Economic Modelling 144 (2025) 106981

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.journals.elsevier.com/economic-modelling

Improving minimum-variance portfolio through shrinkage of large covariance matrices[☆],[☆☆]

Fangquan Shi [a], Lianjie Shu [b], Fangyi He [c], Wenpo Huang [d],[*]

[a] School of Finance, Nanjing Audit University, Jiangsu, China
[b] Faculty of Business Administration, University of Macau, Macao Special Administrative Region of China
[c] School of Finance, Southwestern University of Finance and Economics, Sichuan, China
[d] Experimental Center of Data Science and Intelligent Decision, Hangzhou Dianzi University, Zhejiang, China

ARTICLE INFO

Dataset link: MV Portfolio via Shrinkage of Large CM (Original data)

Keywords:
Portfolio optimization
Covariance matrix
Linear shrinkage
High dimension

ABSTRACT

The global minimum-variance (GMV) portfolio derived from the sample covariance matrix often performs poorly due to large estimation errors. Linear shrinkage covariance estimators have been extensively studied to address this issue. This study proposes an optimal shrinkage intensity selection for the linear shrinkage estimator family using cross-validated negative log-likelihood function minimization. Moreover, we provide theoretical insights into the selection process. Empirical studies have shown that the proposed approach produces more stable covariance matrix estimators than the Frobenius loss minimization method, resulting in improved GMV portfolios. Furthermore, linear shrinkage estimators that use a diagonal matrix or a matrix based on a one-factor model as the target matrix generally achieve the best performance. They also outperform nonlinear shrinkage covariance estimators, especially with a large number of assets. This superiority is evident in terms of out-of-sample variance, turnover, and the Sharpe ratio.

1. Introduction

One of the foundational theories in financial economics is mean–variance analysis (Markowits, 1952), providing a framework for assembling an asset portfolio to optimize expected return for a given level of risk. To implement such portfolios, investors must estimate the expected returns and the covariance matrix of asset returns. Common estimators, such as sample means and sample covariance matrices, often suffer from large estimation errors. These errors can result in suboptimal portfolios that perform poorly in out-of-sample testing (Jobson and Korkie, 1980, 1981; Frost and Savarino, 1986, 1988; Michaud, 1989; Best and Grauer, 1991; Chopra and Ziemba, 1993; Broadie, 1993). Furthermore, the results reveal that estimation errors in the mean returns have a greater impact than those in the covariance matrix (Merton, 1980; Best and Grauer, 1991). Consequently, recent studies have focused on global minimum-variance (GMV) portfolios that exclusively rely on the covariance estimate (Jagannathan and Ma, 2003; DeMiguel et al., 2009a; Ledoit and Wolf, 2017; Jiang et al., 2019; Shi et al., 2020).

Nevertheless, the sample covariance matrix may exhibit substantial estimation errors, especially in high-dimensional environments. Thus, the sample covariance matrix-derived GMV portfolio is still prone to underperformance. The resulting ideal weights tend to take extreme values and become unstable over time. Several approaches have been offered in the literature to address this issue, which can be classified into two categories. The first category involves imposing some limits on portfolio weights. The rationale behind it is that constraining or penalizing portfolio weights can reduce extreme values and stabilize the portfolio weights. This category includes studies such as Jagannathan and Ma (2003), Brodie et al. (2009), DeMiguel et al. (2009a), Fan et al. (2012), Behr et al. (2013), Levy and Levy (2014), Xing et al. (2014), Yen and Yen (2014), Li (2015), Yen (2016), Ao et al. (2019), Kremer et al. (2020), Li et al. (2022), and Tu and Li (2024).

The second category seeks improved estimators for the covariance matrix. A dominant approach in this category is the linear shrinkage estimator, which is based on a weighted combination of the sample covariance matrix and a target matrix. The sample covariance matrix

☆ We thank the editor, the associate editor, and two anonymous reviewers for their insightful comments that help us improve this paper.
☆☆ Funding: F.S. was supported by the National Natural Science Foundation of China (No. 72101119). L.S. was supported by the Department of Science and Technology of Guangdong Province (EF020/FBASLJ/2022/GDSTC) and the University of Macau Research Committee (MYRG2022-00017-FBA, MYRG-GRG2023-00004-FBA). W.H. was supported by the National Natural Science Foundation of China (No. 72171064).
* Corresponding author.
E-mail address: whuang@hdu.edu.cn (W. Huang).

https://doi.org/10.1016/j.econmod.2024.106981
Received 18 April 2024; Received in revised form 20 December 2024; Accepted 20 December 2024
Available online 31 December 2024

is unbiased but has a high variance, whereas the target matrix often has a high bias and a low variance. Moreover, different target matrices have been proposed in the finance literature, including the multiple of the identity matrix (Ledoit and Wolf, 2004b), double constant matrix (Ledoit, 1995), diagonal matrix (Ledoit, 1995), constant correlation (CC) matrix (Ledoit and Wolf, 2004a), covariance matrix based on a one-factor model with the market as the factor (Ledoit and Wolf, 2003), and a combination of a highly structured matrix and an eigenvalue-clipping-based estimator (Deshmukh and Dubey, 2020). In addition to the linear shrinkage estimators, several nonlinear shrinkage estimators have been proposed that use a nonlinear transformation function, such as the nonlinear shrinkage (NLS) covariance matrix estimator (Ledoit and Wolf, 2012), analytical nonlinear shrinkage (ANLS) covariance matrix estimator (Ledoit and Wolf, 2020), and quadratic inverse shrinkage (QIS) estimator (Ledoit and Wolf, 2022).

The portfolios based on linear shrinkage covariance estimators are easy to interpret. They can be considered a hybrid of the sample portfolio (constructed using the sample covariance matrix) and the target portfolio (constructed using the target matrix). Different target matrices yield different portfolios. For example, the equally weighted (EW) portfolio (DeMiguel et al., 2009b) can be viewed as a GMV portfolio constructed with the identity matrix. Meanwhile, the volatility timing portfolio (Kirby and Ostdiek, 2012) is a diagonal-matrix-based GMV portfolio construct. Covariance matrices with factor structures are used to construct the factor model-based portfolios (Chan et al., 1999).

The linear shrinkage covariance estimator involves the selection of the shrinkage intensity. When shrinkage intensity is small, the shrinkage estimator resembles the sample covariance matrix, and the resulting portfolio behaves similarly to the sample portfolio. Meanwhile, a higher shrinkage intensity causes the generated portfolio to behave more similarly to the target portfolio. Therefore, the shrinkage intensity has an important effect on the GMV portfolio performance. How to select the optimal shrinkage intensity is still a practical challenge. This subject, however, has been studied less rigorously. Ledoit (1995) and Ledoit and Wolf (2003, 2004a,b) proposed an intuitive criterion for selecting shrinkage intensity to minimize the distance between the true and estimated covariance matrices using the Frobenius norm. A more natural way of selecting the shrinkage intensity is using cross-validation, which is widely addressed in high-dimensional statistics. The present study aims to develop a formal method for selecting the shrinkage intensity for the family of linear shrinkage estimators of covariance. As our first contribution, we suggest minimizing the cross-validated negative log-likelihood function to determine shrinkage intensity.

The empirical results show that our approach typically selects a higher shrinkage intensity than the intuitive method based on minimizing the Frobenius loss of estimation errors. Hence, the proposed linear shrinkage covariance matrix estimator more closely resembles the target matrix than the Frobenius loss minimization estimator. The target matrix often has small variance than the sample covariance matrix. Thus, the proposed approach yields a more stable covariance matrix estimator of asset returns than the intuitive approach, resulting in improved GMV portfolios with lower variance, higher Sharpe ratios (SR), and lower turnover. This study's second contribution is to provide theoretical insights into the selection of shrinkage intensity. The results show that the proposed linear shrinkage covariance estimator is consistent, since it converges to the true covariance matrix as the sample size approaches infinity. Furthermore, our theoretical findings show that when the target matrix is close to the sample covariance matrix, the optimal shrinkage intensity. Meanwhile, the optimal shrinkage intensity is small when the target matrix differs significantly from the sample covariance matrix.

Our third contribution is a comprehensive comparative study of GMV portfolios based on the family of LS and NLS estimators. The empirical results demonstrate that the proposed linear shrinkage estimators result in improved GMV portfolios than those based on minimizing Frobenius loss. Furthermore, among the linear shrinkage covariance matrix estimators that minimize the cross-validated negative log-likelihood, those that use either a diagonal matrix or a matrix based on a one-factor model as the target matrix outperform the others in terms of out-of-sample portfolio performance. Moreover, these estimators outperform NLS estimators, especially when the number of assets is large.

The remainder of this study is structured as follows. Section 2 explores the shrinkage estimation of the covariance matrix. Section 3 examines the selection of shrinkage intensity and provides some theoretical insights. Section 4 outlines the framework adopted to compare the performance of various portfolio strategies. Section 5 presents empirical results relevant to the out-of-sample performance of the proposed strategy and its competing methodologies. Finally, Section 6 provides the concluding remarks.

2. Shrinkage covariance matrix estimators

Consider the GMV portfolio of p assets, which can be constructed by solving the following optimization problem:

$$\min_{w} \quad w'\hat{\Sigma}w \tag{1}$$

$$\text{subject to} \quad w'\mathbf{1} = 1, \tag{2}$$

where w is a $p \times 1$ vector of portfolio weights, $\hat{\Sigma}$ is a $p \times p$ covariance matrix estimator of assets, and $\mathbf{1}$ is a $p \times 1$ vector of 1 s. The optimal portfolio weights are

$$\hat{w} = \frac{\hat{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\hat{\Sigma}^{-1}\mathbf{1}}. \tag{3}$$

The estimated covariance matrix determines the GMV portfolio weights. A better estimate of the covariance matrix should improve out-of-sample portfolio performance.

2.1. Linear shrinkage of covariance matrices

The linear shrinkage covariance matrix estimator is a weighted sum of the sample covariance matrix and a target matrix, given by

$$\hat{\Sigma}_{\text{LS}} = (1 - \nu)S + \nu\hat{\Sigma}_{\text{target}}, \tag{4}$$

where $0 < \nu < 1$ denotes the shrinkage intensity, and $\hat{\Sigma}_{\text{target}}$ denotes the target matrix toward which the sample covariance matrix is shrunk.

The sample covariance matrix is unbiased yet may exhibit substantial variation, resulting in unstable portfolio weights. For $p > n$, because the sample covariance matrix is not even invertible, it cannot be used to construct optimal portfolios. To ensure the positive definiteness of the covariance estimator $\hat{\Sigma}_{\text{LS}}$, Proposition 1 indicates that the target matrix needs to be positive definite. This is straightforward as the weighted sum of a positive semi-definite matrix and a positive definite matrix is positive definite.

Proposition 1. *The linear shrinkage covariance matrix estimator $\hat{\Sigma}_{LS}$ is positive definite for $0 < \nu < 1$ if the target matrix Σ_{target} is positive definite.*

The extant literature has suggested different target matrices. For instance, Ledoit and Wolf (2004b) studied the simplest target matrix, namely, a multiple of the identity matrix, given by

$$\hat{\Sigma}_{\text{I}} = \bar{s}I \tag{5}$$

where I represents the identity matrix, and $\bar{s}$ denotes the average value of the diagonal elements in the sample covariance matrix S.

The target matrix $\hat{\Sigma}_{\text{I}}$ is highly structured and stable. Nevertheless, it exhibits high bias because it assumes all covariances are zero. By providing an additional parameter to characterize the covariances, Ledoit (1995) proposes the double constant matrix as the target matrix, given by

$$\hat{\Sigma}_{\rm DC} = \bar{s}I + \bar{q}(\mathbf{11}' - I), \tag{6}$$

where $\bar{q} = \frac{2}{p(p-1)}\sum_{i=1}^{p}\sum_{j=1}^{i-1} s_{ij}$ represents the average of the off-diagonal elements of the sample covariance matrix, and $\mathbf{1}$ denotes a conformable column vector of 1 s.

In addition to $\hat{\Sigma}_{\rm DC}$, Ledoit (1995) suggests the utilization of a diagonal target matrix as

$$\hat{\Sigma}_{\rm D} = S_{\rm d} = {\rm Diag}(s_{11}, \ldots, s_{ii}, \ldots, s_{pp}), \tag{7}$$

where s_{ii} denotes the ith diagonal element of the sample covariance matrix S.

The diagonal target matrix assumes that all correlations are zero, which may result in bias. To overcome this risk, the CC matrix can be employed as the target matrix (Ledoit and Wolf, 2004a), given by

$$\hat{\Sigma}_{\rm CC} = S_{\rm d}^{1/2}\hat{R}_{\rm CC}S_{\rm d}^{1/2}, \tag{8}$$

where $\hat{R}_{\rm CC}$ represents a correlation matrix with off-diagonal elements equal to the average value ($\bar{\rho}$) of the off-diagonal elements in the sample correlation matrix C.

The CC matrix asserts that all paired correlations are equal, which is overly restrictive. To estimate the target matrix with varying paired correlations, we use the single-factor model (Ledoit and Wolf, 2003). The single-index model assumes that stock returns are generated by the following.

$$y_t = \alpha + \beta f_t + \varepsilon_t, \tag{9}$$

where $y_t = (y_{1t}, \ldots, y_{pt})'$ denotes a $p \times 1$ vector of asset returns, f_t represents the single factor, and $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{pt})'$ is a $p \times 1$ vector of error items that are independent of each other. Additionally, $\alpha = (\alpha_1, \ldots, \alpha_p)$ stands for a $p \times 1$ vector of abnormal returns, and $\beta = (\beta_1, \ldots, \beta_p)$ represents a $p \times 1$ vector of stock betas. This factor model allows for the computation of the covariance matrix of asset returns as follows:

$$\Sigma_{\rm SF} = \sigma_f^2 \beta\beta' + \Delta, \tag{10}$$

where σ_f^2 denotes the variance of the single factor, and Δ represents a diagonal matrix with it ith diagonal element being the variance of ε_{it}. Denote $\hat{\sigma}_f^2, \hat{\beta}$, and $\hat{\Delta}$ as the estimated counterparts of σ_f^2, β, and Δ, respectively. The target covariance matrix can be estimated by

$$\hat{\Sigma}_{\rm SF} = \hat{\sigma}_f^2 \hat{\beta}\hat{\beta}' + \hat{\Delta}. \tag{11}$$

Proposition 2. *The double constant matrix $\hat{\Sigma}_{DC}$ is almost surely positive definite.*

Proposition 3. *The CC matrix $\hat{\Sigma}_{CC}$ is almost surely positive definite.*

Proposition 4. *If $\bar{q} \geq 0$ and $\bar{\rho} \geq 0$, then the target matrices $\hat{\Sigma}_{DC}$ and $\hat{\Sigma}_{CC}$ can be rewritten as*

$$\hat{\Sigma}_{DC} = \bar{q}\mathbf{11}' + Diag(\bar{s} - \bar{q}, \ldots, \bar{s} - \bar{q}), \tag{12}$$

and

$$\hat{\Sigma}_{CC} = \bar{s}\tilde{\beta}\tilde{\beta}' + Diag((1-\bar{\rho})s_{11}, \ldots, (1-\bar{\rho})s_{pp}), \tag{13}$$

where $\tilde{\beta} = \left(\sqrt{\frac{\bar{\rho}s_{11}}{\bar{s}}}, \ldots, \sqrt{\frac{\bar{\rho}s_{pp}}{\bar{s}}}\right)'$

Propositions 2–3 indicate that $\hat{\Sigma}_{\rm DC}$ and $\hat{\Sigma}_{\rm CC}$ are almost surely positive definite. Proposition 4 further indicates that both $\hat{\Sigma}_{\rm DC}$ and $\hat{\Sigma}_{\rm CC}$ are closely related to the single-index model. Both can be expressed in a form similar to the target matrix $\hat{\Sigma}_{\rm SF}$ when $\bar{q} \geq 0$ and $\bar{\rho} \geq 0$.

Ledoit and Wolf (2003) propose the utilization of the equal-weighted sum of asset returns as the market returns. The target matrix is then given by

$$\hat{\Sigma}_{\rm M} = \hat{\sigma}_{\rm M}^2 \hat{\beta}\hat{\beta}' + \hat{\Delta}, \tag{14}$$

where $\hat{\sigma}_{\rm M}$ represents the sample variance of the equal-weighted index for the p asset returns. However, just averaging the asset returns may not provide a representative factor for these asset returns. To address this issue, one can employ the first principal component as the single index and derive the following target matrix:

$$\hat{\Sigma}_{\rm 1PC} = \hat{\sigma}_{\rm 1PC}^2 \hat{\beta}\hat{\beta}' + \hat{\Delta}, \tag{15}$$

where $\hat{\sigma}_{\rm 1PC}$ denotes the sample variance of the first principal component of the p asset returns.

The family of linear shrinkage covariance estimator can be generalized as

$$\hat{\Sigma}_{\text{LS-X}} = (1-\nu)S + \nu\hat{\Sigma}_{\rm X}, \tag{16}$$

where $\hat{\Sigma}_{\rm X}$ denotes a particular target matrix with X = I, DC, D, CC, M, and 1PC. Proposition 5 describes the relationship between the GMV portfolio based on linear shrinkage covariance matrices and the matrix-norm-constrained GMV portfolio. Proposition 5 restates Proposition 2 from DeMiguel et al. (2009a). Hence, the proof is omitted here. In particular, if the target matrix is $\hat{\Sigma}_{\rm I}$, the GMV portfolio based on $\hat{\Sigma}_{\text{LS-I}}$ is equivalent to the two-norm-constrained portfolio. Shrinking the sample covariance matrix toward the target matrix $\hat{\Sigma}_{\rm X}$ is equivalent to imposing an additional weight constraint $w'\hat{\Sigma}_{\rm X}w \leq \varsigma$ when solving the GMV portfolio optimization problem with the sample covariance matrix.

Proposition 5. *Provided S is nonsingular, for each $0 \leq \nu \leq 1$, there exists a constant ς, such that the solution to the GMV problem in (1)–(2) with the covariance estimator, $\hat{\Sigma}$, replaced by $\hat{\Sigma}_{LS\text{-}X} = (1-\nu)S + \nu\hat{\Sigma}_X$, coincides with the solution to the matrix-norm-constrained GMV portfolio. This is the solution to the GMV problem in (1)–(2) with the covariance estimator, $\hat{\Sigma}$, replaced by the sample covariance matrix S, subject to the additional constraint*

$$w'\hat{\Sigma}_X w \leq \varsigma. \tag{17}$$

2.2. Selection of the shrinkage intensity

The shrinkage intensity ν is crucial to the linear shrinkage estimator and portfolio selection (DeMiguel et al., 2013). It determines the bias–variance trade-off in the estimator. Typically, the sample covariance matrix has minimal bias but relatively high variance, whereas the target matrix has high bias but low variance (Warton, 2008). When the shrinkage intensity is high, the shrinkage covariance matrix estimator behaves more similarly to the target matrix, resulting in increased bias but lower variance. However, when the shrinkage intensity is low, the shrinkage covariance matrix estimator behaves more similarly to the sample covariance matrix, resulting in lower bias but increased variance. In this section, we propose selecting an optimal shrinkage intensity ν by reducing the cross-validated negative log-likelihood function to balance the bias–variance trade-off. Theoretical insights into selecting the shrinkage intensity for linear shrinkage estimator and portfolio selection are presented.

The processes for selecting shrinkage intensity are as follows: (i) partitioning the data; (ii) estimating the linear shrinkage covariance matrix based on different shrinkage intensity values; and (iii) finding the optimal shrinkage intensity. In Step (i), we randomly divide the data Y into K[1] approximately equal groups. The kth group $Y^{[k]}$ serves

[1] We set $K = 10$ in this empirical study, which is the most common choice.

the validation data, whereas the remaining $K-1$ groups $Y^{[-k]}$ constitute the training data. In Step (ii), based on training data, we generate the linear shrinkage covariance matrix $\hat{\Sigma}^{[-k]}$ using a grid of values within the range [0, 1] for the shrinkage intensity. Moreover, we use the validation data to calculate the out-of-sample covariance $S^{[k]}$. Finally, in Step (iii), we choose the optimal shrinkage intensity that minimizes the following cross-validated negative log-likelihood function:

$$\min \sum_{k=1}^{K} \operatorname{tr}\left(\left(\hat{\Sigma}^{[-k]}\right)^{-1} S^{[k]}\right) + \log(\det(\Sigma^{[-k]})). \tag{18}$$

Theorem 1. *Consider increasing sample sizes, such that as $n \to \infty, n_k \to \infty \forall k$, if $S \xrightarrow{p} \Sigma$, then $v^* \xrightarrow{p} 0$ when the target matrix is $\hat{\Sigma}_D$ or a rotation-equivariant matrix $\hat{\Sigma}_{RE}$ (e.g., $\hat{\Sigma}_I$), which has the same eigenvectors as the sample covariance matrix.*

Theorem 2. *Stationary points of the observed likelihood estimated by K-fold cross-validation for $\hat{\Sigma}_{LS\text{-}D}$ have the following form*

$$v^* = \left(\sum_{k=1}^{K}\sum_{i=1}^{p} \omega_{ik}\right)^{-1} \sum_{k=1}^{K}\sum_{i=1}^{p} \omega_{ik} \frac{\psi_{ii}^{[k]} - \gamma_i^{[-k]}}{1-\gamma_i^{[-k]}}, \tag{19}$$

where $\omega_{ik} = \left(1-\gamma_i^{[-k]}\right)^2 \Big/ \left((1-v^*)\gamma_i^{[-k]} + v^*\right)^2$ and $\psi_{ii}^{[k]}$ is the ith diagonal element of the matrix $\Psi^{[k]} = \left(Q^{[-k]}\right)' \left(\hat{\Sigma}_D^{[-k]}\right)^{-1/2} S^{[k]} \left(\hat{\Sigma}_D^{[-k]}\right)^{-1/2} Q^{[-k]}$. The sample covariance matrix $S^{[-k]}$ based on the kth training data has the following correlation variance decomposition: $S^{[-k]} = \left(\hat{\Sigma}_D^{[-k]}\right)^{1/2} C^{[-k]} \left(\hat{\Sigma}_D^{[-k]}\right)^{1/2}$. Meanwhile, the correlation matrix $C^{[-k]}$ has the following eigendecomposition: $C^{[-k]} = Q^{[-k]}\Gamma^{-k}\left(Q^{[-k]}\right)'$ where $\Gamma^{[-k]} = Diag(\gamma_1^{[-k]}, \ldots, \gamma_p^{[-k]})$.

Theorem 3. *If the target matrix $\hat{\Sigma}_{RE}$ is rotation-equivariant (e.g., $\hat{\Sigma}_I$), that is, it has the same eigenvectors as the sample covariance matrix, then stationary points of the observed likelihood estimated by K-fold cross-validation for $\hat{\Sigma}_{RE}$ have the following form:*

$$v^* = \left(\sum_{k=1}^{K}\sum_{i=1}^{p} \omega_{ik}\right)^{-1} \sum_{k=1}^{K}\sum_{i=1}^{p} \omega_{ik} \frac{\pi_{ii}^{[k]} - \lambda_i^{[-k]}}{\tau_i - \lambda_i^{[-k]}}, \tag{20}$$

where $\omega_{ik} = \left(\tau_i - \lambda_i^{[-k]}\right)^2 \Big/ \left((1-v^*)\lambda_i^{[-k]} + v^*\tau_i\right)^2$ and $\pi_{ii}^{[k]}$ is the ith diagonal element of the matrix $\Pi^{[k]} = \left(U^{[-k]}\right)' S^{[k]} U^{[-k]}$. The sample covariance $S^{[-k]}$ based on the kth training data has the following eigendecomposition: $S^{[-k]} = U^{[-k]}\Lambda^{-k}\left(U^{[-k]}\right)'$ where $\Lambda^{[-k]} = Diag(\lambda_1^{[-k]}, \ldots, \lambda_p^{[-k]})$. The target matrix $\hat{\Sigma}_{RE}$ based on the kth training data has the following eigendecomposition: $\hat{\Sigma}_{RE}^{[-k]} = U^{[-k]}T\left(U^{[-k]}\right)'$ where $T = Diag(\tau_1, \ldots, \tau_p)$.

The proof of Theorems 2 and 3 involves expressing the negative log-likelihood function in terms of $\psi_{ii}^{[k]}, \gamma_i^{[-k]}$, and v for Theorem 2 (or $\pi_{ii}^{[k]}, \lambda_i^{[-k]}, \tau_i$, and v for Theorem 3), then solving for the stationary point. Following Warton (2008), we omitted the proof.

Theorem 1 suggests that the $\hat{\Sigma}_{\text{LS-I}}$ and $\hat{\Sigma}_{\text{LS-D}}$ are consistent estimators for Σ because they approach to the sample covariance matrix as $n \to \infty$, which is a consistent estimator itself. This is suitable as the shrinkage intensity should go to zero ($v \to 0$) when an increasing amount of data becomes available ($n \to \infty$).

The term $\psi_{ii}^{[k]}$ in Theorem 2 can be interpreted as the out-of-sample estimate of ith eigenvalue of the sample correlation matrix. Meanwhile, the term $\psi_{ii}^{[k]} - \gamma_i^{[-k]}$ represents the bias in the estimated eigenvalue. The optimal shrinkage intensity v^* is a weighted sum of the ratio of $\psi_{ii}^{k} - \gamma_i^{[-k]}$ to $1 - \gamma_i^{[-k]}$. If the bias is minor, the optimal v tends to be small. This suggests that v^* decreases with the sample size n, because the estimating bias diminishes as the sample size increases.

Theorem 3 underscores the relationship between the optimal shrinkage intensity v^* and the difference between the target and sample covariance matrices. When the target matrix is close to the sample covariance matrix, their eigenvalues are also close to each other. Consequently, the magnitude of the ratio of $\pi_{ii}^{[k]} - \lambda_i^{[-k]}$ to $\tau_i - \lambda_i^{[-k]}$ is large, and the optimal shrinkage intensity v^* tends to be large. However, if the target matrix differs significantly from the sample covariance matrix, the optimal shrinkage intensity v^* is typically modest.

Note that Theorem 3 focuses on rotation-equivariant target matrices with the same eigenvectors as the sample covariance matrix. For more general target matrices, it is difficult to determine the link between the ideal shrinkage intensity and the difference between the sample covariance matrix and the target matrix. Nonetheless, when the gap between the sample covariance matrix and the target matrix decreases, we anticipate that the optimal shrinkage intensity increases, as will be demonstrated in the empirical section.

3. Empirical setup

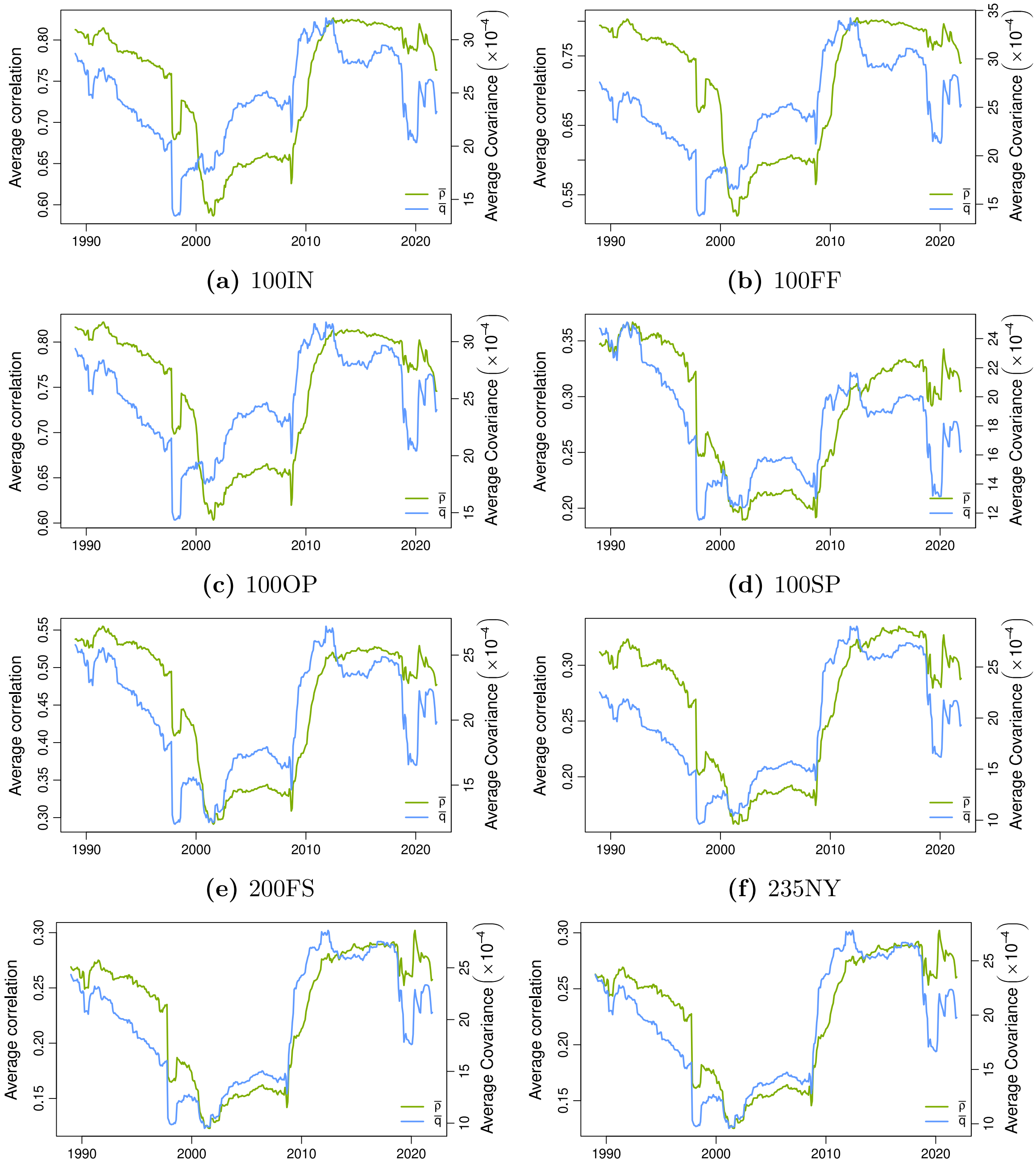
3.1. Data sets and studied portfolios

We obtained monthly stock return data from the Center for Research in Security Prices (CRSP) for the period January 1979–December 2021. Furthermore, we obtained monthly returns for bivariate portfolios based on common factors from Ken French's website. Using this stock return data, we constructed the following data sets, namely, (i) 100 portfolios based on size and investment (100IN), (ii) 100 Fama–French portfolios (100FF), (iii) 100 portfolios based on size and operating profitability (100OP), (iv) 100 randomly selected stocks from the S&P 500 index (100SP), (v) the combination of 100FF and 100SP portfolios (200FS), (vi) 235 stocks from the New York Stock Exchange (235NY), (vii) 300 stocks from the CRSP Stock Database (300CP), and (viii) 500 stocks from the CRSP Database (500CP).

Fig. 1 illustrates the average values of sample correlations and covariances derived using a rolling window procedure with a window size of 120 for different data sets. Fig. 1 shows that average correlations and covariances are consistently positive and exhibit comparable fluctuations across the estimated windows. Furthermore, the 100IN, 100FF, and 100OP data sets have average correlations ranging from 0.55 to 0.8, which are significantly higher than the values (approximately 0.15–0.3) for the 100SP, 235NY, 300CP, and 500CP data sets. This difference is reasonable as each asset in the former data sets indicates a portfolio constructed from individual stocks driven by common factors.

The portfolios evaluated in this study are as follows:

- **GMV-S:** Portfolio (3), where $\hat{\Sigma}$ is replaced by the sample covariance matrix, serves as a benchmark portfolio for evaluating the improvement of a covariance matrix estimator over the sample covariance matrix. It is important to note that this portfolio is unavailable when $p > n$, as the sample covariance is singular in this case.
- **EW:** The EW portfolio (DeMiguel et al., 2009b) avoids optimization or parameter estimation and can be used as another benchmark portfolio to assess the effort of portfolio optimization and covariance matrix estimation.
- **NLS:** Portfolio (3), where $\hat{\Sigma}$ is replaced by the nonlinear shrinkage covariance estimator proposed by Ledoit and Wolf (2012).
- **ANLS:** Portfolio (3), where $\hat{\Sigma}$ is replaced by the ANLS covariance estimator proposed by Ledoit and Wolf (2020).
- **QIS:** Portfolio (3), where $\hat{\Sigma}$ is replaced by the quadratic shrinkage covariance estimator proposed by Ledoit and Wolf (2022).
- **SN$_1$:** Portfolio (3), where $\hat{\Sigma}$ is replaced by the 1-Schatten norm penalized covariance matrix estimator studied by Shi et al. (2020). When the tuning parameters meet a specific criterion, this portfolio corresponds to the LS-I portfolio. Please refer to Shi et al. (2020) for details.



(a) 100IN (b) 100FF (c) 100OP (d) 100SP (e) 200FS (f) 235NY (g) 300CP (h) 500CP

Fig. 1. Average values of sample correlations and covariances for various data sets.
Notes: Fig. 1 displays the average pairwise correlations $\bar{\rho}$ and average covariances $\bar{q}$ over the rolling windows of length 120 for various data sets.

- **RR:** Portfolio (3), where $\hat{\Sigma}$ is replaced by the ridge regularized (RR) correlation matrix estimator proposed by Warton (2008). When the shrinkage intensity is equal, the RR portfolio and the LS-D portfolio coincide.
- **LS-X-F where X is I, DC, D, CC, and M:** Portfolio (3), where $\hat{\Sigma}$ is replaced by the linear shrinkage estimators with target matrices such as identity matrix (Ledoit and Wolf, 2004b), double constant matrix (Ledoit, 1995), diagonal matrix (Ledoit, 1995), CC matrix (Ledoit and Wolf, 2004a), and covariance matrix based on one-factor model with market as the factor (Ledoit and Wolf, 2003). The "**F**" indicates that the shrinkage intensity is determined by minimizing the Frobenius norm between the covariance matrix estimator and the true covariance matrix.[2]
- **LS-X-L where X is I, DC, D, CC, M, and 1PC:** Portfolio (3), where $\hat{\Sigma}$ is replaced by the linear shrinkage estimators with target matrices such as identity matrix, double constant matrix, diagonal matrix, CC matrix, and the covariance matrix based on

[2] It is also possible to select the shrinkage intensity by minimizing the cross-validated Frobenius loss $\sum_{k=1}^{K}\|\hat{\Sigma}^{[-k]} - S^{[k]}\|_F$. We are in favor of the cross-validated negative log-likelihood function in Eq. (18) because the term $\log(\det(\hat{\Sigma}^{[-k]}))$ tends to make the covariance matrix estimator well-conditioned.

one-factor model with market or the first principal component as the factor. The "**L**" indicates that the shrinkage intensity is selected by minimizing the negative log-likelihood function.

3.2. Evaluation criteria

We analyze the out-of-sample performance of different portfolios using the following criteria: (i) out-of-sample portfolio variance (VR), (ii) out-of-sample portfolio SR, (iii) portfolio turnover (TO), and (iv) net portfolio SR (NSR). The VR serves as a metric for evaluating portfolio risk, whereas the SR provides information about the risk-adjusted return on investment. Meanwhile, TO, as measured by trading volume, demonstrates the stability of portfolio selection. A low TO for rebalancing portfolios suggests low transaction costs. Finally, the NSR reveals portfolio returns after controlling for risk and transaction costs.

In accordance with DeMiguel et al. (2009a), we use a "rolling window" procedure to compare out-of-sample performance across different portfolios. The window length, denoted as n, is set to encompass 120 consecutive data points, corresponding to a decadal timescale for monthly data. Optimal weights for different portfolios are calculated for each window using return data. This process iterates on a monthly basis, incorporating data from the following month while discarding data from the earliest month. The rolling window mechanism is used until the dataset is complete. Ultimately, $T-n$ vectors $\hat{w}_t$ of portfolio weights are generated for each sequential out-of-sample period $t = \{n, n+1, \ldots, T-1\}$.

Holding a portfolio with weights $\hat{w}_t$ for a single period yields the out-of-sample return r_{t+1} at time $t+1$. This return is computed as $r_{t+1} = \hat{w}_t' y_{t+1}$, where y_{t+1} is a vector representing asset returns at time $t+1$. Subsequently, the out-of-sample VR, SR, and TO for each portfolio are calculated as follows:

$$\widehat{\mathrm{VR}} = \frac{1}{T-n-1}\sum_{t=n}^{T-1}(r_{t+1}-\hat{\mu})^2, \text{ with } \hat{\mu} = \frac{1}{T-n}\sum_{t=n}^{T-1} r_{t+1} \tag{21}$$

$$\widehat{\mathrm{SR}} = \frac{\hat{\mu}}{\sqrt{\widehat{\mathrm{VR}}}}, \tag{22}$$

$$\widehat{\mathrm{TO}} = \frac{1}{T-n-1}\sum_{t=n}^{T-1}\sum_{j=1}^{p}(|\hat{w}_{j,t+1}-\hat{w}_{j,t^+}|), \tag{23}$$

where $\hat{w}_{j,t}$ represents the portfolio weight of asset j at time t, $\hat{w}_{j,t^+}$ denotes the portfolio weight before-rebalancing at time $t+1$, and $\hat{w}_{j,t+1}$ stands for the optimal portfolio weight after-rebalancing at time $t+1$. The computation of portfolio turnover involves taking the simple average of rebalancing trades across the p available assets throughout the $T-n-1$ trading periods.

An optimal portfolio aims to achieve high risk-adjusted returns while keeping turnover low. Nonetheless, a trade-off exists between risk and turnover. Portfolio rebalancing is frequently required to manage risk, incurring increased transaction costs. Hence, evaluating the NSR, accounting for these expenses, becomes imperative. The NSR is computed using Eq. (22), but out-of-sample returns are adjusted to encompass transaction costs. DeMiguel et al. (2013) calculate the adjusted return at time $t+1$ as follows:

$$r_{t+1}^{\mathrm{adj}-} = (1+\hat{w}_t' y_{t+1})\left(1 - c\sum_{j=1}^{p}|\hat{w}_{j,t+1}-\hat{w}_{j,t^+}|\right) - 1, \tag{24}$$

where the transaction cost c is often set at 50 basis points per trade.

4. Empirical results

4.1. Behavior of shrinkage intensity

Fig. 2 presents the optimal values of the shrinkage intensity ν for linear shrinkage estimators (LS) over rolling windows of length 120 for various data sets based on two selection criteria, namely, minimizing the Frobenius norm between the true and estimated covariance matrices and minimizing the negative log-likelihood function. A relatively large shrinkage intensity is selected based on minimizing the negative log-likelihood over the rolling windows, as opposed to the Frobenius loss minimization selection criterion. This mismatch implies that shrinking the sample covariance matrix using these two different selection methods would result in drastically different out-of-sample portfolio performances. Furthermore, the optimal shrinkage intensity for LS-X-F can significantly vary across the rolling windows. Since a single observation replacement generates substantial changes in the optimal shrinkage intensity, the rapid changes indicate that LS-X-F, particularly LS-CC-F, may not be robust to outliers. However, the optimal shrinkage intensity exhibits smaller fluctuations over the rolling windows.

Table 1 presents the average value and standard deviation (in parentheses) of the optimal shrinkage intensity for linear shrinkage estimators with different target matrices over rolling windows for various data sets. The following are certain interesting observations.

First, for most data sets, the average optimal shrinkage intensity based on minimizing negative log-likelihood is significantly greater than that based on minimizing the Frobenius loss. The 100IN, 100FF, and 100OP data sets are the exceptions, as they use the CC matrix as their target matrix. Furthermore, in most cases, the standard deviation of the optimal shrinkage intensity selected by the negative log-likelihood minimization strategy is less than that of the Frobenius loss minimization approach.

Second, the optimal shrinkage intensity is the smallest for $\hat{\Sigma}_{\text{LS-I-L}}$, followed by $\hat{\Sigma}_{\text{LS-D-L}}$, $\hat{\Sigma}_{\text{LS-DC-L}}$, $\hat{\Sigma}_{\text{LS-CC-L}}$, and $\hat{\Sigma}_{\text{LS-M-L}}$. This trend is consistent with our prediction that the optimal shrinkage intensity is higher for the target matrix with lower bias. Moreover, the sample covariance matrix exhibits no bias but a high variance. The goal of shrinkage is to reduce the variance in the sample covariance matrix at the expense of slightly increased bias. The multiple of the identity matrix has low variance but high bias. Consequently, to obtain a more favorable bias–variance trade-off, the shrinkage intensity should be relatively low. However, the covariance matrix based on the factor model has a low bias but a high variance. Hence, the optimal shrinkage intensity for the target matrix based on the factor model is expected to be relatively large.

Third, the optimal shrinkage intensity increases with the number of stocks p. As p increases for a fixed sample size n, there is less data available to estimate the sample eigenvalues. Consequently, the predicted eigenvalues may be more biased and unstable. In this case, the optimal shrinkage intensity tends to increase, as suggested in Theorems 2 and 3.

4.2. Condition numbers

Table 2 reports the mean and standard deviation (in parentheses) of condition number for different methods across rolling windows for various data sets. The condition number of a square matrix is the ratio of its maximum and smallest eigenvalues. A low condition number implies that the matrix is well-conditioned, suggesting accurate inverse computation. Conversely, a large condition number suggests an ill-conditioned matrix. In such cases, even minor estimation errors in the sample covariance matrix might result in a substantial inaccuracy in the inverse matrix, significantly impacting the determination of optimal portfolio weights. Moreover, if the condition number is infinite, the matrix is singular and not invertible, making it unsuitable for constructing optimal portfolios. This scenario arises when the number of variables p surpasses the sample size n for the sample covariance matrix.

From Table 2 we can observe that the condition number of linear shrinkage estimators based on minimizing the negative log-likelihood is, unsurprisingly, lower than that based on minimizing the Frobenius loss. The condition numbers of the sample covariance matrices are

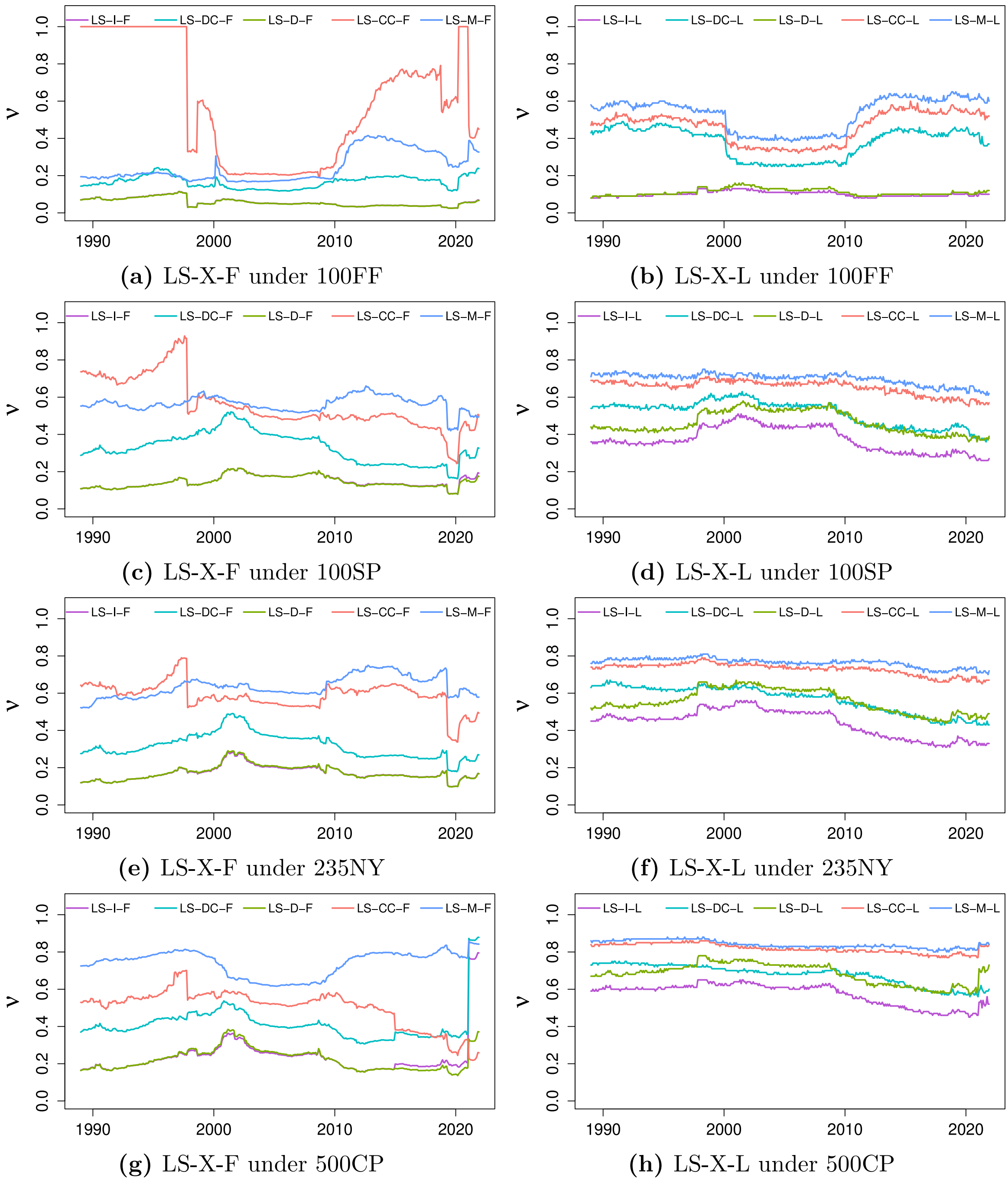


Fig. 2. Shrinkage intensity of different methods for various datasets.
Notes: Fig. 2 displays the optimal values of the shrinkage intensity ν for linear shrinkage estimators (LS) with different target matrices over rolling windows of length 120 for various datasets. The target matrices are the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), and covariance matrix based on the market model (M). Minimizing Frobenius loss (F) or negative log-likelihood function (L) determines the ideal shrinkage intensity.

large for the first four data sets and infinite for the last four data sets ($p > n$), implying that the sample covariance matrix is frequently ill-conditioned. The target matrix frequently has a modest condition number. For example, the identity matrix or its multiple has a condition number of 1. The condition number of the diagonal matrix is equal to the ratio of the maximum to the minimum diagonal elements, both of which are small. Moreover, linear shrinkage estimators with a higher shrinkage intensity would behave more similarly to the target matrix, resulting in a more stable covariance estimate with a lower condition number.

Among the linear shrinkage covariance matrix estimators based on minimizing the negative log-likelihood, $\hat{\Sigma}_{\text{LS-I-L}}$ exhibits the lowest condition numbers, followed by $\hat{\Sigma}_{\text{LS-DC-L}}$, $\hat{\Sigma}_{\text{LS-CC-L}}$, and $\hat{\Sigma}_{\text{LS-M-L}}$. Additionally, the condition numbers of $\hat{\Sigma}_{\text{LS-1PC-L}}$ are smaller than the $\hat{\Sigma}_{\text{LS-M-L}}$ despite the optimal values of shrinkage intensity for both approaches being close. Among the NLS covariance matrix estimators, $\hat{\Sigma}_{\text{NLS}}$ has the smallest condition numbers, followed by $\hat{\Sigma}_{\text{QIS}}$ and $\hat{\Sigma}_{\text{ANLS}}$ for most data sets. Moreover, the condition numbers of the NLS covariance matrix estimators are larger than those of $\hat{\Sigma}_{\text{LS-I-L}}$, $\hat{\Sigma}_{\text{LS-DC-L}}$, and $\hat{\Sigma}_{\text{SN}_1}$, and smaller than those of $\hat{\Sigma}_{\text{LS-CC-L}}$, $\hat{\Sigma}_{\text{LS-M-L}}$, $\hat{\Sigma}_{\text{LS-1PC-L}}$, and $\hat{\Sigma}_{\text{RR}}$.

Table 1
Mean and standard deviation of shrinkage intensity for linear shrinkage covariance estimators.

Method	100IN	100FF	100OP	100SP	200FS	235NY	300CP	500CP
$\hat{\Sigma}_{\text{LS-I-F}}$	0.053	0.057	0.054	0.147	0.096	0.172	0.237	0.236
	(0.024)	(0.021)	(0.023)	(0.031)	(0.022)	(0.038)	(0.114)	(0.102)
$\hat{\Sigma}_{\text{LS-I-L}}$	0.104	0.101	0.106	0.373	0.246	0.441	0.511	0.573
	(0.018)	(0.014)	(0.022)	(0.068)	(0.035)	(0.074)	(0.058)	(0.058)
$\hat{\Sigma}_{\text{LS-DC-F}}$	0.143	0.166	0.178	0.330	0.244	0.318	0.397	0.411
	(0.036)	(0.033)	(0.054)	(0.082)	(0.050)	(0.066)	(0.097)	(0.093)
$\hat{\Sigma}_{\text{LS-DC-L}}$	0.442	0.378	0.424	0.515	0.407	0.567	0.608	0.679
	(0.072)	(0.082)	(0.055)	(0.068)	(0.015)	(0.072)	(0.059)	(0.054)
$\hat{\Sigma}_{\text{LS-D-F}}$	0.053	0.056	0.053	0.145	0.095	0.173	0.218	0.222
	(0.024)	(0.021)	(0.022)	(0.031)	(0.022)	(0.040)	(0.063)	(0.059)
$\hat{\Sigma}_{\text{LS-D-L}}$	0.114	0.111	0.118	0.464	0.286	0.557	0.671	0.683
	(0.024)	(0.018)	(0.028)	(0.061)	(0.040)	(0.064)	(0.054)	(0.056)
$\hat{\Sigma}_{\text{LS-CC-F}}$	0.766	0.596	0.680	0.559	0.271	0.587	0.511	0.501
	(0.277)	(0.316)	(0.242)	(0.135)	(0.056)	(0.072)	(0.112)	(0.101)
$\hat{\Sigma}_{\text{LS-CC-L}}$	0.529	0.463	0.538	0.653	0.534	0.730	0.806	0.819
	(0.065)	(0.083)	(0.039)	(0.038)	(0.023)	(0.031)	(0.022)	(0.025)
$\hat{\Sigma}_{\text{LS-M-F}}$	0.262	0.244	0.276	0.565	0.474	0.636	0.737	0.740
	(0.036)	(0.082)	(0.052)	(0.042)	(0.116)	(0.056)	(0.079)	(0.068)
$\hat{\Sigma}_{\text{LS-M-L}}$	0.589	0.532	0.605	0.699	0.661	0.764	0.832	0.841
	(0.069)	(0.089)	(0.057)	(0.031)	(0.042)	(0.024)	(0.018)	(0.020)
$\hat{\Sigma}_{\text{LS-1PC-L}}$	0.588	0.532	0.600	0.695	0.661	0.760	0.826	0.835
	(0.071)	(0.091)	(0.065)	(0.033)	(0.044)	(0.027)	(0.022)	(0.027)

Notes: Table 1 reports the average and standard deviation (in parentheses) of optimal shrinkage intensity for linear shrinkage estimators (LS) with different target matrices over rolling windows of length 120 for various datasets. The target matrices are the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), covariance matrix based on the one-factor model as the market factor (M), and covariance matrix based on the one-factor model as the first principal component (1PC) as the factor. The optimal shrinkage intensity can be determined by minimizing Frobenius loss (F) or negative log-likelihood function (L).

Table 2
Condition numbers of linear and nonlinear shrinkage covariance matrix estimators.

Method	100IN	100FF	100OP	100SP	200FS	235NY	300CP	500CP
S	87 701	87 584	89 395	13 123	–	–	–	–
	(22 550)	(27 801)	(27 483)	(6495)	(–)	(–)	(–)	(–)
$\hat{\Sigma}_{\text{NLS}}$	782	1067	1387	135	471	203	178	247
	(181)	(1864)	(2203)	(110)	(152)	(95)	(114)	(115)
$\hat{\Sigma}_{\text{ANLS}}$	1458	1487	1504	222	554	228	196	271
	(455)	(520)	(519)	(114)	(177)	(107)	(127)	(127)
$\hat{\Sigma}_{\text{QIS}}$	1477	1482	1519	215	543	219	185	247
	(509)	(592)	(585)	(113)	(175)	(103)	(118)	(116)
$\hat{\Sigma}_{\text{SN}_1}$	794	765	770	110	395	186	156	219
	(174)	(166)	(189)	(37)	(121)	(73)	(55)	(71)
$\hat{\Sigma}_{\text{RR}}$	1092	1037	1066	236	947	1249	1552	2576
	(243)	(233)	(277)	(99)	(314)	(429)	(1172)	(1241)
$\hat{\Sigma}_{\text{LS-I-F}}$	1581	1350	1486	205	915	376	263	434
	(638)	(549)	(555)	(80)	(379)	(156)	(115)	(199)
$\hat{\Sigma}_{\text{LS-DC-F}}$	2115	1514	1609	143	648	292	199	320
	(686)	(459)	(411)	(66)	(300)	(131)	(73)	(130)
$\hat{\Sigma}_{\text{LS-D-F}}$	2118	1853	2113	474	2424	2849	3185	5802
	(816)	(720)	(826)	(218)	(1093)	(1112)	(1295)	(2333)
$\hat{\Sigma}_{\text{LS-CC-F}}$	862	974	958	238	1718	1404	2594	4430
	(154)	(274)	(556)	(110)	(677)	(497)	(2887)	(3472)
$\hat{\Sigma}_{\text{LS-M-F}}$	1904	2069	1936	235	2216	1882	2035	3676
	(461)	(367)	(535)	(86)	(313)	(1339)	(1976)	(2600)
$\hat{\Sigma}_{\text{LS-I-L}}$	681	662	664	62	289	102	74	93
	(163)	(155)	(183)	(28)	(100)	(53)	(33)	(42)
$\hat{\Sigma}_{\text{LS-DC-L}}$	658	652	649	82	351	145	120	172
	(143)	(139)	(160)	(28)	(110)	(55)	(38)	(54)
$\hat{\Sigma}_{\text{LS-D-L}}$	940	899	912	123	656	569	616	913
	(235)	(220)	(262)	(62)	(249)	(261)	(823)	(855)
$\hat{\Sigma}_{\text{LS-CC-L}}$	1015	986	993	202	899	1162	1594	2613
	(246)	(265)	(286)	(80)	(294)	(366)	(1030)	(1078)
$\hat{\Sigma}_{\text{LS-M-L}}$	1013	1097	1042	206	1703	1682	1897	3411
	(219)	(240)	(263)	(79)	(411)	(1261)	(1925)	(2653)
$\hat{\Sigma}_{\text{LS-1PC-L}}$	991	1090	1023	198	1624	1380	1586	2837
	(214)	(269)	(253)	(70)	(429)	(847)	(898)	(1530)

Notes: Table 2 reports the mean and standard deviation (in parenthesis) of condition number for linear shrinkage estimators (LS) with different target matrices over rolling windows of length 120 for various data sets. The target matrices include the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), covariance matrix based on a one-factor model with the market as the factor (M), and covariance matrix based on a one-factor model with the first principal component (1PC) as the factor. Minimizing Frobenius loss (F) or negative log-likelihood function (L) determines the ideal shrinkage intensity. Furthermore, the sample covariance matrix (S), Schatten 1-norm (SN_1) regularized covariance matrix, ridge regularized (RR) correlation matrix estimator, and some NLS covariance matrix estimators are included for comparisons.

4.3. Portfolio performance based on real data sets

Table 3 reports the monthly out-of-sample portfolio variances for different portfolios across real data sets. Panel A reveals that the GMV-S portfolio has the highest out-of-sample variance. Although the EW portfolio yields lower out-of-sample portfolio variation than the GMV-S portfolio, it demonstrates a significantly higher variance than other portfolios. The NLS, ANLS, and QIS portfolios all perform similarly, with lower out-of-sample variances than the LS-I-F portfolio, consistent with Ledoit and Wolf's (2017) findings. Moreover, Panels B and C show that the LS-X-L portfolios had lower out-of-sample variances than their LS-X-F counterparts in almost all data sets, with statistically significant differences in the first five data sets. Among the portfolios that minimize the negative log-likelihood, LS-D-L has the lowest out-of-sample portfolio variance in three of the eight data sets, followed by LS-1PC-L. Furthermore, LS-I-L and LS-DC-L exhibit much greater out-of-sample variances than LS-D-L, LS-M-L, and LS-1PC-L, particularly in the 235NY, 300CP, and 500CP data sets. Notably, LS-X-L achieves lower out-of-sample variances than the NLS, ANLS, QIS, and SN_1 portfolios, with statistically significant differences, especially for data sets with a large number of assets. This finding suggests that simple linear shrinkage covariance estimators with the appropriate shrinkage intensity can improve out-of-sample portfolio performance when compared to more complex NLS covariance estimators. Moreover, the LS-X-L portfolio often demonstrates lower out-of-sample variance than the RR portfolio across most data sets.

Table 4 shows the monthly out-of-sample SR for several portfolios using real data sets. Except for the 100SP data set, the GMV-S portfolio consistently shows the lowest SR, with the EW portfolio coming in second. Meanwhile, portfolios using NLS covariance estimators have greater SR than the LS-I-F portfolio, except in the 100SP and 235NY data sets. For most data sets, the LS-X-L portfolios outperform the LS-X-F portfolios in terms of the SR. Among the portfolios focused on minimizing the negative log-likelihood, LS-D-L, LS-M-L, and LS-1PC-L obtain more excellent SR, notably in the last four data sets where $p > n$. In most data sets, LS-X-L portfolios exhibit higher SR than those based on nonlinear shrinkage estimators, with substantial differences for data sets where $p > n$. Additionally, the RR portfolio exhibits relatively high SR.

Evaluating portfolio stability is critical, as transaction costs are high in practice. Table 5 reports the monthly portfolio turnover of different portfolios using real data sets. The GMV-S portfolio has the highest turnover for each data set due to the sample covariance matrix's extreme instability. By comparison, the EW portfolio has the smallest turnover. Meanwhile, the turnover of ANLS and QIS is uniformly larger than NLS.

Table 2 shows that LS-X-L portfolios have lower turnover than LS-X-F portfolios, indicating greater stability. Moreover, the LS-X-L portfolios had lower turnover than SN_1, RR, and portfolios based on nonlinear shrinkage covariance estimators (NLS, ANLS, QIS) across almost all data sets.

After accounting for risk and transaction costs, Table 6 reports various portfolios' monthly out-of-sample NSR across real data sets. The NSR for the GMV-S portfolio is negative for the first four data sets, primarily owing to excessive turnover. The EW portfolio has higher NSR than the NLS, ANLS, and QIS portfolios during the last six data sets, owing to its low turnover. Meanwhile, the LS-X-L portfolios achieve greater NSR than their LS-X-F counterparts, which is unsurprising given their low turnover and high SR.

Table 3
Monthly out-of-sample portfolio variance based on real data sets.

Method	100IN	100FF	100OP	100SP	200FS	235NY	300CP	500CP
Panel A: Portfolios from existing literature								
GMV-S	69.39***	55.43***	46.02***	40.79***	–	–	–	–
EW	24.62***	25.77***	24.92***	17.73***	20.28***	19.84***	20.07***	19.33***
NLS	11.36	12.56	10.54	10.22	8.94	6.60***	6.01***	5.53***
ANLS	11.71	12.40	10.37	10.33	8.96	6.65***	6.01***	5.60***
QIS	11.53	12.17	10.36	10.15	8.96	6.64***	5.95***	5.59***
SN_1	11.57	12.43	10.57	10.26	9.04*	6.68***	6.30***	5.66***
RR	11.49	12.41	10.43	10.14	8.75	4.59	4.29	3.70
Panel B: Portfolios from linear shrinkage estimator by minimizing Frobenius norm								
LS-I-F	13.39**	13.13**	10.96	11.42***	10.27***	7.07	6.62	5.89
LS-DC-F	13.77***	13.26**	10.95	10.26**	9.35***	6.79	6.20	5.69
LS-D-F	13.72**	13.30**	11.13**	11.74***	10.12***	5.02**	4.70	3.80
LS-CC-F	17.13***	17.26***	13.63***	12.05	9.78**	4.90	5.44	4.38
LS-M-F	12.65**	13.25**	10.96*	10.19	8.59	4.64	4.22	3.78
Panel C: Portfolios from linear shrinkage estimator by minimizing negative log-likelihood								
LS-I-L	11.27	12.28	10.50	9.70	8.73	6.63***	6.09***	5.66***
LS-DC-L	10.69	11.92	10.33	9.55	8.54	6.66***	6.10***	5.72***
LS-D-L	11.11	12.23	10.32	9.49	8.44	4.55	4.30	3.85
LS-CC-L	11.46	12.04	10.43	11.86***	8.99*	4.97	5.28*	4.48
LS-M-L	11.17	12.37	10.24	10.14	8.43	4.72	4.30	3.88
LS-1PC-L	11.16	12.28	10.28	9.88	8.34	4.74	4.27	3.90

Notes: Table 3 shows the monthly out-of-sample portfolio variances (displayed in $\%^2$) of different portfolios under various real data sets. The studied portfolios include equally weighted portfolio (EW), and the GMV portfolios based on different covariance matrix estimators, namely, the sample covariance matrix (GMV-S), nonlinear shrinkage (NLS) covariance matrix estimator (Ledoit and Wolf, 2012), analytical nonlinear shrinkage (ANLS) covariance matrix estimator (Ledoit and Wolf, 2020), quadratic inverse shrinkage (QIS) estimator (Ledoit and Wolf, 2022), the Schatten 1-norm (SN_1) regularized covariance matrix (Shi et al., 2020), ridge regularized (RR) correlation matrix estimator (Warton, 2008), and linear shrinkage (LS) covariance matrix estimators with different target matrices (the multiple of the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), and covariance matrix estimated from the one-factor model with the market (M) or the first principal component (1PC) as the factor) by minimizing Frobenius loss (F) or negative log-likelihood (L). Panel B tests the null hypothesis of no difference in the out-of-sample variance between LS-X-F and LS-X-L, where X = {I, DC, D, CC, M}. Panels A and C test the null hypothesis of no difference in the out-of-sample variance between LS-1PC-L and other approaches. These p-values are computed using the stationary bootstrap technique with bootstrap sampling of B = 1000 and block size of b = 5 (Ledoit and Wolf, 2011). "–" indicates that the GMV portfolio cannot be constructed due to the singularity of the sample covariance matrix.

* Denote significance at the 10% level.

** Denote significance at the 5% level.

*** Denote significance at the 1% level.

Notably, the LS-X-L portfolios have larger NSR than the EW and other portfolios for most data sets, especially the last four data sets with a large number of assets. This observation reveals that the GMV portfolio can potentially improve risk- and transaction-cost-adjusted returns, assuming that the covariance matrix estimate is of adequate quality.

4.4. Characteristics of portfolio weights

Table 7 reports the different portfolios' characteristics over rolling windows of length 120 under various data sets. The reported characteristics of portfolio weights include minimum weight (Min), maximum weight (Max), mean absolute deviation (MAD) from the EW portfolio, the Herfindahl–Hirschman Index (HHI) computed by HHI $= \|\hat{\omega}\|_2^2 = \sum_{i=1}^{p} \hat{\omega}_i^2$, the gross exposure (GE) computed by GE $= \|\hat{\omega}\|_1 = \sum_{i=1}^{p} |\hat{\omega}_i|$, and the proportion of leverage (PL) computed by PL $= \frac{1}{p}\sum_{i=1}^{p} \mathbb{1}_{(\hat{\omega}_i<0)}$, where $\mathbb{1}_{(\cdot)}$ denotes the indicator function. The minimum (Min) and maximum (Max) values provide insights into the portfolio's extreme weights. The MAD is particularly useful for determining the gap between the optimized portfolio and the EW portfolio. The HHI measures the portfolio's concentration level and ranges from $1/p$ to 1. A lower HHI value suggests a more diverse portfolio. Finally, the GE calculates the overall proportions of long and short positions in the portfolio. It serves as a metric for the total exposure to financial markets, providing information about the level of risk that investors are taking.

Table 7 offers some insights. First, the GMV-S portfolio has the highest extreme portfolio weights, the highest MAD from the EW portfolio, the highest HHI, and the highest GE, indicating insufficient diversification and substantial market risk. Second, the LS-X-L portfolios have lower extreme portfolio weights, MAD, HHI, GE, and PL than their LS-X-F counterparts in almost all cases. Third, LS-DC-L has the best characteristics among the LS-X-L portfolios, whereas LS-CC-L has the worst characteristics in most cases. LS-1PC-L exhibits better characteristics than LS-M-L across all data sets. Fourth, LS-X-L portfolios show better portfolio weight characteristics than SN_1, RR, and nonlinear shrinkage estimator-based portfolios.

5. Portfolio performance based on simulated data

To assess the robustness of portfolio performance across different values of p (number of assets) and n (sample size), we compare the performance of various portfolios using simulated data. The data follows a multivariate normal distribution, with means and covariance matrices calibrated using the 500CP data set. Specifically, we generate three data sets, each having 2000 samples and 50, 100, and 500 assets. For example, while simulating asset returns for a portfolio of 50 stocks, we calibrate the multivariate normal distribution's means and covariance matrix using data from 50 randomly chosen stocks from the 500CP dataset. The window sizes are 60, 120, 240, and 1800.

Tables 8–11 present the monthly out-of-sample portfolio variance, SR, turnover, and NSR for different portfolios under simulated data with varying p and n. As sample size n increases, we observe that portfolio variance and turnover decrease, whereas SR and NSR increase. This is to be expected given that a larger sample size allows for a more accurate calculation of the covariance matrix required to construct the GMV portfolio. Furthermore, when n is significantly larger than p, all studied portfolios, except for the EW portfolio, exhibit similar performance. This effect is particularly evident in the case where $p = 50$ and $n = 1800$, as shown in Tables 8–11. This resemblance is also expected because, as the sample size approaches infinity, the studied covariance matrix estimators converge to the population covariance matrix. Furthermore, portfolio variance tends to decrease as p increases, which is sensible because a larger number of assets helps diversify

Table 4
Monthly out-of-sample Sharpe ratio based on real data sets.

Method	100IN	100FF	100OP	100SP	200FS	235NY	300CP	500CP
Panel A: Portfolios from existing literature								
GMV-S	0.074***	0.224**	0.210**	0.151**	–	–	–	–
EW	0.228*	0.215***	0.224**	0.292	0.258**	0.262*	0.279**	0.284**
NLS	0.325	0.412	0.350	0.262	0.315**	0.290*	0.310**	0.321**
ANLS	0.317	0.420*	0.348	0.258	0.318**	0.288*	0.313*	0.317**
QIS	0.318	0.419*	0.348	0.258	0.316**	0.290*	0.317*	0.320**
SN_1	0.320	0.409	0.359	0.264	0.320**	0.304	0.314*	0.316**
RR	0.316	0.399	0.349	0.272	0.354	0.357	0.376	0.393
Panel B: Portfolios from linear shrinkage estimator by minimizing Frobenius norm								
LS-I-F	0.289**	0.419	0.365	0.242**	0.311	0.299	0.308	0.301
LS-DC-F	0.281**	0.413	0.365	0.265*	0.315	0.301	0.321	0.315
LS-D-F	0.281**	0.410	0.349	0.245**	0.346	0.349	0.360	0.386
LS-CC-F	0.244**	0.301**	0.296*	0.215	0.315	0.287	0.269	0.286
LS-M-F	0.295**	0.391	0.347	0.261	0.366	0.342	0.366	0.384
Panel C: Portfolios from linear shrinkage estimator by minimizing negative log-likelihood								
LS-I-L	0.327	0.407	0.359	0.281	0.319**	0.304	0.325	0.332*
LS-DC-L	0.344	0.397	0.350	0.290	0.319**	0.307	0.329	0.337
LS-D-L	0.325	0.396	0.349	0.293	0.351	0.356	0.385	0.397
LS-CC-L	0.328	0.397	0.346	0.222***	0.302***	0.281**	0.263**	0.275**
LS-M-L	0.338	0.377	0.340	0.262	0.357	0.333	0.359	0.377
LS-1PC-L	0.340	0.384	0.343	0.268	0.361	0.336	0.365	0.380

Notes: Table 4 reports the monthly out-of-sample Sharpe ratio (SR) of different portfolios under various real data sets. The studied portfolios include equally weighted portfolio (EW) and the GMV portfolios based on different covariance matrix estimators, namely, the sample covariance matrix (GMV-S), nonlinear shrinkage (NLS) covariance matrix estimator (Ledoit and Wolf, 2012), analytical nonlinear shrinkage (ANLS) covariance matrix estimator (Ledoit and Wolf, 2020), quadratic inverse shrinkage (QIS) estimator (Ledoit and Wolf, 2022), Schatten 1-norm (SN_1) regularized covariance matrix (Shi et al., 2020), ridge regularized (RR) correlation matrix estimator (Warton, 2008), and linear shrinkage (LS) covariance matrix estimators with different target matrices (the multiple of the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), and covariance matrix estimated from the one-factor model with the market (M) or the first principal component (1PC) as the factor) by minimizing Frobenius loss (F) or negative log-likelihood (L). Furthermore, Panel B tests the null hypothesis of no difference in the out-of-sample variance between LS-X-F and LS-X-L, where X = {I, DC, D, CC, M}. Meanwhile, Panels A and C test the null hypothesis of no difference in the out-of-sample variance between LS-1PC-L and other approaches. These p-values of testing difference in SR are computed using the stationary bootstrap technique with bootstrap sampling of B = 1000 and block size of b = 5 (Ledoit and Wolf, 2008). "–" indicates that the GMV portfolio cannot be constructed due to the singularity of the sample covariance matrix.

* Indicate significance at the 10% level.

** Indicate significance at the 5% level.

*** Indicate significance at the 1% level.

Table 5
Monthly portfolio turnover based on real data sets.

Method	100IN	100FF	100OP	100SP	200FS	235NY	300CP	500CP
Panel A: Portfolios from existing literature								
GMV-S	7.600	6.992	6.666	2.669	–	–	–	–
EW	0.022	0.024	0.023	0.050	0.038	0.055	0.061	0.061
NLS	0.640	0.828	0.989	0.278	0.528	0.265	0.233	0.245
ANLS	1.064	1.101	1.056	0.451	0.558	0.284	0.246	0.257
QIS	1.011	1.029	1.002	0.427	0.608	0.315	0.268	0.275
SN_1	0.693	0.699	0.672	0.279	0.496	0.305	0.276	0.290
RR	0.710	0.719	0.689	0.288	0.523	0.295	0.268	0.283
Panel B: Portfolios from linear shrinkage estimator by minimizing Frobenius norm								
LS-I-F	0.984	0.918	0.906	0.378	0.718	0.400	0.332	0.333
LS-DC-F	1.061	0.908	0.873	0.268	0.549	0.328	0.271	0.291
LS-D-F	1.042	0.969	0.963	0.427	0.794	0.418	0.349	0.332
LS-CC-F	0.452	0.546	0.486	0.261	0.642	0.262	0.293	0.321
LS-M-F	0.887	0.927	0.849	0.242	0.534	0.228	0.190	0.220
Panel C: Portfolios from linear shrinkage estimator by minimizing negative log-likelihood								
LS-I-L	0.640	0.655	0.631	0.208	0.426	0.233	0.207	0.225
LS-DC-L	0.464	0.510	0.479	0.181	0.378	0.210	0.191	0.205
LS-D-L	0.650	0.655	0.636	0.194	0.433	0.198	0.167	0.196
LS-CC-L	0.534	0.556	0.527	0.227	0.401	0.210	0.191	0.218
LS-M-L	0.503	0.558	0.511	0.207	0.408	0.186	0.164	0.190
LS-1PC-L	0.494	0.545	0.508	0.194	0.396	0.177	0.157	0.183

Notes: Table 5 reports the monthly portfolio turnover of different portfolios under various real data sets. The studied portfolios include equally weighted (EW) portfolio and the GMV portfolios based on different covariance matrix estimators, namely, the sample covariance matrix (GMV-S), nonlinear shrinkage (NLS) covariance matrix estimator (Ledoit and Wolf, 2012), analytical nonlinear shrinkage (ANLS) covariance matrix estimator (Ledoit and Wolf, 2020), quadratic inverse shrinkage (QIS) estimator (Ledoit and Wolf, 2022), the Schatten 1-norm (SN_1) regularized covariance matrix (Shi et al., 2020), ridge regularized (RR) correlation matrix estimator (Warton, 2008), and linear shrinkage (LS) covariance matrix estimators with different target matrices (the multiple of the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), and covariance matrix estimated from the one-factor model with the market (M) or the first principal component (1PC) as the factor) by minimizing the Frobenius loss (F) or negative log-likelihood (L). "–" indicates that the GMV portfolio cannot be constructed due to the singularity of the sample covariance matrix.

individual risk. However, turnover tends to increase with larger p-values. The patterns for the SR and NSR become less evident as p increases.

Panels B and C of Table 8 reveal that the LS-X-L portfolios exhibit lower out-of-sample variances compared to their LS-X-F counterparts when $p = 50$. However, when $p = 500$, the LS-X-L portfolios usually exhibit higher out-of-sample variances than their LS-X-F counterparts. The LS-D-L, LS-M-L, and LS-1PC-L portfolios produce lower out-of-sample variances across varied p and n. Furthermore, these portfolios frequently exhibit lower out-of-sample variances than portfolios based

Table 6
Monthly out-of-sample net Sharpe ratio based on real data sets.

Method	100IN	100FF	100OP	100SP	200FS	235NY	300CP	500CP
Panel A: Portfolios from existing literature								
GMV-S	−0.361***	−0.242***	−0.275***	−0.059***	–	–	–	–
EW	0.226**	0.213*	0.222*	0.286	0.254	0.256	0.272	0.277
NLS	0.226**	0.288	0.193***	0.218	0.225***	0.238**	0.262**	0.268**
ANLS	0.159***	0.261**	0.181***	0.187***	0.223***	0.232**	0.263**	0.262**
QIS	0.167***	0.268**	0.189***	0.191**	0.213***	0.228**	0.262**	0.261**
SN_1	0.215***	0.308	0.253	0.220	0.236**	0.244*	0.258**	0.254**
RR	0.208***	0.294	0.239	0.227	0.264	0.286	0.310	0.319
Panel B: Portfolios from linear shrinkage estimator by minimizing Frobenius norm								
LS-I-F	0.151***	0.287	0.223**	0.186***	0.197***	0.223**	0.242*	0.231*
LS-DC-F	0.135***	0.284*	0.228**	0.222***	0.224**	0.237*	0.265	0.252*
LS-D-F	0.137***	0.272	0.200***	0.182***	0.219***	0.255**	0.278*	0.299**
LS-CC-F	0.187*	0.232**	0.227	0.177	0.210	0.227	0.205	0.208
LS-M-F	0.167***	0.261*	0.216***	0.223	0.272	0.289	0.318	0.326
Panel C: Portfolios from linear shrinkage estimator by minimizing negative log-likelihood								
LS-I-L	0.229**	0.311	0.258	0.248	0.246**	0.258	0.283	0.284
LS-DC-L	0.270	0.320	0.273	0.260	0.253**	0.265	0.290	0.293
LS-D-L	0.224***	0.299	0.247	0.261	0.275	0.309	0.344	0.346
LS-CC-L	0.246	0.315	0.262	0.189***	0.233***	0.234**	0.221***	0.223***
LS-M-L	0.260	0.295	0.257	0.229	0.285	0.289	0.319	0.328
LS-1PC-L	0.264	0.304	0.261	0.237	0.291	0.295	0.327	0.333

Notes: Table 6 shows the monthly out-of-sample net Sharpe ratio of different portfolios under various real data sets. The studied portfolios include equally weighted portfolio (EW) and the GMV portfolios based on different covariance matrix estimators, namely, the sample covariance matrix (GMV-S), nonlinear shrinkage (NLS) covariance matrix estimator (Ledoit and Wolf, 2012), analytical nonlinear shrinkage (ANLS) covariance matrix estimator (Ledoit and Wolf, 2020), quadratic inverse shrinkage (QIS) estimator (Ledoit and Wolf, 2022), Schatten 1-norm (SN_1) regularized covariance matrix (Shi et al., 2020), ridge regularized (RR) correlation matrix estimator (Warton, 2008), and linear shrinkage (LS) covariance matrix estimators with different target matrices (the multiple of the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), and covariance matrix estimated from the one-factor model with the market (M) or the first principal component (1PC) as the factor) by minimizing Frobenius loss (F) or negative log-likelihood (L). Panel B tests the null hypothesis that there is no difference in the out-of-sample variance between LS-X-F and LS-X-L, where X = {I, DC, D, CC, M}. Panels A and C test the null hypothesis that LS-1PC-L and other techniques have the same out-of-sample variance. The p-values for testing difference in Sharpe ratios are calculated using the stationary bootstrap technique, with bootstrap sampling of B = 1000 and block size of b = 5 (Ledoit and Wolf, 2008). "–" indicates that the GMV portfolio cannot be constructed due to the singularity of the sample covariance matrix.

* Indicate significance at the 10% level.

** Indicate significance at the 5% level.

*** Indicate significance at the 1% level.

on nonlinear shrinkage estimators. Table 10 shows that LS-X-L portfolios had lower turnover than their LS-X-F counterparts and other portfolios, except for the EW portfolio. Table 11 shows that the LS-D-L, LS-M-L, and LS-1PC-L portfolios have greater NSR than other portfolios.

6. Conclusion

In practice, the covariance matrix estimate significantly impacts GMV portfolio performance. The sample covariance matrix often entails substantial estimation error, resulting in suboptimal portfolios with elevated risk and turnover, especially when there are a large number of assets. Due to its simplicity and efficiency, the LS of sample covariance matrix has been widely recommended as an alternative to improve the GMV portfolio performance. The shrinkage intensity strikes a balance between bias and variance, which is critical for LS covariance matrix estimators. This study discusses the selection of shrinkage intensity based on minimizing the cross-validated negative log-likelihood function. Furthermore, theoretical insights into the shrinkage intensity are provided.

Given the same target matrix, we show that the shrinkage intensity selected based on minimizing the negative log-likelihood is always greater than that based on minimizing the Frobenius loss. Consequently, the LS estimator with intensity chosen using the former criterion tends to exhibit better stability (i.e., smaller condition number) than the same estimator with intensity chosen using the latter criterion. Within the family of LS estimators, the target matrix based on the diagonal matrix or the matrix based on the one-factor model outperforms other target matrices.

We also show that GMV portfolios based on LS estimators with shrinkage intensity selected based on the suggested criterion outperform those based on NLS estimators in terms of reduced variance, turnover, and enhanced SR. Moreover, the portfolio performance improvement is statistically significant, especially when there are a large number of assets.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the editor, the associate editor, and two anonymous reviewers for their insightful comments that help us improve this paper.

Appendix A. Proof of Propositions 2 and 3

Let $S = (s_{ij})_{p\times p}$ denote the sample covariance matrix. Since S is semi-positive definite, we have $\mathbf{1}'S\mathbf{1} \geq 0$, which implies $\sum_{i=1}^{p}\sum_{i=1}^{p} s_{ij} \geq 0$. Thus, we have $\sum_{i\neq j} s_{ij} \geq -\sum_{i=1}^{p} s_{ii}$ or

$$\bar{q} = \frac{1}{p^2-p}\sum_{i\neq j} s_{ij} \geq -\frac{1}{p(p-1)}\sum_{i=1}^{p} s_{ii} = -\frac{1}{p-1}\bar{s} \tag{25}$$

The matrix $\hat{\Sigma}_{\mathrm{DC}}$ can be expressed as

$$\hat{\Sigma}_{\mathrm{DC}} = \bar{q}\mathbf{1}\mathbf{1}' + \mathrm{Diag}(\bar{s}-\bar{q},\ldots,\bar{s}-\bar{q}). \tag{26}$$

The ith ($i = 1,\ldots,p$) principal minor of $\hat{\Sigma}_{\mathrm{DC}}$ is

$$|D_i| = (\bar{s}-\bar{q})^i + i\bar{q}(1-\bar{q})^{i-1} \tag{27}$$

$$= (\bar{s}-\bar{q})^{i-1}(\bar{s}+(i-1)\bar{q}) \tag{28}$$

$$\geq (\bar{s}-\bar{q})^{i-1}\left(\bar{s}-\frac{i-1}{p-1}\bar{s}\right) \tag{29}$$

The first $p-1$ principal minors are positive, that is $|D_i| > 0$ for $i = 1,\ldots,p-1$. $|D_p| \geq 0$ and equality holds when $\sum_{i=1}^{p}\sum_{i=1}^{p} s_{ij} = 0$, which has a probability of zero. Therefore, we have proved that the matrix $\hat{\Sigma}_{\mathrm{DC}}$ is almost surely positive definite.

Table 7
Characteristics of different portfolios for various data sets.

Method	Min	Max	MAD	HHI	GE	PL	Min	Max	MAD	HHI	GE	PL
	Panel A: 100FF						*Panel B: 100SP*					
GMV-S	−0.8083	0.8659	0.2541	10.5843	25.4108	0.4951	−0.3276	0.3636	0.0903	1.4569	9.0506	0.4701
EW	0.0100	0.0100	0.0000	0.0100	1.0000	0.0000	0.0100	0.0100	0.0000	0.0100	1.0000	0.0000
NLS	−0.1914	0.2337	0.0635	0.6648	6.3775	0.4578	−0.0529	0.0822	0.0230	0.0929	2.3953	0.3853
ANLS	−0.1964	0.2384	0.0648	0.6907	6.5092	0.4585	−0.0564	0.0869	0.0243	0.1023	2.5163	0.3903
QIS	−0.1899	0.2309	0.0626	0.6460	6.2949	0.4571	−0.0549	0.0852	0.0236	0.0973	2.4522	0.3875
SN_1	−0.1811	0.2283	0.0628	0.6480	6.3237	0.4523	−0.0547	0.0852	0.0237	0.0988	2.4715	0.3853
RR	−0.1506	0.2836	0.0637	0.6922	6.3382	0.4870	−0.0589	0.1170	0.0250	0.1178	2.5106	0.4241
LS-I-F	−0.2124	0.2652	0.0744	0.9003	7.4651	0.4651	−0.0705	0.1036	0.0290	0.1444	2.9769	0.4089
LS-DC-F	−0.2140	0.2639	0.0744	0.9069	7.4680	0.4620	−0.0523	0.0843	0.0230	0.0947	2.3937	0.3810
LS-D-F	−0.1991	0.3228	0.0776	1.0092	7.7396	0.4907	−0.0917	0.1420	0.0323	0.1898	3.2439	0.4399
LS-CC-F	−0.1197	0.2673	0.0558	0.5387	5.5223	0.4971	−0.0464	0.1262	0.0256	0.1239	2.5232	0.4469
LS-M-F	−0.2092	0.2976	0.0769	0.9993	7.6665	0.4870	−0.0534	0.1013	0.0230	0.1013	2.3534	0.3978
LS-I-L	−0.1725	0.2180	0.0597	0.5877	6.0152	0.4498	−0.0428	0.0694	0.0190	0.0674	2.0314	0.3510
LS-DC-L	−0.1495	0.1883	0.0504	0.4206	5.0828	0.4444	−0.0383	0.0626	0.0170	0.0562	1.8569	0.3298
LS-D-L	−0.1398	0.2725	0.0602	0.6222	5.9803	0.4873	−0.0338	0.0947	0.0188	0.0731	1.9039	0.3850
LS-CC-L	−0.1313	0.2723	0.0580	0.5735	5.7246	0.5097	−0.0418	0.1213	0.0243	0.1128	2.3962	0.4423
LS-M-L	−0.1513	0.2300	0.0574	0.5578	5.6919	0.4943	−0.0468	0.0957	0.0212	0.0881	2.1621	0.3903
LS-1PC-L	−0.1508	0.2253	0.0562	0.5415	5.5788	0.4869	−0.0448	0.0906	0.0200	0.0803	2.0642	0.3753
	Panel C: 235NY						*Panel D: 500CP*					
GMV-S	–	–	–	–	–	–	–	–	–	–	–	–
EW	0.0043	0.0043	0.0000	0.0043	1.0000	0.0000	0.0033	0.0033	0.0000	0.0033	1.0000	0.0000
NLS	−0.0299	0.0435	0.0096	0.0407	2.3848	0.3667	−0.0216	0.0293	0.0067	0.0251	2.1673	0.3452
ANLS	−0.0307	0.0452	0.0099	0.0428	2.4475	0.3720	−0.0221	0.0300	0.0068	0.0261	2.2085	0.3485
QIS	−0.0304	0.0451	0.0099	0.0426	2.4413	0.3717	−0.0219	0.0298	0.0068	0.0258	2.1966	0.3479
SN_1	−0.0322	0.0486	0.0105	0.0474	2.5677	0.3810	−0.0239	0.0329	0.0073	0.0296	2.3362	0.3585
RR	−0.0298	0.1488	0.0114	0.0893	2.5861	0.4651	−0.0212	0.1185	0.0082	0.0633	2.3585	0.4589
LS-I-F	−0.0384	0.0578	0.0125	0.0665	3.0291	0.4024	−0.0267	0.0367	0.0081	0.0365	2.5637	0.3704
LS-DC-F	−0.0336	0.0509	0.0110	0.0521	2.6751	0.3860	−0.0232	0.0320	0.0071	0.0284	2.2799	0.3495
LS-D-F	−0.0418	0.1647	0.0139	0.1203	3.1991	0.4685	−0.0292	0.1245	0.0096	0.0786	2.7745	0.4705
LS-CC-F	−0.0257	0.1723	0.0122	0.1141	2.6814	0.5171	−0.0212	0.1517	0.0099	0.0954	2.7795	0.5287
LS-M-F	−0.0247	0.1235	0.0100	0.0708	2.2677	0.4392	−0.0170	0.1016	0.0070	0.0475	2.0046	0.4222
LS-I-L	−0.0262	0.0400	0.0087	0.0339	2.1753	0.3484	−0.0190	0.0264	0.0060	0.0207	1.9748	0.3272
LS-DC-L	−0.0241	0.0367	0.0080	0.0295	2.0351	0.3346	−0.0178	0.0247	0.0056	0.0187	1.8828	0.3179
LS-D-L	−0.0194	0.1284	0.0090	0.0643	2.0102	0.4439	−0.0112	0.1015	0.0062	0.0421	1.7621	0.4193
LS-CC-L	−0.0206	0.1665	0.0113	0.1059	2.4349	0.5241	−0.0114	0.1445	0.0083	0.0766	2.2374	0.5381
LS-M-L	−0.0207	0.1158	0.0091	0.0622	2.0380	0.4319	−0.0146	0.0983	0.0066	0.0438	1.8818	0.4164
LS-1PC-L	−0.0183	0.1113	0.0087	0.0570	1.9590	0.4223	−0.0127	0.0947	0.0063	0.0405	1.8082	0.4079

Notes: Table 7 shows the characteristics of several portfolios over rolling windows of length 120 under different data sets. Portfolio weights' reported characteristics include minimum weight (Min), maximum weight (Max), mean absolute deviation (MAD) from the equally weighted (EW) portfolio, Herfindahl–Hirschman Index (HHI), gross exposure (GE), and leverage proportion (PL). The studied portfolios include EW portfolio and the GMV portfolio based on different covariance matrix estimators, namely, the sample covariance matrix (GMV-S), nonlinear shrinkage (NLS) covariance matrix estimator (Ledoit and Wolf, 2012), analytical nonlinear shrinkage (ANLS) covariance matrix estimator (Ledoit and Wolf, 2020), quadratic inverse shrinkage (QIS) estimator (Ledoit and Wolf, 2022), Schatten 1-norm (SN_1) regularized covariance matrix (Shi et al., 2020), ridge regularized (RR) correlation matrix estimator (Warton, 2008), and linear shrinkage (LS) covariance matrix estimators with different target matrices (the multiple of the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), and covariance matrix estimated from the one-factor model with the market (M) or the first principal component (1PC) as the factor) by minimizing the Frobenius loss (F) or negative log-likelihood (L).

Similarly, we can prove that $\hat{R}_{\text{CC}}$ is almost surely positive definite. Let $C = (c_{ij})_{p\times p}$ denote the sample correlation matrix. Since C is semi-positive definite, we have $\mathbf{1}'C\mathbf{1} \geq 0$, which implies $\sum_{i=1}^{p}\sum_{i=1}^{p} c_{ij} \geq 0$, thus $\sum_{i\neq j} c_{ij} + p \geq 0$ or

$$\bar{\rho} = \frac{1}{p^2 - p}\sum_{i\neq j} c_{ij} \geq -\frac{1}{p-1}. \tag{30}$$

The matrix $\hat{R}_{\text{CC}}$ can be expressed as

$$\hat{R}_{\text{CC}} = \bar{\rho}\mathbf{1}\mathbf{1}' + \text{Diag}(1-\bar{\rho}, \dots, 1-\bar{\rho}). \tag{31}$$

The ith ($i = 1, \dots, p$) principal minor of $\hat{R}_{\text{CC}}$ is

$$|D_i| = (1-\bar{\rho})^i + i\bar{\rho}(1-\bar{\rho})^{i-1} \tag{32}$$

$$= (1-\bar{\rho})^{i-1}(1 + (i-1)\bar{\rho}) \tag{33}$$

$$\geq (1-\bar{\rho})^{i-1}\left(1 - \frac{i-1}{p-1}\right) \tag{34}$$

The first $p-1$ principal minors are positive, that is $|D_i| > 0$ for $i = 1, \dots, p-1$. $|D_p| \geq 0$ and equality holds when $\sum_{i=1}^{p}\sum_{i=1}^{p} c_{ij} = 0$, which has a probability of zero. Therefore, we have proved that the matrix $\hat{R}_{\text{CC}}$ is almost surely positive definite.

Appendix B. Proof of Theorem 1

The matrices $S^{[k]}$ and $S^{[-k]}$ represent the sample covariance matrices computed using the kth and remaining $K-1$ data groups , respectively. The actual covariance matrix is $\Sigma = PHP'$, where H is a diagonal matrix with eigenvalues on the diagonal and P is the matrix whose columns are the corresponding eigenvectors.

When the target matrix is a rotation-equivariant matrix $\hat{\Sigma}_{\text{RE}}$, the shrinkage estimator based on the training data is expressed as $\hat{\Sigma}^{[-k]}_{\text{LS-RE}} = (1-\nu)S^{[-k]} + \nu\hat{\Sigma}^{[-k]}_{\text{RE}}$. $\hat{\Sigma}^{[-k]}_{\text{RE}}$ has the following eigendecomposition $\hat{\Sigma}^{[-k]}_{\text{RE}} = U^{[-k]}\mathrm{T}\left(U^{[-k]}\right)'$, where the columns of $U^{[-k]}$ are the eigenvectors of $S^{[-k]}$ and $\mathrm{T} = \text{Diag}(\tau_1, \dots, \tau_p)$. The K-fold cross-validated negative log-likelihood can be expressed as

$$\sum_{k=1}^{K} -\log L\left(\hat{\Sigma}^{[-k]}_{\text{LS-I}}, S^{[k]}\right) \tag{35}$$

$$= \sum_{k=1}^{K}\left[\text{tr}\left(\left(\hat{\Sigma}^{[-k]}_{\text{LS-I}}\right)^{-1} S^{[k]}\right) + \log\left|\hat{\Sigma}^{[-k]}_{\text{LS-I}}\right|\right] \tag{36}$$

$$= \sum_{k=1}^{K}\left[\text{tr}\left(\left((1-\nu)S^{[-k]} + \nu\hat{\Sigma}^{[-k]}_{\text{RE}}\right)^{-1} S^{[k]}\right) + \log\left|(1-\nu)S^{[-k]} + \nu\hat{\Sigma}^{[-k]}_{\text{RE}}\right|\right]. \tag{37}$$

Table 8
Monthly out-of-sample portfolio variance based on simulated data.

Method	$p = 50$				$p = 100$				$p = 500$			
	$n = 60$	$n = 120$	$n = 240$	$n = 1800$	$n = 60$	$n = 120$	$n = 240$	$n = 1800$	$n = 60$	$n = 120$	$n = 240$	$n = 1800$
Panel A: Portfolios from existing literature												
GMV-S	55.00***	14.44***	10.65***	7.72	–	28.99***	8.02***	6.21	–	–	–	0.15***
EW	23.04***	23.07***	22.89***	22.52***	21.41***	21.49***	21.53***	21.37***	20.68***	20.79***	20.73***	24.83***
NLS	11.87***	10.82***	9.79	7.69	10.18***	8.68***	6.90***	6.12	7.18***	5.14***	3.61***	0.49**
ANLS	12.25***	10.83***	9.80	7.69	10.21***	8.65***	6.84***	6.14	7.17***	5.09***	3.56***	32.19***
QIS	12.31***	10.84***	9.79	7.70	10.25***	8.63***	6.78***	6.15	7.24***	5.13***	3.57***	0.15***
SN_1	12.48***	11.10***	9.90***	7.69	10.55***	8.88***	7.12***	6.16	7.23***	5.12***	3.59***	0.27***
RR	11.57***	10.68***	9.76	7.70	8.11***	7.16**	6.22	6.12	5.28***	3.89***	2.80***	0.35***
Panel B: Portfolios from linear shrinkage estimator by minimizing Frobenius norm												
LS-I-F	12.22***	11.08***	9.92***	7.69	10.76**	9.21***	7.07*	6.17	7.23***	5.17	3.60	0.48***
LS-DC-F	12.09***	10.98***	9.87**	7.69	10.40***	8.83	7.14***	6.16	7.25***	5.13**	3.57***	0.60***
LS-D-F	11.44***	10.77***	9.85***	7.71	8.54***	7.95***	6.68***	6.18	5.27***	3.91**	2.84	0.39***
LS-CC-F	12.66**	11.30***	10.07***	7.66	8.25*	7.35***	6.38***	6.21	5.81***	4.38***	3.09***	0.46***
LS-M-F	10.93*	10.37	9.70	7.66	7.68*	6.94	6.20	6.12	5.28***	3.96	2.89	0.78***
Panel C: Portfolios from linear shrinkage estimator by minimizing negative log-likelihood												
LS-I-L	11.95***	10.85***	9.81*	7.69	10.44***	8.82***	7.24***	6.16	7.38***	5.20***	3.67***	0.47***
LS-DC-L	11.94***	10.84***	9.81*	7.68	10.58***	8.86***	7.27***	6.16	7.46***	5.23***	3.69***	0.44***
LS-D-L	11.10	10.39	9.63	7.70	8.17***	7.11**	6.21	6.10	5.48*	4.00	2.90	0.35***
LS-CC-L	12.87***	11.75***	10.48***	7.65	8.34***	7.49***	6.52***	6.24*	6.16***	4.64***	3.33***	0.33
LS-M-L	10.85	10.40	9.72*	7.66	7.63	6.96	6.22	6.12	5.34*	4.01	2.90	0.33
LS-1PC-L	10.94	10.32	9.63	7.64	7.67	6.97	6.22	6.12	5.40	4.04	2.92	0.33

Notes: Table 8 shows the out-of-sample portfolio variances (displayed in $\%^2$) of different portfolios under simulated data under various p and n. The studied portfolios include equally weighted portfolio (EW) and the GMV portfolios based on different covariance matrix estimators, namely, the sample covariance matrix (GMV-S), nonlinear shrinkage (NLS) covariance matrix estimator (Ledoit and Wolf, 2012), analytical nonlinear shrinkage (ANLS) covariance matrix estimator (Ledoit and Wolf, 2020), quadratic inverse shrinkage (QIS) estimator (Ledoit and Wolf, 2022), Schatten 1-norm (SN_1) regularized covariance matrix (Shi et al., 2020), ridge regularized (RR) correlation matrix estimator (Warton, 2008), and linear shrinkage (LS) covariance matrix estimators with different target matrices (the multiple of the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), and covariance matrix estimated from the one-factor model with the market (M) or the first principal component (1PC) as the factor) by minimizing Frobenius loss (F) or negative log-likelihood (L). Panel B tests the null hypothesis that there is no change in the out-of-sample variance between LS-X-F and LS-X-L, where X = {I, DC, D, CC, M}. In Panels A and C, we test the null hypothesis that LS-1PC-L and other techniques have the same out-of-sample variance. These p-values for testing difference in variance are calculated using the stationary bootstrap technique with bootstrap sampling of B = 1000 and block size of b = 5 (Ledoit and Wolf, 2011). "–" indicates that the GMV portfolio cannot be constructed due to the singularity of the sample covariance matrix.

* Indicate significance at the 10% level, in the testing result.

** Indicate significance at the 10% level, in the testing result.

*** Indicate significance at the 1% level, in the testing result.

As $n_k \xrightarrow{p} \infty$, we have $S^{[k]} \xrightarrow{p} \Sigma$, $S^{[-k]} \xrightarrow{p} \Sigma$, and $\hat{\Sigma}_{\text{RE}}^{[-k]} \xrightarrow{p} P\mathrm{T}P'$. It follows that

$$\sum_{k=1}^{K} -\log L\left(\hat{\Sigma}_{\text{LS-I}}^{[-k]}, S^{[k]}\right) \tag{38}$$

$$\xrightarrow{p} \sum_{k=1}^{K}\left[\operatorname{tr}\left(\left((1-\nu)\Sigma+\nu P\mathrm{T}P'\right)^{-1}\Sigma\right)+\log\left|(1-\nu)\Sigma+\nu P\mathrm{T}P'\right|\right] \tag{39}$$

$$= \sum_{k=1}^{K}\left[\operatorname{tr}\left(P\left((1-\nu)H+\nu\mathrm{T}\right)^{-1}P'\Sigma\right)+\log\left|P((1-\nu)H+\nu\mathrm{T})P'\right|\right] \tag{40}$$

$$= \sum_{k=1}^{K}\left[\operatorname{tr}\left(((1-\nu)H+\nu\mathrm{T})^{-1}P'\Sigma P\right)+\log\left|((1-\nu)H+\nu\mathrm{T})P'P\right|\right] \tag{41}$$

$$= \sum_{k=1}^{K}\left[\operatorname{tr}\left(((1-\nu)H+\nu\mathrm{T})^{-1}H\right)+\log\left|((1-\nu)H+\nu\mathrm{T})\right|\right] \tag{42}$$

$$= K\left[\sum_{i=1}^{p}\left(\frac{\eta_i}{(1-\nu)\eta_i+\nu\tau_i}+\log((1-\nu)\eta_i+\nu\tau_i)\right)\right], \tag{43}$$

where η_i is the ith diagonal element of H. Differentiating $\sum_{k=1}^{K} -\log L\left(\hat{\Sigma}_{\text{LS-I}}^{[-k]}, S^{[k]}\right)$ with respect to ν leads to

$$\frac{\partial\left(\sum_{k=1}^{K} -\log L\left(\hat{\Sigma}_{\text{LS-I}}^{[-k]}, S^{[k]}\right)\right)}{\partial\nu}$$

$$\xrightarrow{p} K\sum_{i=1}^{p}\frac{\partial\left(\frac{\eta_i}{(1-\nu)\eta_i+\nu\tau_i}+\log((1-\nu)\eta_i+\nu\tau_i)\right)}{\partial\nu} \tag{44}$$

$$= K\sum_{i=1}^{p}\left(\frac{-\eta_i(1-\eta_i)}{((1-\nu)\eta_i+\nu\tau_i)^2}+\frac{1-\eta_i}{(1-\nu)\eta_i+\nu\tau_i}\right) \tag{45}$$

$$= K\nu\sum_{i=1}^{p}\frac{(\tau_i-\eta_i)^2}{((1-\nu)\eta_i+\nu\tau_i)^2}, \tag{46}$$

which is positive for $\nu > 0$, $K > 0$ and $\eta_i \neq \tau_i\ \exists i$. If the true covariance matrix Σ has a full rank ($\eta_i > 0\ \forall i$) and its eigenvalues are not all the same as those of $\hat{\Sigma}_{\text{RE}}$ ($\eta_i \neq \tau_i\ \exists i$), then the negative log-likelihood function in Eq. (44) has a unique minimum at $\nu^* = 0$.

When the target matrix is the diagonal matrix $\hat{\Sigma}_{\text{D}}$, the shrinkage estimator based on the training data is expressed as

$$\hat{\Sigma}_{\text{LS-D}}^{[-k]} = (1-\nu)S^{[-k]}+\nu\hat{\Sigma}_{\text{D}} = \left(\hat{\Sigma}_{\text{D}}^{[-k]}\right)^{1/2}\left((1-\nu)C^{[-k]}+\nu I\right)\left(\hat{\Sigma}_{\text{D}}^{[-k]}\right)^{1/2}, \tag{47}$$

where $C^{[-k]}$ is sample correlation matrix of the $K-1$ groups of data. The K-fold cross-validated negative log-likelihood can be expressed as

$$\sum_{k=1}^{K} -\log L\left(\hat{\Sigma}_{\text{LS-D}}^{[-k]}, S^{[k]}\right) \tag{48}$$

$$= \sum_{k=1}^{K}\left[\operatorname{tr}\left(\left(\hat{\Sigma}_{\text{LS-D}}^{[-k]}\right)^{-1}S^{[k]}\right)+\log\left|\hat{\Sigma}_{\text{LS-D}}^{[-k]}\right|\right] \tag{49}$$

$$= \sum_{k=1}^{K}\left[\operatorname{tr}\left(\left(\hat{\Sigma}_{\text{D}}^{[-k]}\right)^{-1/2}\left((1-\nu)C^{[-k]}+\nu I\right)^{-1}\left(\hat{\Sigma}_{\text{D}}^{[-k]}\right)^{-1/2}S^{[k]}\right)\right. \tag{50}$$

$$\left.+\log\left|\left(\hat{\Sigma}_{\text{D}}^{[-k]}\right)^{1/2}\left((1-\nu)C^{[-k]}+\nu I\right)\left(\hat{\Sigma}_{\text{D}}^{[-k]}\right)^{1/2}\right|\right] \tag{51}$$

$$= \sum_{k=1}^{K}\left[\operatorname{tr}\left(\left((1-\nu)C^{[-k]}+\nu I\right)^{-1}\left(\hat{\Sigma}_{\text{D}}^{[-k]}\right)^{-1/2}S^{[k]}\left(\hat{\Sigma}_{\text{D}}^{[-k]}\right)^{-1/2}\right)\right. \tag{52}$$

$$\left.+\log\left|\left((1-\nu)C^{[-k]}+\nu I\right)\right|+\log\left|\hat{\Sigma}_{\text{D}}^{[-k]}\right|\right] \tag{53}$$

Let Σ_{D} denote a diagonal matrix whose diagonal elements are same as the corresponding elements in the true covariance matrix Σ. Let $R = \Sigma_{\text{D}}^{-1/2}\Sigma\Sigma_{\text{D}}^{-1/2}$ denote the true correlation matrix, which has the following eigendecomposition $R = Q\Gamma Q'$, where Γ is a diagonal matrix whose diagonal entries are the eigenvalues and Q is the matrix whose

Table 9
Monthly out-of-sample Sharpe ratio based on simulated data.

Method	$p = 50$				$p = 100$				$p = 500$			
	$n = 60$	$n = 120$	$n = 240$	$n = 1800$	$n = 60$	$n = 120$	$n = 240$	$n = 1800$	$n = 60$	$n = 120$	$n = 240$	$n = 1800$
Panel A: Portfolios from existing literature												
GMV-S	0.191***	0.321***	0.399***	0.504	–	0.177***	0.324***	0.334***	–	–	–	1.342***
EW	0.291***	0.291***	0.295***	0.397	0.288***	0.287***	0.276***	0.168**	0.289***	0.289***	0.288***	0.277***
NLS	0.382**	0.391*	0.418	0.505	0.386**	0.404**	0.416	0.342**	0.422*	0.476	0.531	0.615***
ANLS	0.376**	0.391*	0.417	0.505	0.384***	0.403**	0.416	0.346	0.421*	0.477	0.538	0.248***
QIS	0.377**	0.390**	0.418	0.505	0.384***	0.403**	0.415	0.347	0.422*	0.476	0.536	1.395
SN_1	0.375**	0.386**	0.416	0.506	0.375***	0.397**	0.415	0.355	0.417**	0.471	0.530	1.450
RR	0.392	0.397	0.422	0.504	0.395**	0.409*	0.410*	0.347*	0.439	0.485	0.548**	1.388***
Panel B: Portfolios from linear shrinkage estimator by minimizing Frobenius norm												
LS-I-F	0.378	0.386*	0.416	0.506	0.370**	0.383***	0.400***	0.346**	0.416	0.463**	0.518	1.222***
LS-DC-F	0.380	0.389	0.417	0.506	0.383	0.403	0.417	0.358	0.417	0.469	0.526	1.132***
LS-D-F	0.394	0.394*	0.419*	0.504	0.380***	0.377***	0.377***	0.338**	0.438	0.477*	0.537	1.346***
LS-CC-F	0.360	0.376	0.410	0.498	0.344	0.365	0.374	0.342	0.337	0.372	0.429**	1.166***
LS-M-F	0.400	0.406	0.425	0.500	0.406	0.419	0.421	0.356	0.440	0.480	0.534	1.022***
Panel C: Portfolios from linear shrinkage estimator by minimizing negative log-likelihood												
LS-I-L	0.382**	0.393*	0.418	0.507	0.384**	0.407	0.421	0.355	0.418**	0.476	0.530	1.236***
LS-DC-L	0.382**	0.393*	0.419	0.507	0.384**	0.407	0.422	0.359	0.418**	0.476	0.530	1.263***
LS-D-L	0.402	0.409	0.429	0.504	0.406	0.423	0.425	0.353	0.440	0.487*	0.544**	1.388***
LS-CC-L	0.360***	0.371***	0.402***	0.491**	0.343***	0.360***	0.372***	0.348	0.332***	0.364***	0.411***	1.359***
LS-M-L	0.404	0.407	0.425	0.499	0.410	0.420	0.424	0.357	0.439	0.479	0.533	1.424
LS-1PC-L	0.399	0.404	0.424	0.502	0.410	0.421	0.425	0.358	0.440	0.481	0.534	1.423

Notes: Table 9 reports the out-of-sample Sharpe ratio of different portfolios under simulated data under various p and n. The studied portfolios include equally weighted portfolio (EW) and the GMV portfolios based on different covariance matrix estimators, namely, the sample covariance matrix (GMV-S), nonlinear shrinkage (NLS) covariance matrix estimator (Ledoit and Wolf, 2012), analytical nonlinear shrinkage (ANLS) covariance matrix estimator (Ledoit and Wolf, 2020), quadratic inverse shrinkage (QIS) estimator (Ledoit and Wolf, 2022), Schatten 1-norm (SN_1) regularized covariance matrix (Shi et al., 2020), ridge regularized (RR) correlation matrix estimator (Warton, 2008), and linear shrinkage (LS) covariance matrix estimators with different target matrices (the multiple of the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), and covariance matrix estimated from the one-factor model with the market (M) or the first principal component (1PC) as the factor) by minimizing Frobenius loss (F) or negative log-likelihood (L). Panel B tests the null hypothesis that there is no difference in the out-of-sample variance between LS-X-F and LS-X-L, where X = {I, DC, D, CC, M}. Panels A and C test the null hypothesis that LS-1PC-L and other techniques have the same out-of-sample variance. The stationary bootstrap technique is used to obtain these p-values for evaluating differences in Sharpe ratios, with bootstrap sampling of B = 1000 and block size of b = 5 (Ledoit and Wolf, 2008). "–" indicates that the GMV portfolio cannot be constructed due to the singularity of the sample covariance matrix.

* Indicate significance at the 10% level, in the testing result.

** Indicate significance at the 5% level, in the testing result.

*** Indicate significance at the 1% level, in the testing result.

Table 10
Monthly portfolio turnover based on simulated data.

Method	$p = 50$				$p = 100$				$p = 500$			
	$n = 60$	$n = 120$	$n = 240$	$n = 1800$	$n = 60$	$n = 120$	$n = 240$	$n = 1800$	$n = 60$	$n = 120$	$n = 240$	$n = 1800$
Panel A: Portfolios from existing literature												
GMV-S	2.065	0.355	0.183	0.110	–	1.923	0.362	0.155	–	–	–	2.192
EW	0.078	0.078	0.078	0.077	0.067	0.067	0.067	0.068	0.070	0.070	0.070	0.070
NLS	0.224	0.174	0.141	0.107	0.247	0.266	0.369	0.183	0.293	0.292	0.343	32.815
ANLS	0.414	0.170	0.134	0.107	0.295	0.411	0.198	0.148	0.315	0.311	0.356	0.093
QIS	0.394	0.178	0.137	0.107	0.328	0.389	0.204	0.148	0.332	0.326	0.373	2.105
SN_1	0.247	0.180	0.139	0.107	0.333	0.254	0.196	0.146	0.362	0.348	0.336	0.768
RR	0.228	0.166	0.132	0.106	0.296	0.225	0.178	0.143	0.359	0.329	0.309	0.547
Panel B: Portfolios from linear shrinkage estimator by minimizing Frobenius norm												
LS-I-F	0.221	0.177	0.140	0.108	0.362	0.312	0.230	0.150	0.374	0.405	0.448	0.494
LS-DC-F	0.205	0.167	0.136	0.107	0.258	0.225	0.188	0.145	0.351	0.367	0.386	0.444
LS-D-F	0.214	0.173	0.139	0.107	0.368	0.328	0.244	0.151	0.375	0.390	0.432	0.512
LS-CC-F	0.204	0.158	0.130	0.106	0.256	0.212	0.181	0.147	0.410	0.396	0.401	0.464
LS-M-F	0.173	0.135	0.118	0.105	0.227	0.178	0.153	0.139	0.312	0.278	0.264	0.346
Panel C: Portfolios from linear shrinkage estimator by minimizing negative log-likelihood												
LS-I-L	0.181	0.152	0.130	0.110	0.222	0.196	0.173	0.157	0.303	0.293	0.300	0.504
LS-DC-L	0.176	0.149	0.128	0.106	0.203	0.185	0.169	0.144	0.286	0.282	0.293	0.522
LS-D-L	0.136	0.113	0.105	0.102	0.182	0.152	0.138	0.137	0.291	0.265	0.263	0.547
LS-CC-L	0.167	0.132	0.114	0.103	0.213	0.173	0.153	0.143	0.329	0.297	0.293	0.589
LS-M-L	0.163	0.131	0.116	0.105	0.209	0.167	0.146	0.138	0.287	0.263	0.261	0.589
LS-1PC-L	0.148	0.119	0.109	0.103	0.199	0.161	0.143	0.137	0.280	0.258	0.259	0.588

Notes: Table 10 shows the portfolio turnover of various portfolios using simulated data with varying p and n. The studied portfolios include equally weighted portfolio (EW), and the GMV portfolios based on different covariance matrix estimators, namely, the sample covariance matrix (GMV-S), nonlinear shrinkage (NLS) covariance matrix estimator (Ledoit and Wolf, 2012), analytical nonlinear shrinkage (ANLS) covariance matrix estimator (Ledoit and Wolf, 2020), quadratic inverse shrinkage (QIS) estimator (Ledoit and Wolf, 2022), the Schatten 1-norm (SN_1) regularized covariance matrix (Shi et al., 2020), ridge regularized (RR) correlation matrix estimator (Warton, 2008), and linear shrinkage (LS) covariance matrix estimators with different target matrices (the multiple of the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), and covariance matrix estimated from the one-factor model with the market (M) or the first principal component (1PC) as the factor) by minimizing Frobenius loss (F) or negative log-likelihood (L). "–" indicates that the GMV portfolio cannot be constructed because the sample covariance matrix is singular.

Table 11
Monthly out-of-sample net Sharpe ratio based on simulated data.

Method	$p = 50$				$p = 100$				$p = 500$			
	$n = 60$	$n = 120$	$n = 240$	$n = 1800$	$n = 60$	$n = 120$	$n = 240$	$n = 1800$	$n = 60$	$n = 120$	$n = 240$	$n = 1800$
Panel A: Portfolios from existing literature												
GMV-S	0.051***	0.274***	0.372***	0.484	–	−0.003***	0.260***	0.303***	–	–	–	−1.291***
EW	0.283***	0.283***	0.287***	0.389	0.280***	0.280***	0.269***	0.160**	0.281***	0.281***	0.280***	0.270***
NLS	0.350***	0.365***	0.396**	0.486	0.348***	0.359***	0.345***	0.305***	0.369	0.413	0.442*	−1.253***
ANLS	0.317***	0.366***	0.396**	0.486	0.338***	0.333***	0.379***	0.316*	0.364*	0.409	0.444	0.240***
QIS	0.321***	0.364***	0.396**	0.486	0.334***	0.337***	0.376***	0.317*	0.361**	0.405	0.439**	−1.173***
SN_1	0.341***	0.359***	0.394**	0.487	0.325***	0.355***	0.379**	0.326	0.351***	0.395**	0.442*	0.703***
RR	0.360**	0.372**	0.401	0.485	0.344***	0.368***	0.375***	0.318*	0.363*	0.402**	0.457	0.918**
Panel B: Portfolios from linear shrinkage estimator by minimizing Frobenius norm												
LS-I-F	0.347*	0.360**	0.394	0.487	0.316***	0.332***	0.357***	0.316*	0.347***	0.375***	0.401***	0.861
LS-DC-F	0.351	0.364	0.396	0.487	0.344***	0.366***	0.383**	0.329	0.353***	0.389***	0.425***	0.843
LS-D-F	0.364***	0.369***	0.397***	0.485	0.318***	0.319***	0.330***	0.307***	0.357***	0.380***	0.410***	0.926
LS-CC-F	0.332	0.353	0.390	0.479	0.299**	0.327	0.338	0.313	0.252***	0.277***	0.315**	0.821
LS-M-F	0.375	0.385	0.406	0.481	0.366**	0.386	0.391	0.328	0.374	0.412	0.458	0.824
Panel C: Portfolios from linear shrinkage estimator by minimizing negative log-likelihood												
LS-I-L	0.356**	0.370**	0.398	0.487	0.351**	0.375*	0.389	0.324	0.364	0.413	0.453	0.862
LS-DC-L	0.357**	0.371**	0.399	0.488	0.354*	0.377	0.391	0.330	0.367	0.416	0.455	0.862
LS-D-L	0.383	0.392	0.412	0.486	0.375	0.396	0.399	0.325	0.380	0.423	0.468	0.918**
LS-CC-L	0.338***	0.353***	0.385***	0.473***	0.306***	0.329***	0.343***	0.319	0.266***	0.296***	0.331***	0.835***
LS-M-L	0.380	0.388	0.407	0.480	0.373	0.389	0.395	0.329	0.378	0.415	0.458	0.897
LS-1PC-L	0.377	0.386	0.407	0.483	0.375	0.391	0.397	0.330	0.381	0.418	0.460	0.897

Notes: Table 11 shows the out-of-sample net Sharpe ratio for different portfolios using simulated data under various p and n. The studied portfolios include equally weighted portfolio (EW) and the GMV portfolios based on different covariance matrix estimators, namely, the sample covariance matrix (GMV-S), nonlinear shrinkage (NLS) covariance matrix estimator (Ledoit and Wolf, 2012), analytical nonlinear shrinkage (ANLS) covariance matrix estimator (Ledoit and Wolf, 2020), quadratic inverse shrinkage (QIS) estimator (Ledoit and Wolf, 2022), Schatten 1-norm (SN_1) regularized covariance matrix (Shi et al., 2020), ridge regularized (RR) correlation matrix estimator (Warton, 2008), and linear shrinkage (LS) covariance matrix estimators with different target matrices (the multiple of the identity matrix (I), double constant matrix (DC), diagonal matrix (D), constant correlation matrix (CC), and covariance matrix estimated from the one-factor model with the market (M) or the first principal component (1PC) as the factor) by minimizing Frobenius loss (F) or negative log-likelihood (L). Panel B tests the null hypothesis that there is no change in the out-of-sample variance between LS-X-F and LS-X-L, where X = {I, DC, D, CC, M}. Panels A and C test the null hypothesis that LS-1PC-L and other techniques have the same out-of-sample variance. These p-values of testing difference in Sharpe ratios are calculated using the stationary bootstrap technique, with bootstrap sampling of B = 1000 and block size of b = 5 (Ledoit and Wolf, 2008). "–" indicates that the GMV portfolio cannot be constructed due to the singularity of the sample covariance matrix.

* Indicate significance at the 10% level.

** Indicate significance at the 5% level.

*** Indicate significance at the 1% level.

columns are the corresponding eigenvectors. As $n_k \xrightarrow{p} \infty$, we have $S^{[k]} \xrightarrow{p} \Sigma$, $C^{[-k]} \xrightarrow{p} R$, and $\hat{\Sigma}_{\mathrm{D}}^{[-k]} \xrightarrow{p} \Sigma_{\mathrm{D}}$. It follows that

$$\sum_{k=1}^{K} -\log L\left(\hat{\Sigma}_{\text{LS-D}}^{[-k]}, S^{[k]}\right) \tag{54}$$

$$\xrightarrow{p} K\left[\mathrm{tr}\left(((1-\nu)R+\nu I)^{-1} R\right) + \log|((1-\nu)R+\nu I)| + \log|\Sigma_{\mathrm{D}}|\right] \tag{55}$$

$$= K\left[\mathrm{tr}\left(Q\left((1-\nu)\Gamma+\nu I\right)^{-1} Q' R\right) + \log|Q((1-\nu)\Gamma+\nu I)Q'| + \log|\Sigma_{\mathrm{D}}|\right] \tag{56}$$

$$= K\left[\mathrm{tr}\left(((1-\nu)\Gamma+\nu I)^{-1} Q' R Q\right) + \log|((1-\nu)\Gamma+\nu I)Q'Q| + \log|\Sigma_{\mathrm{D}}|\right] \tag{57}$$

$$= K\left[\mathrm{tr}\left(((1-\nu)\Gamma+\nu I)^{-1} \Gamma\right) + \log|((1-\nu)\Gamma+\nu I)| + \log|\Sigma_{\mathrm{D}}|\right] \tag{58}$$

$$= K\left[\sum_{i=1}^{p}\left(\frac{\gamma_i}{(1-\nu)\gamma_i+\nu} + \log((1-\nu)\gamma_i+\nu)\right) + \log|\Sigma_{\mathrm{D}}|\right], \tag{59}$$

Differentiating $\sum_{k=1}^{K} -\log L\left(\hat{\Sigma}_{\text{LS-D}}^{[-k]}, S^{[k]}\right)$ with respect to ν leads to

$$\frac{\partial\left(\sum_{k=1}^{K} -\log L\left(\hat{\Sigma}_{\text{LS-D}}^{[-k]}, S^{[k]}\right)\right)}{\partial\nu} \tag{60}$$

$$\xrightarrow{p} K\sum_{i=1}^{p}\frac{\partial\left(\frac{\gamma_i}{(1-\nu)\gamma_i+\nu} + \log((1-\nu)\gamma_i+\nu) + \log|\Sigma_{\mathrm{D}}|\right)}{\partial\nu} \tag{61}$$

$$= K\sum_{i=1}^{p}\left(\frac{-\gamma_i(1-\gamma_i)}{((1-\nu)\gamma_i+\nu)^2} + \frac{1-\gamma_i}{(1-\nu)\gamma_i+\nu}\right) \tag{62}$$

$$= K\nu\sum_{i=1}^{p}\frac{(1-\gamma_i)^2}{((1-\nu)\gamma_i+\nu)^2}, \tag{63}$$

which is positive for $\nu > 0$, $K > 0$ and $\gamma_i \neq 1\ \exists i$. If the true correlation matrix R has a full rank ($\gamma_i > 0\ \forall i$) and is not the identity matrix ($\gamma_i \neq 1\ \exists i$), then the negative log-likelihood function in Eq. (60) has a unique minimum at $\nu^* = 0$.

Data availability

Data will be made available on request.

MV Portfolio via Shrinkage of Large CM (Original data) (Mendeley Data)

References

Ao, M., Yingying, L., Zheng, X., 2019. Approaching mean-variance efficiency for large portfolios. Rev. Financ. Stud. 32 (7), 2890–2919.

Behr, P., Guettler, A., Miebs, F., 2013. On portfolio optimization: Imposing the right constraints. J. Bank. Financ. 37 (4), 1232–1242.

Best, M.J., Grauer, R.R., 1991. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. Rev. Financ. Stud. 4 (2), 315–342.

Broadie, M., 1993. Computing efficient frontiers using estimated parameters. Ann. Oper. Res. 45, 21–58.

Brodie, J., Daubechies, I., De Mol, C., Giannone, D., Loris, I., 2009. Sparse and stable Markowitz portfolios. Proc. Natl. Acad. Sci. 106 (30), 12267–12272.

Chan, L.K., Karceski, J., Lakonishok, J., 1999. On portfolio optimization: Forecasting covariances and choosing the risk model. Rev. Financ. Stud. 12 (5), 937–974.

Chopra, V.K., Ziemba, W.T., 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. J. Portfolio Manag. 19 (2), 6–11.

DeMiguel, V., Garlappi, L., Nogales, F.J., Uppal, R., 2009a. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. Manage. Sci. 55 (5), 798–812.

DeMiguel, V., Garlappi, L., Uppal, R., 2009b. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? Rev. Financ. Stud. 22 (5), 1915–1953.

DeMiguel, V., Martin-Utrera, A., Nogales, F.J., 2013. Size matters: Optimal calibration of shrinkage estimators for portfolio selection. J. Bank. Financ. 37 (8), 3018–3034.

Deshmukh, S., Dubey, A., 2020. Improved covariance matrix estimation with an application in portfolio optimization. IEEE Signal Process. Lett. 27, 985–989.

Fan, J., Zhang, J., Yu, K., 2012. Vast portfolio selection with gross-exposure constraints. J. Amer. Statist. Assoc. 107 (498), 592–606.

Frost, P.A., Savarino, J.E., 1986. An empirical Bayes approach to efficient portfolio selection. J. Financ. Quant. Anal. 21 (3), 293–305.

Frost, P.A., Savarino, J.E., 1988. For better performance: Constrain portfolio weights. J. Portfolio Manag. 15 (1), 29.

Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. J. Finance 58 (4), 1651–1683.

Jiang, C., Du, J., An, Y., 2019. Combining the minimum-variance and equally-weighted portfolios: Can portfolio performance be improved? Econ. Model. 80, 260–274.

Jobson, J.D., Korkie, B., 1980. Estimation for Markowitz efficient portfolios. J. Amer. Statist. Assoc. 75 (371), 544–554.

Jobson, J.D., Korkie, R.M., 1981. Putting Markowitz theory to work. J. Portfolio Manag. 7 (4), 70–74.

Kirby, C., Ostdiek, B., 2012. It's all in the timing: simple active portfolio strategies that outperform naive diversification. J. Financ. Quant. Anal. 47 (2), 437–467.

Kremer, P.J., Lee, S., Bogdan, M., Paterlini, S., 2020. Sparse portfolio selection via the sorted ℓ_1-norm. J. Bank. Financ. 110, 105687.

Ledoit, O.R.H., 1995. Essays on Risk and Return in the Stock Market (Ph.D. thesis). Massachusetts Institute of Technology.

Ledoit, O., Wolf, M., 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. J. Empir. Financ. 10 (5), 603–621.

Ledoit, O., Wolf, M., 2004a. Honey, I shrunk the sample covariance matrix. J. Portfolio Manag. 30 (4), 110–119.

Ledoit, O., Wolf, M., 2004b. A well-conditioned estimator for large-dimensional covariance matrices. J. Multivariate Anal. 88 (2), 365–411.

Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the sharpe ratio. J. Empir. Financ. 15 (5), 850–859.

Ledoit, O., Wolf, M., 2011. Robust performances hypothesis testing with the variance. Wilmott 2011 (55), 86–89.

Ledoit, O., Wolf, M., 2012. Nonlinear shrinkage estimation of large-dimensional covariance matrices. Ann. Statist. 40 (2), 1024–1060.

Ledoit, O., Wolf, M., 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. Rev. Financ. Stud. 30 (12), 4349–4388.

Ledoit, O., Wolf, M., 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. Ann. Statist. 48 (5), 3043–3065.

Ledoit, O., Wolf, M., 2022. Quadratic shrinkage for large covariance matrices. Bernoulli 28 (3), 1519–1547.

Levy, H., Levy, M., 2014. The benefits of differential variance-based constraints in portfolio optimization. European J. Oper. Res. 234 (2), 372–381.

Li, J., 2015. Sparse and stable portfolio selection with parameter uncertainty. J. Bus. Econom. Statist. 33 (3), 381–392.

Li, G., Huang, L., Yang, J., Zhang, W., 2022. A synthetic regression model for large portfolio allocation. J. Bus. Econom. Statist. 40 (4), 1665–1677.

Markowits, H.M., 1952. Portfolio selection. J. Finance 7 (1), 71–91.

Merton, R.C., 1980. On estimating the expected return on the market: An exploratory investigation. J. Financ. Econ. 8 (4), 323–361.

Michaud, R.O., 1989. The Markowitz optimization enigma: Is 'optimized'optimal? Financ. Anal. J. 45 (1), 31–42.

Shi, F., Shu, L., Yang, A., He, F., 2020. Improving minimum-variance portfolios by alleviating overdispersion of eigenvalues. J. Financ. Quant. Anal. 55 (8), 2700–2731.

Tu, X., Li, B., 2024. Robust portfolio selection with smart return prediction. Econ. Model. 135, 106719.

Warton, D.I., 2008. Penalized normal likelihood and ridge regularization of correlation and covariance matrices. J. Amer. Statist. Assoc. 103 (481), 340–349.

Xing, X., Hu, J., Yang, Y., 2014. Robust minimum variance portfolio with L-infinity constraints. J. Bank. Financ. 46, 107–117.

Yen, Y.-M., 2016. Sparse weighted-norm minimum variance portfolios. Rev. Finance 20 (3), 1259–1287.

Yen, Y.-M., Yen, T.-J., 2014. Solving norm constrained portfolio optimization via coordinate-wise descent algorithms. Comput. Statist. Data Anal. 76, 737–759.