sigma}$ is proportional to β. Therefore,

$$\begin{aligned}
\mathbb{C}\text{orr}(v_1^\top r_t, w_{ew}^\top r_t) &= \mathbb{C}\text{orr}(\beta^\top r_t, w_{ew}^\top r_t) \\
&= \mathbb{C}\text{orr}\left(\frac{1}{N}\sum_{i=1}^N \beta_i^2 f_t + \beta_i \varepsilon_{it}, \frac{1}{N}\sum_{i=1}^N \beta_i f_t + \varepsilon_{it}\right) \\
&= \frac{\mathbb{C}\text{ov}(\bar\beta_2 f_t, \bar\beta_1 f_t) + \mathbb{C}\text{ov}\left(\frac{1}{N}\sum_{i=1}^N \beta_i\varepsilon_{it}, \frac{1}{N}\sum_{i=1}^N \varepsilon_{it}\right)}{\sqrt{\left(\sigma_f^2\bar\beta_2^2 + \sigma_\varepsilon^2\bar\beta_2/N\right)\left(\sigma_f^2\bar\beta_1^2 + \sigma_\varepsilon^2/N\right)}}.
\end{aligned} \tag{IA99}$$

Now, since the $(\varepsilon_{it})_{1\le i\le N}$ are i.i.d., we have that

$$\mathbb{C}\text{ov}\left(\frac{1}{N}\sum_{i=1}^N \beta_i\varepsilon_{it}, \frac{1}{N}\sum_{i=1}^N \varepsilon_{it}\right) = \frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N \beta_i \mathbb{C}\text{ov}(\varepsilon_{it}, \varepsilon_{jt}) = \frac{1}{N}\sigma_\varepsilon^2\bar\beta_1. \tag{IA100}$$

Therefore, the correlation in (IA99) becomes

$$\mathbb{C}\text{orr}(v_1^\top r_t, w_{ew}^\top r_t) = \frac{\sigma_f^2\bar\beta_1\bar\beta_2 + \sigma_\varepsilon^2\bar\beta_1/N}{\sqrt{\left(\sigma_f^2\bar\beta_2^2 + \sigma_\varepsilon^2\bar\beta_2/N\right)\left(\sigma_F^2\bar\beta_1^2 + \sigma_\varepsilon^2/N\right)}}, \tag{IA101}$$

which is the desired result. It is straightforward to check the two conditions under which this correlation converges to one, which concludes the proof.

Proof of Proposition IA.5

Under the assumption that the covariance matrix $\mathbf{\Sigma}$ is known, $\hat\mu$ is the only element in the portfolio that suffers from parameter uncertainty. Then, assuming that excess returns are i.i.d. normal, the sample vector of means $\hat\mu$ follows a multivariate normal distribution with mean μ and covariance matrix $\mathbf{\Sigma}/T$. Thus, it is straightforward to show that the shrinkage portfolio $\hat w^\star(\delta,\kappa)$ is distributed

as

$$\hat{w}^\star(\delta,\kappa) \sim \mathcal{N}\left(w_{sg} + \frac{\kappa}{\gamma} w_{sz}, \frac{1}{T}\frac{\kappa^2}{\gamma^2}\mathbf{S}\mathbf{B}_{sh}\right). \tag{IA102}$$

We now derive expressions for the three components of the OOSU variance in (IA42). First, the variance of out-of-sample mean return is

$$\mathbb{V}\left[\hat{w}^\star(\delta,\kappa)^\top \mu\right] = \mu^\top \mathbb{V}[\hat{w}^\star(\delta,\kappa)]\mu = \frac{1}{T}\frac{\kappa^2}{\gamma^2}\mu^\top \mathbf{S} w_{sz}. \tag{IA103}$$

Second, given that $\hat{w}^\star(\delta,\kappa)$ is normally distributed, we have from Rencher and Schaalje (2008, Theorem 5.2.) that the variance of out-of-sample return variance is

$$\mathbb{V}\left[\hat{w}^\star(\delta,\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\delta,\kappa)\right] = 2\mathrm{Tr}\Big((\mathbf{\Sigma}\mathbb{V}[\hat{w}^\star(\delta,\kappa)])^2\Big) + 4\mathbb{E}[\hat{w}^\star(\delta,\kappa)]^\top \mathbf{\Sigma}\mathbb{V}[\hat{w}^\star(\delta,\kappa)]\mathbf{\Sigma}\mathbb{E}[\hat{w}^\star(\delta,\kappa)], \tag{IA104}$$

where

$$\mathrm{Tr}\Big((\mathbf{\Sigma}\mathbb{V}[\hat{w}^\star(\delta,\kappa)])^2\Big) = \frac{1}{T^2}\frac{\kappa^4}{\gamma^4}\mathrm{Tr}(\mathbf{S}^4), \tag{IA105}$$

and

$$\begin{aligned}
&\mathbb{E}[\hat{w}^\star(\delta,\kappa)]^\top \mathbf{\Sigma}\mathbb{V}[\hat{w}^\star(\delta,\kappa)]\mathbf{\Sigma}\mathbb{E}[\hat{w}^\star(\delta,\kappa)] \\
&= \frac{1}{T}\frac{\kappa^2}{\gamma^2}\left(w_{sg} + \frac{\kappa}{\gamma} w_{sz}\right)^\top \mathbf{\Sigma}\mathbf{S}^2\left(w_{sg} + \frac{\kappa}{\gamma} w_{sz}\right) \\
&= \frac{1}{T}\frac{\kappa^2}{\gamma^2}\left(w_{sg}^\top \mathbf{\Sigma}\mathbf{S}^2 w_{sg} + \frac{2\kappa}{\gamma} w_{sg}^\top \mathbf{\Sigma}\mathbf{S}^2 w_{sz} + \frac{\kappa^2}{\gamma^2} w_{sz}^\top \mathbf{\Sigma}\mathbf{S}^2 w_{sz}\right).
\end{aligned} \tag{IA106}$$

Third, given that $\hat{w}^\star(\delta,\kappa)$ is normally distributed, we have from Rencher and Schaalje (2008, Theorem 5.2.) that the covariance between out-of-sample mean return and variance is

$$\begin{aligned}
\mathbb{C}\mathrm{ov}\left[\hat{w}^\star(\delta,\kappa)^\top \mu, \hat{w}^\star(\delta,\kappa)^\top \mathbf{\Sigma}\hat{w}^\star(\delta,\kappa)\right] &= 2\mu^\top \mathbb{V}[\hat{w}^\star(\delta,\kappa)]\mathbf{\Sigma}\mathbb{E}[\hat{w}^\star(\delta,\kappa)] \\
&= \frac{2}{T}\frac{\kappa^2}{\gamma^2}\left(\mu^\top \mathbf{S}^2 w_{sg} + \frac{\kappa}{\gamma}\mu^\top \mathbf{S}^2 w_{sz}\right).
\end{aligned} \tag{IA107}$$

Plugging (IA103), (IA104), and (IA107) in (IA42) gives the expression for OOSU variance in (IA25), which concludes the proof.

Proof of Proposition IA.6

To prove the results in this proposition, we use Okhrin and Schmid (2006, Theorem 1) to find that, under Assumptions 1 and 2, the mean and covariance matrix of the shrinkage portfolio $\hat{w}^\star(\kappa)$ in (11) are

$$\mathbb{E}[\hat{w}^\star(\kappa)] = w_g + \frac{\kappa}{\gamma}\frac{T}{T-N-1}\mathbf{B}\mu, \tag{IA108}$$

$$\begin{aligned}\mathbb{V}[\hat{w}^\star(\kappa)] =& \left(\frac{\sigma_g^2}{T-N-1} + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)+T^2\psi^2}{(T-N)(T-N-1)(T-N-3)}\right)\mathbf{B} \\ &+ \frac{\kappa^2}{\gamma^2}\frac{T^2(T-N+1)}{(T-N)(T-N-1)^2(T-N-3)}\mathbf{B}\mu\mu^\top\mathbf{B}.\end{aligned} \tag{IA109}$$

Moreover, we use the following useful properties: $\mathbf{B}e = \mathbf{0}$, $\mathbf{B}\boldsymbol{\Sigma}\mathbf{B} = \mathbf{B}$, $\mu^\top\mathbf{B}\boldsymbol{\Sigma}w_{ew} = \mu_{ew} - \mu_g$, and $w_{ew}^\top\boldsymbol{\Sigma}\mathbf{B}\boldsymbol{\Sigma}w_{ew} = \sigma_{ew}^2 - \sigma_g^2$.

Part 1. The OOSU mean of the shrinkage portfolio $\hat{w}^\star(\pi,\kappa)$ in (IA29) is

$$\begin{aligned}\mathbb{E}[U(\hat{w}^\star(\pi,\kappa))] =& (1-\pi)\mu_{ew} + \pi\mathbb{E}\left[\hat{w}^\star(\kappa)^\top\mu\right] - \frac{\gamma}{2}\Big((1-\pi)^2\sigma_{ew}^2 \\ &+ \pi^2\mathbb{E}\left[\hat{w}^\star(\kappa)^\top\boldsymbol{\Sigma}\hat{w}^\star(\kappa)\right] + 2\pi(1-\pi)\mathbb{E}\left[w_{ew}^\top\boldsymbol{\Sigma}\hat{w}^\star(\kappa)\right]\Big).\end{aligned} \tag{IA110}$$

From Kan et al. (2022b, Lemma 1), we have

$$\mathbb{E}\left[\hat{w}^\star(\kappa)^\top\mu\right] = \mu_g + \frac{\kappa}{\gamma}\frac{T}{T-N-1}\psi^2, \tag{IA111}$$

$$\mathbb{E}\left[\hat{w}^\star(\kappa)^\top\boldsymbol{\Sigma}\hat{w}^\star(\kappa)\right] = \frac{T-2}{T-N-1}\sigma_g^2 + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)(N-1)+T^2(T-2)\psi^2}{(T-N)(T-N-1)(T-N-3)}. \tag{IA112}$$

Moreover, using Equation (IA108), we find that

$$\mathbb{E}\left[w_{ew}^\top\boldsymbol{\Sigma}\hat{w}^\star(\kappa)\right] = \sigma_g^2 + \frac{\kappa}{\gamma}\frac{T}{T-N-1}(\mu_{ew}-\mu_g). \tag{IA113}$$

Plugging (IA111)–(IA113) into (IA110), we find that the OOSU mean is given by Equation (IA31).

Part 2. We derive the expressions for the three components of the OOSU variance in (IA32).

First, the variance of out-of-sample mean return is

$$\mathbb{V}\left[\hat{w}^\star(\pi,\kappa)^\top\mu\right] = \pi^2\mathbb{V}\left[\hat{w}^\star(\kappa)^\top\mu\right],$$

where $\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \mu\Big]$ is given by (IA44), which results in Equation (IA33).

Second, the variance of out-of-sample return variance is

$$\begin{aligned}\mathbb{V}\Big[\hat{w}^\star(\pi,\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\pi,\kappa)\Big] &= \pi^4 \mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big] + 4\pi^2(1-\pi)^2 w_{ew}^\top \mathbf{\Sigma} \mathbb{V}[\hat{w}^\star(\kappa)] \mathbf{\Sigma} w_{ew} \\ &\quad + 4\pi^3(1-\pi)\mathbb{C}\text{ov}\Big[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} w_{ew}, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big]. \end{aligned} \tag{IA114}$$

The first term, $\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big]$, is given by (IA59). Using Equation (IA109), the second term is

$$\begin{aligned} w_{ew}^\top \mathbf{\Sigma} \mathbb{V}[\hat{w}^\star(\kappa)] \mathbf{\Sigma} w_{ew} &= (\sigma_{ew}^2 - \sigma_g^2)\left(\frac{\sigma_g^2}{T-N-1} + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)+T^2\psi^2}{(T-N)(T-N-1)(T-N-3)}\right) \\ &\quad + \frac{\kappa^2}{\gamma^2}\frac{T^2(T-N+1)}{(T-N)(T-N-1)^2(T-N-3)}(\mu_{ew}-\mu_g)^2. \end{aligned}$$

The third term is similar to (IA68) and is

$$\begin{aligned}\mathbb{C}\text{ov}\Big[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} w_{ew}, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big] &= \frac{2\kappa}{\gamma}(\mu_{ew}-\mu_g)\Bigg(\sigma_g^2 \frac{T(T-2)}{(T-N-1)^2(T-N-3)} \\ &+ \frac{\kappa^2}{\gamma^2}\frac{T^2(T-2)(T+N-3+2T\psi^2)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}\Bigg). \end{aligned} \tag{IA115}$$

Putting these three terms together into (IA114) gives Equation (IA34).

Third, the covariance between out-of-sample mean return and variance is

$$\begin{aligned}\mathbb{C}\text{ov}\Big[\hat{w}^\star(\pi,\kappa)^\top \mu, \hat{w}^\star(\pi,\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\pi,\kappa)\Big] &= \pi^3 \mathbb{C}\text{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big] \\ &\quad + 2\pi^2(1-\pi)\mathbb{C}\text{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} w_{ew}\Big]. \end{aligned} \tag{IA116}$$

The first term, $\mathbb{C}\text{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big]$, is given by (IA68). Using Equation (IA109), the second term is

$$\begin{aligned}&\mathbb{C}\text{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} w_{ew}\Big] = \mu^\top \mathbb{V}[\hat{w}^\star(\kappa) \mathbf{\Sigma} w_{ew} \\ &= (\mu_{ew}-\mu_g)\left(\frac{\sigma_g^2}{T-N-1} + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)(T-N-1)+2T^2(T-N)\psi^2}{(T-N)(T-N-1)^2(T-N-3)}\right). \end{aligned}$$

Putting these two terms together into (IA116) results in the final expression in Equation (IA35) and concludes the proof.

Proof of Proposition IA.7

To prove this proposition, we use Okhrin and Schmid (2006, Theorem 1), who show that if $T > N$, $N \geq 2$, and Assumption 2 holds, then the mean and covariance matrix of the SGMV portfolio $\hat{w}_g$ are

$$\mathbb{E}[\hat{w}_g] = w_g \quad \text{and} \quad \mathbb{V}[\hat{w}_g] = \frac{\sigma_g^2}{T-N-1}\mathbf{B}.$$

Using this result, the mean and covariance matrix of the shrinkage portfolio $\hat{w}(\nu) = \nu w_{ew} + (1-\nu)\hat{w}_g$ are

$$\mathbb{E}[\hat{w}(\nu)] = \nu\mu_{ew} + (1-\nu)\mu_g \quad \text{and} \quad \mathbb{V}[\hat{w}(\nu)] = (1-\nu)^2\frac{\sigma_g^2}{T-N-1}\mathbf{B}.$$

Therefore, the mean squared error $\mathbb{E}[(\hat{w}(\nu)^\top\mu - \mu_g)^2]$ is given by

$$\mathbb{E}\Big[(\hat{w}(\nu)^\top\mu - \mu_g)^2\Big] = \nu^2(\mu_{ew} - \mu_g)^2 + (1-\nu)^2\frac{\sigma_g^2\psi^2}{T-N-1}.$$

Taking the derivative with respect to ν and setting it to zero yields the final result (IA37).

Proof of Proposition IA.8

Given the distribution of $\hat{w}^\star(\delta,\kappa)$ in (IA102), the expected out-of-sample mean return and variance of $\hat{w}^\star(\delta,\kappa)$ are

$$\mathbb{E}[\hat{w}^\star(\delta,\kappa)^\top\mu] = w_{sg}^\top\mu + \frac{\kappa}{\gamma}w_{sz}^\top\mu, \tag{IA117}$$

$$\mathbb{E}[\hat{w}^\star(\delta,\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\delta,\kappa)] = w_{sg}^\top\mathbf{\Sigma}w_{sg} + \frac{2\kappa}{\gamma}w_{sg}^\top\mathbf{\Sigma}w_{sz} + \frac{\kappa^2}{\gamma^2}\left(w_{sz}^\top\mathbf{\Sigma}w_{sz} + \frac{1}{T}\mathrm{Tr}(\mathbf{S}^2)\right), \tag{IA118}$$

which gives the expression for OOSU mean in (IA39) and completes the proof.

Proof of Proposition IA.9

Using the fact that sample mean returns are distributed as $\hat{\mu} \sim \mathcal{N}(\mu, \mathbf{\Sigma}/T)$, the shrinkage portfolio that takes $\mathbf{\Sigma}$ as given is distributed as $\hat{w}^\star(\kappa) \sim \mathcal{N}(\mathbb{E}[\hat{w}^\star(\kappa)], \mathbb{V}[\hat{w}^\star(\kappa)])$ with

$$\mathbb{E}[\hat{w}^\star(\kappa)] = w_g + \frac{\kappa}{\gamma}\mathbf{B}\mu, \tag{IA119}$$

$$\mathbb{V}[\hat{w}^\star(\kappa)] = \frac{\kappa^2}{\gamma^2}\frac{\mathbf{B}}{T}. \tag{IA120}$$

Using this result and the formulas for the mean, variance, and covariance of quadratic forms available in Rencher and Schaalje (2008), we can find closed-form expressions for the different moments that define the mean-risk OOSU measure:

$$\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top \mu\Big] = \mathbb{E}[\hat{w}^\star(\kappa)]^\top \mu = \mu_g + \frac{\kappa}{\gamma}\psi^2. \tag{IA121}$$

$$\begin{aligned}\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top \boldsymbol{\Sigma} \hat{w}^\star(\kappa)\Big] &= \mathrm{Tr}(\boldsymbol{\Sigma}\mathbb{V}[\hat{w}^\star(\kappa)]) + \mathbb{E}[\hat{w}^\star(\kappa)]^\top \boldsymbol{\Sigma}\mathbb{E}[\hat{w}^\star(\kappa)] \\ &= \sigma_g^2 + \frac{\kappa^2}{\gamma^2}\left(\psi^2 + \frac{N-1}{T}\right),\end{aligned} \tag{IA122}$$

$$\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \mu\Big] = \mu^\top \mathbb{V}[\hat{w}^\star(\kappa)]\mu = \frac{\kappa^2}{\gamma^2}\frac{\psi^2}{T}, \tag{IA123}$$

$$\begin{aligned}\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \boldsymbol{\Sigma} \hat{w}^\star(\kappa)\Big] &= 2\mathrm{Tr}((\boldsymbol{\Sigma}\mathbb{V}[\hat{w}^\star(\kappa)])^2) + 4\mathbb{E}[\hat{w}^\star(\kappa)]^\top \boldsymbol{\Sigma}\mathbb{V}[\hat{w}^\star(\kappa)]\boldsymbol{\Sigma}\mathbb{E}[\hat{w}^\star(\kappa)] \\ &= \frac{\kappa^4}{\gamma^4}\frac{2}{T}\left(2\psi^2 + \frac{N-1}{T}\right),\end{aligned} \tag{IA124}$$

$$\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \boldsymbol{\Sigma} \hat{w}^\star(\kappa)\Big] = 2\mu^\top \mathbb{V}[\hat{w}^\star(\kappa)]\boldsymbol{\Sigma}\mathbb{E}[\hat{w}^\star(\kappa)] = \frac{\kappa^3}{\gamma^3}\frac{2\psi^2}{T}. \tag{IA125}$$

Using (IA121)–(IA122), the OOSU mean is

$$\begin{aligned}\mathbb{E}[U(\hat{w}^\star(\kappa))] &= \mathbb{E}\Big[\hat{w}^\star(\kappa)^\top \mu\Big] - \frac{\gamma}{2}\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top \boldsymbol{\Sigma} \hat{w}^\star(\kappa)\Big] \\ &= \mu_g - \frac{\gamma}{2}\sigma_g^2 + \frac{1}{\gamma}\left(\kappa\psi^2 - \frac{\kappa^2}{2}\left(\psi^2 + \frac{N-1}{T}\right)\right).\end{aligned} \tag{IA126}$$

Moreover, using (IA123)–(IA125), the OOSU variance is

$$\begin{aligned}\mathbb{V}[U(\hat{w}^\star(\kappa))] &= \mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \mu\Big] + \frac{\gamma^2}{4}\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \boldsymbol{\Sigma} \hat{w}^\star(\kappa)\Big] - \gamma\,\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \boldsymbol{\Sigma} \hat{w}^\star(\kappa)\Big] \\ &= \frac{\kappa^2}{\gamma^2}\frac{\psi^2}{T}\left((1-\kappa)^2 + \kappa^2\frac{N-1}{2T\psi^2}\right).\end{aligned} \tag{IA127}$$

Using (IA126)–(IA127), the mean-risk OOSU defined in Equation (20) simplifies to (IA40).

To conclude the proof, we need to prove that the shrinkage intensity $\kappa_R^\star$ maximizing (IA40) is smaller than the intensity $\kappa_E^\star$ that maximizes the OOSU mean in (IA126). Note that the OOSU mean decreases with κ if $\kappa \geq \kappa_E^\star$. Therefore, to prove that any $\kappa \geq \kappa_E^\star$ delivers a smaller mean-risk OOSU than $\kappa_E^\star$, and thus that $\kappa_R^\star \leq \kappa_E^\star$, we need to prove that the OOSU variance in (IA127) increases with κ if $\kappa \geq \kappa_E^\star$. Taking the derivative of the OOSU variance with respect to κ, this is the case if

$$4\kappa^2\left(1 + \frac{N-1}{2T\psi^2}\right) - 6\kappa + 2 \geq 0 \tag{IA128}$$

for $\kappa \geq \kappa_E^\star$. If $\frac{4(N-1)}{T\psi^2} \geq 1$, the polynomial on the left-hand side of inequality (IA128) has no real roots and it is positive for any κ. Otherwise, it has only one positive real root and we need to prove that it is smaller than $\kappa_E^\star$. That is, it must hold that

$$\frac{6+\sqrt{36-32\left(1+\frac{N-1}{2T\psi^2}\right)}}{8\left(1+\frac{N-1}{2T\psi^2}\right)} \leq \frac{\psi^2}{\psi^2+\frac{N-1}{T}} \tag{IA129}$$

for $\frac{4(N-1)}{T\psi^2} \leq 1$. Inequality (IA129) simplifies to showing that

$$\sqrt{1-\frac{4(N-1)}{T\psi^2}} \leq \frac{\psi^2-\frac{N-1}{T}}{\psi^2+\frac{N-1}{T}}$$

holds for $0 \leq \frac{N-1}{T} \leq \psi^2/4$, which is indeed the case and concludes the proof.

References in Internet Appendix

Ao, M., Li, Y., Zheng, X., 2019. Approaching mean-variance efficiency for large portfolios. The Review of Financial Studies 32, 2890–2919.

Cochrane, J., 2005. Asset Pricing: Revised Edition. Princeton University Press.

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? The Review of Financial Studies 22, 1915–1953.

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

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

Goldfarb, D., Iyengar, G., 2003. Robust portfolio selection problems. Mathematics of Operations Research 28, 1–38.

Hansen, L. P., Jagannathan, R., 1997. Assessing specification errors in stochastic discount factor models. The Journal of Finance 52, 557–590.

Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. The Journal of Finance 58, 1651–1684.

Kan, R., Wang, X., 2023. Optimal portfolio choice with unknown benchmark efficiency. Management Science. *Forthcoming.*

Kan, R., Wang, X., Zheng, X., 2022a. In-sample and out-of-sample Sharpe ratios of multi-factor asset pricing models. SSRN working paper.

Kan, R., Wang, X., Zhou, G., 2022b. Optimal portfolio choice with estimation risk: No risk-free asset case. Management Science 68, 2047–2068.

Kan, R., Zhou, G., 2007. Optimal portfolio choice with parameter uncertainty. Journal of Financial and Quantitative Analysis 42, 621–656.

Kozak, S., Nagel, S., Santosh, S., 2020. Shrinking the cross-section. Journal of Financial Economics 135, 271–292.

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

Ledoit, O., Wolf, M., 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. The Review of Financial Studies 30, 4349–4388.

Ledoit, O., Wolf, M., 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. The Annals of Statistics 48, 3043–3065.

Okhrin, Y., Schmid, W., 2006. Distributional properties of portfolio weights. Journal of Econometrics 134, 235–256.

Rencher, A., Schaalje, G. B., 2008. Linear Models in Statistics (2nd ed.). John Wiley & Sons.

Tu, J., Zhou, G., 2011. Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. Journal of Financial Economics 99, 204–215.