A Robust Approach to Optimal Portfolio Choice with Parameter Uncertainty*

Nathan Lassance Alberto Martín-Utrera Majeed Simaan

December 20, 2021

Abstract

Many studies show that mean-variance portfolios perform poorly, delivering suboptimal *average* out-of-sample utility. A lesser-known fact we characterize is that their out-of-sample utility is also very *volatile*. Using our analytical characterization of performance volatility, we propose a robustness measure that balances out-of-sample utility mean and volatility and show that neither mean-variance portfolios nor minimum-variance portfolios offer maximal robust performance. Our robustness measure serves as a portfolio framework to construct strategies that achieve the optimal tradeoff between out-of-sample utility mean and volatility. These strategies are resilient to estimation errors and outperform portfolios that ignore parameter uncertainty or out-of-sample utility risk.

Keywords: estimation risk, 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`. N. Lassance is supported by the Fonds de la Recherche Scientifique (FNRS) under Grant number J.0115.22. We thank Vitali Alekseev, Stephen Brown, Victor DeMiguel, Kristoffer Glover, Raymond Kan, Yueliang Lu, Harry Scheule, Frédéric Vrins, Xiaolu Wang, Alex Weissensteiner, Guofu Zhou, as well as seminar and conference participants at UCLouvain, University of Technology Sydney, 15th International Conference on Computational and Financial Econometrics, Financial Management Association Annual Meeting, INFORMS Annual Meeting, and Northern Finance Association Annual Meeting for their comments.

1 Introduction

Since the seminal work of Markowitz (1952), academics have extensively documented the sensitivity of optimal mean-variance portfolios to estimation errors in the vector of means and the covariance matrix of stock returns and have shown that estimated optimal portfolios tend to deliver poor out-of-sample performance (DeMiguel, Garlappi, and Uppal, 2009; Ledoit and Wolf, 2017). An influential paper in this literature is Kan and Zhou (2007), which characterizes the *average* out-of-sample utility (OOSU) of mean-variance investors. Kan and Zhou use this metric to optimally combine the risk-free asset, the sample tangency portfolio, and the sample global-minimum-variance portfolio, which results in an improved investment strategy that generates substantial average out-of-sample utility gains.

While the analytical characterization of the average OOSU is an essential step toward having a complete understanding of the stochastic nature of portfolio performance, two key questions remain unanswered: 1) *what is the OOSU risk of estimated optimal portfolios?* and 2) *why is it relevant to characterize OOSU risk of estimated optimal portfolios?* To address the first question, we characterize in closed form the OOSU standard deviation of the sample mean-variance (SMV) portfolio, the sample global-minimum-variance (SGMV) portfolio, and any combination of the two portfolios.

We argue that there are two main reasons why studying the OOSU volatility of estimated optimal portfolios is relevant. First, our analytical closed-form expression of OOSU risk allows us to uncover an essential aspect of the stochastic nature of portfolio performance. For example, the OOSU two-sigma interval of the SMV portfolio calibrated from a dataset of 25 portfolios of stocks sorted on size and book-to-market (25SBTM) with 120 monthly return observations and a risk-aversion coefficient of three is $[-12.4\%, 0.94\%]$, which is a paramount concern of SMV portfolios for investors who deem performance uncertainty as a critical variable of their investment-decision process.

The second reason why characterizing OOSU risk is essential is because it allows us to develop a *novel* portfolio robustness metric defined as the difference between OOSU mean and a multiple of OOSU risk. Then, we exploit our proposed robustness metric to combine the SMV and SGMV portfolios optimally as in Kan, Wang, and Zhou (2021).[1] We show theoretically that neither the SMV portfolio nor the SGMV portfolio offers the maximal robust performance individually and that one must combine both to achieve a better tradeoff between OOSU mean and OOSU volatility. In addition, the empirical analysis suggests that the combination of the SMV and SGMV portfolios that optimizes our proposed robustness metric delivers better out-of-sample performance than those strategies that ignore parameter uncertainty or OOSU risk. In particular, compared to the combination of portfolios that only maximizes OOSU mean, our robust combination delivers larger out-of-sample certainty-equivalent return and Sharpe ratio while attaining a lower downside risk, both before and after transaction costs.

Our manuscript makes four contributions to the existing literature on parameter uncertainty and portfolio selection. First, we provide the exact closed-form expression for the OOSU variance of the SMV portfolio, the SGMV portfolio, and any combination of these two portfolios. Using this analytical result, we document that the SMV portfolio OOSU volatility is substantially larger than that of the SGMV portfolio. Take, for instance, the 25SBTM dataset and a risk-aversion coefficient of three. For this case, 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 case, the SMV portfolio requires an unrealistically large sample size of over 13,000 monthly observations –more than 1,000 years of data– to deliver a performance as stable as that of the SGMV portfolio.[2]

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. Our measure of portfolio robustness explicitly accounts for the impact of estimation error on the performance of estimated portfolios by fully characterizing the first two moments (i.e., mean and variance) of out-of-sample performance. In our view, a robust portfolio should deliver a stable out-of-sample utility that, on average, performs well. The robustness measure we

[1]Our portfolio robustness metric can accommodate a wider range of portfolio combinations. Indeed, in Section IA.2.5 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).

[2]In unreported results, we show that the required sample size for the SMV portfolio to deliver an out-of-sample utility as stable as that of the SGMV portfolio is of similar magnitude across different datasets.

propose is in the spirit of this view. Using the analytical characterization of our proposed measure, we compare the robustness of the SMV and SGMV portfolios. We show that the SMV portfolio is much less robust than the SGMV portfolio. In particular, for the 25SBTM dataset and a risk-aversion coefficient of three, the SMV portfolio requires over 696 monthly observations –equivalent to 58 years of data– to be as robust as the SGMV portfolio.

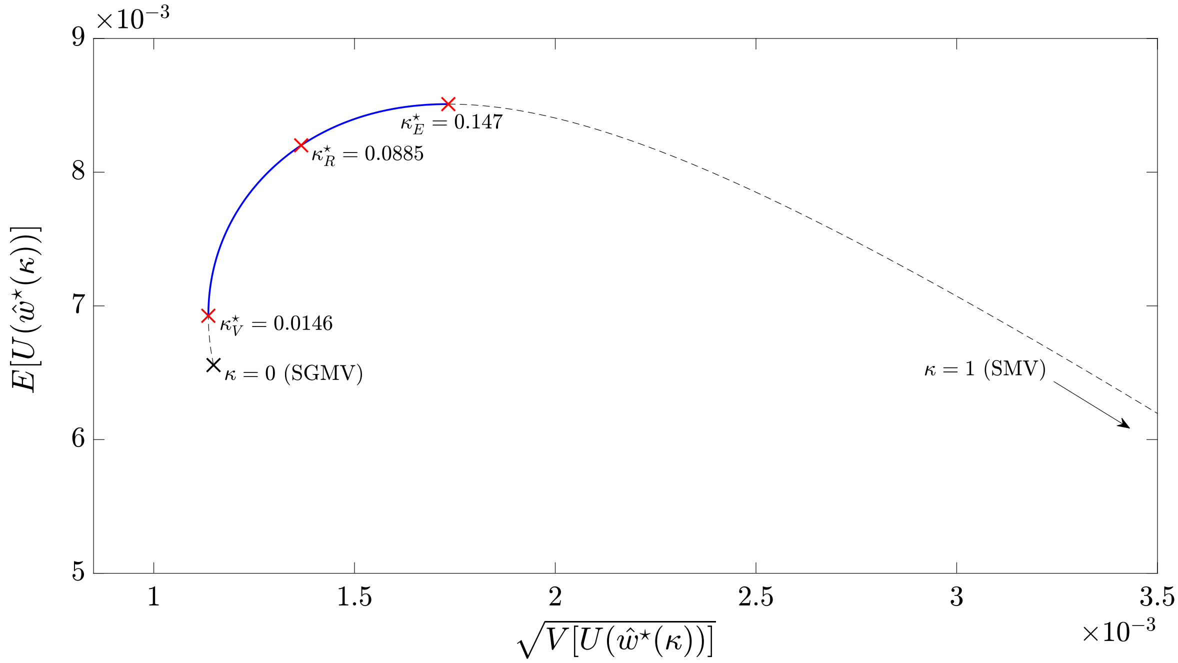
Our third contribution is to propose a novel calibration criterion for combining the SMV portfolio with the SGMV portfolio that maximizes our measure of portfolio robustness. We demonstrate that the shrinkage portfolio that optimizes our robustness metric assigns a larger tilt toward the SGMV portfolio than that of the shrinkage portfolio that maximizes OOSU mean. However, we show that it is not optimal to fully tilt toward the SGMV portfolio and, therefore, one must optimally combine both sample portfolios to attain maximal robustness. While the optimal shrinkage intensity is unfeasible to the investor because it depends on the true, but *unknown*, vector of means and covariance matrix, we propose a feasible consistent estimator of the optimal shrinkage intensity that maximizes our proposed robustness measure.

Our robust portfolio framework sacrifices a small out-of-sample average performance to achieve substantially more stable performance. Figure 1 illustrates the main idea of our proposed method. The vertical axis depicts the OOSU mean of the shrinkage portfolio, and the horizontal axis depicts the OOSU standard deviation. The parameter κ is the shrinkage intensity that determines the optimal combination between the SMV and SGMV portfolios. The shrinkage portfolio exploiting $\kappa_E^\star$ maximizes OOSU mean, the shrinkage portfolio exploiting $\kappa_V^\star$ minimizes OOSU standard deviation, 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 3.6% reduction in OOSU mean relative to the shrinkage portfolio exploiting $\kappa_E^\star$.[3]

It is important to note that exploiting our robustness measure to combine portfolios resembles the diversification idea behind the mean-variance efficient frontier of Markowitz (1952). In our case, instead of obtaining the optimal combination of stocks that minimizes

[3]While Figure 1 considers the true shrinkage intensities, our simulation results show that our proposed shrinkage intensity $\kappa_R^\star$ delivers a more significant improvement in portfolio performance compared to $\kappa_E^\star$ when the shrinkage intensities are unknown and must be estimated.

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 returns are calibrated from the monthly return data of the 25 portfolios of stocks sorted on size and book-to-market. For the estimation of the shrinkage portfolios, we use a sample size of $T = 120$ months and a risk-aversion coefficient of $\gamma = 3$. The solid blue line in the figure 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 4 with $\lambda = 2$.

portfolio risk for a given level of expected portfolio return, we obtain the optimal combination of the SMV and SGMV portfolios that minimizes OOSU risk for a given level of OOSU mean.

Our fourth contribution is to evaluate the out-of-sample performance of our proposed robust shrinkage portfolio relative to several benchmarks. 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 our robust shrinkage portfolio also delivers a larger OOSU mean than the portfolios that are specifically designed to maximize OOSU mean in those cases where the data is not Gaussian, and the shrinkage intensities are estimated from the data. This suggests that our proposed portfolio framework to construct robust investment strategies is resilient to estimation errors.

We also study the performance of our methodology on seven empirical datasets of monthly return data: six datasets of characteristic-sorted portfolios and one dataset of individual stocks from CRSP. Compared to the optimal shrinkage portfolio calibrated under the OOSU mean criterion, we document that the proposed robust shrinkage portfolio offers a lower turnover while delivering a better certainty-equivalent return, Sharpe ratio, and downside risk, even net of transaction costs. In addition, we use three-year non-overlapping windows to measure the volatility of out-of-sample certainty-equivalent returns of the considered shrinkage portfolios and show that the performance of our robust portfolio is more stable over time than that of the shrinkage portfolio that only maximizes OOSU mean. Our proposed shrinkage portfolio also delivers better out-of-sample certainty-equivalent returns than those of the equally weighted portfolio, the reward-to-risk timing strategy of Kirby and Ostdiek (2012), and the SGMV portfolio in six out of seven datasets.[4] Overall, our results highlight the importance of accounting for OOSU risk to construct mean-variance optimal portfolios.

Our work is closely related to the literature that exploits *shrinkage estimators* to mitigate the impact of parameter uncertainty.[5] Such estimators are traditionally applied to alleviate the impact of parameter uncertainty affecting the inputs of the portfolio problem, like the mean (Jorion, 1986; Barroso and Saxena, 2021) and the covariance matrix (Ledoit and Wolf, 2003, 2004, 2017, 2020a). A prominent example is the work of Ledoit and Wolf (2003, 2004) who focus on the optimal linear combination between two covariance matrices so that the resulting combination minimizes the mean squared error of the estimated matrix. Unlike these papers, we focus on combining *portfolios* to attain an optimal tradeoff between OOSU mean and OOSU volatility.

We build on the literature pioneered by Kan and Zhou (2007) that considers the out-of-sample utility *mean* as a criterion to analytically determine the optimal shrinkage portfolio

[4]Even though our theoretical results and base case empirical methodology consider sample estimates of the mean and covariance matrix of stock returns, we show in Appendix IA.2 that the results are robust to using the shrinkage estimator of the covariance matrix of Ledoit and Wolf (2020a).

[5]There is a large number of papers that propose different approaches to combat the impact of parameter uncertainty on portfolio selection. Some of these papers use Bayesian statistics (Jorion, 1986; Avramov and Zhou, 2010), factor models (De Nard, Ledoit, and Wolf, 2019), forward-looking information (DeMiguel, Plyakha, Uppal, and Vilkov, 2013), model misspecification (Rapponi, Uppal, and Zaffaroni, 2021), weight constraints (Jagannathan and Ma, 2003; DeMiguel, Garlappi, Nogales, and Uppal, 2009), robust optimization (Goldfarb and Iyengar, 2003), and sparse estimation (Goto and Xu, 2015; Ao, Li, and Zheng, 2019).

under the assumption that stock returns are iid Gaussian.[6] A large literature exploits the OOSU mean criterion introduced by Kan and Zhou (2007) in the construction of alternative portfolio strategies that help mitigate the impact of parameter uncertainty.[7] In contrast to these papers, our work also accounts for the OOSU *standard deviation* in the calibration of shrinkage portfolios, which offers a more robust performance than the portfolios that ignore OOSU risk or parameter uncertainty.

Our shrinkage portfolio approach shares fundamental elements with regularization techniques, which are one of the most common machine learning approaches adopted in the recent asset pricing literature (Giglio, Kelly, and Xiu, 2021). Like regularization, the shrinkage portfolio considered 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 shrinkage portfolio approach considered in this manuscript is not only a practical approach to alleviating the impact of statistical errors on the performance of estimated portfolios, but it is also an economically sound method that is related to the investment-decision problem of ambiguity-averse investors. In particular, we characterize the exact relationship between the shrinkage portfolio that combines the SMV and SGMV portfolios and the ambiguity-averse portfolios considered by Garlappi, Uppal, and Wang (2007). We

[6]Earlier studies consider out-of-sample utility 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 analytically characterize the expected out-of-sample utility losses from parameter uncertainty under the assumption of iid Gaussian returns.

[7]For instance, Zhou (2008) applies the framework of Kan and Zhou to solve the investment-decision problem of an active portfolio manager. DeMiguel, Garlappi, and Uppal (2009) derive the critical sample size for which the SMV portfolio delivers a larger OOSU mean than the equally weighted (EW) portfolio. Frahm and Memmel (2010) derive the combination between the SGMV portfolio and the EW portfolio that minimizes the mean of the out-of-sample variance. Tu and Zhou (2011) consider combinations between several estimates of the mean-variance portfolio and the EW portfolio that maximize OOSU mean. DeMiguel, Martín-Utrera, and Nogales (2013) provide several calibration criteria for shrinkage portfolios such as the mean of the out-of-sample variance, utility, and Sharpe ratio. DeMiguel, Martín-Utrera, and Nogales (2015) generalize the analysis of Kan and Zhou to a multiperiod setting with transaction costs. Branger, Lučivjanská, and Weissensteiner (2019) derive an optimal grouping of EW sub-portfolios that maximizes OOSU mean. Kan and Wang (2021) derive the combination of a set of benchmark portfolios and positive-alpha test assets that maximizes OOSU mean. Finally, Kan, Wang, and Zhou (2021) consider the common framework with no risk-free asset and derive the combination of the SMV and SGMV portfolios that maximizes OOSU mean. In addition to this literature, a number of papers propose to maximize the average out-of-sample performance of an estimated portfolio in a data-driven way instead of using the parametric approach introduced by Kan and Zhou (2007); see DeMiguel, Martín-Utrera, and Nogales (2013) and Kircher and Rosch (2021) for a bootstrap application, and Füss, Koeppel, and Miebs (2021) for a jackknife application.

show in closed form that a larger degree of ambiguity in mean returns leads the ambiguity-averse investor to apply a larger tilt toward the SGMV portfolio. In line with the theoretical relationship between shrinkage and ambiguity-averse portfolios, our manuscript proposes a method to establish the degree of ambiguity in mean returns that provides a robust out-of-sample performance.

Our work is also related to 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." Similarly, we show that the proposed criterion for combining portfolios that maximizes the difference between OOSU mean and a multiple of OOSU standard deviation is equivalent to a robust optimization problem where the investor maximizes the worst-case scenario of the unknown OOSU mean. Therefore, our proposed shrinkage portfolio explicitly accounts for statistical errors affecting the estimation of OOSU mean.

Finally, our work is also related to several papers that study the *distribution* of out-of-sample portfolio performance measures. Kan and Smith (2008) and Kan, Wang, and Zhou (2021) derive the distribution of the out-of-sample mean return and variance of efficient portfolios. Similarly, Kan, Wang, and Zheng (2021) derive the distribution of the out-of-sample Sharpe ratio of the tangency portfolio and use this result to explain why many asset-pricing models underperform the market portfolio out of sample. Unlike these papers, we use the OOSU mean and volatility to assess 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 OOSU volatility in combining portfolios.

2 Mean-variance portfolios and parameter uncertainty

In this section, we define the mean-variance portfolio framework in the presence of parameter uncertainty. In Section 2.1, we provide details about the main theoretical assumptions. In Section 2.2, we discuss the properties of mean-variance portfolios without parameter uncertainty. In Section 2.3, we present the optimal shrinkage portfolio proposed by Kan, Wang, and Zhou (2021) that maximizes OOSU mean in the presence of parameter uncertainty.

Finally, in Section 2.4, we explain why considering shrinkage portfolios is an economically sound method to mitigating the impact of estimation errors.

2.1 Notation and 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 with mean μ and positive-definite covariance matrix $\mathbf{\Sigma}$, and w is a vector of portfolio weights. 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. Using historical return data over the past T months $(r_1, \dots, r_T)$, the investor estimates the vector of means μ and covariance matrix of stock returns $\mathbf{\Sigma}$ with their sample counterparts:[8]

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

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) through time.*

The condition $T > N + 7$ in Assumption 1 is needed to ensure that the out-of-sample utility variance derived in Section 3 exists. Assumption 2 is a standard assumption in the literature used for analytical tractability (Kan and Zhou, 2007; Ao, Li, and Zheng, 2019). Under Assumption 2, $\hat{\mu}$ and $\widehat{\mathbf{\Sigma}}$ are independent and follow a multivariate Gaussian distribution and Wishart distribution, respectively.

While it is unlikely that the empirical data follow a Gaussian distribution, there are several reasons why Assumption 2 does not compromise the performance of the portfolio strategies that rely on this assumption. First, even when stock returns are non-Gaussian, there is a close relationship between expected utility and the mean-variance framework (Kroll, Levy,

[8]In Appendix IA.2.4, we assume that the covariance matrix is known to the investor as in Garlappi, Uppal, and Wang (2007), and therefore parameter uncertainty only stems from the vector of means.

and Markowitz, 1984; Markowitz, 2014). Second, 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). Third, our empirical results show that the shrinkage portfolios calibrated under the assumption of iid Gaussian returns deliver good out-of-sample performance even for datasets with real data where returns are likely not Gaussian.

2.2 Mean-variance portfolios without parameter uncertainty

We first consider the classical Markowitz (1952) portfolio problem where the investor knows the true distributional properties of stock returns, i.e., μ and $\mathbf{\Sigma}$. Let $\gamma > 0$ denote the investor's risk-aversion coefficient. Then, the optimal mean-variance portfolio is the solution to the following quadratic program

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

where $U(w)$ is the *utility* of portfolio w. The solution to problem (2) is

$$w^\star = w_g + \frac{1}{\gamma} w_z, \tag{3}$$

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

$$w_g = \mathbf{\Sigma}^{-1} e (e^\top \mathbf{\Sigma}^{-1} e)^{-1}, \tag{4}$$

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

$$w_z = \mathbf{B}\mu, \quad \mathbf{B} = \mathbf{\Sigma}^{-1}(\mathbf{I} - e w_g^\top). \tag{5}$$

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

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

$$\sigma_g^2 = w_g^\top \mathbf{\Sigma} w_g = (e^\top \mathbf{\Sigma}^{-1} e)^{-1}, \tag{7}$$

$$\psi^2 = w_z^\top \mathbf{\Sigma} w_z = \mu^\top \mathbf{\Sigma}^{-1}\mu - \mu_g^2/\sigma_g^2, \tag{8}$$

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 other words, ψ^2 can also be interpreted as the Sharpe-ratio inefficiency of the GMV portfolio.

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

$$U(w^\star) = U(w_g) + \frac{\psi^2}{2\gamma}. \tag{9}$$

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 the 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, as documented in prior literature (Merton, 1980; Chopra and Ziemba, 1993). Therefore, it is essential to account for parameter uncertainty in the construction of investment strategies, which we address in the next section.

2.3 Optimizing out-of-sample utility mean

In practice, investors do not know the true vector of means and the covariance matrix of stock returns, and instead, they estimate these parameters from historical return data using the sample estimates provided in Equation (1). Accordingly, the sample mean-variance portfolio, 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 global minimum-variance 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 $\widehat{\mathbf{\Sigma}}$.

The estimation risk affecting the SMV portfolio leads to suboptimal performance as noted by DeMiguel, Garlappi, and Uppal (2009). To combat the impact of parameter uncertainty, we consider shrinkage techniques, which help reduce the impact of statistical errors on the performance of mean-variance portfolios.[9] Indeed, Kan, Wang, and Zhou (2021) show that one can improve *average* out-of-sample performance by *combining* the SMV portfolio $\hat{w}^\star$ with the SGMV portfolio $\hat{w}_g$. Similarly, we consider a linear combination between these two portfolios, where the shrinkage intensity $\kappa \in [0,1]$ establishes the optimal combination:

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

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 $\hat{w}$ as

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

Following the literature pioneered by Kan and Zhou (2007), one can construct a portfolio that mitigates the impact of estimation risk on portfolio performance by optimizing the OOSU *mean* of the estimated portfolio. In the following proposition, we review some of the main results of Kan, Wang, and Zhou (2021) for the shrinkage portfolio $\hat{w}^\star(\kappa)$ in Equation (11).[10]

Proposition 1 (Kan, Wang, and Zhou (2021))**.** *Let Assumptions 1 and 2 hold. Then,*

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

$$\mathbb{E}[U(\hat{w}_g)] = \mu_g - \frac{\gamma}{2}\frac{T-2}{T-N-1}\sigma_g^2. \tag{13}$$

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

$$\mathbb{E}[U(\hat{w}^\star(\kappa))] = \mathbb{E}[U(\hat{w}_g)] + \Theta(\kappa), \tag{14}$$

where

$$\Theta(\kappa) = \frac{1}{\gamma}\frac{T}{T-N-1}\left(\kappa\psi^2 - \kappa^2\left(\psi^2 + \frac{N-1}{T}\right)\frac{T(T-2)}{2(T-N)(T-N-3)}\right). \tag{15}$$

[9] In Section 2.4, we argue that this method is economically sound because of the explicit connection between the shrinkage portfolio and the ambiguity-averse portfolio of Garlappi, Uppal, and Wang (2007).

[10] All proofs are available in the Internet Appendix.

3. The shrinkage intensity $\kappa_E^\star$ maximizing out-of-sample utility mean in (14) is

$$\kappa_E^\star = \frac{(T-N)(T-N-3)}{T(T-2)} \frac{\psi^2}{\psi^2 + \frac{N-1}{T}} \in [0,1]. \tag{16}$$

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

A number of comments are in order. First, $\kappa_E^\star$ is an oracle estimator that depends on the unknown parameter ψ^2. Kan, Wang, and Zhou (2021) 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 a wide range of benchmark strategies. Second, the optimal shrinkage intensity $\kappa_E^\star$ increases with ψ^2 and decreases with the ratio N/T. Third, $\kappa_E^\star$ also corresponds to the shrinkage intensity that minimizes the bias of the investor's out-of-sample utility. More formally,

$$\kappa_E^\star = \arg\min_{\kappa \in [0,1]} \mathbb{E}[U(w^\star) - U(\hat{w}^\star(\kappa))]. \tag{17}$$

While $\kappa_E^\star$ delivers the least-biased out-of-sample utility, the OOSU variance can still be large. In Section 3, we extend the analysis of Proposition 1 to study the OOSU variance of several sample portfolios, and we utilize this result to construct optimal portfolios that balance OOSU mean and volatility in Section 4.

2.4 Relation to ambiguity-averse portfolios

In this section, we argue that a shrinkage portfolio that combines the SMV and SGMV portfolios is not only a useful technique to mitigate the impact of estimation risk, but it is also an economically sound approach. In particular, we show that there is an explicit relationship between the shrinkage portfolio considered in this manuscript and the optimal portfolio of an ambiguity-averse investor. The insights provided in this section build on the work of Garlappi, Uppal, and Wang (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

$$\max_{w:w^\top e=1} \min_{\mu} \; w^\top \mu - \frac{\gamma}{2} w^\top \widehat{\mathbf{\Sigma}} w \quad \text{subject to } (\hat{\mu}-\mu)^\top \widehat{\mathbf{\Sigma}}^{-1}(\hat{\mu}-\mu) \leq \varepsilon^2, \tag{18}$$

where $(\hat{\mu}-\mu)^\top \widehat{\mathbf{\Sigma}}^{-1}(\hat{\mu}-\mu)$ 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 (18). Garlappi, Uppal, and Wang (2007) show that the closed-form solution of this problem is

$$\hat{w}^\star(\varepsilon) = \frac{1}{\gamma} \widehat{\mathbf{\Sigma}}^{-1} \left(\frac{1}{1+\varepsilon/(\gamma \sigma_P^\star)} \right) \left(\hat{\mu} - \frac{B - \gamma(1+\varepsilon/(\gamma\sigma_P^\star))}{A} e \right), \tag{19}$$

where $A = e^\top \widehat{\mathbf{\Sigma}}^{-1} e$, $B = \hat{\mu}^\top \widehat{\mathbf{\Sigma}}^{-1} e$, and parameter $\sigma_P^\star$ is the unique positive real root to a specific fourth-degree polynomial that is monotonically decreasing in ε. Garlappi, Uppal, and Wang (2007) show that the ambiguity-averse portfolio $\hat{w}^\star(\varepsilon)$ converges to the SMV portfolio when $\varepsilon \to 0$, and to the SGMV portfolio when $\varepsilon \to \infty$. Intuitively, for $0 < \varepsilon < \infty$ the ambiguity-averse portfolio $\hat{w}^\star(\varepsilon)$ combines the SMV and the SGMV portfolios, similarly to the shrinkage portfolio $\hat{w}^\star(\kappa)$ in (11). In the next proposition, we characterize the intensity κ of the shrinkage portfolio $\hat{w}^\star(\kappa)$ in (11) as a function of the ambiguity-aversion parameter ε.

Proposition 2. *The shrinkage portfolio $\hat{w}^\star(\kappa)$ is equal to the ambiguity-averse portfolio $\hat{w}^\star(\varepsilon)$ in Equation (19) when*

$$\kappa = \left(1 + \frac{\varepsilon}{\gamma \sigma_P^\star}\right)^{-1}, \tag{20}$$

where the ratio $\varepsilon/\sigma_P^\star$ is monotonically increasing in ε.

Proposition 2 provides the explicit link between the shrinkage portfolio considered in this manuscript and the ambiguity-averse portfolio of Garlappi, Uppal, and Wang (2007). In particular, Equation (20) 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. Given an optimally calibrated shrinkage intensity κ, Equation (20) allows us to determine the equivalent degree of ambiguity in mean returns that results in an ambiguity-averse portfolio that delivers robust out-of-sample performance.

We show in Section 4 that our proposed robustness criterion establishes a larger tilt toward the SGMV portfolio and thus a more considerable degree of ambiguity in mean returns than the traditional shrinkage criterion that only maximizes OOSU mean.[11] In other words, our proposed shrinkage criterion based on the portfolio robustness measure introduced in Section 4 delivers portfolios that are less sensitive to the estimation errors in mean returns.

3 Out-of-sample utility variance

In this section, we characterize the OOSU variance of sample portfolios. For notational simplicity, in Section 3.1 we derive the closed-form analytical expression of the OOSU variance for the shrinkage portfolio that combines the SMV and SGMV portfolios, which contains as particular cases the OOSU variance of the individual SMV and SGMV portfolios.[12] In Section 3.2, we study the monotonicity properties of the OOSU variance.

3.1 Out-of-sample utility variance of shrinkage portfolios

We first define in the following lemma the OOSU variance of any random portfolio, which we use to obtain the OOSU variance of any combination of the SMV and SGMV portfolios.

Lemma 1. *The out-of-sample utility variance of a random vector of portfolio weights $\hat{w}$ is*

$$\mathbb{V}[U(\hat{w})] = \mathbb{V}\left[\hat{w}^\top \mu\right] + \frac{\gamma^2}{4}\mathbb{V}\left[\hat{w}^\top \mathbf{\Sigma} \hat{w}\right] - \gamma\,\mathbb{C}\mathrm{ov}\left[\hat{w}^\top \mu, \hat{w}^\top \mathbf{\Sigma} \hat{w}\right]. \tag{21}$$

Kan, Wang, and Zhou (2021, Proposition 1) derive a stochastic representation of the out-of-sample mean return and variance of any combination between the SMV and SGMV portfolios. In the following proposition, we use this result to provide the closed-form analytical expression of the OOSU variance of the shrinkage portfolio defined in Equation (11).[13]

[11] 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 $\varepsilon = 0.76$, whereas our proposed shrinkage intensity $\kappa_R^\star = 0.0885$ maximizing the difference between OOSU mean and twice the OOSU standard deviation corresponds to a larger value of ε equal to 1.35.

[12] In Appendix IA.2.5, we extend our analysis to considering a shrinkage portfolio that combines the SMV portfolio, the SGMV portfolio, and the EW portfolio as in Tu and Zhou (2011).

[13] We are thankful to Raymond Kan for his helpful feedback, which greatly helped us obtain our main result in this section.

Proposition 3. *Let Assumptions 1 and 2 hold. Then, the out-of-sample utility variance of the shrinkage portfolio $\hat{w}^\star(\kappa)$ is*

$$\mathbb{V}[U(\hat{w}^\star(\kappa))] = \mathbb{V}[U(\hat{w}_g)] + \Delta(\kappa), \tag{22}$$

where

$$\mathbb{V}[U(\hat{w}_g)] = \frac{\sigma_g^2\psi^2}{T-N-1} + \frac{\gamma^2\sigma_g^4(N-1)(T-2)}{2(T-N-1)^2(T-N-3)} \tag{23}$$

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

$$\Delta(\kappa) = a_1\kappa^4 + a_2\kappa^3 + a_3\kappa^2 + a_4\kappa, \tag{24}$$

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

$$a_1 = \frac{1}{2\gamma^2}\frac{T^2(T-2)C(T,N,\psi^2)}{(T-N)^2(T-N-1)^2(T-N-2)(T-N-3)^2(T-N-5)(T-N-7)}, \tag{25}$$

$$a_2 = -\frac{2\psi^2}{\gamma^2}\frac{T^2(T-2)(T+N-3+2T\psi^2)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}, \tag{26}$$

$$\begin{aligned} a_3 = &\frac{\psi^2}{\gamma^2}\frac{2T(N+1)+T^2(T-N-3+2(T-N)\psi^2)}{(T-N)(T-N-1)^2(T-N-3)} \\ &+ \sigma_g^2\frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}, \end{aligned} \tag{27}$$

$$a_4 = -2\sigma_g^2\psi^2\frac{T(T-2)}{(T-N-1)^2(T-N-3)}, \tag{28}$$

where

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

Note from Proposition 3 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 use the analytical expression of OOSU variance in Proposition 3 to obtain the following Corollary.

Corollary 1. *Provided that ψ^2 is strictly positive, there is a nonzero shrinkage intensity $0 < \kappa < 1$ 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 even though the SGMV portfolio does not require estimating the vector of means, this estimated portfolio does not deliver the lowest OOSU variance, and it is still optimal to combine the SMV and SGMV portfolios to minimize OOSU variance in (22). We illustrate this point in Figure 1, where we depict in the horizontal axis the OOSU standard deviation of different shrinkage portfolios using the closed-form expression obtained in Proposition 3. We see that for the case considered in Figure 1, the shrinkage intensity that minimizes OOSU variance is $\kappa_V^\star = 0.0146 > 0$.

3.2 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 4. *The out-of-sample utility variance of the shrinkage portfolio $\hat{w}^\star(\kappa)$*

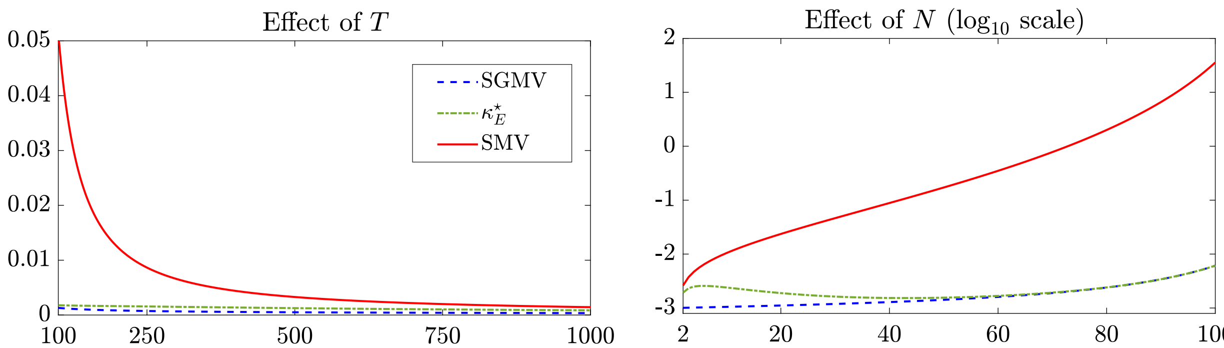
1. *decreases with the sample size T and converges to zero as $T \to \infty$,*

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 $\kappa \geq \kappa_E^\star$.*[14]

The results in Proposition 4 are intuitive because increasing N or decreasing T increases the statistical errors affecting the estimated moments of stock returns, and thus, they increase the OOSU volatility of the shrinkage portfolio. 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 4 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. Therefore, κ needs to be smaller than $\kappa_E^\star$ in order to reduce OOSU volatility.

[14]This is a sufficient but not necessary condition.

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 (dashed blue line), the shrinkage portfolio maximizing out-of-sample utility mean ($\kappa = \kappa_E^\star$, dash-dotted green line), and the SMV portfolio ($\kappa = 1$, solid red line). The population vector of means and covariance matrix of stock returns are calibrated from the monthly return data of the 25 portfolios of stocks sorted on size and book-to-market. 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.

Figure 2 illustrates the analytical results in Proposition 4. For the sake of 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. This gives a value of $\sigma_g = 0.0436$ and a value of $\psi^2 = 0.0625$. 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 study the OOSU volatility of three portfolios: the SGMV portfolio, the SMV portfolio, and the shrinkage portfolio maximizing OOSU mean with $\kappa = \kappa_E^\star$.

The left Panel in Figure 2 shows that the OOSU standard deviation of the SMV portfolio is substantially larger than that of the shrinkage portfolio exploiting $\kappa_E^\star$ and the SGMV portfolio. Specifically, for a realistic sample size of $T = 120$ observations, the OOSU standard deviation of the SMV portfolio is 29 times larger than that of the SGMV portfolio. In comparison, the OOSU standard deviation of the shrinkage portfolio is 1.51 times larger than that of the SGMV portfolio. Additionally, consistent with Proposition 4, 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 $T = 13{,}400$ 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 4.[15] The effect of the number of stocks N is particularly large for the SMV portfolio, for which we see that OOSU volatility increases much more rapidly than for the SGMV portfolio and the shrinkage portfolio exploiting $\kappa_E^\star$.

The analysis presented in this section indicates that the OOSU standard deviation of sample mean-variance portfolios can be substantial. Unlike SMV portfolios, a robust portfolio should offer a stable out-of-sample performance as well as a favorable average out-of-sample performance. This is the objective of the robustness measure introduced in the next section.

4 A new portfolio robustness measure

We now use the results in Section 3 to propose a new portfolio robustness measure defined as the difference between OOSU mean and a multiple of OOSU standard deviation. For notational simplicity, our presentation focuses on the shrinkage portfolio that combines the SMV and the SGMV portfolios. However, our results can be easily adapted to the SMV and SGMV portfolios by setting the shrinkage intensity κ to one and zero, respectively. Section 4.1 introduces the robustness measure. Section 4.2 studies the shrinkage portfolio that optimizes the proposed robustness measure. Section 4.3 explains how we estimate the shrinkage intensities. Section 4.4 describes the monotonicity properties of the robustness measure. Finally, Section 4.5 relates our proposed metric to the literature on robust optimization.

4.1 A new robustness measure

In our view, a robust portfolio should not only deliver good average performance but also a stable performance. In line with this view, we define our portfolio robustness measure as the difference between OOSU mean and a multiple of OOSU standard deviation for any estimated portfolio $\hat{w}$:

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

[15]In Figure 2, there is a range of values of N for which the OOSU standard deviation of the shrinkage portfolio exploiting $\kappa_E^\star$ decreases with N. This does not contradict the result in Proposition 4 because the results in this proposition consider a fixed κ. On the contrary $\kappa_E^\star$ in Figure 2 varies with N.

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 this robustness measure as the *mean-risk OOSU*. Our proposed mean-risk OOSU metric captures our view of portfolio robustness, and maximizing this metric delivers an efficient trade-off between OOSU mean and standard deviation. In Section 4.5, we show that the mean-risk OOSU criterion is equivalent to a robust optimization problem.

4.2 The robust optimal portfolio

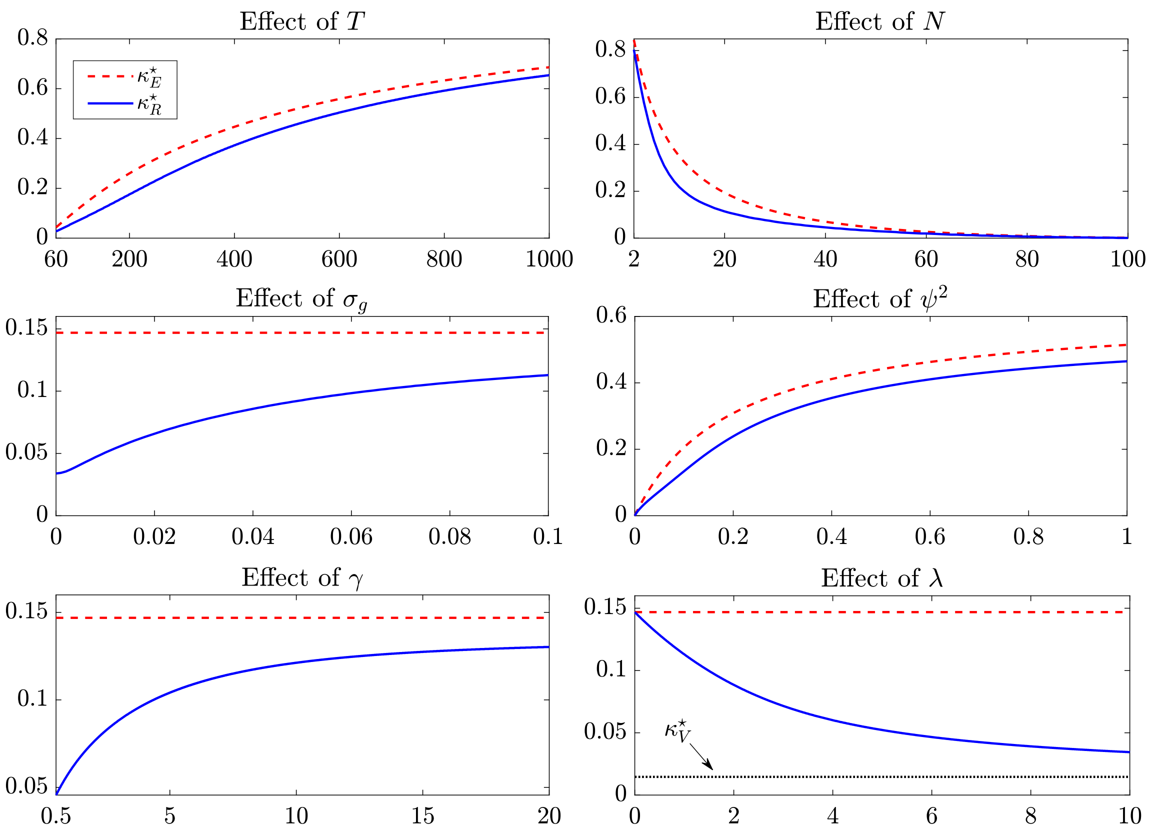
We now define our robust shrinkage portfolio that combines the SMV and SGMV portfolios to maximize the mean-risk OOSU defined in Section 4.1. Formally, the intensity of the robust shrinkage portfolio is the solution of the following problem:

$$\kappa_R^\star = \arg\max_{\kappa \in [0,1]} R(\hat{w}^\star(\kappa)), \tag{30}$$

which can be easily solved numerically using the analytical expressions for the OOSU mean in (14) and for the OOSU variance in (22). 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 $\kappa_V^\star$ is the shrinkage intensity minimizing OOSU variance.

Our measure of portfolio robustness resembles the efficient frontier of Markowitz (1952). In our case, instead of obtaining the optimal combination of stocks that minimizes portfolio risk for a given level of expected portfolio return, we obtain the optimal combination of the SMV and SGMV portfolios that minimizes OOSU risk for a given level of OOSU mean. Figure 1 depicts the OOSU efficient frontier for the 25SBTM dataset, $T = 120$, and $\gamma = 3$. Note that in this figure, the shrinkage intensity $\kappa_E^\star$ proposed by Kan, Wang, and Zhou (2021) delivers one of the multiple efficient shrinkage portfolios. In addition, we observe that the shrinkage portfolio that maximizes the mean-risk OOSU metric 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 OOSU mean of about 3.6% relative to the shrinkage portfolio that exploits $\kappa_E^\star$. Accordingly, our proposed robust shrinkage approach can deliver portfolios with a stable out-of-sample performance that perform well on average.

Figure 3: Monotonicity properties of optimal shrinkage intensities



Notes. This figure depicts the shrinkage intensity $\kappa_E^\star$ maximizing out-of-sample utility mean (dotted red line), and the shrinkage intensity $\kappa_R^\star$ maximizing the portfolio robustness measure in (29) (solid blue line), for different values of the six parameters that define the analytical expression of the portfolio robustness measure in Section 4. The population vector of means and covariance matrix of stock returns are calibrated from the monthly return data of the 25 portfolios of stocks sorted on size and book-to-market. The base-case values of the six parameters are $T = 120$, $N = 25$, $\sigma_g = 0.0436$, $\psi^2 = 0.0625$, $\gamma = 3$, and $\lambda = 2$. In each plot, we change the value of one of these six parameters while keeping the other five equal to the base-case value. In the bottom-right plot, $\kappa_V^\star$ is the shrinkage intensity minimizing out-of-sample utility variance.

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

Proposition 5. *Let Assumptions 1 and 2 hold. Then, the shrinkage intensity $\kappa_R^\star$ solving (30) 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 5 shows that our proposed robust shrinkage portfolio is asymptotically optimal. The second result of Proposition 5 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 further reduce OOSU variance and enhance portfolio robustness. Using the insight in Corollary 1 that $\kappa_V^\star > 0$ if $\psi^2 > 0$, the second result of Proposition 5 also implies that neither the SMV portfolio nor the SGMV portfolio optimizes our proposed measure of portfolio robustness for finite samples, and hence it is necessary to combine them using intensity $\kappa_R^\star$ to achieve maximal robust performance.

In Figure 3, we illustrate the monotonicity properties of the optimal shrinkage intensity $\kappa_R^\star$ that maximizes the mean-risk OOSU metric. We calibrate the parameters required to obtain the optimal shrinkage intensity using the 25SBTM dataset. In particular, we have $N = 25$, $\psi^2 = 0.0625$, and $\sigma_g = 0.0436$. In addition, we set $T = 120$, $\gamma = 3$, and $\lambda = 2$. We then change one parameter at a time to study the monotonicity properties of $\kappa_R^\star$.

We observe from Figure 3 that both $\kappa_E^\star$ and $\kappa_R^\star$ increase with T and decrease with N. This result is intuitive because statistical errors affecting the estimated moments of stock returns decrease with the ratio T/N and, thus, less shrinkage toward the SGMV portfolio is necessary when this ratio increases. Also, the difference between $\kappa_E^\star$ and $\kappa_R^\star$ becomes smaller as T increases because both κ's converge to one as T goes to infinity. Second, while $\kappa_E^\star$ is independent of σ_g, the proposed shrinkage intensity $\kappa_R^\star$ increases with σ_g because, as the return volatility of the GMV portfolio increases, shrinking toward the SGMV portfolio becomes less attractive in terms of OOSU risk. Third, we observe that both shrinkage intensities increase with ψ^2, but $\kappa_R^\star$ increases less rapidly because, as shown in Proposition 4, the OOSU standard deviation of the shrinkage portfolio deteriorates with ψ^2. Fourth, while $\kappa_E^\star$ is independent of γ, the proposed shrinkage intensity $\kappa_R^\star$ increases with γ and gets closer to $\kappa_E^\star$. This is because, as γ increases, the exposure to sample mean returns decreases, and this has the effect of reducing OOSU standard deviation, which gives more relevance to the OOSU mean in the mean-risk OOSU metric. Fifth, as the coefficient λ increases, the OOSU standard deviation of the shrinkage portfolio becomes a more relevant element of the mean-risk OOSU criterion. Therefore, the shrinkage intensity $\kappa_R^\star$ converges to $\kappa_V^\star$ when $\lambda \to \infty$. This insight is consistent with our second result in Proposition 5.

4.3 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 the OOSU mean of the shrinkage portfolio depends on parameter ψ^2. Since these parameters are unknown, it is important to obtain estimates of the optimal shrinkage intensities that are statistically consistent, which we propose in Appendix IA.1. Specifically, we estimate ψ^2 via the adjusted estimator of Kan and Zhou (2007) and we estimate σ_g^2 via the shrinkage estimator of Frahm and Memmel (2010). In the rest of the manuscript, we denote the estimated shrinkage intensities as $\hat{\kappa}_E^\star$ and $\hat{\kappa}_R^\star$.[16]

4.4 Monotonicity properties of the robustness measure

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

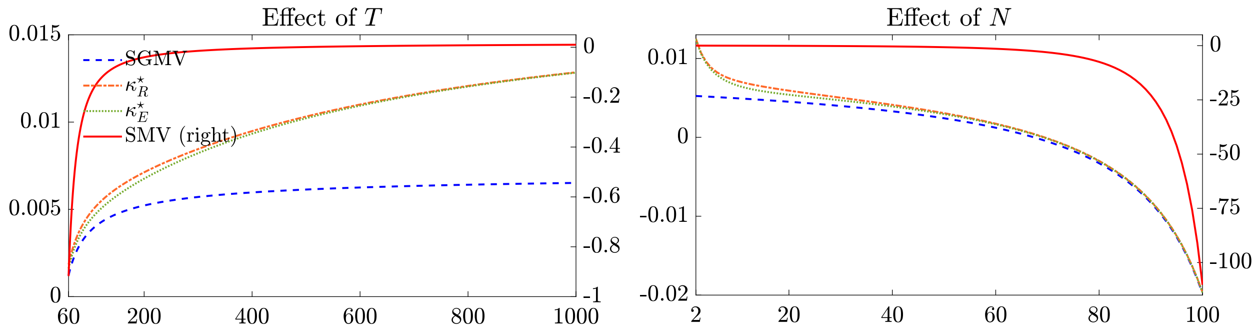
Proposition 6. *The mean-risk out-of-sample utility of the shrinkage portfolio $\hat{w}^\star(\kappa)$*

1. *increases with the sample size T and the mean return of the GMV portfolio μ_g,*
2. *decreases with the number of stocks N, the return variance of the GMV portfolio σ_g^2, and the shrinkage intensity κ if $\kappa \geq \kappa_E^\star$.*

Proposition 6 demonstrates that the mean-risk OOSU metric 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 metric 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

[16]Note that the result in Part 2 of Proposition 5 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.

Figure 4: Monotonicity properties of portfolio robustness measure



Notes. This figure depicts the portfolio robustness measure in (29) of four different portfolios: the SGMV portfolio (dashed blue), the shrinkage portfolio maximizing the robustness measure in Section 4 ($\kappa = \kappa_R^\star$, dash-dotted orange), the shrinkage portfolio maximizing out-of-sample utility mean ($\kappa = \kappa_E^\star$, dotted green), and the SMV portfolio (solid red, right y-axis). The population vector of means and covariance matrix of stock returns are calibrated from the monthly return data of the 25 portfolios of stocks sorted on size and book-to-market. We set a risk-aversion coefficient of $\gamma = 3$ and a coefficient $\lambda = 2$ for the robustness measure. We vary the sample size T in the left plot while keeping a fixed $N = 25$, and we vary the number of stocks N in the right plot while keeping a fixed $T = 120$.

shown in Proposition 4. Proposition 6 also demonstrates that allocating more weight to the SMV portfolio than $\kappa_E^\star$ necessarily deteriorates the mean-risk OOSU metric, which is why $\kappa_R^\star \leq \kappa_E^\star$ as highlighted in Proposition 5.[17]

Figure 4 illustrates the main insights highlighted in Proposition 4. For the sake of conciseness, we only show the results for the sample size T and the number of stocks N. The distributional parameters of stock returns are calibrated from the 25SBTM dataset. We assume that the risk-aversion coefficient is $\gamma = 3$, and the mean-risk OOSU metric is defined for $\lambda = 2$. In the left Panel of Figure 4, we vary the sample size T while keeping a fixed number of stocks $N = 25$, and in the right Panel, we vary N while keeping a fixed $T = 120$. We study the mean-risk OOSU metric of four portfolios: the SGMV portfolio, the SMV portfolio, the shrinkage portfolio maximizing OOSU mean with $\kappa = \kappa_E^\star$ in (16), and the proposed robust shrinkage portfolio maximizing the mean-risk OOSU metric with $\kappa = \kappa_R^\star$ in (30).

The left Panel in Figure 4 shows that the robustness measure improves with the sample

[17]Note that, for simplicity, we do not discuss the monotonicity properties of the robustness measure with respect to ψ^2. In unreported results, we show that the effect of ψ^2 on the mean-risk OOSU robustness metric is more nuanced. On the one hand, OOSU standard deviation always increases with ψ^2 as shown in Proposition 4. On the other hand, OOSU mean decreases with ψ^2 when the shrinkage intensity is large, i.e., when $\kappa \geq \frac{2(T-N)(T-N-3)}{T(T-2)}$, in which case the mean-risk OOSU metric also decreases with ψ^2. However, when κ is below this threshold, the impact that ψ^2 has on the mean-risk OOSU metric depends on the value of λ.

size. Also, we see that the shrinkage portfolio that maximizes OOSU mean does not deliver the largest portfolio robustness. For example, when $T = 120$, the mean-risk OOSU delivered by the shrinkage portfolio maximizing OOSU mean is 0.504% while the mean-risk OOSU metric achieved by our robust shrinkage portfolio is 0.547%.[18] Moreover, we show that the SMV portfolio is much less robust than the SGMV portfolio and it requires a sample size of $T = 696$ monthly observations to deliver a mean-risk OOSU metric as large as that of the SGMV portfolio.

The right Panel in Figure 4 shows that the portfolio robustness measure decreases with the number of stocks, and this reduction of portfolio robustness is particularly severe for the SMV portfolio. Additionally, when N is not too large, our proposed shrinkage portfolio exploiting $\kappa_R^\star$ delivers a substantially better mean-risk OOSU metric than that of the SGMV portfolio. In particular, for the base-case value of $N = 25$, the mean-risk OOSU delivered by $\kappa_R^\star$ is 0.547% while that of the SGMV portfolio is 0.426%.

4.5 Relation to robust optimization

In this Section, we interpret our proposed mean-risk OOSU criterion through the lens of robust optimization.[19] 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:

$$\mathcal{S}(\lambda, \kappa) = \widehat{\mathbb{E}}[U(\hat{w}^\star(\kappa))] \pm \lambda\sqrt{\widehat{\mathbb{V}}[U(\hat{w}^\star(\kappa))]}, \tag{31}$$

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 (31) can be interpreted as a confidence interval similar to Garlappi, Uppal, and Wang (2007). Accordingly, an ambiguity-

[18]The simulation study of Section 5.1 shows that the outperformance is more pronounced when using the estimated shrinkage intensities, $\hat{\kappa}_E^\star$ and $\hat{\kappa}_R^\star$.

[19]There is extensive literature on robust optimization and portfolio theory. Two of the most prominent papers in this literature are Goldfarb and Iyengar (2003) and Garlappi, Uppal, and Wang (2007).

averse mean-variance investor who wants to maximize OOSU mean solves the robust optimization problem

$$\max_{\kappa\in[0,1]} \min_{\mathcal{S}(\lambda,\kappa)} \mathbb{E}[U(\hat{w}^\star(\kappa))] = \max_{\kappa\in[0,1]} \widehat{\mathbb{E}}[U(\hat{w}^\star(\kappa))] - \lambda\sqrt{\widehat{\mathbb{V}}[U(\hat{w}^\star(\kappa))]}. \tag{32}$$

The above formulation corresponds to the optimization problem that delivers the estimated robust shrinkage intensity introduced in this section, i.e., $\hat{\kappa}^\star_R$. Therefore, our proposed shrinkage portfolio implicitly accounts for the estimation errors in the OOSU mean. This is important because estimation errors in the OOSU mean contaminate the estimated shrinkage intensity that maximizes OOSU mean, $\hat{\kappa}^\star_E$, and as Kan and Wang (2021) show, those estimation errors can severely affect the out-of-sample performance of shrinkage portfolios. We confirm this finding in the simulation analysis of Section 5.1, 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.

5 Empirical analysis

In this section, we characterize the economic benefits from exploiting our measure of portfolio robustness in the construction of shrinkage portfolios. For comparison purposes, we compare the performance of our robust shrinkage portfolio with that of several other benchmark portfolio strategies. We consider simulated return data in Section 5.1 and real return data in Section 5.2.

5.1 Simulated return data

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 iid Gaussian, and instead, we simulate data using the bootstrap method of Efron (1979). For the construction of the simulated data, we use the monthly returns of the six datasets considered in Kan, Wang, and Zhou (2021). The first four datasets are downloaded from Kenneth French's website: (i) 10 momentum portfolios (*10MOM*) from

January 1927 through December 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, (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.[20]

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

$$\mathbb{E}[U(\hat{w}^\star(\kappa))] \approx \frac{1}{M}\sum_{m=1}^{M} U(\hat{w}^\star_m(\kappa)), \tag{33}$$

$$\mathbb{V}[U(\hat{w}^\star(\kappa))] \approx \frac{1}{M}\sum_{m=1}^{M} \left(U(\hat{w}^\star_m(\kappa)) - \mathbb{E}[U(\hat{w}^\star(\kappa))]\right)^2, \tag{34}$$

$$R(\hat{w}^\star(\kappa)) \approx \mathbb{E}[U(\hat{w}^\star(\kappa))] - \lambda\sqrt{\mathbb{V}[U(\hat{w}^\star(\kappa))]}, \tag{35}$$

where $U(\hat{w}^\star_m(\kappa))$ is the investor's out-of-sample utility defined as in Equation (12) of the estimated shrinkage portfolio $\hat{w}^\star_m(\kappa)$ obtained from the mth simulated dataset. We set the risk-aversion coefficient to $\gamma = 3$ as in Kan and Zhou (2007) and Kan, Wang, and Zhou (2021). We also set the coefficient λ to $\lambda = 2$, which corresponds to a two-sigma uncertainty set around the estimated OOSU mean in (31).[21]

The simulation with iid Gaussian data is interesting because the theoretical results rely on

[20] We thank Kenneth French and Robert Novy-Marx for making their data publicly available.

[21] In the performance analysis with real return data in Section 5.2, we also consider the case with $\lambda = 4$ and with a cross-validated λ.

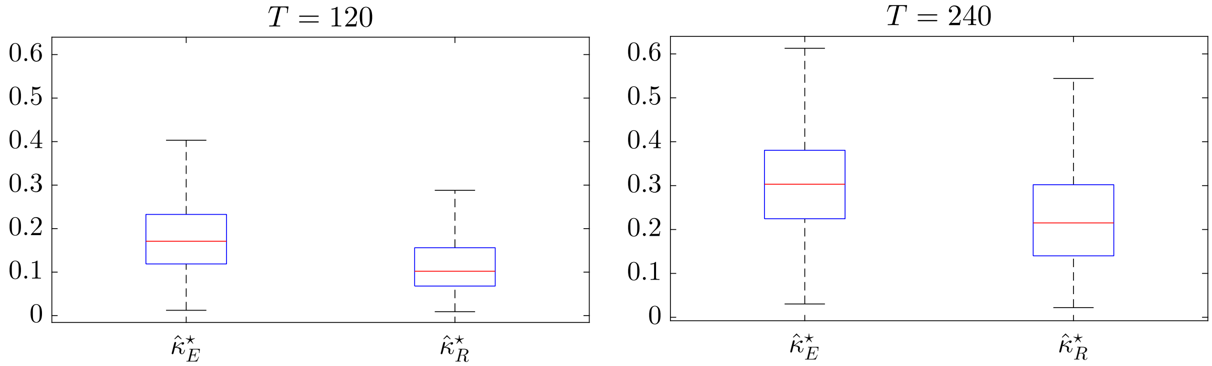
the assumption that stock returns are iid multivariate Gaussian; see Assumption 2. However, this assumption does not hold in practice, and therefore, the second type of simulated data is obtained by bootstrapping return data from the original sample. In particular we create 1,000 bootstrap samples of $2T$ 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 (33)–(35) from the 1,000 OOSU observations obtained from the 1,000 bootstrap samples.

5.1.1 Discussion of results for simulated return data

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 mean-risk OOSU measure with $\lambda = 2$. Panel A reports the performance of the estimated shrinkage portfolios that exploit the estimated intensities defined in Appendix IA.1, and Panel B reports the performance loss (in percentage) of the estimated shrinkage portfolios that exploit the estimated intensities defined in Appendix IA.1 relative to the performance of the estimated shrinkage portfolios that exploit the optimal but unfeasible shrinkage intensities.

We observe that for sample sizes of $T = 180$ and 240, the shrinkage portfolio that exploits $\hat{\kappa}_E^\star$ has a similar or larger OOSU mean than that of the shrinkage portfolio that exploits $\hat{\kappa}_R^\star$. This is reasonable because $\hat{\kappa}_E^\star$ is estimated to maximize OOSU mean under the assumption that stock returns are iid Gaussian, which is satisfied in this part of the analysis. However, when the sample size decreases to $T = 120$, 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. In unreported results, we also show that the shrinkage portfolio constructed with $\hat{\kappa}_R^\star$ systematically delivers an OOSU mean larger than that of the SMV and SGMV portfolios.

Figure 5: Boxplots of shrinkage intensities in Gaussian simulated data



Notes. These boxplots depict the estimated shrinkage intensity of the portfolio maximizing out-of-sample utility mean ($\hat{\kappa}^\star_E$) and the estimated shrinkage intensity of the portfolio maximizing the robustness measure in Section 4 ($\hat{\kappa}^\star_R$). The boxplots depict the estimated intensities from 100,000 simulated samples of T observations drawn from a multivariate Gaussian distribution whose moments are calibrated from the dataset of 25 portfolios of stocks sorted on size and book-to-market. We consider a sample size of $T = 120$ and 240, a risk-aversion coefficient of $\gamma = 3$, and a coefficient $\lambda = 2$ for the portfolio robustness measure.

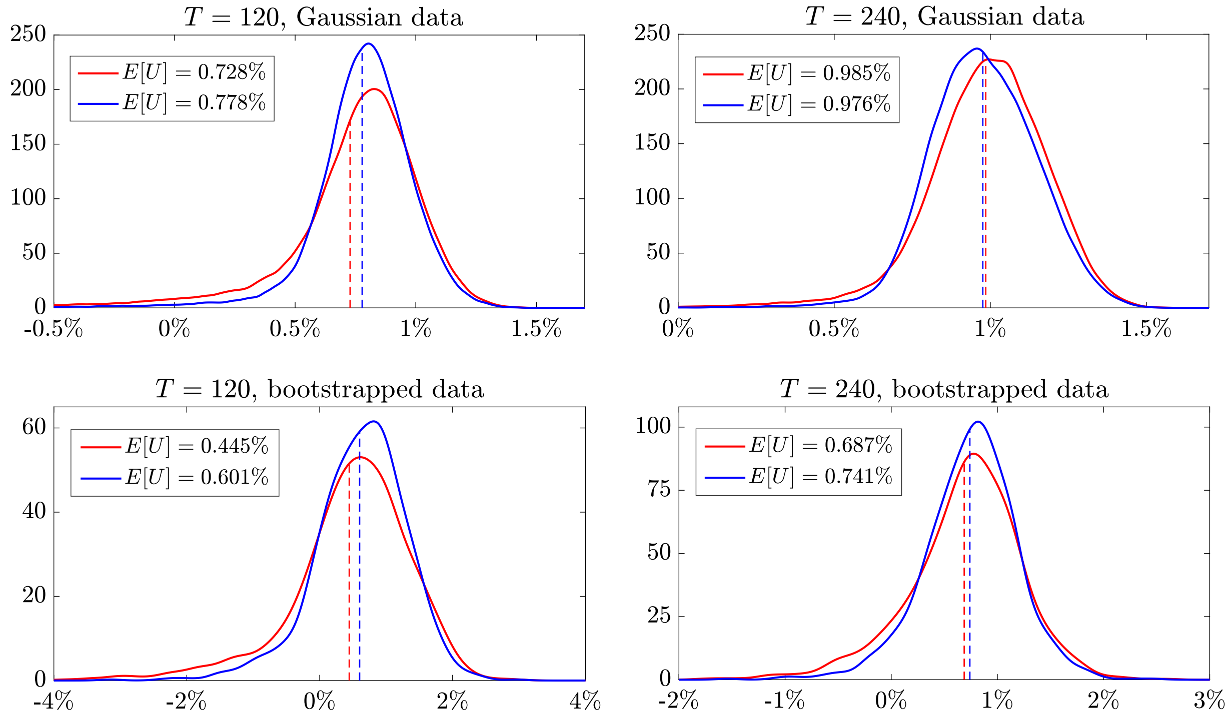
In addition, we observe that the shrinkage portfolio that exploits $\hat{\kappa}^\star_E$ delivers an OOSU that is notably more volatile than that of the shrinkage portfolio that exploits $\hat{\kappa}^\star_R$. The difference is particularly large for the case with a sample size of $T = 120$ observations.[22] Accordingly, the shrinkage portfolio that exploits the intensity $\hat{\kappa}^\star_E$ delivers a smaller mean-risk OOSU than that obtained by the shrinkage portfolio that exploits $\hat{\kappa}^\star_R$.

Panel B of Table 1 shows that the loss in OOSU mean and mean-risk OOSU relative to the optimal shrinkage intensity is generally larger for the shrinkage portfolio exploiting $\hat{\kappa}^\star_E$ than for the shrinkage portfolio exploiting $\hat{\kappa}^\star_R$. This is because the estimated $\hat{\kappa}^\star_E$ is more affected by statistical errors and is more unstable than the estimated $\hat{\kappa}^\star_R$. To see this, Figure 5 depicts the boxplots of the estimated shrinkage intensities $\hat{\kappa}^\star_E$ and $\hat{\kappa}^\star_R$ for the *25SBTM* dataset across all the simulated samples. We observe that $\hat{\kappa}^\star_E$ has a larger volatility, particularly for $T = 120$, which confirms that this shrinkage intensity is more sensitive to estimation errors.

Table 2 reports the performance results of the two shrinkage portfolios for the bootstrap data. In terms of OOSU mean, we see that the shrinkage portfolio exploiting $\hat{\kappa}^\star_R$ performs remarkably better. Specifically, it outperforms the shrinkage portfolio exploiting $\hat{\kappa}^\star_E$ for all sample sizes for five out of six datasets. Therefore, the performance of the shrinkage portfolio exploiting $\hat{\kappa}^\star_E$ deteriorates relative to that of the robust shrinkage portfolio once the data is

[22]On average across the six datasets, $\hat{\kappa}^\star_E$ yields a monthly OOSU volatility of 0.56%, versus 0.43% for $\hat{\kappa}^\star_R$.

Figure 6: Density of monthly out-of-sample utility in simulated data



Notes. This figure depicts the density of monthly out-of-sample utilities of the estimated shrinkage portfolios maximizing out-of-sample utility mean ($\hat{\kappa}^\star_E$, in red) and the portfolio robustness measure in Section 4 ($\hat{\kappa}^\star_R$, 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 25 portfolios of stocks sorted on size and book-to-market. The bottom two plots 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.

not Gaussian and the shrinkage intensities are estimated with error. In addition, the shrinkage intensity $\hat{\kappa}^\star_E$ yields an OOSU volatility that is on average across all datasets 31%, 25%, and 21% larger than that delivered by $\hat{\kappa}^\star_R$ for $T = 120$, 180, and 240 months, respectively.

Figure 6 summarizes the results from the simulation analysis. It 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 $T = 120$ or when the data is not iid Gaussian, the shrinkage portfolio that exploits $\hat{\kappa}^\star_R$ has a larger OOSU mean than that of the shrinkage portfolio that exploits $\hat{\kappa}^\star_E$. In addition, the shrinkage portfolio exploiting $\hat{\kappa}^\star_R$ yields a smaller OOSU volatility than that of the shrinkage portfolio exploiting $\hat{\kappa}^\star_E$, both for

$T = 120$ and 240. Finally, the OOSU density of the shrinkage portfolio exploiting $\hat{\kappa}_E^{\star}$ has a significantly heavier left tail than that of the shrinkage portfolio exploiting $\hat{\kappa}_R^{\star}$. This is not only true for the non-Gaussian bootstrap data but also when stock returns are Gaussian. The lower downside risk offered by our robust shrinkage portfolio represents an additional advantage of our proposed method.

5.2 Real return data

We now evaluate the out-of-sample performance of the two shrinkage portfolios considered in this manuscript and several benchmark portfolios using real return data. Real return data typically does not satisfy the Gaussian assumption that we use to develop the theoretical results. Therefore, this empirical analysis allows us to test the robustness of our considered portfolios to non-Gaussian data. The results presented in this section demonstrate that the shrinkage portfolio maximizing our proposed mean-risk OOSU criterion emerges as a robust portfolio that delivers favorable average out-of-sample performance *and* a stable out-of-sample performance even when stock returns are not Gaussian. We describe the performance evaluation methodology in Section 5.2.1, and we discuss the results in Section 5.2.2. We provide additional robustness tests in Appendix IA.2.

5.2.1 Portfolio strategies and performance evaluation

We use the same six datasets of characteristic and industry-sorted portfolios as those considered for the simulated return data in Section 5.1. In Appendix IA.2.3, we also consider a dataset of 50 individual stocks from the CRSP database.

We study the performance of eight portfolio strategies. First, the equally weighted (EW) portfolio. Second, the reward-to-risk (RTR) timing strategy of Kirby and Ostdiek (2012, Equation (13)) that is designed to outperform the EW portfolio while keeping a small turnover. 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, Wang, and Zhou (2021). The last three portfolio strategies correspond with different versions of the proposed shrinkage portfolio that exploits the intensity $\hat{\kappa}_R^\star$ maximizing the mean-risk OOSU criterion that we introduce in Section 4. We consider two fixed values of parameter λ and a cross-validated λ that is determined by the data. The two fixed values are $\lambda = 2$ and $\lambda = 4$. For the cross-validated λ, we use a three-fold cross-validation method similar to that used in Hastie, Tibshirani, and Friedman (2009), DeMiguel et al. (2009), or Ao, Li, and Zheng (2019).[23] We denote the cross-validated parameter as $\hat{\lambda}_{cv}$. We set the risk-aversion coefficient to $\gamma = 3$ as in Kan and Zhou (2007) and Kan, Wang, and Zhou (2021).[24]

Similar to DeMiguel, Garlappi, and Uppal (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 monthly returns in the dataset and T the sample size used to estimate the portfolios. Then, starting in month $T+1$, we estimate portfolio w using an estimation window that uses the first T monthly returns of our sample, and compute its 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 proceed similarly until the end of the sample, resulting in a time series of $\tau - T$ out-of-sample returns; i.e., $\tilde{p}_t$, $t = T+1, \dots, \tau$. Our experiments consider estimation windows of size $T = 120$ and 240 monthly observations.

We compute the portfolio turnover over the out-of-sample period as

$$\text{Turnover}_t = \sum_{i=1}^{N} |w_{i,t} - w_{i,(t-1)^+}|, \quad t = T+1, \dots, \tau, \tag{36}$$

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

$$p_{T+1} = \tilde{p}_{T+1} \quad \text{and} \quad p_t = (1+\tilde{p}_t)(1 - c \times \text{Turnover}_{t-1}) - 1, \quad t = T+2, \dots, \tau, \tag{37}$$

[23] We use a three-fold cross-validation method because is computationally fast and the method delivers good performance. However, we find in unreported results that using five-fold cross-validation delivers similar results.

[24] In Appendix IA.2.1, we also consider risk-aversion coefficients of $\gamma = 1$ and 5. In addition, in Appendix IA.2.4 and IA.2.5, we also consider portfolios that exploit the nonlinear shrinkage estimator of the covariance matrix of Ledoit and Wolf (2020b), and a shrinkage portfolio that combines the SMV, SGMV, and EW portfolios as in Tu and Zhou (2011).

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 similar to the level of proportional transaction costs considered by Kan, Wang, and Zhou (2021).[25]

Given the time series of out-of-sample returns net of proportional transaction costs p_t, we then compute the out-of-sample mean return and variance as

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

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

$$\text{OOS CER} = 12 \times \left(\mu_p - \frac{\gamma}{2}\sigma_p^2\right), \tag{38}$$

$$\text{OOS SR} = \sqrt{12} \times \mu_p / \sigma_p, \tag{39}$$

$$\text{Turnover} = \frac{1}{\tau - T} \sum_{t=T+1}^{\tau} \text{Turnover}_t. \tag{40}$$

We also test the null hypothesis that the OOS CER or SR delivered by $\hat{\kappa}_R^\star$ are equal to those delivered by $\hat{\kappa}_E^\star$, against the alternative hypothesis that $\hat{\kappa}_R^\star$ yields larger OOS CER or SR. We compute the test p-values using the block bootstrap approach of Politis and Romano (1994) with a block size of five and 1,000 bootstrap samples.

In addition to the aforementioned performance measures, we also assess the downside risk of the shrinkage portfolios. It is important to compare portfolio strategies in terms of their downside risk because it is a relevant dimension of portfolio performance for investors (Ang, Chen, and Xing, 2006; Bali, Demirtas, and Levy, 2009). We measure downside risk with the 1% and 5% Value-at-Risk of the out-of-sample portfolio returns. Finally, we also compute the mean-risk OOSU metric introduced in Section 4 of the two shrinkage portfolios. In particular, we divide the out-of-sample portfolio returns into non-overlapping three-year

[25] In Appendix IA.2.2, we show that our conclusions are robust to considering a larger value of $c = 30$ basis points.

windows,[26] and compute the OOS CER in each window. We then compute the mean and standard deviation of the OOS CER across all windows to obtain a measure of the portfolio mean-risk OOSU. For simplicity, we focus on the case with $\lambda = 2$; however, in unreported results, we confirm that our conclusions are robust to considering different λ's.

5.2.2 Discussion of results for real return data

Tables 3 and 4 report the main empirical results for portfolios constructed with window sizes of $T = 120$ and $T = 240$ monthly observations, respectively. First, we observe that the EW portfolio is not the optimal strategy. Indeed, it is only for the *49IND* dataset that it delivers the largest OOS CER and Sharpe ratio. For the five other datasets, the EW portfolio is outperformed by the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$. In comparison, the RTR portfolio achieves its intended goal because it systematically outperforms the EW portfolio while keeping a similar turnover. However, the RTR portfolio is also outperformed by the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ except for the *49IND* dataset.

Second, the SMV portfolio delivers the worst performance among all considered portfolio strategies. In particular, the SMV portfolio is systematically outperformed by the SGMV portfolio. However, we observe that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ has a larger OOS CER than the SGMV portfolio, both before and after transaction costs. Moreover, the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ has a comparable Sharpe ratio to that of the SGMV portfolio. In contrast, the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ generally outperforms the SGMV portfolio before transaction costs, but this outperformance often disappears after accounting for transaction costs because of its large turnover. This empirical evidence suggests that our proposed mean-risk OOSU criterion is a valuable and robust metric for combining the SMV and SGMV portfolios that fares well both in the absence and in the presence of transaction costs.

Third, we compare the proposed shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ to that of Kan, Wang, and Zhou (2021) exploiting $\hat{\kappa}_E^\star$. The results show that the shrinkage portfolio utilizing intensity $\hat{\kappa}_R^\star$, which accounts for both OOSU mean and volatility, delivers better out-of-sample certainty-equivalent return and Sharpe ratio, with the difference being statistically signifi-

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

cant. Looking at the case $\lambda = 2$, the shrinkage intensity $\hat{\kappa}_R^\star$ systematically delivers a better performance than $\hat{\kappa}_E^\star$. Moreover, it also delivers a lower turnover and, thus, an even greater improvement net of transaction costs. We also observe that these conclusions are robust to using a different fixed value of $\lambda = 4$. One may argue that keeping λ fixed for each estimation window is restrictive because the optimal value may change over time. This is why, in Tables 3 and 4, we also consider a cross-validated $\hat{\lambda}_{cv}$ that changes over time in a data-driven way. The results suggest that this helps improve out-of-sample performance. Specifically, the robust shrinkage portfolio that exploits $\hat{\lambda}_{cv}$ outperforms the robust shrinkage portfolio exploiting $\lambda = 2$ in nine out of 12 cases in terms of OOS CER,[27] and 10 out of 12 cases in terms of OOS Sharpe ratio, both before and after transaction costs.

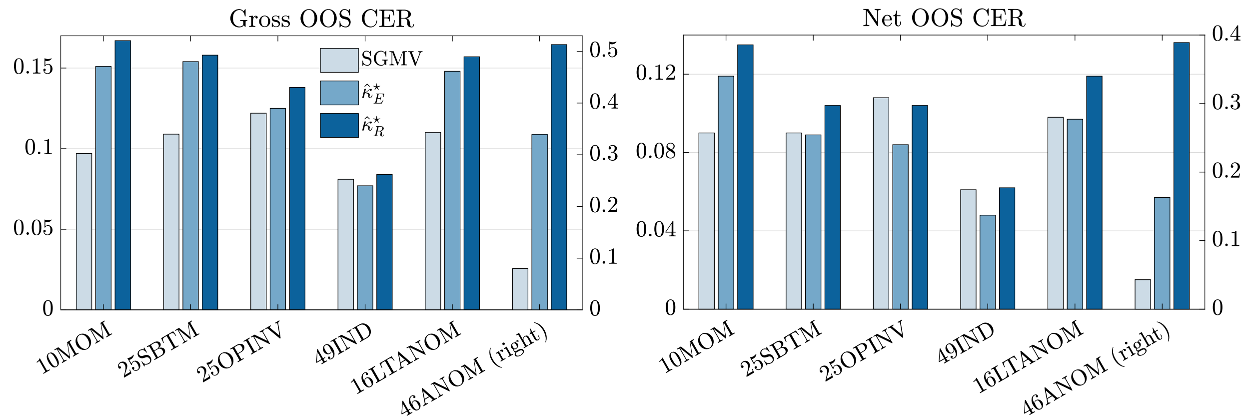
The previous results are illustrated in Figure 7 that depicts, for the case with $T = 120$, the out-of-sample CER delivered by the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ with a cross-validated parameter λ. For comparison, the figure also depicts the performance of the SGMV portfolio and the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$, which are the two benchmark portfolios that perform best in our analysis. The figure shows that the shrinkage portfolio using $\hat{\kappa}_E^\star$ delivers good performance relative to the SGMV portfolio before transaction costs, but this outperformance does not hold in general after accounting for transaction costs. In comparison, the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ yields a greater OOS CER than that of the SGMV portfolio and the shrinkage portfolio using $\hat{\kappa}_E^\star$, both before and after transaction costs.

To gauge the economic magnitude of the outperformance delivered by the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$, we report in Table 5 the cumulative wealth obtained by investing one dollar in the shrinkage portfolios using $\hat{\kappa}_E^\star$ and $\hat{\kappa}_R^\star$.[28] For this part of the analysis, 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 $\hat{\kappa}_E^\star$. For example, when $T = 120$, the cumulative wealth increases by 28% when $\lambda = 2$, 59% when $\lambda = 4$, and 75% when $\lambda = \hat{\lambda}_{cv}$, on average

[27]The 12 cases correspond with the results across the six datasets for the two sample sizes considered in the analysis.

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

Figure 7: Out-of-sample certainty-equivalent return in real data



Notes. This figure depicts the annualized out-of-sample certainty-equivalent return of the SGMV portfolio (first bar), the shrinkage portfolio maximizing out-of-sample utility mean ($\hat{\kappa}_E^\star$), and the shrinkage portfolio maximizing the portfolio robustness measure in Section 4 ($\hat{\kappa}_R^\star$) for the six datasets discussed in Section 5.1. We report the results of the robust shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ obtained from a cross-validated coefficient λ. We use a sample size of $T = 120$ monthly observations and a risk-aversion coefficient of $\gamma = 3$. The certainty-equivalent return is either gross (left plot) or net of proportional transaction costs of 20 basis points (right plot). The right y-axis reports the performance for the 46ANOM dataset for visibility.

across the six datasets.

Our fourth result is that, in addition to outperforming the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ in terms of OOS CER and Sharpe ratio, the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ delivers a more stable out-of-sample performance. In Table 6, we report the mean, standard deviation, and the mean-risk OOSU metric with $\lambda = 2$ of the out-of-sample CER performance measure introduced in Section 5.2.1. We find that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ delivers an OOS CER that is both larger on average and more stable over time. This result is consistent with our definition of portfolio robustness in the presence of parameter uncertainty.

Finally, in Table 7, we report the Value-at-Risk of the shrinkage portfolios exploiting $\hat{\kappa}_E^\star$ and $\hat{\kappa}_R^\star$. We observe that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ delivers a substantially lower downside risk than that of the portfolio exploiting $\hat{\kappa}_E^\star$, which is an additional empirical feature offered by our proposed robust shrinkage portfolio.

In Appendix IA.2, we assess the robustness of our results to considering different risk-aversion coefficients, a higher level of transaction costs, using a dataset of individual stocks, relying on a shrinkage estimator of the covariance matrix, and a different combination that exploits the SMV portfolio, the SGMV portfolio, and the equally weighted portfolio as in

Tu and Zhou (2011). We confirm that the insights from this section are robust to these alternative experimental settings.

6 Conclusion

Kan and Zhou (2007) analytically characterize the substantial out-of-sample utility (OOSU) losses that mean-variance investors experience, on average, due to parameter uncertainty. In this manuscript, we instead characterize the OOSU volatility of the sample mean-variance and global-minimum-variance portfolios, and show that SMV portfolios need unrealistically large sample sizes —for some datasets over 1,000 years of monthly return data— to deliver an out-of-sample performance as stable as that of SGMV portfolios. We 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 shrinkage portfolios that optimize our proposed measure of portfolio robustness tend to deliver higher certainty-equivalent returns, Sharpe ratios, and cumulative wealth with lower turnover and downside risk. Our framework can be applied to a broader range of settings than the one considered in the main body of the manuscript. For instance, we also utilize our framework to construct shrinkage portfolios that combine the SMV portfolio, the SGMV portfolio, and the equally weighted portfolio as in Tu and Zhou (2011). Regardless of the combination used in the empirical analysis, we show that our methodology provides robust portfolios that are more resilient to estimation errors and exhibit favorable out-of-sample performance.

Tables

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

		$\mathbb{E}[U(\hat{w}^\star(\kappa))]$			$\sqrt{\mathbb{V}[U(\hat{w}^\star(\kappa))]}$			$R(\hat{w}^\star(\kappa))$		
	T	120	180	240	120	180	240	120	180	240
Panel A: Estimated shrinkage intensities (in %)										
10MOM	$\hat{\kappa}^\star_E$	0.90	0.99	1.04	0.27	0.18	0.15	0.36	0.62	0.75
	$\hat{\kappa}^\star_R$	0.93	0.99	1.03	0.20	0.15	0.13	0.54	0.70	0.77
25SBTM	$\hat{\kappa}^\star_E$	0.73	0.89	0.99	0.37	0.25	0.20	0.00	0.40	0.58
	$\hat{\kappa}^\star_R$	0.78	0.89	0.98	0.24	0.19	0.17	0.29	0.52	0.63
25OPINV	$\hat{\kappa}^\star_E$	1.05	1.23	1.36	0.38	0.27	0.23	0.29	0.69	0.90
	$\hat{\kappa}^\star_R$	1.07	1.22	1.33	0.26	0.22	0.21	0.55	0.77	0.90
49IND	$\hat{\kappa}^\star_E$	0.71	0.93	1.08	0.45	0.29	0.23	-0.20	0.36	0.62
	$\hat{\kappa}^\star_R$	0.78	0.95	1.07	0.26	0.20	0.19	0.25	0.55	0.69
16LTANOM	$\hat{\kappa}^\star_E$	1.52	1.78	1.96	0.42	0.34	0.29	0.68	1.11	1.39
	$\hat{\kappa}^\star_R$	1.48	1.73	1.92	0.37	0.34	0.30	0.73	1.05	1.31
46ANOM	$\hat{\kappa}^\star_E$	6.20	8.40	9.68	1.46	1.12	0.91	3.28	6.17	7.86
	$\hat{\kappa}^\star_R$	6.06	8.33	9.63	1.26	1.03	0.86	3.55	6.26	7.91
Panel B: Percentage loss relative to optimal shrinkage intensities (in %)										
10MOM	$\hat{\kappa}^\star_E$	-9.25	-5.75	-4.48	107	49.4	28.9	-51.0	-22.6	-13.4
	$\hat{\kappa}^\star_R$	-3.22	-2.59	-3.18	109	64.4	43.7	-30.6	-16.8	-12.6
25SBTM	$\hat{\kappa}^\star_E$	-14.5	-8.54	-6.58	110	49.0	28.1	-100	-38.2	-21.4
	$\hat{\kappa}^\star_R$	-5.16	-4.37	-4.87	78.8	47.6	36.2	-47.1	-24.0	-18.5
25OPINV	$\hat{\kappa}^\star_E$	-10.8	-7.02	-5.34	82.6	34.6	20.7	-62.0	-25.1	-14.7
	$\hat{\kappa}^\star_R$	-4.15	-4.93	-5.20	73.4	43.7	35.4	-32.6	-20.5	-17.1
49IND	$\hat{\kappa}^\star_E$	-17.5	-9.68	-6.97	160	66.1	33.0	-139	-47.7	-23.9
	$\hat{\kappa}^\star_R$	-5.56	-4.55	-4.82	90.8	54.5	39.4	-54.0	-25.6	-18.8
16LTANOM	$\hat{\kappa}^\star_E$	-8.31	-5.38	-3.73	30.6	18.3	15.2	-33.3	-15.6	-9.84
	$\hat{\kappa}^\star_R$	-7.64	-6.78	-4.98	43.0	34.7	31.2	-32.0	-22.3	-15.7
46ANOM	$\hat{\kappa}^\star_E$	-3.97	-1.56	-0.80	14.9	6.79	4.19	-16.2	-4.27	-1.89
	$\hat{\kappa}^\star_R$	-3.82	-1.60	-0.83	18.5	9.18	5.50	-15.1	-4.71	-2.10

Notes. This table reports the out-of-sample performance of estimated shrinkage portfolios across six different types of simulated Gaussian 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 4. We report the results for the estimated shrinkage portfolio maximizing out-of-sample utility mean ($\hat{\kappa}^\star_E$) and for the estimated shrinkage portfolio maximizing the proposed robustness measure ($\hat{\kappa}^\star_R$). For each simulation, we define the population parameters of a multivariate Gaussian distribution with the sample moments of each of the six datasets described in Section 5.1, 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 (33)-(35). 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. Panel A reports the performance of the estimated shrinkage portfolios that exploit the estimated intensities defined in Appendix IA.1, and Panel B reports the performance loss (in percentage) of the estimated shrinkage portfolios that exploit the estimated intensities defined Appendix IA.1 relative to the performance of the estimated shrinkage portfolios that exploit the optimal but unfeasible shrinkage intensities.

Table 2: Out-of-sample performance of shrinkage portfolios in bootstrap data

		$\mathbb{E}[U(\hat{w}^\star(\kappa))]$			$\sqrt{\mathbb{V}[U(\hat{w}^\star(\kappa))]}$			$R(\hat{w}^\star(\kappa))$		
	T	120	180	240	120	180	240	120	180	240
10MOM	$\hat{\kappa}^\star_E$	0.73	0.82	0.86	0.86	0.70	0.53	-0.99	-0.58	-0.19
	$\hat{\kappa}^\star_R$	0.80	0.85	0.88	0.72	0.60	0.47	-0.64	-0.34	-0.05
25SBTM	$\hat{\kappa}^\star_E$	0.45	0.55	0.69	1.07	0.78	0.57	-1.69	-1.02	-046
	$\hat{\kappa}^\star_R$	0.60	0.65	0.74	0.80	0.60	0.49	-0.99	-0.56	-0.19
25OPINV	$\hat{\kappa}^\star_E$	0.78	0.88	0.96	0.81	0.46	0.41	-0.84	-0.10	0.14
	$\hat{\kappa}^\star_R$	0.88	0.92	0.97	0.61	0.38	0.33	-0.34	0.15	0.32
49IND	$\hat{\kappa}^\star_E$	0.44	0.53	0.60	0.70	0.44	0.34	-0.96	-0.34	-0.08
	$\hat{\kappa}^\star_R$	0.58	0.62	0.66	0.51	0.33	0.26	-0.43	-0.04	0.15
16LTANOM	$\hat{\kappa}^\star_E$	1.22	1.40	1.50	0.80	0.59	0.48	-0.38	0.21	0.54
	$\hat{\kappa}^\star_R$	1.21	1.36	1.47	0.63	0.49	0.41	-0.05	0.37	0.64
46ANOM	$\hat{\kappa}^\star_E$	3.85	5.36	6.32	3.36	2.65	1.97	-2.88	0.06	2.37
	$\hat{\kappa}^\star_R$	4.25	5.64	6.52	2.46	2.17	1.67	-0.67	1.30	3.18

Notes. This table reports the out-of-sample performance of estimated shrinkage portfolios across six different types of bootstrap 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 4. We report the results for the estimated shrinkage portfolio maximizing out-of-sample utility mean ($\hat{\kappa}^\star_E$) and for the estimated shrinkage portfolio maximizing the proposed robustness measure ($\hat{\kappa}^\star_R$). For each simulation, we construct 1,000 bootstrap samples of size $2T$ monthly observations from each of the six datasets described in Section 5.1. For each simulated bootstrap sample, we construct the two shrinkage portfolios using the first T observations, and compute the out-of-sample utility of the portfolio in the remaining T observations. Finally, we use the out-of-sample utilities of the 1,000 simulated bootstrap samples to construct our performance metrics as in (33)-(35). 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 3: Out-of-sample performance in real data with a sample size $T = 120$

		Benchmark strategies					*Proposed strategies* ($\hat{\kappa}_R^\star$)		
		EW	RTR	SGMV	SMV	$\hat{\kappa}_E^\star$	$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
10MOM	Gross OOS CER	0.069	0.082	0.097	-0.048	0.151	0.157	0.147	**0.167**
	Net OOS CER	0.068	0.081	0.090	-0.108	0.119	0.130**	0.123	**0.135**
	Gross OOS SR	0.659	0.755	0.876	0.868	0.963	0.970	0.939	1.009*
	Net OOS SR	0.653	0.748	0.829	0.765	0.876	0.888	0.861	0.902
	Turnover	0.040	0.047	0.300	2.704	1.360	1.158	1.022	1.331
	Average $\hat{\kappa}$	/	/	0	1	0.395	0.296	0.210	0.259
25SBTM	Gross OOS CER	0.080	0.086	0.109	-0.708	0.154	0.157	0.149	**0.158**
	Net OOS CER	0.079	0.085	0.090	-0.890	0.089	0.107**	**0.111***	0.104
	Gross OOS SR	0.705	0.744	0.994	0.711	0.962	1.002**	1.022*	1.084***
	Net OOS SR	0.700	0.738	0.857	0.457	0.748	0.801***	0.845**	0.820*
	Turnover	0.045	0.047	0.783	10.06	2.776	2.163	1.619	2.236
	Average $\hat{\kappa}$	/	/	0	1	0.214	0.145	0.089	0.104
25OPINV	Gross OOS CER	0.089	0.097	0.122	-0.814	0.125	0.135	0.135	**0.138**
	Net OOS CER	0.088	0.096	0.108	-0.952	0.084	0.103**	**0.113***	0.104
	Gross OOS SR	0.800	0.866	1.073	0.554	0.872	0.958***	1.054***	1.025**
	Net OOS SR	0.794	0.859	0.976	0.376	0.712	0.809***	0.922***	0.843**
	Turnover	0.040	0.048	0.569	6.982	1.730	1.293	0.920	1.395
	Average $\hat{\kappa}$	/	/	0	1	0.191	0.120	0.067	0.083
49IND	Gross OOS CER	0.092	**0.096**	0.081	-4.667	0.077	0.082	0.083	0.084
	Net OOS CER	0.091	**0.095**	0.061	-4.745	0.048	0.058*	0.062	0.062
	Gross OOS SR	0.816	0.864	0.788	0.244	0.704	0.763*	0.793*	0.801
	Net OOS SR	0.809	0.856	0.647	0.042	0.539	0.610**	0.649**	0.646
	Turnover	0.049	0.051	0.822	15.34	1.205	0.977	0.878	0.928
	Average $\hat{\kappa}$	/	/	0	1	0.043	0.024	0.016	0.008
16LTANOM	Gross OOS CER	0.078	0.087	0.110	-0.274	0.148	0.152	0.143	**0.157**
	Net OOS CER	0.077	0.086	0.098	-0.394	0.097	0.110	0.108	**0.119**
	Gross OOS SR	0.703	0.759	1.003	0.725	0.943	0.981	0.983	1.119**
	Net OOS SR	0.697	0.752	0.916	0.538	0.771	0.817**	0.826	0.917*
	Turnover	0.044	0.055	0.502	5.657	2.167	1.751	1.450	1.605
	Average $\hat{\kappa}$	/	/	0	1	0.309	0.218	0.151	0.131
46ANOM	Gross OOS CER	0.056	0.084	0.080	-14.74	0.339	0.541***	**0.602*****	0.513
	Net OOS CER	0.055	0.083	0.043	-11.46	0.163	0.363***	**0.435*****	0.389*
	Gross OOS SR	0.581	0.759	0.798	1.953	2.030	2.038	2.039	1.927
	Net OOS SR	0.575	0.750	0.523	1.612	1.777	1.795**	1.806*	1.616*
	Turnover	0.044	0.055	1.520	48.03	13.57	11.55	10.50	5.377
	Average $\hat{\kappa}$	/	/	0	1	0.250	0.207	0.181	0.063

Notes. This table reports the out-of-sample performance of the portfolio strategies introduced in Section 5.2.1 for the six datasets discussed in Section 5.1. 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), the annualized out-of-sample Sharpe ratio (OOS SR), and the monthly out-of-sample turnover of the portfolio strategies. For the two return-performance metrics (i.e., OOS CER and OOS SR), 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 $\star, \star\star, \star\star\star$ for the OOS CER and SR of the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ 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 numbers in bold font identify the best portfolio in terms of OOS CER.

Table 4: Out-of-sample performance in real data with a sample size $T = 240$

		Benchmark strategies					*Proposed strategies* ($\hat{\kappa}_R^\star$)		
		EW	RTR	SGMV	SMV	$\hat{\kappa}_E^\star$	$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
10MOM	Gross OOS CER	0.079	0.094	0.109	0.086	0.171	0.177⋆	0.179	**0.192**
	Net OOS CER	0.078	0.093	0.105	0.054	0.150	0.157⋆⋆	0.160⋆	**0.172**⋆
	Gross OOS SR	0.739	0.857	0.993	0.946	1.019	1.031⋆⋆	1.035	1.096⋆⋆⋆
	Net OOS SR	0.733	0.851	0.961	0.881	0.961	0.975⋆⋆	0.981⋆	1.029⋆⋆
	Turnover	0.039	0.037	0.184	1.416	0.915	0.838	0.790	0.831
	Average $\hat{\kappa}$	/	/	0	1	0.546	0.471	0.422	0.389
25SBTM	Gross OOS CER	0.092	0.098	0.127	-0.062	0.179	0.178	0.170	**0.181**
	Net OOS CER	0.091	0.097	0.117	-0.154	0.135	0.141	0.137	**0.147**
	Gross OOS SR	0.803	0.850	1.195	0.894	1.038	1.055⋆	1.050	1.183⋆⋆⋆
	Net OOS SR	0.797	0.845	1.110	0.738	0.900	0.922⋆⋆	0.922	1.021⋆⋆⋆
	Turnover	0.041	0.039	0.442	4.337	1.881	1.618	1.395	1.413
	Average $\hat{\kappa}$	/	/	0	1	0.364	0.289	0.217	0.187
25OPINV	Gross OOS CER	0.085	0.093	0.137	-0.184	0.159	**0.164**	0.163	0.154
	Net OOS CER	0.084	0.092	0.131	-0.246	0.133	0.144	**0.147**	0.133
	Gross OOS SR	0.786	0.859	1.225	0.679	1.009	1.099⋆⋆⋆	1.202⋆⋆⋆	1.095⋆
	Net OOS SR	0.780	0.853	1.176	0.576	0.908	1.002⋆⋆⋆	1.114⋆⋆⋆	0.985
	Turnover	0.039	0.036	0.271	2.772	1.048	0.848	0.637	0.871
	Average $\hat{\kappa}$	/	/	0	1	0.320	0.232	0.147	0.161
49IND	Gross OOS CER	0.080	**0.082**	0.061	-1.184	0.052	0.061	0.062	0.063
	Net OOS CER	0.079	**0.081**	0.052	-1.262	0.037	0.048⋆	0.052	0.053
	Gross OOS SR	0.752	0.793	0.678	0.162	0.563	0.642⋆⋆	0.678⋆	0.694
	Net OOS SR	0.745	0.787	0.607	0.050	0.471	0.556⋆⋆	0.600⋆	0.611⋆
	Turnover	0.048	0.040	0.361	4.332	0.634	0.503	0.421	0.429
	Average $\hat{\kappa}$	/	/	0	1	0.087	0.054	0.035	0.011
16LTANOM	Gross OOS CER	0.068	0.080	0.129	0.047	0.211	0.217	**0.221**	0.203
	Net OOS CER	0.067	0.079	0.122	-0.014	0.175	0.184	**0.191**	0.172
	Gross OOS SR	0.655	0.732	1.191	0.960	1.131	1.167⋆⋆	1.207⋆⋆	1.261⋆⋆
	Net OOS SR	0.650	0.726	1.139	0.854	1.024	1.061⋆⋆	1.100⋆⋆	1.119⋆
	Turnover	0.042	0.040	0.279	2.578	1.504	1.363	1.254	1.277
	Average $\hat{\kappa}$	/	/	0	1	0.522	0.449	0.388	0.309
46ANOM	Gross OOS CER	0.051	0.078	0.127	-6.346	-1.044	-0.781⋆⋆⋆	-0.665⋆⋆⋆	**0.327**⋆⋆⋆
	Net OOS CER	0.050	0.077	0.109	-5.933	-1.101	-0.850⋆⋆⋆	-0.737⋆⋆⋆	**0.237**⋆⋆⋆
	Gross OOS SR	0.551	0.723	1.173	1.364	1.428	1.436⋆⋆	1.441⋆⋆	1.482
	Net OOS SR	0.545	0.717	1.031	1.220	1.293	1.303⋆⋆⋆	1.309⋆⋆	1.331
	Turnover	0.046	0.042	0.769	15.52	8.488	7.912	7.626	4.401
	Average $\hat{\kappa}$	/	/	0	1	0.511	0.471	0.447	0.215

Notes. This table reports the out-of-sample performance of the portfolio strategies introduced in Section 5.2.1 for the six datasets discussed in Section 5.1. Each estimated portfolio is constructed using a sample 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), the annualized out-of-sample Sharpe ratio (OOS SR), and the monthly out-of-sample turnover of the portfolio strategies. For the two return-performance metrics (i.e., OOS CER and OOS SR), 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 $\star, \star\star, \star\star\star$ for the OOS CER and SR of the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ 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 numbers in bold font identify the best portfolio in terms of OOS CER.

Table 5: Cumulative wealth net of transaction costs of shrinkage portfolios in real data

	$\hat{\kappa}^\star_E$	$\hat{\kappa}^\star_R$		
		$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
Panel A: $T = 120$ (July 1983 – December 2013)				
10MOM	88.3	98.2 (+11%)	101 (+14%)	117 (+32%)
25SBTM	146	189 (+29%)	219 (+49%)	370 (+153%)
25OPINV	34.2	57.1 (+67%)	101 (+195%)	52.1 (+52%)
49IND	5.85	7.48 (+28%)	8.83 (+51%)	8.28 (+41%)
16LTANOM	27.6	33.7 (+22%)	35.6 (+29%)	88.6 (+221%)
46ANOM	2995	3248 (+8%)	3432 (+15%)	1411 (-53%)
Panel B: $T = 240$ (July 1993 – December 2013)				
10MOM	5.87	6.35 (+8%)	6.94 (+18%)	7.27 (+24%)
25SBTM	18.3	19.3 (+6%)	19.0 (+4%)	32.7 (+78%)
25OPINV	9.87	12.4 (+26%)	16.5 (+67%)	10.6 (+7%)
49IND	2.25	2.96 (+32%)	3.37 (+50%)	3.51 (+56%)
16LTANOM	8.78	10.0 (+14%)	11.8 (+35%)	13.8 (+57%)
46ANOM	46.5	47.9 (+3%)	48.8 (+5%)	52.2 (+12%)

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 ($\hat{\kappa}^\star_E$), and the estimated shrinkage portfolio that maximizes the portfolio robustness measure in Section 4 ($\hat{\kappa}^\star_R$). We report the cumulative wealth of the shrinkage portfolios for the six datasets discussed in Section 5.1. The out-of-sample returns 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). The figures in parenthesis report the percentage difference in cumulative wealth between the shrinkage portfolio exploiting $\hat{\kappa}^\star_R$ and the shrinkage portfolio exploiting $\hat{\kappa}^\star_E$. We use a risk-aversion coefficient of $\gamma = 3$, and coefficients $\lambda = 2$, $\lambda = 4$, and a cross-validated λ for the shrinkage portfolio that maximizes the proposed robustness measure.

Table 6: Out-of-sample certainty-equivalent returns of shrinkage portfolios in real data

		Mean OOS CER		Std dev OOS CER		Mean-risk OOS CER	
		$T = 120$	$T = 240$	$T = 120$	$T = 240$	$T = 120$	$T = 240$
Panel A: Without transaction costs (in %)							
10MOM	$\hat{\kappa}^\star_E$	1.332	1.523	2.510	2.739	-3.688	-3.956
	$\hat{\kappa}^\star_R$	1.372	1.562	2.273	2.577	-3.175	-3.592
25SBTM	$\hat{\kappa}^\star_E$	1.335	1.552	1.632	1.675	-1.929	-1.798
	$\hat{\kappa}^\star_R$	1.349	1.543	1.508	1.397	-1.667	-1.252
25OPINV	$\hat{\kappa}^\star_E$	0.986	1.229	1.321	1.327	-1.656	-1.425
	$\hat{\kappa}^\star_R$	1.087	1.326	1.096	1.125	-1.105	-0.925
49IND	$\hat{\kappa}^\star_E$	0.706	0.487	1.138	0.908	-1.570	-1.328
	$\hat{\kappa}^\star_R$	0.735	0.552	1.045	0.782	-1.355	-1.013
16LTANOM	$\hat{\kappa}^\star_E$	1.219	1.755	1.565	2.534	-1.911	-3.312
	$\hat{\kappa}^\star_R$	1.259	1.801	1.304	2.351	-1.349	-2.901
46ANOM	$\hat{\kappa}^\star_E$	5.880	-4.728	9.954	4.270	-14.03	-13.27
	$\hat{\kappa}^\star_R$	6.674	-3.009	8.903	3.676	-11.13	-10.36
Panel B: Net of transaction costs (in %)							
10MOM	$\hat{\kappa}^\star_E$	1.060	1.344	2.477	2.730	-3.894	-4.115
	$\hat{\kappa}^\star_R$	1.140	1.398	2.237	2.567	-3.333	-3.735
25SBTM	$\hat{\kappa}^\star_E$	0.780	1.182	1.433	1.706	-2.086	-2.230
	$\hat{\kappa}^\star_R$	0.914	1.224	1.311	1.424	-1.707	-1.624
25OPINV	$\hat{\kappa}^\star_E$	0.630	1.019	1.328	1.312	-2.025	-1.605
	$\hat{\kappa}^\star_R$	0.820	1.157	1.083	1.110	-1.346	-1.063
49IND	$\hat{\kappa}^\star_E$	0.463	0.364	1.113	0.930	-1.763	-1.496
	$\hat{\kappa}^\star_R$	0.538	0.454	1.018	0.799	-1.498	-1.144
16LTANOM	$\hat{\kappa}^\star_E$	0.801	1.466	1.546	2.507	-2.291	-3.548
	$\hat{\kappa}^\star_R$	0.919	1.541	1.276	2.322	-1.633	-3.103
46ANOM	$\hat{\kappa}^\star_E$	3.641	-5.681	8.857	4.298	-14.07	-14.28
	$\hat{\kappa}^\star_R$	4.657	-3.992	7.824	3.720	-10.99	-11.43

Notes. This table reports the out-of-sample certainty-equivalent return (OOS CER) mean, standard deviation, and mean-risk, defined as the difference between the mean and twice the standard deviation of shrinkage portfolios. We consider the estimated shrinkage portfolio that maximizes out-of-sample utility mean ($\hat{\kappa}^\star_E$), and the estimated shrinkage portfolio that maximizes the portfolio robustness measure in Section 4 ($\hat{\kappa}^\star_R$). We report the performance results of the shrinkage portfolios for the six datasets discussed in Section 5.1. 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 (33)-(35) 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.

Table 7: Value-at-risk of shrinkage portfolios in real data

		OOS 1% VaR				OOS 5% VaR			
		$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$			$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$		
			$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$		$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
Panel A: Without transaction costs									
$T = 120$	*10MOM*	0.241	0.216	0.203	0.207	0.116	0.101	0.095	0.091
	25SBTM	0.202	0.164	0.149	0.131	0.100	0.081	0.067	0.068
	25OPINV	0.203	0.180	0.116	0.138	0.089	0.073	0.058	0.060
	49IND	0.134	0.116	0.106	0.101	0.070	0.065	0.066	0.062
	16LTANOM	0.190	0.165	0.157	0.116	0.103	0.088	0.079	0.070
	46ANOM	0.522	0.453	0.426	0.159	0.220	0.186	0.163	0.074
$T = 240$	*10MOM*	0.288	0.272	0.266	0.213	0.129	0.121	0.112	0.102
	25SBTM	0.215	0.193	0.170	0.152	0.094	0.083	0.070	0.069
	25OPINV	0.198	0.159	0.130	0.138	0.100	0.081	0.066	0.078
	49IND	0.142	0.106	0.094	0.092	0.068	0.063	0.059	0.057
	16LTANOM	0.180	0.163	0.148	0.108	0.113	0.100	0.092	0.080
	46ANOM	0.962	0.873	0.817	0.500	0.381	0.358	0.331	0.156
Panel B: Net of transaction costs									
$T = 120$	*10MOM*	0.244	0.219	0.205	0.208	0.116	0.106	0.097	0.095
	25SBTM	0.216	0.168	0.155	0.138	0.109	0.085	0.071	0.073
	25OPINV	0.208	0.183	0.118	0.142	0.091	0.075	0.062	0.064
	49IND	0.136	0.118	0.108	0.102	0.073	0.067	0.068	0.063
	16LTANOM	0.193	0.170	0.162	0.119	0.109	0.093	0.081	0.073
	46ANOM	0.542	0.473	0.445	0.163	0.238	0.204	0.174	0.091
$T = 240$	*10MOM*	0.291	0.276	0.270	0.215	0.130	0.124	0.114	0.103
	25SBTM	0.218	0.196	0.174	0.153	0.100	0.088	0.073	0.071
	25OPINV	0.200	0.160	0.131	0.139	0.103	0.083	0.067	0.060
	49IND	0.143	0.107	0.095	0.093	0.069	0.064	0.060	0.057
	16LTANOM	0.182	0.165	0.151	0.115	0.116	0.102	0.094	0.083
	46ANOM	0.963	0.876	0.820	0.505	0.391	0.367	0.354	0.158

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 4 ($\hat{\kappa}_R^\star$) for the six datasets discussed in Section 5.1. We estimate the shrinkage portfolios using sample sizes of $T = 120$ and $T = 240$ monthly observations, a risk-aversion coefficient of $\gamma = 3$, and coefficients $\lambda = 2$, $\lambda = 4$, and a cross-validated λ 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 returns. Panel A considers the case without transaction costs, and Panel B considers the case with proportional transaction costs of 20 basis points.

References

Ang, A., J. Chen, and Y. Xing. 2006. Downside risk. *The Review of Financial Studies* 19:1191–239.

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

Avramov, D., and G. Zhou. 2010. Bayesian portfolio analysis. *The Annual Review of Financial Economics* 2:25–47.

Bali, T., O. Demirtas, and H. Levy. 2009. Is there an intertemporal relation between downside risk and expected returns? *Journal of Financial and Quantitative Analysis* 44:883–909.

Barroso, P., and K. Saxena. 2021. Lest we forget: using out-of-sample forecast errors in portfolio optimization. *The Review of Financial Studies* (forthcoming).

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

Brown, S. 1976. *Optimal portfolio choice under uncertainty.* Ph.D. Thesis, University of Chicago.

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

De Nard, G., O. Ledoit, and M. Wolf. 2019. Factor models for portfolio selection in large dimensions: the good, the better and the ugly. *Journal of Financial Econometrics* 1–22.

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

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

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

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

DeMiguel, V., Y. Plyakha, R. Uppal, and G. Vilkov. 2013. Improving portfolio selection using option-implied volatility and skewness. *Journal of Financial and Quantitative Analysis* 48:1813–45.

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

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

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

Füss, R., C. Koeppel, and F. Miebs. 2021. Diversifying estimation errors: An efficient averaging rule for portfolio optimization. *University of St.Gallen, School of Finance Research Paper No.2021/05*.

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

Giglio, S., B. T. Kelly, and D. Xiu. 2021. Factor models, machine learning, and asset pricing. *SSRN working paper* .

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

Goto, S., and Y. Xu. 2015. Improving mean variance optimization through sparse hedging restrictions. *Journal of Financial and Quantitative Analysis* 50:1415–41.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The elements of statistical learning: Prediction, inference and data mining (2nd ed.)*. Springer.

Jagannathan, R., and T. Ma. 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance* 58:1651–84.

Jorion, P. 1986. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* 21:279–92.

Kan, R., and D. Smith. 2008. The distribution of the sample minimum-variance frontier. *Management Science* 54:1364–80.

Kan, R., and X. Wang. 2021. Optimal portfolio choice with benchmark. *Rotman School of Management Working Paper No. 3760640*.

Kan, R., X. Wang, and X. Zheng. 2021. In-sample and out-of-sample sharpe ratios of multi-factor asset pricing models. *Available at SSRN 3454628*.

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

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

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

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

Kroll, Y., H. Levy, and H. Markowitz. 1984. Mean-variance versus direct utility maximization. *The Journal of Finance* 39:47–61.

Ledoit, O., and M. Wolf. 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10:603–21.

———. 2004. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis* 88:365–411.

———. 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies* 30:4349–88.

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

———. 2020b. The power of (non-)linear shrinking: A review and guide to covariance matrix estimation. *Journal of Financial Econometrics* 1–32.

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

———. 2014. Mean–variance approximations to expected utility. *European Journal of Operational Research* 234:346–55.

Merton, R. C. 1980. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* 8:323–61.

Novy-Marx, R., and M. Velikov. 2016. A taxonomy of anomalies and their trading costs. *The Review of Financial Studies* 29:104–47.

Politis, D., and J. Romano. 1994. The stationary bootstrap. *Journal of the American Statistical Association* 89:1303–13.

Rapponi, V., R. Uppal, and P. Zaffaroni. 2021. Robust portfolio choice. *Available at SSRN 3933063*.

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

———. 2011. Markowitz meets talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics* 99:204–15.

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

A Robust Approach to Optimal Portfolio Choice with Parameter Uncertainty

In Section IA.1, we provide details for the feasible estimators we use in the empirical analysis. In Section IA.2, we check the robustness of our main results to several variations of the main experimental setting considered in the main body of the manuscript. In Section IA.3, we report the proofs for all the theoretical results in the manuscript.

IA.1 Feasible estimators of shrinkage intensities

The shrinkage intensities $\kappa_E^\star$ and $\kappa_R^\star$ in Equations (16) and (30) are unfeasible because they depend on the unknown distributional parameters of stock returns. In this section, we provide details for the feasible estimators we use in the empirical analysis.

For the return variance of the zero-cost portfolio, which is defined in (8) 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

$$\hat{\psi}_{kz}^2 = \frac{(T-N-1)\hat{\psi}^2 - (N-1)}{T} + \frac{2(\hat{\psi}^2)^{\frac{N-1}{2}}(1+\hat{\psi}^2)^{-\frac{T-2}{2}}}{T \times B\left(\frac{\hat{\psi}^2}{1+\hat{\psi}^2}; \frac{N-1}{2}, \frac{T-N+1}{2}\right)}, \tag{IA1}$$

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 (7), we rely on the shrinkage portfolio estimator proposed by Frahm and Memmel (2010, Theorem 2), which provides smaller mean out-of-sample variance than the SGMV portfolio. Specifically, we estimate σ_g^2 as

$$\hat{\sigma}_g^2 = \hat{w}_{fm}^\top \widehat{\mathbf{\Sigma}} \hat{w}_{fm}, \tag{IA2}$$

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

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

with a shrinkage intensity

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

IA.2 Robustness tests of empirical results

We now assess the robustness of our results to considering different risk-aversion coefficients, a higher level of transaction costs, using a dataset of individual stocks, relying on a shrinkage estimator of the covariance matrix, and combining the SMV and SGMV portfolios with the equally weighted portfolio.

IA.2.1 Different risk-aversion coefficients

Tables IA.1 and IA.2 replicate the results in Tables 3 and 4 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.1, 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. The cross-validated λ generally delivers the best OOS CER before transaction costs, but it is often outperformed by the constant $\lambda = 4$ after transaction costs because fixing λ yields a smaller turnover. 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.2, 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. Further, we observe that the outperformance delivered by $\hat{\kappa}_R^\star$ is not too sensitive to the value of λ that is considered. However, the cross-validated λ tends to deliver the best out-of-sample performance. Overall, the results presented in this section confirm that the insights shown in the main body of the manuscript are robust to considering different risk-aversion parameters.

IA.2.2 Higher level of transaction costs

In Table IA.3, we replicate the results in Tables 3 and 4 using a higher level of proportional transaction cost of 30 basis points in Equation (37). 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.

The main insights of the manuscript are robust to considering a higher level of proportional transaction costs. Because the shrinkage portfolio exploiting $\hat{\kappa}^{\star}_{E}$ has a larger turnover than that of the shrinkage portfolio exploiting $\hat{\kappa}^{\star}_{R}$, the impact of a higher level of transaction costs on the performance of the shrinkage portfolio exploiting $\hat{\kappa}^{\star}_{E}$ is more severe than that of the shrinkage portfolio exploiting $\hat{\kappa}^{\star}_{R}$. Comparing the SGMV portfolio and the shrinkage portfolio exploiting $\hat{\kappa}^{\star}_{R}$, we find that $\hat{\kappa}^{\star}_{R}$ continues to deliver a better performance net of transaction costs in nearly all cases. However, the outperformance is not as large as in the case with 20 basis points because the SGMV portfolio has a smaller turnover. Finally, the shrinkage portfolio exploiting intensity $\hat{\kappa}^{\star}_{R}$ maintains a superior performance to that of the RTR portfolio, except for the *49IND* dataset.

IA.2.3 Dataset of individual stocks

The six datasets considered in the manuscript are industry-sorted and characteristic-sorted portfolios. We now assess the out-of-sample performance for a dataset of 50 individual stocks. We construct the dataset following the methodology in Jagannathan and Ma (2003) and DeMiguel, Garlappi, and Uppal (2009), among others. We download monthly stock returns from the Center for Research in Security Prices (CRSP) spanning September 1966 through December 2019. Starting from September 1986, and then every year, we identify all the stocks that have at least 80 monthly returns available during the past 20 years and the next year. Among those, we select the 50 stocks with the largest market capitalization to form our dataset for the next year, as in Barroso and Saxena (2021).

In Table IA.4, we report the out-of-sample performance of the six portfolio strategies considered in Tables 3 and 4. We observe that, for this dataset, the benefits from combining the SGMV portfolio with the SMV portfolio are more limited. Specifically, the SGMV portfolio has a better out-of-sample performance than the two shrinkage portfolios both before and after transaction costs. This is because, with individual stocks, the sample mean is largely contaminated by estimation risk. Moreover, Campbell, Lo, and Mackinlay (1997) explain that individual security returns display much less autocorrelation than portfolio returns and therefore past returns are not as good predictors of future returns. This finding is consistent with Barroso and Saxena (2021) who also find that, for datasets of individual stocks, minimum-variance portfolios outperform mean-variance portfolios. Our methodology better captures the challenges faced with individual stocks and assigns a shrinkage intensity $\hat{\kappa}_R^\star$ that is close to zero. In contrast, the shrinkage intensity $\hat{\kappa}_E^\star$ is larger, which results in a shrinkage portfolio that delivers a worse out-of-sample performance than that exploiting the shrinkage intensity $\hat{\kappa}_R^\star$ as well as the EW and RTR portfolios. In addition, we observe that the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ systematically outperforms the EW and RTR portfolios.

IA.2.4 Shrinkage estimator of the covariance matrix

The theoretical results in the manuscript consider the sample estimator of the covariance matrix in Equation (1) because its distribution is known under the assumption of iid Gaussian stock returns. However, it is well known that this estimator contains large statistical errors, especially when the number of stocks is large relative to the sample size. In this section, we explain how to use our methodology when we use a *shrinkage* estimator of the covariance matrix instead of the sample covariance matrix.

We use the latest developments in the field of covariance matrix estimation and rely on the nonlinear shrinkage estimator of Ledoit and Wolf (2020a). It is optimal for the minimum-variance loss function as in Ledoit and Wolf (2017) but is computationally much faster because it has an analytical solution. Denoting the estimator by $\widehat{\mathbf{\Sigma}}_{lw}$, the combination of the estimated mean-variance and minimum-variance portfolios is

$$\hat{w}_{lw}^\star(\kappa) = \frac{\widehat{\mathbf{\Sigma}}_{lw}^{-1} e}{e^\top \widehat{\mathbf{\Sigma}}_{lw}^{-1} e} + \frac{\kappa}{\gamma} \widehat{\mathbf{\Sigma}}_{lw}^{-1} \left(\mathbf{I} - \frac{e e^\top \widehat{\mathbf{\Sigma}}_{lw}^{-1}}{e^\top \widehat{\mathbf{\Sigma}}_{lw}^{-1} e} \right) \hat{\mu}. \tag{IA3}$$

Analytical expressions for the OOSU mean and variance of portfolio $\hat{w}_{lw}^\star(\kappa)$ are not available and, thus, the optimal shrinkage intensities maximizing OOSU mean and mean-risk OOSU are unknown. As a remedy, Kan, Wang, and Zhou (2021) propose to deploy the optimal κ derived under the *sample* covariance matrix. However, this approach overstates the impact of estimation errors affecting the covariance matrix. Therefore, 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.

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.1. *Let $N \geq 2$, $T > N$, and Assumption 2 holds. When the covariance matrix $\mathbf{\Sigma}$ is known, the mean-risk out-of-sample utility of the shrinkage portfolio $\hat{w}^\star(\kappa)$ is*

$$R(\hat{w}^\star(\kappa)) = \mu_g - \frac{\gamma}{2}\sigma_g^2 + \frac{1}{\gamma}\left(\kappa\psi^2 - \frac{\kappa^2}{2}\left(\psi^2 + \frac{N-1}{T}\right) - \lambda\frac{\kappa\psi}{\sqrt{T}}\sqrt{(1-\kappa)^2 + \kappa^2\frac{N-1}{2T\psi^2}}\right). \quad \text{(IA4)}$$

Moreover, the optimal shrinkage intensity $\kappa_R^\star \leq \kappa_E^\star$, 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, contrary to the case that exploits the sample covariance matrix instead of the true covariance matrix, $\kappa_R^\star$ does not depend on the parameters γ and σ_g^2.

In Table IA.5, we report the out-of-sample performance of the shrinkage portfolio $\hat{w}_{lw}^\star(\kappa)$ in (IA3) using the estimated shrinkage intensities $\hat{\kappa}_E^\star$ and $\hat{\kappa}_R^\star$ obtained from Proposition IA.1.[29] The results in Table IA.5 confirm that our main insights are robust to considering the Ledoit and Wolf (2020a) shrinkage covariance matrix. The SMV portfolio remains by far the worst strategy and is outperformed by the SGMV portfolio. However, it is still useful to combine the two portfolios to enhance out-of-sample performance. The shrinkage portfolio exploiting intensity $\hat{\kappa}_E^\star$ often improves the gross OOS CER relative to the SGMV portfolio, but this improvement often disappears net of transaction costs. In contrast, our robust methodology to determine the optimal shrinkage intensity delivers, in general, better performance. For example, when $\lambda = 2$, the shrinkage portfolio exploiting $\hat{\kappa}_R^\star$ outperforms the shrinkage portfolio exploiting $\hat{\kappa}_E^\star$ in terms of OOS CER in all cases, both before and after transaction costs. Moreover, similar to the results in the main body of

[29] Note that it is straightforward to obtain the optimal value of $\hat{\kappa}_E^\star$ from Equation (IA4) by setting $\lambda = 0$.

the manuscript, we find that the cross-validated λ generates even superior performance. Compared to the SGMV portfolio, the cross-validated λ yields a greater OOS CER in 11 out of 12 cases before transaction costs and eight out of 12 cases after transaction costs. This is a remarkable result because GMV portfolios estimated with shrinkage estimators of the covariance matrix are notoriously difficult to outperform in practice.

IA.2.5 Combining with the equally weighted portfolio

Motivated by the finding of DeMiguel, Garlappi, and Uppal (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

$$\hat{w}^\star(\delta, \kappa) = (1-\delta)w_{ew} + \delta \hat{w}^\star(\kappa) = (1-\delta)w_{ew} + \delta((1-\kappa)\hat{w}_g + \kappa \hat{w}^\star), \tag{IA5}$$

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

$$\mu_{ew} = w_{ew}^\top \mu \quad \text{and} \quad \sigma_{ew}^2 = w_{ew}^\top \mathbf{\Sigma} w_{ew}, \tag{IA6}$$

for the mean return and variance of the EW portfolio.

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

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

$$\mathbb{E}[U(\hat{w}^\star(\delta, \kappa))] = (1-\delta)\mu_{ew} + \delta\left(\mu_g + \frac{\kappa}{\gamma}\frac{T}{T-N-1}\psi^2\right) - \frac{\gamma}{2}\left((1-\delta)^2\sigma_{ew}^2\right.$$

$$+ \delta^2\left(\frac{T-2}{T-N-1}\sigma_g^2 + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)(T\psi^2+N-1)}{(T-N)(T-N-1)(T-N-3)}\right)$$
$$+ 2\delta(1-\delta)\left(\sigma_g^2 + \frac{\kappa}{\gamma}\frac{T}{T-N-1}(\mu_{ew}-\mu_g)\right)\Bigg). \tag{IA7}$$

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

$$\mathbb{V}[U(\hat{w}^\star(\delta,\kappa))] = \mathbb{V}\Big[\hat{w}^\star(\delta,\kappa)^\top\mu\Big] + \frac{\gamma^2}{4}\mathbb{V}\Big[\hat{w}^\star(\delta,\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\delta,\kappa)\Big]$$
$$- \gamma\,\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\delta,\kappa)^\top\mu, \hat{w}^\star(\delta,\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\delta,\kappa)\Big], \tag{IA8}$$

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

$$\mathbb{V}\Big[\hat{w}^\star(\delta,\kappa)^\top\mu\Big] = \delta^2\left(\frac{\sigma_g^2\psi^2}{T-N-1} + \frac{\kappa^2\psi^2}{\gamma^2}\frac{2T(N+1)+T^2(T-N-3+2(T-N)\psi^2)}{(T-N)(T-N-1)^2(T-N-3)}\right), \tag{IA9}$$

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

$$\mathbb{V}\Big[\hat{w}^\star(\delta,\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\delta,\kappa)\Big] = \delta^4\Bigg(\frac{2\sigma_g^4(N-1)(T-2)}{(T-N-1)^2(T-N-3)}$$
$$+ \frac{4\kappa^2\sigma_g^2}{\gamma^2}\frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}$$
$$+ \frac{2\kappa^4}{\gamma^4}\frac{T^2(T-2)C(T,N,\psi^2)}{(T-N)^2(T-N-1)^2(T-N-2)(T-N-3)^2(T-N-5)(T-N-7)}\Bigg)$$
$$+ 4\delta^2(1-\delta)^2\Bigg((\sigma_{ew}^2-\sigma_g^2)\left(\frac{\sigma_g^2}{T-N-1} + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)+T^2\psi^2}{(T-N)(T-N-1)(T-N-3)}\right)$$
$$+ \frac{\kappa^2}{\gamma^2}\frac{T^2(T-N+1)}{(T-N)(T-N-1)^2(T-N-3)}(\mu_{ew}-\mu_g)^2\Bigg) + 8\delta^3(1-\delta)\frac{\kappa}{\gamma}(\mu_{ew}-\mu_g)$$
$$\left(\sigma_g^2\frac{T(T-2)}{(T-N-1)^2(T-N-3)} + \frac{\kappa^2}{\gamma^2}\frac{T^2(T-2)(T+N-3+2T\psi^2)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}\right), \tag{IA10}$$

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

$$\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\delta,\kappa)^\top\mu, \hat{w}^\star(\delta,\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\delta,\kappa)\Big] = 2\delta^3\frac{\kappa}{\gamma}\Bigg(\sigma_g^2\psi^2\frac{T(T-2)}{(T-N-1)^2(T-N-3)}$$

$$+\frac{\kappa^2\psi^2}{\gamma^2}\frac{T^2(T-2)(T+N-3+2T\psi^2)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}\Bigg)+2\delta^2(1-\delta)(\mu_{ew}-\mu_g)$$
$$\left(\frac{\sigma_g^2}{T-N-1}+\frac{\kappa^2}{\gamma^2}\frac{T(T-2)(T-N-1)+2T^2(T-N)\psi^2}{(T-N)(T-N-1)^2(T-N-3)}\right). \tag{IA11}$$

Using the result in Proposition IA.2, we can find numerically the shrinkage intensities $(\delta_E^\star, \kappa_E^\star)$ maximizing OOSU mean and the shrinkage intensities $(\delta_R^\star, \kappa_R^\star)$ maximizing our proposed mean-risk OOSU robustness metric in (29). 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.1. We estimate μ_{ew} and σ_{ew}^2 via the plug-in estimators

$$\hat{\mu}_{ew}=w_{ew}^\top\hat{\mu} \quad \text{and} \quad \hat{\sigma}_{ew}^2=w_{ew}^\top\widehat{\boldsymbol{\Sigma}}w_{ew}, \tag{IA12}$$

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 π of this shrinkage portfolio as the parameter π that minimizes the mean squared error of the out-of-sample mean return of the portfolio.

Proposition IA.3. *Let Assumptions 1 and 2 hold. Then, the shrinkage portfolio $\hat{w}(\pi) = \pi w_{ew} + (1-\pi)\hat{w}_g$ that minimizes the mean squared error $\mathbb{E}[(\hat{w}(\pi)'\mu-\mu_g)^2]$ is obtained for*

$$\pi=\frac{\sigma_g^2\psi^2}{\sigma_g^2\psi^2+(\mu_{ew}-\mu_g)(T-N-1)}. \tag{IA13}$$

Using Proposition IA.3, we estimate μ_g as

$$\hat{\mu}_g=\hat{w}(\hat{\pi})^\top\hat{\mu} \quad \text{with} \quad \hat{\pi}=\frac{\hat{\sigma}_g^2\hat{\psi}_{kz}^2}{\hat{\sigma}_g^2\hat{\psi}_{kz}^2+(\hat{\mu}_{ew}-\hat{w}_g^\top\hat{\mu})(T-N-1)}, \tag{IA14}$$

where $\hat{\psi}_{kz}^2$ is defined in (IA1), $\hat{\sigma}_g^2$ in (IA2), and $\hat{\mu}_{ew}$ in (IA12). Finally, using those estimators of the distributional parameters of stock returns, we can obtain the estimated shrinkage intensities $(\hat{\delta}_E^\star, \hat{\kappa}_E^\star)$ and $(\hat{\delta}_R^\star, \hat{\kappa}_R^\star)$ numerically using the results in Proposition IA.2.

In Tables IA.6 and IA.7, we report the out-of-sample performance of the shrinkage portfolio $\hat{w}^\star(\delta,\kappa)$ in (IA5) that combines the SMV, SGMV, and EW portfolios using the intensities $(\hat{\delta}^\star_E, \hat{\kappa}^\star_E)$ and $(\hat{\delta}^\star_R, \hat{\kappa}^\star_R)$. As in Tables 3 and 4, we consider constant values of $\lambda = 2$ and $\lambda = 4$, as well as a three-fold cross-validated λ.

We observe that the results in the manuscript are robust to considering the EW portfolio in the shrinkage portfolios. Specifically, the robust shrinkage portfolio optimized for $\lambda = 2$ yields 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. As in Tables 3 and 4, we also find that the cross-validated λ yields strong out-of-sample performance. In particular, the cross-validated robust shrinkage portfolio outperforms the SMV, SGMV, and EW portfolios in terms of OOS CER in all cases before transaction costs and in nearly all cases after transaction costs.

Overall, the results presented in this section 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.

Table IA.1: Out-of-sample performance in real data with lower risk aversion

	$T = 120$						$T = 240$					
	RTR	SGMV	$\hat{\kappa}^{\star}_{E}$	$\hat{\kappa}^{\star}_{R}$			RTR	SGMV	$\hat{\kappa}^{\star}_{E}$	$\hat{\kappa}^{\star}_{R}$		
				$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$				$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
10MOM												
Gross OOS CER	0.107	0.120	0.274	0.289	0.204	**0.328**	0.115	0.128	0.308	0.324	0.332	**0.380**
Net OOS CER	0.106	0.112	0.186	0.217	0.147	**0.239**	0.114	0.124	0.248	0.270	0.283	**0.324**
Gross OOS SR	0.755	0.876	0.783	0.773	0.655	0.811	0.857	0.993	0.819	0.822	0.824	0.876
Net OOS SR	0.748	0.829	0.693	0.690	0.580	0.694	0.851	0.961	0.760	0.766	0.770	0.805
Turnover	0.047	0.300	3.955	3.219	2.571	3.779	0.037	0.184	2.681	2.391	2.212	2.372
Average $\hat{\kappa}$	/	0	0.395	0.273	0.147	0.249	/	0	0.546	0.454	0.397	0.374
25SBTM												
Gross OOS CER	0.120	0.128	0.265	0.269	0.211	**0.281**	0.124	0.143	0.292	0.290	0.243	**0.301**
Net OOS CER	0.119	0.109	0.085	0.137	**0.140**	0.138	0.123	0.133	0.171	0.188	0.158	**0.208**
Gross OOS SR	0.744	0.994	0.734	0.742	0.689	0.849	0.850	1.195	0.774	0.762	0.698	0.844
Net OOS SR	0.738	0.857	0.505	0.529	0.539	0.534	0.845	1.110	0.633	0.626	0.567	0.666
Turnover	0.047	0.783	7.920	5.728	3.070	5.910	0.039	0.442	5.416	4.536	3.727	3.866
Average $\hat{\kappa}$	/	0	0.214	0.131	0.053	0.099	/	0	0.364	0.280	0.183	0.185
25OPINV												
Gross OOS CER	0.121	0.142	0.152	0.174	0.155	**0.187**	0.115	0.156	0.226	**0.238**	0.214	0.206
Net OOS CER	0.120	0.128	0.037	0.096	**0.124**	0.102	0.114	0.149	0.154	0.180	**0.181**	0.150
Gross OOS SR	0.866	1.073	0.562	0.607	0.865	0.695	0.859	1.225	0.673	0.738	0.896	0.705
Net OOS SR	0.859	0.976	0.386	0.439	0.716	0.463	0.853	1.176	0.559	0.618	0.783	0.569
Turnover	0.048	0.569	4.899	3.244	1.259	3.519	0.036	0.271	3.025	2.374	1.360	2.351
Average $\hat{\kappa}$	/	0	0.191	0.104	0.025	0.079	/	0	0.320	0.221	0.092	0.156
49IND												
Gross OOS CER	**0.119**	0.100	0.091	0.098	0.105	0.112	**0.102**	0.076	0.047	0.078	0.081	0.082
Net OOS CER	**0.117**	0.080	0.027	0.064	0.083	0.079	**0.101**	0.067	0.010	0.055	0.070	0.066
Gross OOS SR	0.864	0.788	0.436	0.595	0.792	0.786	0.793	0.678	0.312	0.469	0.672	0.684
Net OOS SR	0.856	0.647	0.251	0.421	0.645	0.584	0.787	0.607	0.208	0.365	0.587	0.562
Turnover	0.051	0.822	2.676	1.430	0.902	1.355	0.040	0.361	1.552	0.937	0.474	0.669
Average $\hat{\kappa}$	/	0	0.043	0.016	0.006	0.006	/	0	0.087	0.041	0.015	0.007
16LTANOM												
Gross OOS CER	0.118	0.129	0.241	0.245	0.188	**0.271**	0.107	0.146	0.380	0.398	**0.417**	0.369
Net OOS CER	0.117	0.117	0.095	0.132	0.096	**0.162**	0.106	0.139	0.273	0.301	**0.329**	0.277
Gross OOS SR	0.759	1.003	0.702	0.703	0.625	0.874	0.732	1.191	0.877	0.893	0.926	0.959
Net OOS SR	0.752	0.916	0.514	0.523	0.440	0.608	0.726	1.139	0.764	0.779	0.813	0.789
Turnover	0.055	0.502	6.336	4.836	3.927	4.479	0.040	0.279	4.515	4.087	3.636	3.745
Average $\hat{\kappa}$	/	0	0.309	0.201	0.117	0.130	/	0	0.522	0.443	0.368	0.299
46ANOM												
Gross OOS CER	0.111	0.097	0.833	1.442	**1.619**	1.410	0.104	0.144	-3.363	-2.576	-2.227	**0.753**
Net OOS CER	0.109	0.061	0.575	1.045	**1.234**	1.035	0.103	0.126	-2.975	-2.335	-2.043	**0.518**
Gross OOS SR	0.759	0.798	1.973	1.974	1.968	1.840	0.723	1.173	1.363	1.367	1.370	1.357
Net OOS SR	0.750	0.523	1.683	1.702	1.719	1.531	0.717	1.031	1.218	1.223	1.227	1.202
Turnover	0.055	1.520	40.39	34.19	30.40	15.25	0.042	0.769	25.33	23.58	22.71	12.93
Average $\hat{\kappa}$	/	0	0.250	0.205	0.175	0.065	/	0	0.511	0.471	0.446	0.217

Notes. This table reports the out-of-sample performance of the portfolio strategies introduced in Section 5.2.1 for the six datasets discussed in Section 5.1. 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), the annualized out-of-sample Sharpe ratio (OOS SR), and the monthly out-of-sample turnover of the portfolio strategies. For the two return-performance metrics (i.e., OOS CER and OOS SR), 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. The numbers in bold font identify the best portfolio in terms of OOS CER.

Table IA.2: Out-of-sample performance in real data with higher risk aversion

	$T = 120$						$T = 240$					
	RTR	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$			RTR	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$		
				$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$				$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
10MOM												
Gross OOS CER	0.057	0.075	0.108	0.112	0.110	**0.117**	0.072	0.090	0.129	0.132	0.133	**0.139**
Net OOS CER	0.056	0.068	0.088	0.094	0.093	**0.097**	0.071	0.085	0.115	0.119	0.121	**0.126**
Gross OOS SR	0.755	0.876	1.059	1.063	1.049	1.081	0.857	0.993	1.139	1.149	1.152	1.187
Net OOS SR	0.748	0.829	0.978	0.986	0.975	0.986	0.851	0.961	1.085	1.096	1.100	1.127
Turnover	0.047	0.300	0.862	0.759	0.695	0.853	0.037	0.184	0.572	0.534	0.512	0.528
Average $\hat{\kappa}$	/	0	0.395	0.312	0.251	0.265	/	0	0.546	0.483	0.443	0.398
25SBTM												
Gross OOS CER	0.052	0.090	0.116	**0.119**	0.117	0.118	0.072	0.111	0.143	0.143	0.140	**0.144**
Net OOS CER	0.051	0.071	0.074	0.084	**0.087**	0.081	0.071	0.101	0.115	0.118	0.118	**0.121**
Gross OOS SR	0.744	0.994	1.079	1.103	1.110	1.127	0.850	1.195	1.202	1.219	1.221	1.299
Net OOS SR	0.738	0.857	0.880	0.917	0.936	0.904	0.845	1.110	1.071	1.093	1.099	1.158
Turnover	0.047	0.783	1.806	1.501	1.281	1.543	0.039	0.442	1.201	1.065	0.960	0.946
Average $\hat{\kappa}$	/	0	0.214	0.155	0.114	0.109	/	0	0.364	0.298	0.242	0.192
25OPINV												
Gross OOS CER	0.074	0.102	0.104	0.110	0.111	**0.112**	0.072	0.119	0.131	**0.135**	**0.135**	0.130
Net OOS CER	0.073	0.088	0.076	0.087	**0.092**	0.088	0.071	0.113	0.115	0.121	**0.124**	0.115
Gross OOS SR	0.866	1.073	1.021	1.070	1.100	1.099	0.859	1.225	1.168	1.220	1.259	1.214
Net OOS SR	0.859	0.976	0.878	0.938	0.978	0.950	0.853	1.176	1.081	1.139	1.184	1.125
Turnover	0.048	0.569	1.143	0.938	0.795	0.992	0.036	0.271	0.670	0.571	0.485	0.589
Average $\hat{\kappa}$	/	0	0.191	0.132	0.094	0.089	/	0	0.320	0.244	0.183	0.168
49IND												
Gross OOS CER	**0.073**	0.061	0.059	0.062	0.063	0.063	**0.062**	0.046	0.041	0.045	0.046	0.047
Net OOS CER	**0.072**	0.042	0.036	0.041	0.042	0.043	**0.061**	0.037	0.029	0.035	0.037	0.038
Gross OOS SR	0.864	0.788	0.768	0.788	0.796	0.801	0.793	0.678	0.642	0.671	0.683	0.690
Net OOS SR	0.856	0.647	0.615	0.642	0.653	0.654	0.787	0.607	0.558	0.592	0.608	0.613
Turnover	0.051	0.822	0.981	0.894	0.859	0.868	0.040	0.361	0.483	0.430	0.399	0.390
Average $\hat{\kappa}$	/	0	0.043	0.029	0.022	0.012	/	0	0.087	0.061	0.047	0.018
16LTANOM												
Gross OOS CER	0.056	0.091	0.114	0.117	0.115	**0.119**	0.053	0.112	0.164	0.167	**0.169**	0.156
Net OOS CER	0.055	0.079	0.082	0.090	0.090	**0.093**	0.052	0.105	0.142	0.147	**0.150**	0.136
Gross OOS SR	0.759	1.003	1.068	1.092	1.094	1.146	0.732	1.191	1.290	1.321	1.347	1.347
Net OOS SR	0.752	0.916	0.914	0.947	0.955	0.987	0.726	1.139	1.192	1.225	1.252	1.231
Turnover	0.055	0.502	1.369	1.159	1.025	1.058	0.040	0.279	0.915	0.837	0.786	0.806
Average $\hat{\kappa}$	/	0	0.309	0.232	0.179	0.134	/	0	0.522	0.455	0.405	0.319
46ANOM												
Gross OOS CER	0.057	0.062	0.226	0.346	**0.383**	0.320	0.052	0.110	-0.594	-0.436	-0.366	**0.227**
Net OOS CER	0.056	0.025	0.104	0.229	**0.272**	0.239	0.051	0.092	-0.652	-0.497	-0.427	**0.170**
Gross OOS SR	0.759	0.798	2.074	2.086	2.089	1.913	0.723	1.173	1.488	1.499	1.506	1.575
Net OOS SR	0.750	0.523	1.825	1.842	1.852	1.593	0.717	1.031	1.355	1.367	1.375	1.424
Turnover	0.055	1.520	8.265	7.096	6.535	3.514	0.042	0.769	5.141	4.799	4.632	2.743
Average $\hat{\kappa}$	/	0	0.250	0.208	0.185	0.063	/	0	0.511	0.472	0.448	0.214

Notes. This table reports the out-of-sample performance of the portfolio strategies introduced in Section 5.2.1 for the six datasets discussed in Section 5.1. 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), the annualized out-of-sample Sharpe ratio (OOS SR), and the monthly out-of-sample turnover of the portfolio strategies. For the two return-performance metrics (i.e., OOS CER and OOS SR), 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. The numbers in bold font identify the best portfolio in terms of OOS CER.

Table IA.3: Out-of-sample performance in real data with higher transaction costs

	$T = 120$						$T = 240$					
	RTR	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$			RTR	SGMV	$\hat{\kappa}_E^\star$	$\hat{\kappa}_R^\star$		
				$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$				$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
10MOM												
Gross OOS CER	0.082	0.097	0.151	0.157	0.147	**0.167**	0.094	0.109	0.171	0.177	0.179	**0.192**
Net OOS CER	0.080	0.087	0.103	0.116	0.111	**0.120**	0.092	0.102	0.139	0.148	0.151	**0.163**
Gross OOS SR	0.755	0.876	0.963	0.970	0.939	1.009	0.857	0.993	1.019	1.031	1.035	1.096
Net OOS SR	0.745	0.804	0.833	0.847	0.822	0.847	0.848	0.945	0.932	0.947	0.953	0.995
Turnover	0.047	0.300	1.360	1.158	1.022	1.331	0.037	0.184	0.915	0.838	0.790	0.831
25SBTM												
Gross OOS CER	0.086	0.109	0.154	0.157	0.149	**0.158**	0.098	0.127	0.179	0.178	0.170	**0.181**
Net OOS CER	0.084	0.080	0.057	0.081	**0.092**	0.077	0.096	0.111	0.113	0.122	0.120	**0.130**
Gross OOS SR	0.744	0.994	0.962	1.002	1.022	1.084	0.850	1.195	1.038	1.055	1.050	1.183
Net OOS SR	0.735	0.789	0.638	0.698	0.753	0.685	0.842	1.068	0.830	0.854	0.857	0.940
Turnover	0.047	0.783	2.776	2.163	1.619	2.236	0.039	0.442	1.881	1.618	1.395	1.413
25OPINV												
Gross OOS CER	0.097	0.122	0.125	0.135	0.135	**0.138**	0.093	0.137	0.159	0.164	**0.163**	0.154
Net OOS CER	0.096	0.101	0.063	0.088	**0.102**	0.087	0.092	0.128	0.121	0.133	**0.140**	0.123
Gross OOS SR	0.866	1.073	0.872	0.958	1.054	1.025	0.859	1.225	1.009	1.099	1.202	1.095
Net OOS SR	0.855	0.928	0.631	0.734	0.856	0.751	0.850	1.152	0.858	0.953	1.070	0.930
Turnover	0.048	0.569	1.730	1.293	0.920	1.395	0.036	0.271	1.048	0.848	0.637	0.871
49IND												
Gross OOS CER	**0.096**	0.081	0.077	0.082	0.083	0.084	**0.082**	0.061	0.052	0.061	0.062	0.063
Net OOS CER	**0.094**	0.051	0.034	0.046	0.052	0.050	**0.081**	0.048	0.029	0.042	0.047	0.047
Gross OOS SR	0.864	0.788	0.704	0.763	0.793	0.801	0.793	0.678	0.563	0.642	0.678	0.694
Net OOS SR	0.852	0.576	0.456	0.533	0.576	0.568	0.783	0.571	0.425	0.513	0.560	0.569
Turnover	0.051	0.822	1.205	0.977	0.878	0.928	0.040	0.361	0.634	0.503	0.421	0.429
16LTANOM												
Gross OOS CER	0.087	0.110	0.148	0.152	0.143	**0.157**	0.080	0.129	0.211	0.217	**0.221**	0.203
Net OOS CER	0.085	0.092	0.071	0.090	0.091	**0.099**	0.078	0.119	0.157	0.168	**0.176**	0.156
Gross OOS SR	0.759	1.003	0.943	0.981	0.983	1.119	0.732	1.191	1.131	1.167	1.207	1.261
Net OOS SR	0.748	0.872	0.684	0.733	0.746	0.813	0.724	1.112	0.970	1.007	1.047	1.047
Turnover	0.055	0.502	2.167	1.751	1.450	1.605	0.040	0.279	1.504	1.363	1.254	1.277
46ANOM												
Gross OOS CER	0.084	0.080	0.339	0.541	**0.602**	0.513	0.078	0.127	-1.044	-0.781	-0.665	**0.327**
Net OOS CER	0.082	0.025	0.058	0.263	**0.342**	0.325	0.076	0.099	-1.136	-0.889	-0.777	**0.191**
Gross OOS SR	0.759	0.798	2.030	2.038	2.039	1.927	0.723	1.173	1.428	1.436	1.441	1.482
Net OOS SR	0.746	0.385	1.633	1.656	1.673	1.446	0.714	0.959	1.222	1.232	1.238	1.253
Turnover	0.055	1.520	13.57	11.55	10.50	5.377	0.042	0.769	8.488	7.912	7.626	4.401

Notes. This table reports the out-of-sample performance of the portfolio strategies introduced in Section 5.2.1 for the six datasets discussed in Section 5.1. 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), the annualized out-of-sample Sharpe ratio (OOS SR), and the monthly out-of-sample turnover of the portfolio strategies. For the two return-performance metrics (i.e., OOS CER and OOS SR), 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. The numbers in bold font identify the best portfolio in terms of OOS CER.

Table IA.4: Out-of-sample performance in a dataset of 50 individual stocks

	Benchmark strategies					*Proposed strategies* ($\hat{\kappa}_R^\star$)		
	EW	RTR	SGMV	SMV	$\hat{\kappa}_E^\star$	$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
Gross OOS CER	0.056	0.054	**0.074**	-1.420	0.054	0.065	0.069	0.063
Net OOS CER	0.055	0.052	**0.064**	-1.480	0.044	0.055	0.059	0.053
Gross OOS SR	0.596	0.592	0.721	-0.077	0.580	0.657	0.686	0.642
Net OOS SR	0.587	0.578	0.656	-0.156	0.515	0.593	0.622	0.577
Turnover	0.057	0.084	0.407	3.049	0.436	0.405	0.399	0.415
Average $\hat{\kappa}$	/	/	0	1	0.064	0.038	0.026	0.029

Notes. This table reports the out-of-sample performance of the portfolio strategies introduced in Section 5.2.1 for the dataset of 50 individual stocks discussed in Appendix IA.2.3. 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), the annualized out-of-sample Sharpe ratio (OOS SR), and the monthly out-of-sample turnover of the portfolio strategies. For the two return-performance metrics (i.e., OOS CER and OOS SR), 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.

Table IA.5: Out-of-sample performance with shrinkage estimator of the covariance matrix

	$T = 120$						$T = 240$					
	RTR	SGMV	$\hat{\kappa}^{\star}_{E}$	$\hat{\kappa}^{\star}_{R}$			RTR	SGMV	$\hat{\kappa}^{\star}_{E}$	$\hat{\kappa}^{\star}_{R}$		
				$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$				$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
10MOM												
Gross OOS CER	0.096	0.022	0.144	0.146	0.113	**0.174**	0.109	0.104	0.168	0.170	0.173	**0.195**
Net OOS CER	0.090	-0.030	0.113	0.117	0.085	**0.143**	0.105	0.074	0.146	0.150	0.154	**0.175**
Gross OOS SR	0.874	0.917	0.966	0.952	0.848	1.027	0.994	0.971	1.019	1.020	1.023	1.096
Net OOS SR	0.831	0.823	0.886	0.876	0.767	0.926	0.964	0.910	0.964	0.966	0.970	1.033
Turnover	0.265	2.308	1.386	1.240	1.220	1.323	0.174	1.324	0.933	0.875	0.834	0.819
Average $\hat{\kappa}$	/	1	0.475	0.331	0.185	0.313	0	1	0.597	0.494	0.441	0.412
25SBTM												
Gross OOS CER	0.108	-0.237	0.156	0.161	0.135	**0.166**	0.126	0.016	0.168	0.168	0.153	**0.183**
Net OOS CER	0.095	-0.373	0.091	0.107	0.095	**0.110**	0.117	-0.062	0.124	0.129	0.119	**0.149**
Gross OOS SR	1.013	0.829	0.974	0.984	0.914	1.050	1.191	0.933	1.008	1.005	0.962	1.128
Net OOS SR	0.915	0.620	0.791	0.809	0.753	0.821	1.119	0.793	0.884	0.885	0.844	0.986
Turnover	0.54	6.561	2.857	2.366	1.791	2.352	0.373	3.569	1.916	1.698	1.503	1.425
Average $\hat{\kappa}$	/	1	0.346	0.230	0.099	0.185	0	1	0.456	0.362	0.243	0.243
25OPINV												
Gross OOS CER	0.125	-0.306	0.130	0.135	0.124	**0.140**	0.139	-0.099	0.159	**0.168**	0.156	0.155
Net OOS CER	**0.116**	-0.405	0.088	0.102	0.101	0.106	0.134	-0.149	0.134	**0.146**	0.140	0.135
Gross OOS SR	1.129	0.684	0.882	0.927	0.980	0.995	1.253	0.724	0.990	1.078	1.182	1.079
Net OOS SR	1.063	0.533	0.740	0.788	0.848	0.827	1.212	0.633	0.899	0.985	1.091	0.979
Turnover	0.369	4.572	1.765	1.367	0.944	1.412	0.224	2.259	1.054	0.881	0.650	0.820
Average $\hat{\kappa}$	/	1	0.310	0.183	0.046	0.141	0	1	0.401	0.290	0.103	0.191
49IND												
Gross OOS CER	**0.099**	-1.399	0.076	0.089	0.090	**0.099**	0.075	-0.748	0.045	0.065	**0.077**	0.073
Net OOS CER	**0.090**	-1.505	0.051	0.072	0.079	0.080	0.069	-0.806	0.031	0.055	**0.070**	0.065
Gross OOS SR	1.014	0.312	0.676	0.813	0.918	0.898	0.827	0.141	0.522	0.678	0.839	0.799
Net OOS SR	0.938	0.177	0.556	0.703	0.824	0.768	0.774	0.050	0.444	0.603	0.778	0.731
Turnover	0.378	6.01	1.053	0.697	0.484	0.793	0.253	2.864	0.592	0.439	0.292	0.326
Average $\hat{\kappa}$	/	1	0.125	0.043	0.005	0.032	0	1	0.139	0.058	0.005	0.007
16LTANOM												
Gross OOS CER	0.114	-0.105	0.146	0.149	0.130	**0.161**	0.133	0.095	0.212	0.218	**0.227**	0.210
Net OOS CER	0.104	-0.201	0.094	0.105	0.096	**0.122**	0.127	0.041	0.175	0.184	**0.195**	0.180
Gross OOS SR	1.043	0.770	0.937	0.951	0.899	1.110	1.228	0.993	1.128	1.154	1.198	1.280
Net OOS SR	0.971	0.602	0.778	0.795	0.760	0.916	1.180	0.892	1.027	1.051	1.095	1.145
Turnover	0.412	4.330	2.209	1.875	1.457	1.625	0.254	2.260	1.522	1.427	1.325	1.243
Average $\hat{\kappa}$	/	1	0.417	0.284	0.166	0.177	0	1	0.603	0.527	0.440	0.343
46ANOM												
Gross OOS CER	0.084	-2.940	-1.062	-0.984	-0.936	**0.862**	0.122	-3.783	-2.379	-2.322	-2.282	**0.143**
Net OOS CER	0.066	-2.760	-0.988	-0.905	-0.846	**0.731**	**0.109**	-3.649	-2.324	-2.268	-2.228	0.072
Gross OOS SR	0.901	2.291	2.315	2.313	2.307	2.290	1.204	1.496	1.529	1.531	1.532	1.552
Net OOS SR	0.742	2.113	2.170	2.172	2.172	2.122	1.092	1.386	1.424	1.427	1.429	1.446
Turnover	0.765	18.22	13.68	13.31	12.92	7.027	0.549	10.42	8.760	8.663	8.581	4.606
Average $\hat{\kappa}$	/	1	0.673	0.640	0.598	0.281	0	1	0.788	0.771	0.755	0.348

Notes. This table reports the out-of-sample performance of the portfolio strategies introduced in Section 5.2.1 and Appendix IA.2.4 for the six datasets discussed in Section 5.1. Each estimated portfolio is constructed using a sample size of $T = 120$ and 240 monthly observations. The covariance matrix is estimated using the nonlinear shrinkage estimator of Ledoit and Wolf (2020a). 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), the annualized out-of-sample Sharpe ratio (OOS SR), and the monthly out-of-sample turnover of the portfolio strategies. For the two return-performance metrics (i.e., OOS CER and OOS SR), 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. The numbers in bold font identify the best portfolio in terms of OOS CER.

Table IA.6: Out-of-sample performance exploiting equally weighted portfolio with $T = 120$

		Benchmark strategies				*Proposed strategies* $(\hat{\delta}^{\star}_R, \hat{\kappa}^{\star}_R)$		
		EW	SGMV	SMV	$(\hat{\delta}^{\star}_E, \hat{\kappa}^{\star}_E)$	$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
10MOM	Gross OOS CER	0.069	0.097	-0.048	0.144	0.147	0.119	**0.155**
	Net OOS CER	0.068	0.090	-0.108	0.113	0.120	0.096	**0.124**
	Gross OOS SR	0.659	0.876	0.868	0.944	0.939	0.848	0.971
	Net OOS SR	0.653	0.829	0.765	0.858	0.859	0.772	0.864
	Turnover	0.040	0.300	2.704	1.357	1.144	0.978	1.323
	Average $\hat{\delta}$	0	/	/	0.708	0.670	0.560	0.609
	Average $\hat{\kappa}$	/	0	1	0.625	0.474	0.250	0.489
25SBTM	Gross OOS CER	0.080	0.109	-0.708	0.157	0.157	0.139	**0.159**
	Net OOS CER	0.079	0.090	-0.890	0.093	0.103	0.103	**0.109**
	Gross OOS SR	0.705	0.994	0.711	0.971	0.995	0.964	1.078
	Net OOS SR	0.700	0.857	0.457	0.763	0.785	0.802	0.839
	Turnover	0.045	0.783	10.06	2.717	2.257	1.479	2.090
	Average $\hat{\delta}$	0	/	/	0.648	0.628	0.538	0.543
	Average $\hat{\kappa}$	/	0	1	0.428	0.316	0.103	0.261
25OPINV	Gross OOS CER	0.089	0.122	-0.814	0.119	0.126	0.111	**0.127**
	Net OOS CER	0.088	**0.108**	-0.952	0.077	0.095	0.091	0.093
	Gross OOS SR	0.800	1.073	0.554	0.850	0.928	0.957	0.980
	Net OOS SR	0.794	0.976	0.376	0.684	0.773	0.827	0.789
	Turnover	0.040	0.569	6.982	1.749	1.286	0.826	1.423
	Average $\hat{\delta}$	0	/	/	0.821	0.787	0.619	0.697
	Average $\hat{\kappa}$	/	0	1	0.270	0.169	0.069	0.137
49IND	Gross OOS CER	0.092	0.081	-4.667	0.081	0.088	0.092	**0.097**
	Net OOS CER	**0.091**	0.061	-4.745	0.056	0.072	0.083	0.083
	Gross OOS SR	0.816	0.788	0.244	0.726	0.817	0.855	0.875
	Net OOS SR	0.809	0.647	0.042	0.581	0.710	0.793	0.782
	Turnover	0.049	0.822	15.34	1.067	0.663	0.370	0.590
	Average $\hat{\delta}$	0	/	/	0.431	0.365	0.262	0.239
	Average $\hat{\kappa}$	/	0	1	0.231	0.132	0.012	0.030
16LTANOM	Gross OOS CER	0.078	0.110	-0.274	0.144	0.148	0.130	**0.157**
	Net OOS CER	0.077	0.098	-0.394	0.093	0.107	0.094	**0.118**
	Gross OOS SR	0.703	1.003	0.725	0.931	0.969	0.933	1.112
	Net OOS SR	0.697	0.916	0.538	0.758	0.802	0.764	0.908
	Turnover	0.044	0.502	5.657	2.186	1.761	1.510	1.636
	Average $\hat{\delta}$	0	/	/	0.784	0.785	0.717	0.707
	Average $\hat{\kappa}$	/	0	1	0.441	0.300	0.199	0.188
46ANOM	Gross OOS CER	0.056	0.080	-14.74	0.332	0.535	**0.596**	0.492
	Net OOS CER	0.055	0.043	-11.46	0.156	0.358	**0.429**	0.365
	Gross OOS SR	0.581	0.798	1.953	2.028	2.037	2.036	1.851
	Net OOS SR	0.575	0.523	1.612	1.776	1.793	1.803	1.538
	Turnover	0.044	1.520	48.03	13.60	11.59	10.53	5.531
	Average $\hat{\delta}$	0	/	/	0.722	0.709	0.698	0.553
	Average $\hat{\kappa}$	/	0	1	0.499	0.459	0.434	0.260

Notes. This table reports the out-of-sample performance of the portfolio strategies introduced in Section 5.2.1 and Appendix IA.2.5 for the six datasets discussed in Section 5.1. Each estimated portfolio is constructed using a sample size of $T = 120$ monthly observations. The covariance matrix is estimated using the nonlinear shrinkage estimator of Ledoit and Wolf (2020a). 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), the annualized out-of-sample Sharpe ratio (OOS SR), and the monthly out-of-sample turnover of the portfolio strategies. For the two return-performance metrics (i.e., OOS CER and OOS SR), 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 $\hat{\delta}$ and $\hat{\kappa}$ 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.

Table IA.7: Out-of-sample performance exploiting equally weighted portfolio with $T = 240$

		Benchmark strategies				*Proposed strategies* $(\hat{\delta}^{\star}_R, \hat{\kappa}^{\star}_R)$		
		EW	SGMV	SMV	$(\hat{\delta}^{\star}_E, \hat{\kappa}^{\star}_E)$	$\lambda = 2$	$\lambda = 4$	$\lambda = \hat{\lambda}_{cv}$
10MOM	Gross OOS CER	0.079	0.109	0.086	0.170	0.175	0.175	**0.189**
	Net OOS CER	0.078	0.105	0.054	0.148	0.155	0.156	**0.169**
	Gross OOS SR	0.739	0.993	0.946	1.013	1.024	1.023	1.084
	Net OOS SR	0.733	0.961	0.881	0.955	0.968	0.969	1.017
	Turnover	0.039	0.184	1.416	0.906	0.836	0.787	0.837
	Average $\hat{\delta}$	0	/	/	0.781	0.781	0.766	0.731
	Average $\hat{\kappa}$	/	0	1	0.726	0.629	0.545	0.597
25SBTM	Gross OOS CER	0.092	0.127	-0.062	**0.182**	0.180	0.166	**0.182**
	Net OOS CER	0.091	0.117	-0.154	0.139	0.144	0.136	**0.150**
	Gross OOS SR	0.803	1.195	0.894	1.047	1.060	1.038	1.184
	Net OOS SR	0.797	1.110	0.738	0.913	0.933	0.920	1.031
	Turnover	0.041	0.442	4.337	1.836	1.558	1.297	1.352
	Average $\hat{\delta}$	0	/	/	0.662	0.657	0.628	0.619
	Average $\hat{\kappa}$	/	0	1	0.601	0.472	0.326	0.371
25OPINV	Gross OOS CER	0.085	0.137	-0.184	0.153	**0.158**	0.156	0.147
	Net OOS CER	0.084	0.131	-0.246	0.128	0.138	**0.141**	0.125
	Gross OOS SR	0.786	1.225	0.679	0.988	1.074	1.170	1.056
	Net OOS SR	0.780	1.176	0.576	0.886	0.975	1.085	0.943
	Turnover	0.039	0.271	2.772	1.059	0.853	0.616	0.898
	Average $\hat{\delta}$	0	/	/	0.864	0.863	0.809	0.806
	Average $\hat{\kappa}$	/	0	1	0.383	0.267	0.153	0.210
49IND	Gross OOS CER	**0.080**	0.061	-1.184	0.064	0.075	0.078	**0.080**
	Net OOS CER	**0.079**	0.052	-1.262	0.051	0.066	0.073	0.071
	Gross OOS SR	0.752	0.678	0.162	0.634	0.745	0.798	0.811
	Net OOS SR	0.745	0.607	0.050	0.556	0.683	0.760	0.747
	Turnover	0.048	0.361	4.332	0.549	0.362	0.205	0.349
	Average $\hat{\delta}$	0	/	/	0.415	0.390	0.335	0.321
	Average $\hat{\kappa}$	/	0	1	0.248	0.131	0.052	0.033
16LTANOM	Gross OOS CER	0.068	0.129	0.047	0.204	0.209	**0.213**	0.191
	Net OOS CER	0.067	0.122	-0.014	0.168	0.176	**0.182**	0.160
	Gross OOS SR	0.655	1.191	0.960	1.112	1.144	1.181	1.215
	Net OOS SR	0.650	1.139	0.854	1.003	1.035	1.072	1.071
	Turnover	0.042	0.279	2.578	1.508	1.384	1.270	1.273
	Average $\hat{\delta}$	0	/	/	0.837	0.858	0.869	0.808
	Average $\hat{\kappa}$	/	0	1	0.673	0.574	0.492	0.447
46ANOM	Gross OOS CER	0.051	0.127	-6.346	-0.977	-0.723	-0.609	**0.340**
	Net OOS CER	0.050	0.109	-5.933	-1.037	-0.795	-0.683	**0.249**
	Gross OOS SR	0.551	1.173	1.364	1.434	1.442	1.447	1.491
	Net OOS SR	0.545	1.031	1.220	1.297	1.306	1.312	1.336
	Turnover	0.046	0.769	15.52	8.478	7.921	7.625	4.446
	Average $\hat{\delta}$	0	/	/	0.793	0.787	0.788	0.534
	Average $\hat{\kappa}$	/	0	1	0.704	0.669	0.640	0.587

Notes. This table reports the out-of-sample performance of the portfolio strategies introduced in Section 5.2.1 and Appendix IA.2.5 for the six datasets discussed in Section 5.1. Each estimated portfolio is constructed using a sample size of $T = 240$ monthly observations. The covariance matrix is estimated using the nonlinear shrinkage estimator of Ledoit and Wolf (2020a). 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), the annualized out-of-sample Sharpe ratio (OOS SR), and the monthly out-of-sample turnover of the portfolio strategies. For the two return-performance metrics (i.e., OOS CER and OOS SR), 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 $\hat{\delta}$ and $\hat{\kappa}$ 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.

IA.3 Proofs of all 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

$$\hat{w}^\star(\kappa) = \hat{w}_g + \frac{\kappa}{\gamma}\hat{w}_z. \tag{IA15}$$

Proof of Proposition 1

The proof of this proposition is in Kan, Wang, and Zhou (2021).

Proof of Proposition 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 (19) 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, Uppal, and Wang (2007) show that a higher ε implies a higher exposure to the SGMV portfolio and, thus, a smaller portfolio-return volatility. Consequently, the ratio $\varepsilon/\sigma_P^\star$ is monotonically increasing in ε.

Proof of Lemma 1

Equation (21) is directly obtained from the definition of OOSU in (12) and the formula for the variance of a sum of two correlated random variables.

Proof of Proposition 3

Kan, Wang, and Zhou (2021, Proposition 1) derive a stochastic representation for the out-of-sample mean return and variance of the shrinkage portfolio $\hat{w}(\kappa)$ that combines the SMV and SGMV portfolios, i.e., for the two random variables $\hat{w}^\star(\kappa)^\top\mu$ and $\hat{w}^\star(\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\kappa)$. Using this result, we can find analytical expressions for the three terms composing the OOSU variance in (21). Specifically, the variance of the out-of-sample mean return is

$$\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top\mu\Big] = \frac{\sigma_g^2\psi^2}{T-N-1} + \frac{\kappa^2\psi^2}{\gamma^2}\frac{2T(N+1)+T^2(T-N-3+2(T-N)\psi^2)}{(T-N)(T-N-1)^2(T-N-3)}, \tag{IA16}$$

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

$$\begin{aligned}\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\kappa)\Big] &= \frac{2\sigma_g^4(N-1)(T-2)}{(T-N-1)^2(T-N-3)}\\ &+ \frac{4\kappa^2\sigma_g^2}{\gamma^2}\frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}\\ &+ \frac{2\kappa^4}{\gamma^4}\frac{T^2(T-2)C(T,N,\psi^2)}{(T-N)^2(T-N-1)^2(T-N-2)(T-N-3)^2(T-N-5)(T-N-7)},\end{aligned} \tag{IA17}$$

where $C(T,N,\psi^2)$ is defined in Proposition 3, and the covariance between the out-of-sample mean return and variance is

$$\begin{aligned}\mathbb{C}\text{ov}\Big[\hat{w}^\star(\kappa)^\top\mu, \hat{w}^\star(\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\kappa)\Big] &= \frac{2\kappa\sigma_g^2\psi^2}{\gamma}\frac{T(T-2)}{(T-N-1)^2(T-N-3)}\\ &+ \frac{2\kappa^3\psi^2}{\gamma^3}\frac{T^2(T-2)(T+N-3+2T\psi^2)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}.\end{aligned} \tag{IA18}$$

We find the final expression for the OOSU variance in Proposition 3 by plugging (IA16)–(IA18) into (21).

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 (22) with respect to κ, evaluated at $\kappa = 0$, is

$$\left.\frac{\partial\mathbb{V}[U(\hat{w}^\star(\kappa))]}{\partial\kappa}\right|_{\kappa=0} = a_4. \tag{IA19}$$

Now, observe that a_4 in Equation (28) 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 5 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 4

Parameters T and N. From the closed-form expression of $\mathbb{V}[U(\hat{w}^\star(\kappa))]$ in Proposition 3, 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

$$\begin{aligned}\frac{\partial\mathbb{V}[U(\hat{w}^\star(\kappa))]}{\partial\sigma_g^2} &= \frac{\psi^2}{T-N-1} + \sigma_g^2\gamma^2\frac{(N-1)(T-2)}{(T-N-1)^2(T-N-3)} \\ &+ \kappa^2\frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} - 2\kappa\psi^2\frac{T(T-2)}{(T-N-1)^2(T-N-3)}.\end{aligned} \tag{IA20}$$

The objective is to show that the derivative in (IA20) 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

$$\begin{aligned}&\frac{\psi^2}{T-N-1} + \kappa^2\frac{T(T-2)(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} \\ &- 2\kappa\psi^2\frac{T(T-2)}{(T-N-1)^2(T-N-3)} \geq 0.\end{aligned} \tag{IA21}$$

Notice that the left-hand side of (IA21) is a second-degree polynomial in κ. Because the coefficient in front of κ^2 is positive, we can prove that inequality (IA21) holds by showing that the polynomial discriminant is always negative. That is, after some simplifications,

$$\psi^2 \frac{T(T-2)}{(T-N-1)(T-N-3)} \leq \frac{(T+N-3)(T\psi^2+N-1)}{(T-N)(T-N-5)}. \tag{IA22}$$

Notice that the right-hand side of inequality (IA22) is of the form $a + b\psi^2$ with $a > 0$. Therefore, we can prove the inequality by showing that the coefficient in front of 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

$$\frac{T(T+N-3)}{(T-N)(T-N-5)} \geq \frac{T(T-2)}{(T-N-1)(T-N-3)}, \tag{IA23}$$

which holds under Assumption 1.

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

$$\begin{aligned}
&\frac{\partial \mathbb{V}[U(\hat{w}^\star(\kappa))]}{\partial \psi^2} = \frac{\sigma_g^2}{T-N-1} \\
&+ \frac{\kappa^4}{2\gamma^2} \frac{T^2(T-2)\frac{\partial C(T,N,\psi^2)}{\partial \psi^2}}{(T-N)^2(T-N-1)^2(T-N-2)(T-N-3)^2(T-N-5)(T-N-7)} \\
&- \frac{2\kappa^3}{\gamma^2} \frac{4\psi^2 T^3(T-2) + T^2(T-2)(T+N-3)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} \\
&+ \frac{\kappa^2}{\gamma^2} \frac{2T(N+1) + T^2(T-N-3) + 4T^2(T-N)\psi^2}{(T-N)(T-N-1)^2(T-N-3)} \\
&+ \kappa^2\sigma_g^2 \frac{T^2(T-2)(T+N-3)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} \\
&- 2\kappa\sigma_g^2 \frac{T(T-2)}{(T-N-1)^2(T-N-3)}.
\end{aligned} \tag{IA24}$$

First, we show that the derivative (IA24) is increasing in σ_g^2 and thus that it suffices to show that it is positive for $\sigma_g^2 = 0$. We have

$$\begin{aligned}
\frac{\partial}{\partial \sigma_g^2}\left(\frac{\partial \mathbb{V}[U(\hat{w}^\star(\kappa))]}{\partial \psi^2}\right) = {}& \frac{1}{T-N-1} + \kappa^2 \frac{T^2(T-2)(T+N-3)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)} \\
&- 2\kappa \frac{T(T-2)}{(T-N-1)^2(T-N-3)}.
\end{aligned} \tag{IA25}$$

Following a similar strategy to the case with σ_g^2 as a parameter, the derivative (IA25) 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 (IA24) is positive for $\sigma_g^2 = 0$. That is, after some simplifications,

$$\begin{aligned} &\frac{\kappa^2}{2} \frac{T^2(T-2)\frac{\partial C(T,N,\psi^2)}{\partial \psi^2}}{(T-N)(T-N-2)(T-N-3)(T-N-7)} \\ &- 2\kappa\Big(4T^3(T-2)\psi^2 + T^2(T-2)(T+N-3)\Big) \\ &+ (T-N-5)\Big(2T(N+1) + T^2(T-N-3) + 4T^2(T-N)\psi^2\Big) \geq 0. \end{aligned} \tag{IA26}$$

We find that the derivative of (IA26) with respect to ψ^2 is positive if

$$\frac{\partial^2 C(T,N,\psi^2)}{\partial (\psi^2)^2} \geq \frac{8T^2(T-2)(T-N-2)(T-N-3)(T-N-7)}{(T-N-5)}, \tag{IA27}$$

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 (IA27) holds under Assumption 1. Therefore, we can now prove the result of the proposition by showing that the derivative in (IA26) is positive for $\psi^2 = 0$. That is,

$$\begin{aligned} &\frac{\kappa^2}{2} \frac{T^2(T-2)\frac{\partial C(T,N,\psi^2)}{\partial \psi^2}\Big|_{\psi^2=0}}{(T-N)(T-N-2)(T-N-3)(T-N-7)} - 2\kappa T^2(T-2)(T+N-3) \\ &+ (T-N-5)(2T(N+1) + T^2(T-N-3)) \geq 0. \end{aligned} \tag{IA28}$$

As usual, we prove inequality (IA28) by showing that the polynomial discriminant is negative, which amounts to showing, after some simplifications, that

$$\frac{\partial C(T,N,\psi^2)}{\partial \psi^2}\Big|_{\psi^2=0} \geq \frac{2T(T-2)(T-N)(T-N-2)(T-N-3)(T+N-3)^2}{(T-N-5)(2(N+1) + T(T-N-3))}, \tag{IA29}$$

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,

$$4a_1\kappa^3 + 3a_2\kappa^2 + 2a_3\kappa + a_4 \geq 0 \tag{IA30}$$

if $\kappa \geq \kappa_E^\star$. As we show below, the derivative in (IA30) decreases with γ when $\kappa \geq \kappa_E^\star$. Therefore, we can derive a sufficient condition on the value of κ for which inequality (IA30) holds by considering the case $\gamma \to \infty$. In that case, the condition in (IA30) 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

$$\kappa \geq \kappa_E^\star \frac{(T-2)(T-N-5)}{(T+N-3)(T-N-3)}, \tag{IA31}$$

which is satisfied when $\kappa \geq \kappa_E^\star$ because the right-hand side of (IA31) is smaller than $\kappa_E^\star$ under Assumption 1.

The only step missing now is to show that the left-hand side of (IA30) is decreasing in γ when $\kappa \geq \kappa_E^\star$. To prove this result, it is useful to introduce the notation

$$\begin{aligned}
\overline{a}_1 &= a_1\gamma^2, \\
\overline{a}_2 &= a_2\gamma^2, \\
\overline{a}_{3,1} &= \psi^2 \frac{2T(N+1) + T^2(T-N-3+2(T-N)\psi^2)}{(T-N)(T-N-1)^2(T-N-3)},
\end{aligned}$$

which are all independent of γ. Now, it is straightforward to show that the left-hand side of (IA30) is decreasing in γ when $\kappa \geq \kappa_E^\star$ if

$$4\overline{a}_1\kappa^2 + 3\overline{a}_2\kappa + 2\overline{a}_{3,1} \geq 0 \tag{IA32}$$

when $\kappa \geq \kappa_E^\star$. Since $\overline{a}_1 \geq 0$, inequality (IA32) 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 (IA32) is smaller than $\kappa_E^\star$. That is,

$$\frac{-3\overline{a}_2 + \sqrt{9\overline{a}_2^2 - 32\overline{a}_1\overline{a}_{3,1}}}{8\overline{a}_1} \leq \kappa_E^\star. \tag{IA33}$$

After some simplifications, proving inequality (IA33) is equivalent to showing that

$$4\overline{a}_1(\kappa_E^\star)^2 + 3\overline{a}_2\kappa_E^\star + 2\overline{a}_{3,1} \geq 0. \tag{IA34}$$

We can reformulate inequality (IA34) as

$$\begin{aligned}
&\frac{\psi^2 C(T, N, \psi^2)}{(T-2)(T-N-2)(T-N-5)(T-N-7)} \\
&- \frac{3\psi^2(T\psi^2 + N - 1)(T + N - 3 + 2T\psi^2)}{T-N-5} \\
&+ \frac{2T(N+1) + T^2(T-N-3+2(T-N)\psi^2)}{(T-N)(T-N-3)}\left(\psi^2 + \frac{N-1}{T}\right)^2 \geq 0.
\end{aligned} \tag{IA35}$$

Notice that inequality (IA35) holds when $\psi^2 = 0$. Therefore, we can prove (IA35) by showing that the derivative of the left-hand side with respect to ψ^2 is positive. That is,

$$\begin{aligned}
&\frac{1}{(T-2)(T-N-2)(T-N-5)(T-N-7)}\Big[(4T\psi^2 + N - 1)(N^4 + N^3T - 3N^3 \\
&- 4N^2T^2 + 22N^2T - 31N^2 + NT^3 - 7NT^2 + 13NT - 5N + T^4 - 12T^3 + 53T^2 - 100T \\
&+ 70) + 3T^2\psi^4(N^3 + 2N^2T - 6N^2 - 7NT^2 + 40NT - 53N + 4T^3 - 34T^2 + 88T - 70)\Big] \\
&- \frac{3T}{T-N-5}\left[(T+N-3)\left(2\psi^2 + \frac{N-1}{T}\right) + 2T\left(3\psi^4 + 2\psi^2\frac{N-1}{T}\right)\right] \\
&+ \frac{2T}{(T-N)(T-N-3)}\Bigg[(2(N+1) + T(T-N-3))\left(\psi^2 + \frac{N-1}{T}\right) + T(T-N) \\
&\left(3\psi^4 + 4\psi^2\frac{N-1}{T} + \left(\frac{N-1}{T}\right)^2\right)\Bigg] \geq 0.
\end{aligned} \tag{IA36}$$

One can check that inequality (IA36) holds when $\psi^2 = 0$ under Assumption 1. Therefore, we can prove (IA36) by showing as before that the derivative of the left-hand side with respect to ψ^2 is positive. That is,

$$\frac{2T}{(T-2)(T-N-2)(T-N-5)(T-N-7)}\Big[2(N^4+N^3T-3N^3-4N^2T^2+22N^2T \\ -31N^2+NT^3-7NT^2+13NT-5N+T^4-12T^3+53T^2-100T+70)+3T\psi^2(N^3 \\ +2N^2T-6N^2-7NT^2+40NT-53N+4T^3-34T^2+88T-70)\Big]-\frac{6T}{T-N-5} \\ \Big[T+N-3+2T\Big(3\psi^2+\frac{N-1}{T}\Big)\Big]+\frac{2T}{(T-N)(T-N-3)}\Big[2(N+1)+T(T-N-3) \\ +2T(T-N)\Big(3\psi^2+2\frac{N-1}{T}\Big)\Big]\geq 0. \tag{IA37}$$

Again, one can check that inequality (IA37) holds when $\psi^2=0$ under Assumption 1. Therefore, we prove as usual that the derivative of the left-hand side of (IA37) with respect to ψ^2 is positive. That is,

$$\frac{N^3+2N^2T-6N^2-7NT^2+40NT-53N+4T^3-34T^2+88T-70}{(T-2)(T-N-2)(T-N-5)(T-N-7)} \\ -\frac{6}{T-N-5}+\frac{2}{T-N-3}\geq 0. \tag{IA38}$$

This last inequality holds under Assumption 1, which concludes the demonstration of inequality (IA32) for $\kappa\geq\kappa_E^\star$.

Proof of Proposition 5

Part 1. The proof is direct because, as shown in Proposition 3, the OOSU variance $\mathbb{V}[U(\hat{w}^\star(\kappa))]\to 0$ as $T\to\infty$ and, thus, the shrinkage intensity $\kappa_R^\star$ corresponds to $\kappa_E^\star$ as $T\to\infty$, and as shown in Proposition 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 (14) 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 (29) 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 4 that OOSU variance increases with κ if $\kappa\geq\kappa_E^\star$.

Moreover, OOSU mean in (14) 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 6

The proof is direct because, on the one hand, the OOSU mean in (14) 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 4 that OOSU standard deviation decreases with T and κ if $\kappa \geq \kappa_E^\star$, increases with N and σ_g^2, and also is independent of μ_g.

Proof of Proposition IA.1

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

$$\mathbb{E}[\hat{w}^\star(\kappa)] = w_g + \frac{\kappa}{\gamma}\mathbf{B}\mu, \tag{IA39}$$

$$\mathbb{V}[\hat{w}^\star(\kappa)] = \frac{\kappa^2}{\gamma^2}\frac{\mathbf{B}}{T}. \tag{IA40}$$

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:

$$\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top \mu\Big] = \mathbb{E}[\hat{w}^\star(\kappa)]^\top \mu = \mu_g + \frac{\kappa}{\gamma}\psi^2. \tag{IA41}$$

$$\begin{aligned}\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma}\hat{w}^\star(\kappa)\Big] &= \mathrm{Tr}(\mathbf{\Sigma}\mathbb{V}[\hat{w}^\star(\kappa)]) + \mathbb{E}[\hat{w}^\star(\kappa)]^\top \mathbf{\Sigma}\mathbb{E}[\hat{w}^\star(\kappa)] \\ &= \sigma_g^2 + \frac{\kappa^2}{\gamma^2}\Big(\psi^2 + \frac{N-1}{T}\Big),\end{aligned} \tag{IA42}$$

$$\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \mu\Big] = \mu^\top \mathbb{V}[\hat{w}^\star(\kappa)]\mu = \frac{\kappa^2}{\gamma^2}\frac{\psi^2}{T}, \tag{IA43}$$

$$\begin{aligned}\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma}\hat{w}^\star(\kappa)\Big] &= 2\mathrm{Tr}((\mathbf{\Sigma}\mathbb{V}[\hat{w}^\star(\kappa)])^2) + 4\mathbb{E}[\hat{w}^\star(\kappa)]^\top \mathbf{\Sigma}\mathbb{V}[\hat{w}^\star(\kappa)]\mathbf{\Sigma}\mathbb{E}[\hat{w}^\star(\kappa)] \\ &= \frac{\kappa^4}{\gamma^4}\frac{2}{T}\Big(2\psi^2 + \frac{N-1}{T}\Big),\end{aligned} \tag{IA44}$$

$$\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma}\hat{w}^\star(\kappa)\Big] = 2\mu^\top \mathbb{V}[\hat{w}^\star(\kappa)]\mathbf{\Sigma}\mathbb{E}[\hat{w}^\star(\kappa)] = \frac{\kappa^3}{\gamma^3}\frac{2\psi^2}{T}. \tag{IA45}$$

Using (IA41)–(IA42), the OOSU mean is

$$
\begin{aligned}
\mathbb{E}[U(\hat{w}^{\star}(\kappa))] &= \mathbb{E}\Big[\hat{w}^{\star}(\kappa)^{\top}\mu\Big] - \frac{\gamma}{2}\mathbb{E}\Big[\hat{w}^{\star}(\kappa)^{\top}\boldsymbol{\Sigma}\hat{w}^{\star}(\kappa)\Big] \\
&= \mu_g - \frac{\gamma}{2}\sigma_g^2 + \frac{1}{\gamma}\left(\kappa\psi^2 - \frac{\kappa^2}{2}\left(\psi^2 + \frac{N-1}{T}\right)\right).
\end{aligned} \tag{IA46}
$$

Moreover, using (IA43)–(IA45), the OOSU variance is

$$
\begin{aligned}
\mathbb{V}[U(\hat{w}^{\star}(\kappa))] &= \mathbb{V}\Big[\hat{w}^{\star}(\kappa)^{\top}\mu\Big] + \frac{\gamma^2}{4}\mathbb{V}\Big[\hat{w}^{\star}(\kappa)^{\top}\boldsymbol{\Sigma}\hat{w}^{\star}(\kappa)\Big] - \gamma\,\mathbb{C}\mathrm{ov}\Big[\hat{w}^{\star}(\kappa)^{\top}\mu, \hat{w}^{\star}(\kappa)^{\top}\boldsymbol{\Sigma}\hat{w}^{\star}(\kappa)\Big] \\
&= \frac{\kappa^2}{\gamma^2}\frac{\psi^2}{T}\left((1-\kappa)^2 + \kappa^2\frac{N-1}{2T\psi^2}\right).
\end{aligned} \tag{IA47}
$$

Using (IA46)–(IA47), the mean-risk OOSU defined in (29) simplifies to Equation (IA4).

To conclude the proof, we need to prove that the shrinkage intensity $\kappa_R^{\star}$ maximizing (IA4) is smaller than the intensity $\kappa_E^{\star}$ that maximizes the OOSU mean in (IA46). 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 (IA47) 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

$$
4\kappa^2\left(1 + \frac{N-1}{2T\psi^2}\right) - 6\kappa + 2 \geq 0 \tag{IA48}
$$

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 (IA48) 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

$$
\frac{6 + \sqrt{36 - 32\left(1 + \frac{N-1}{2T\psi^2}\right)}}{8\left(1 + \frac{N-1}{2T\psi^2}\right)} \leq \frac{\psi^2}{\psi^2 + \frac{N-1}{T}} \tag{IA49}
$$

for $\frac{4(N-1)}{T\psi^2} \leq 1$. Inequality (IA49) simplifies to

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

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

Proof of Proposition IA.2

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

$$\mathbb{E}[\hat{w}^\star(\kappa)] = w_g + \frac{\kappa}{\gamma}\frac{T}{T-N-1}\mathbf{B}\mu, \tag{IA50}$$

$$\begin{aligned}\mathbb{V}[\hat{w}^\star(\kappa)] =& \left(\frac{\sigma_g^2}{T-N-1} + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)+T^2\psi^2}{(T-N)(T-N-1)(T-N-3)}\right)\mathbf{B} \\ &+ \frac{\kappa^2}{\gamma^2}\frac{T^2(T-N+1)}{(T-N)(T-N-1)^2(T-N-3)}\mathbf{B}\mu\mu^\top\mathbf{B}.\end{aligned} \tag{IA51}$$

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

Part 1. The OOSU mean of the shrinkage portfolio $\hat{w}^\star(\delta,\kappa)$ in (IA5) is

$$\begin{aligned}\mathbb{E}[U(\hat{w}^\star(\delta,\kappa))] = &(1-\delta)\mu_{ew} + \delta\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top\mu\Big] - \frac{\gamma}{2}\Bigg((1-\delta)^2\sigma_{ew}^2 \\ &+ \delta^2\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\kappa)\Big] + 2\delta(1-\delta)\mathbb{E}\Big[w_{ew}^\top\mathbf{\Sigma}\hat{w}^\star(\kappa)\Big]\Bigg).\end{aligned} \tag{IA52}$$

From Kan, Wang, and Zhou (2021, Lemma 1), we have

$$\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top\mu\Big] = \mu_g + \frac{\kappa}{\gamma}\frac{T}{T-N-1}\psi^2, \tag{IA53}$$

$$\mathbb{E}\Big[\hat{w}^\star(\kappa)^\top\mathbf{\Sigma}\hat{w}^\star(\kappa)\Big] = \frac{T-2}{T-N-1}\sigma_g^2 + \frac{\kappa^2}{\gamma^2}\frac{T(T-2)(N-1)+T^2(T-2)\psi^2}{(T-N)(T-N-1)(T-N-3)}. \tag{IA54}$$

Moreover, using Equation (IA50), we find that

$$\mathbb{E}\Big[w_{ew}^\top\mathbf{\Sigma}\hat{w}^\star(\kappa)\Big] = \sigma_g^2 + \frac{\kappa}{\gamma}\frac{T}{T-N-1}(\mu_{ew}-\mu_g). \tag{IA55}$$

Plugging (IA53)–(IA55) into (IA52), we find that the OOSU mean is given by Equation (IA7).

Part 2. We derive the expressions for the three components of the OOSU variance in (IA8).

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

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

where $\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \mu\Big]$ is given by (IA16), which results in Equation (IA9).

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

$$\begin{aligned}\mathbb{V}\Big[\hat{w}^\star(\delta,\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\delta,\kappa)\Big] &= \delta^4 \mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big] + 4\delta^2(1-\delta)^2 w_{ew}^\top \mathbf{\Sigma} \mathbb{V}[\hat{w}^\star(\kappa)] \mathbf{\Sigma} w_{ew} \\ &\quad + 4\delta^3(1-\delta)\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} w_{ew}, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big]. \end{aligned} \tag{IA56}$$

The first term, $\mathbb{V}\Big[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big]$, is given by (IA17). Using Equation (IA51), 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 (IA18) and is

$$\begin{aligned} \mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\kappa)^\top \mathbf{\Sigma} w_{ew}, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big] &= \frac{2\kappa}{\gamma}(\mu_{ew}-\mu_g)\Bigg(\sigma_g^2 \frac{T(T-2)}{(T-N-1)^2(T-N-3)} \\ &\quad + \frac{\kappa^2}{\gamma^2}\frac{T^2(T-2)(T+N-3+2T\psi^2)}{(T-N)(T-N-1)^2(T-N-3)(T-N-5)}\Bigg). \end{aligned} \tag{IA57}$$

Putting these three terms together into (IA56) gives Equation (IA10).

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

$$\begin{aligned} \mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\delta,\kappa)^\top \mu, \hat{w}^\star(\delta,\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\delta,\kappa)\Big] &= \delta^3 \mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big] \\ &\quad + 2\delta^2(1-\delta)\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} w_{ew}\Big]. \end{aligned} \tag{IA58}$$

The first term, $\mathbb{C}\mathrm{ov}\Big[\hat{w}^\star(\kappa)^\top \mu, \hat{w}^\star(\kappa)^\top \mathbf{\Sigma} \hat{w}^\star(\kappa)\Big]$, is given by (IA18). Using Equation (IA51), the second term is

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

Putting these two terms together into (IA58) results in the final expression in Equation (IA11) and concludes the proof.

Proof of Proposition IA.3

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

$$\begin{aligned}
\mathbb{E}[\hat{w}(\pi)] &= \pi\mu_{ew} + (1-\pi)\mu_g, \\
\mathbb{V}[\hat{w}(\pi)] &= (1-\pi)^2\frac{\sigma_g^2}{T-N-1}\mathbf{B}.
\end{aligned}$$

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

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

Taking the derivative of $\mathbb{E}[(\hat{w}(\pi)^\top\mu - \mu_g)^2]$ with respect to π and setting it to zero yields the final expression for π in (IA13).

References in Internet Appendix

Barroso, P., and K. Saxena. 2021. Lest we forget: using out-of-sample forecast errors in portfolio optimization. *The Review of Financial Studies* (forthcoming).

Campbell, J., A. Lo, and C. Mackinlay. 1997. *The econometrics of financial markets.* Princeton University Press.

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

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

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

Jagannathan, R., and T. Ma. 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance* 58:1651–84.

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

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

Ledoit, O., and M. Wolf. 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies* 30:4349–88.

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

Okhrin, Y., and W. Schmid. 2006. Distributional properties of portfolio weights. *Journal of Econometrics* 134:235–56.

Rencher, A., and G. B. Schaalje. 2008. *Linear models in statistics (2nd ed.).* John Wiley & Sons.

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