

Research Article

Stable Portfolio Selection Strategy for Mean-Variance-CVaR Model under High-Dimensional Scenarios

Yu Shi , Xia Zhao , Fengwei Jiang , and Yipin Zhu

School of Statistics and Information, Shanghai University of International Business and Economics, Shanghai 201620, China

Correspondence should be addressed to Xia Zhao; zhaoxia-w@163.com

Received 29 April 2020; Accepted 29 May 2020; Published 15 July 2020

Guest Editor: Wenguang Yu

Copyright © 2020 Yu Shi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper aims to study stable portfolios with mean-variance-CVaR criteria for high-dimensional data. Combining different estimators of covariance matrix, computational methods of CVaR, and regularization methods, we construct five progressive optimization problems with short selling allowed. The impacts of different methods on out-of-sample performance of portfolios are compared. Results show that the optimization model with well-conditioned and sparse covariance estimator, quantile regression computational method for CVaR, and reweighted L_1 norm performs best, which serves for stabilizing the out-of-sample performance of the solution and also encourages a sparse portfolio.

1. Introduction

Mean-risk models are widely used and play an important role in financial risk management. The classical and revolutionary work is mean-variance (MV) optimization model proposed by Markowitz [1], in which variance is used to measure risk. Afterwards, more and more researchers are devoted to the study of this field. Kolm et al. [2] review the development, challenges, and trends of MV optimization problems in recent six decades. Considering different risk measures focus on different characteristics of risk, some risk measures other than variance are incorporated into the mean-risk framework. For example, Konno and Yamazaki [3] and Ogryczak and Ruszcynski [4] use absolute deviation and semideviation to measure risk, respectively, and construct mean-risk model for portfolio selection. Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) have been employed as the risk measure to conduct asset allocation (see Consigli [5], Alexander and Baptista [6], Xu et al. [7], and Quaranta and Zaffaroni [8] for more details).

Since different risk measures delineate different information of risk, the combination of two risk measures is used to control risk in mean-risk models. For instance, Konno et al. [9] and Konno and Suzuki [10] construct mean-absolute deviation-skewness model and mean-variance-skewness model,

respectively, in which the skewness is indicative of unidirectional movement (bearish or bullish) of the stock market. Robert and Philip [11] propose a mean-variance-skewness-kurtosis portfolio optimization model and show that higher-order moments of return can significantly change optimal portfolio construction. Roman et al. [12] construct an optimization model based on mean-variance-CVaR criterion, in which variance and CVaR are combined to obtain a range of balanced solutions that are generally discarded by both mean-variance and mean-CVaR models. Additionally, the personal preference between variance and CVaR of the decision-maker can be considered. Gao et al. [13] extend it to dynamic scenario in financial field and Shi et al. [14] give the discussion in insurance investment under regulatory constraints. However, the study of the stability of optimal strategy's out-of-sample performance is ready to explore under mean-variance-CVaR criterion. Thus, the adaptability of optimal investment strategy in practice is questionable, especially in high-dimensional scenarios (dimensionality of assets is comparable to or even larger than the number of observations). Here, we will study this problem from the point of view of model-solving procedure.

To solve mean-variance-CVaR optimization problem, the computation of variance and CVaR for portfolios is a central and fundamental work. For variance term, estimating covariance matrix of the return variables of assets is

necessary. Generally, sample covariance matrix is a good estimator when sample size is large sufficiently. But under high-dimensional scenarios, it usually delivers the presence of poor out-of-sample performance (see, e.g., Green and Hollifield [15]; Chopra and Ziemba [16]; DeMiguel et al. [17]). To deal with such instable out-of-sample performance, researchers make various contributions. Based on different thresholding function, thresholding covariance matrix estimator is one of the mainstreams of improving sample covariance matrix by considering sparsity (e.g., Tibshirani [18]; Bickel and Levin'a [19]; Cai et al. [20]). Moreover, to guarantee the presence of a convex optimization problem, Rothman et al. [21] propose the positive definite sparse covariance estimator (PDSCE). While considering the property of well condition, Ledoit and Wolf [22] propose an estimator of covariance matrix as a linear combination of sample covariance and identity matrix. Maurya [23] develops a well-conditioned and sparse estimator of covariance matrix in high-dimensional setting. This estimator has the properties of well-conditioned, sparsity, and positive definite; especially, the property of positive definite guarantees a convex optimization problem and it performs better than several other popular methods used in literatures (e.g., graphical lasso in Friedman et al. [24], PDSCE in Rothman [21], and Bickel and Levin'a thresholding estimator in Bickel and Levin'a [19]).

For CVaR item, Rockafeller and Uryasev [25, 26] calculate CVaR through a convex programming problem, which pave the way of portfolio selection with CVaR under nonnormal assumption. Bassett et al. [27] bridge the gap between quantile regression and calculation of CVaR without distribution assumption of returns. However, it has also been investigated that the out-of-sample performance is unstable in mean-CVaR model (see Lim et al. [28], Takeda and Kanamori [29], and Kondor et al. [30] for more details). A popular way to make mean-CVaR model stable is regularization technique which can lead to a sparse portfolio simultaneously. Xu et al. [7] incorporate penalty function into mean-CVaR model under the large scale sample scenarios to get a sparse portfolio. Gao and Wu [31] combine penalty function and variance item to improve the out-of-sample performance of mean-CVaR model. Additionally, regularization method has wide applications in constructing mean-variance model to find stable optimal portfolios with better out-of-sample performance. For example, Brodie et al. [32] reformulate classical Markowitz's mean-variance model as a constrained least-squares regression problem. They add a L_1 -regularization term to the objective function and show that this penalty regularizes the optimization problem and encourages sparse portfolios. Other works about regularization technique used in mean-variance or mean-CVaR model to promote stable solutions can be seen in literatures (e.g., Gotoh and Takeda [33] and Fastrich et al. [34]).

Motivated by the above discussion, in this paper, we try to find portfolios with more stable out-of-sample performance for mean-variance-CVaR model under high-dimensional scenarios. Five progressive optimization problems are designed carefully based on mean-variance-CVaR criteria through combining two estimators for covariance matrix, two computing methods for CVaR, and two penalty functions in different ways. Simulation is conducted to compare the out-of-sample performance of portfolios obtained from these optimization problems and the impact of underlying methods on optimal strategy is analyzed. Based on the historical data from the constituent stocks in Shanghai Stock Exchange 50 Index, an empirical study is considered.

The remaining of this paper is organized as follows. Section 2 describes the related methods used in this paper. Section 3 shows the optimization problems we formulate based on the mean-variance-CVaR criterion. Section 4 provides simulation and result comparisons. An empirical study is carried out in Section 5. Finally, Section 6 concludes the paper.

2. Methods

In this section, we will briefly introduce risk measures used in this paper, their estimating methods, and some regularization methods. Suppose that we have the opportunity to invest in p assets with returns X_k , $k = 1, 2, \dots, p$. Let $\mathbf{X} = (X_1, X_2, \dots, X_p)'$ represent return vector with weight vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)'$, in which β_k is the portfolio allocation weight for X_k . Thus, the total return is $Y = \mathbf{X}'\boldsymbol{\beta}$.

2.1. Variance. Variance describes the degree of dispersion of a random variable and hence can measure the fluctuation of investment return. Here, we use $\mathcal{V}(\mathbf{X}'\boldsymbol{\beta})$ to denote the variance of a portfolio return variable Y since $Y = \mathbf{X}'\boldsymbol{\beta}$, and $\mathcal{V}(\mathbf{X}'\boldsymbol{\beta})$ can be expressed as

$$\mathcal{V}(\mathbf{X}'\boldsymbol{\beta}) = \sum_{k=1}^p \sum_{j=1}^p \beta_k \beta_j \sigma_{kj} = \boldsymbol{\beta}' \Sigma \boldsymbol{\beta}, \quad (1)$$

where $\Sigma = (\sigma_{kj})_{p \times p}$ is the covariance matrix of \mathbf{X} and σ_{kj} is the covariance of X_k and X_j .

Generally, Σ is estimated by sample covariance matrix (called as classical estimator); that is, $\hat{\Sigma}_c = (\hat{\sigma}_{kj})_{p \times p}$ and $\hat{\sigma}_{kj} = (1/n - 1) \sum_{i=1}^n (x_{ki} - \bar{x}_k)(x_{ji} - \bar{x}_j)$ denote the sample covariance of X_k and X_j , where n is the sample size.

Since the well-conditioned and sparse estimator proposed by Maurya [23], noted as $\hat{\Sigma}_w$, has been proved its superiority of property. Therefore, in this paper, we employ the well-conditioned and sparse estimator $\hat{\Sigma}_w$ directly. Following the notation in Maurya [23], we have

$$\hat{\Sigma}_w(\lambda_1, \tau_1) = \operatorname{argmin}_{\Sigma=\Sigma^T} \left(\|\Sigma - \hat{\Sigma}_c\|_2^2 + \lambda_1 \|\Sigma^{-}\|_1 + \tau_1 \sum_{k=1}^p (\sigma_i(\Sigma) - \bar{\sigma}_\Sigma)^2 \right), \quad (2)$$

where $\Sigma^- = \Sigma - \Sigma^+$ and Σ^+ denote a diagonal matrix with the same diagonal as Σ , $\sigma_i(\Sigma)$ is the i th largest eigenvalue of matrix Σ , and $\bar{\sigma}_\Sigma$ is the mean of eigenvalues of Σ , and the way to identify the best pair (λ_1, τ_1) can also be found in Maurya [23].

2.2. CVaR. For financial assets or portfolio, VaR refers to the greatest possible loss over specific holding period, at a certain confidence level $100(1 - \alpha)\%$, $\alpha \in (0, 1)$. If the cumulative distribution function for return Y is F , VaR can be defined as

$$\text{VaR}_{1-\alpha} = -F^{-1}(\alpha). \quad (3)$$

As we know, compared with VaR, a key advantage of CVaR is that CVaR satisfies the four coherence axioms of Artzner et al. [35], whereas VaR is not coherent. The definition of CVaR is as follows:

$$\text{CVaR}_{1-\alpha} = E[-Y | -Y \geq \text{VaR}_{1-\alpha}]. \quad (4)$$

Evidently, $\text{CVaR}_{1-\alpha}$, the conditional expectation of losses exceeding $\text{VaR}_{1-\alpha}$, is more informative about the tail of the distribution than $\text{VaR}_{1-\alpha}$. Two popular ways to compute CVaR without distribution constraints have been employed recently. The first one is a linear programming model for optimizing CVaR proposed by Rockafellar and Uryasev [25, 26], generally called as Rockafellar-Uryasev's approximation. They have proved that CVaR can be calculated by solving a convex optimization problem. In other words, CVaR can be formulated as

$$\text{CVaR}_{1-\alpha}^1 = \min_{\beta, \xi} \xi + \alpha^{-1} E[-\mathbf{X}'\beta - \xi]^+, \quad (5)$$

$$\lambda \sum_{k=1}^p \text{Pen}_1(\beta_k) = \sum_{k=1}^p \left\{ \lambda |\beta_k| I(|\beta_k| \leq \lambda) + \frac{-\beta_k^2 + 2\zeta\lambda|\beta_k| - \lambda^2}{2(\zeta-1)} I(\lambda < |\beta_k| \leq \zeta\lambda) + \frac{(\zeta+1)\lambda^2}{2} I(|\beta_k| > \zeta\lambda) \right\}, \quad (8)$$

where $\zeta (\zeta < 2)$ and λ are tuning parameters. It is well known that SCAD is continuous and singular at the origin, penalizes large coefficients equally, and has no bias, and the reweighted L_1 norm is as follows:

$$\lambda \sum_{k=1}^p \text{Pen}_2(\beta_k) = \lambda \sum_{k=1}^p \omega_k |\beta_k|, \quad (9)$$

where λ is a tuning parameters and the value of ω_k can be identified by iteration method based on the value of $|\beta_k|$. Reweighted L_1 norm has a better out-of-sample performance compared with L_1 norm penalty [34], since it gives every component in β a corresponding weight so that the penalization is more accurate.

where ξ is the α th quantile of Y . This method has been widely used in many literatures (e.g., Roman et al. [12], Lim et al. [28], and Gao and Wu [31]).

The second way is quantile regression method proposed in Bassett et al. [27]. Xu et al. [7] have proved that it has faster computational speed compared with Rockafellar-Uryasev's approximation method. According to Bassett et al. [27], the CVaR of a portfolio return Y could be calculated by

$$\text{CVaR}_{1-\alpha}^2 = \alpha^{-1} \min_{\xi} E[\rho_\alpha(Y - \xi)] - E[Y], \quad (6)$$

where ξ is the α th quantile of Y and $\rho_\alpha(u) = u(\alpha - I(u < 0))$ is the check function, in which $I(\cdot)$ is an indicator function that takes on the value one whenever its argument is true and zero otherwise.

Since $Y = \mathbf{X}'\beta$, we can have $Y = X_1 - \sum_{j=2}^p (X_1 - X_j)\beta_j$ under the constraint $\mathbf{1}'\beta = 1$. Let $\widetilde{X}_j = X_1 - X_j (j = 2, 3, \dots, p)$, we can convert equation (6) into

$$\text{CVaR}_{1-\alpha}^2 = \alpha^{-1} \min_{\beta, \xi} E \left[\rho_\alpha \left(X_1 - \sum_{j=2}^p \widetilde{X}_j \beta_j - \xi \right) \right] - E[\mathbf{X}'\beta]. \quad (7)$$

2.3. Regularization Method. To encourage a stable and sparse solution in optimization problem, penalty item $\text{Pen}(\cdot)$ is frequently used, which is regarded as regularization method. $\text{Pen}(\cdot)$ is a general penalty function that allows shrinking the components in β to zero. In this paper, we focus on smoothly clipped absolute deviation (SCAD) penalty proposed by Fan et al. [36] and reweighted L_1 norm penalty proposed by Emmanuel et al. [37], since they have oracle properties.

The SCAD penalty is formulated as follows:

3. Model Formation

Motivated by Roman et al. [12], we consider the following mean-variance-CVaR model:

$$\mathcal{P}_{\text{mvc}}: \begin{cases} \min_{\beta} & \eta \mathcal{V}(\mathbf{X}'\beta) + (1 - \eta) \text{CVaR}_{1-\alpha} \\ \text{s.t.} & \mathbf{1}'\beta = 1 \\ & E[\mathbf{X}'\beta] = \mu_0, \end{cases} \quad (10)$$

where $0 \leq \eta \leq 1$ is a weighting parameter, which represents the attitude that investors hold towards the two risk measures and hence balances the importance of them.

Remark 1. Letting $\eta = 0$ or $\eta = 1$, respectively, the above optimization problem is reduced to the mean-CVaR model or mean-variance model.

Remark 2. Different from Roman et al. [12], in this paper, we incorporate the variance term and CVaR term into the objective function of optimization model simultaneously and short selling is allowed.

3.1. Optimization Problems. In this section, we present the following five optimization problems based on different

methods of calculating variance, CVaR, and different regularization methods in order to obtain optimal portfolio for \mathcal{P}_{mvc} .

To illustrate clearly, we outline the optimization problems in Table 1. Here, “√” means the underlying method is incorporated in the corresponding optimization problem. For example, the second row tells us that the classical estimation $\hat{\Sigma}_c$ of variance term and Rockafellar–Uryasev’s approximation CVaR $^1_{1-\alpha}$ for CVaR term are used in problem \mathcal{P}_1 .

Mathematically, optimization problems can be written in the following forms:

$$\begin{aligned} \mathcal{P}_1: & \left\{ \begin{array}{ll} \min_{\beta, \xi} & \eta \beta' \hat{\Sigma}_c \beta + (1 - \eta) (\xi + \alpha^{-1} E[-\mathbf{X}' \beta - \xi]^+) \\ \text{s.t.} & \mathbf{1}' \beta = 1 \\ & E[\mathbf{X}' \beta] = \mu_0, \end{array} \right. \\ \mathcal{P}_2: & \left\{ \begin{array}{ll} \min_{\beta, \xi} & \eta \beta' \hat{\Sigma}_w \beta + (1 - \eta) (\xi + \alpha^{-1} E[-\mathbf{X}' \beta - \xi]^+) \\ \text{s.t.} & \mathbf{1}' \beta = 1 \\ & E[\mathbf{X}' \beta] = \mu_0, \end{array} \right. \\ \mathcal{P}_3: & \left\{ \begin{array}{ll} \min_{\beta, \xi} & \eta \beta' \hat{\Sigma}_w \beta + (1 - \eta) \alpha^{-1} E \left[\rho_\alpha \left(X_1 - \sum_{j=2}^p \widetilde{X}_j \beta_j - \xi \right) \right] \\ \text{s.t.} & E[\mathbf{X}' \beta] = \mu_0, \end{array} \right. \\ \mathcal{P}_4: & \left\{ \begin{array}{ll} \min_{\beta, \xi} & \eta \beta' \hat{\Sigma}_w \beta + (1 - \eta) \alpha^{-1} E \left[\rho_\alpha \left(X_1 - \sum_{j=2}^p \widetilde{X}_j \beta_j - \xi \right) \right] + \lambda \sum_{i=1}^p \text{Pen}_1(\beta_i) \\ \text{s.t.} & E[\mathbf{X}' \beta] = \mu_0, \end{array} \right. \\ \mathcal{P}_5: & \left\{ \begin{array}{ll} \min_{\beta, \xi} & \eta \beta' \hat{\Sigma}_w \beta + (1 - \eta) \alpha^{-1} E \left[\rho_\alpha \left(X_1 - \sum_{j=2}^p \widetilde{X}_j \beta_j - \xi \right) \right] + \lambda \sum_{i=1}^p \text{Pen}_2(\beta_i) \\ \text{s.t.} & E[\mathbf{X}' \beta] = \mu_0. \end{array} \right. \end{aligned} \quad (11)$$

Why do we design such five optimization problems in the foregoing way? In \mathcal{P}_1 , the classical method of variance term and Rockafellar–Uryasev’s approximation for CVaR term are used to find the solution. To construct \mathcal{P}_2 , we replace the classical estimator $\hat{\Sigma}_c$ in the variance term in \mathcal{P}_1 with $\hat{\Sigma}_w$, and the CVaR part remains the same so that we could find if $\hat{\Sigma}_w$ could mitigate the instability. Further, we replace CVaR $^1_{1-\alpha}$ in \mathcal{P}_2 with CVaR $^2_{1-\alpha}$ to get \mathcal{P}_3 , so that the difference between the two computational methods for CVaR could be compared. In $\mathcal{P}_3 - \mathcal{P}_5$, $\mathbf{1}' \beta = 1$ can be absorbed into the CVaR item based on quantile regression method. To examine the performance of SCAD and reweighted L_1 norm penalty function, we design \mathcal{P}_4 and \mathcal{P}_5 , in which quantile computational method for CVaR is used because of the convenience of selecting tuning parameter in

penalty function and faster computational speed [7]. Through such progressive problem design, we can figure out the performance of different methods and find the most effective one in our setting.

Remark 3. The estimation methods used in \mathcal{P}_1 for variance and CVaR are the same as that in Roman et al. [12].

Remark 4. If $\eta = 0$, \mathcal{P}_4 degenerates into the model in Xu et al. [7]. However, the constraint $E[\mathbf{X}' \beta] = \mu_0$ is missing in their paper. According to Bassett et al. [27], the $E[\mathbf{X}' \beta]$ item in CVaR part (see (7)) in objective function can be neglected only under the constraint that the expected return of portfolio is a constant.

TABLE 1: Methods in optimization problems.

| | $\widehat{\Sigma}_c$ | $\widehat{\Sigma}_w$ | $\text{CVaR}_{1-\alpha}^1$ | $\text{CVaR}_{1-\alpha}^2$ | Pen_1 | Pen_2 |
|-----------------|----------------------|----------------------|----------------------------|----------------------------|----------------|----------------|
| \mathcal{P}_1 | ✓ | | ✓ | | | |
| \mathcal{P}_2 | | ✓ | ✓ | | | |
| \mathcal{P}_3 | ✓ | | | ✓ | | |
| \mathcal{P}_4 | ✓ | | | ✓ | ✓ | |
| \mathcal{P}_5 | ✓ | | ✓ | ✓ | | ✓ |

3.2. *Selection of Tuning Parameter.* For \mathcal{P}_4 and \mathcal{P}_5 , we need to search the best tuning parameter for penalty item.

$$\widehat{\boldsymbol{\theta}} = \operatorname{argmin}_{\boldsymbol{\theta}} \ln \left\{ \eta \cdot \widehat{\beta} \widehat{\Sigma}_w \widehat{\beta} + (1 - \eta) \alpha^{-1} \sum_{i=1}^n \rho_\alpha \left(x_{i,1} - \sum_{j=2}^p \tilde{x}_{i,j} \widehat{\beta}_{j,\theta} - \widehat{\xi}_{\theta} \right) \right\} + df \cdot \frac{\ln n}{2n} \cdot \ln p, \quad (12)$$

where $x_{1,1}, x_{2,1}, \dots, x_{n,1}$ are random samples of X_1 , $\tilde{x}_{1,j}, \tilde{x}_{2,j}, \dots, \tilde{x}_{n,j}$ are random samples of \tilde{X}_j for $j = 2, 3, \dots, p$, and df is the effective dimension of fitting model. Additionally, $df = |\mathcal{E}|$ indicates the number of points in the set \mathcal{E} with $\mathcal{E} = \{j: \widehat{\beta}_{j,\theta} \neq 0, 2 \leq j \leq p\}$. Furthermore, the iterative algorithm for conducting \mathcal{P}_4 can be seen in Wu and Liu [40].

In \mathcal{P}_5 , λ is the nonnegative regularization parameter and $\omega_k, k = 1, 2, \dots, p$ is the tuning parameter in reweighted L_1 norm penalty. For $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_p)'$, we apply the following iterative procedure for every given λ to change the weighting parameter ω_k dynamically and adaptively [37].

- (1) Let the iteration count be h . For any given $\boldsymbol{\omega}^{(h)}$, we solve the problem \mathcal{P}_5 , which gives the solution $\widehat{\beta}^{(h)}$. If the stopping criteria are satisfied, for example, compared with every component in $\widehat{\beta}^{(h-1)}$, the maximum change does not exceed some small constant, we stop the iteration. Otherwise, go to Step (2).
- (2) Use $\widehat{\beta}^{(h)}$ to construct the new weighting parameter $\omega_k^{(h+1)} = 1/(|\widehat{\beta}_k^{(h)}| + \varepsilon)$, where ε can be a small positive number and let $h = h + 1$. Go to Step (1).

Finally, here, $\boldsymbol{\theta} = (\lambda, \omega_1, \omega_2, \dots, \omega_p)'$ can be searched by iteration based on the modified Bayesian Information Criterion equation (12).

4. Simulation and Discussion

4.1. *Data Generation.* To evaluate the out-of-sample performance of these five portfolio optimization problems, the dataset is generated by simulation approach with parameters being estimated from real historical price data of stock index.

4.1.1. *Construction of Return Variable.* Following the idea in Lim et al. [28], the random returns are captured by a hybrid distribution with multivariate normal distribution and exponential distribution. So all assets might suffer a perfectly correlated exponential-tail loss in a small probability. Let z be the exponential random variable with parameter q and we

In \mathcal{P}_4 , the nonnegative regularization parameter λ drives the relevance of the SCAD penalty and ζ is a tuning parameter belonging to SCAD penalty. Here, we search the best pair $\boldsymbol{\theta} = (\lambda, \zeta)'$ over two-dimension grids following the criterion mentioned in the following. Motivated by Lee et al. [38] and Wang et al. [39], we use the modified Bayesian Information Criterion as follows:

set $q = 10$ for simplicity. So the return variable is constructed as follows:

$$\mathbf{X} \sim B(\tau)(z\mathbf{1} + \mathbf{c}) + (1 - B(\tau))N, \quad (13)$$

where $B(\tau)$ is the Bernoulli random variable with parameter τ , N follows multivariate normal distribution with mean vector $\widetilde{\mathbf{u}}$ and covariance matrix $\widetilde{\Sigma}$, and $\mathbf{c} = (c_1, c_2, \dots, c_p)'$ with $c_k = \widetilde{\mu}_k - \sqrt{\widetilde{\Sigma}_{kk}}/2$, where $\widetilde{\Sigma}_{kk}$ is the k -th diagonal element of the matrix $\widetilde{\Sigma}$. The parameter τ controls the tail loss of the distribution.

4.1.2. *Determination of Parameter Values.* We choose 142 stocks which in Shanghai Stock Exchange Constituent (SSEC) index download the data from Joinquant Data platform. The sample period spans from Jan 3, 2015, to Dec 31, 2015. We calculate the daily yield return and estimate their mean return vector $\widetilde{\mathbf{u}} = (\mu_1, \mu_2, \dots, \mu_p)'$ and covariance matrix $\widetilde{\Sigma}$. (Note that the 180 constituent stocks in SSEC are always changing in the sample period, since the sample is adjusted every half year; therefore, we choose 142 stocks which continuously stay in the index over the period.)

4.1.3. *Data Generation.* As we know, when the dimensionality of assets is comparable to the number of observations, the data set belongs to high-dimensional scenarios. Therefore, in order to simulate a high-dimensional data set, we set the size of n close to $p = 142$. Letting $n = 150$ and $\tau = 0.05$, we simulate 101 groups of data, one group as in-sample data, and 100 group as out-of-sample data.

4.2. *Optimization Results.* In this section, we will calculate simulation results based on the five optimization problems, respectively, and give the comparisons from the point view of the out-of-sample performance. Without loss of generality, we set $\mu_0 = 0.01$. All the calculations follow three steps:

Step 1: calculate the weight of assets in portfolio selection model with different $\eta = 0, 0.25, 0.5, 0.75$, and 1 for in-sample data.

Step 2: based on the weights in Step 1, we calculate the variance, CVaR with 0.95 confidence level, and expected return for other 100 groups of out-of-sample data.

Step 3: summarize their mean value and standard deviation for every case. For example, under the case of \mathcal{P}_1 at $\eta = 1$, the mean value and the standard deviation for 100 values of CVaR are calculated to be 10.0812 and (0.9708), respectively.

Steps 1 and 2 mean that we hold the optimal portfolio generated from in-sample data in the out-of-sample period, in which the data follow the same distribution as in-sample data. Therefore, if the optimal portfolio is stable, the performance of portfolio in 100 groups of out-of-sample data should be similar, and hence, we calculate the standard deviation in Step 3 to measure the similarity of the performance in 100 groups.

The results are shown in Table 2.

From Table 2, the following results are concluded:

- (1) Compared \mathcal{P}_1 with \mathcal{P}_2 – \mathcal{P}_5 , the mean for 100 values of variance, CVaR in \mathcal{P}_1 are larger, which means that the risks of portfolios in out-of-sample data get out of control and the expected returns are also higher. This is consistent with the truth of “high risk; high return.” The inherent reason is that sample covariance estimator lost its theoretical support (Law of Large Numbers) in high-dimensional scenarios.
- (2) Compared \mathcal{P}_2 with \mathcal{P}_1 , we can find that all the counterparts in \mathcal{P}_2 decrease, which tells that the use of well-conditioned and sparse estimator ($\hat{\Sigma}_w$) could control the risk significantly and improves the stability of solutions’ out-of-sample performance as well. When $\eta = 0$, the model is reduced to mean-CVaR model, whose solution is irrelevant to the estimation of Σ , so the values of CVaR and expected return should remain the same. The data in Table 2 really illustrate this fact.
- (3) The results in \mathcal{P}_2 and \mathcal{P}_3 share identical in-sample optimal portfolio values and out-of-sample performance. This means that the impact of Rockafellar–Uryasev’s approximation and quantile regression method of CVaR on the solution of optimization problem shows no significant difference under our setting.
- (4) The values in \mathcal{P}_4 further become smaller than ones in \mathcal{P}_3 , which says that the use of SCAD penalty technique can improve the stability and the risk-controlling ability for out-of-sample performance.
- (5) Compared \mathcal{P}_5 with \mathcal{P}_4 , the values of variance and CVaR in \mathcal{P}_5 show a slight decrease, which means that the reweighted L_1 norm penalty contributes to a more stable out-of-sample performance than SCAD penalty, even though the improvement is fractional.
- (6) Bigger η represents a higher proportion for variance item in objective function. Since variance is a more conservative risk measure compared with CVaR,

TABLE 2: Out-of-sample performance of \mathcal{P}_1 – \mathcal{P}_5 .

| | | $\eta = 1$ | $\eta = 0.75$ | $\eta = 0.5$ | $\eta = 0.25$ | $\eta = 0$ |
|-----------------|------------|---------------------|--------------------|--------------------|--------------------|--------------------|
| \mathcal{P}_1 | Variance | 5.0235 (0.3525) | 4.6044 (0.3281) | 4.3821 (0.3163) | 4.2897 (0.3111) | 4.9986 (0.3121) |
| | CVaR | 10.0812 (0.9708) | 9.2110 (0.9307) | 8.7280 (0.9299) | 8.5244 (0.9324) | 9.9947 (0.9786) |
| | Exp return | 0.2389 (0.4227) | 0.2076 (0.3950) | 0.1996 (0.3855) | 0.2205 (0.3894) | 0.31 (0.4571) |
| | Variance | 1.3405 (0.1919) | 1.5978 (0.2613) | 1.9141 (0.3365) | 2.3783 (0.4369) | 5.1578 (0.9564) |
| \mathcal{P}_2 | CVaR | 2.5983 (0.2618) | 3.0591 (0.3213) | 3.6541 (0.3635) | 4.5752 (0.4496) | 9.9947 (0.9786) |
| | Exp return | 0.1939 (0.1226) | 0.1817 (0.1438) | 0.1974 (0.1729) | 0.1690 (0.2176) | 0.31 (0.4571) |
| | Variance | 1.3405 (0.1919) | 1.5978 (0.2613) | 1.9141 (0.3365) | 2.3783 (0.4369) | 5.1578 (0.9564) |
| | CVaR | 2.5983 (0.2618) | 3.0591 (0.3213) | 3.6541 (0.3635) | 4.5752 (0.4496) | 9.9947 (0.9786) |
| \mathcal{P}_3 | Exp return | 0.1939 (0.1226) | 0.1817 (0.1438) | 0.1974 (0.1729) | 0.1690 (0.2176) | 0.31 (0.4571) |
| | Variance | 1.1758 (0.0837) | 1.1934 (0.0815) | 1.2232 (0.0795) | 1.2782 (0.0802) | 1.2769 (0.0764) |
| | CVaR | 2.4739 (0.25) | 2.4997 (0.2670) | 2.5401 (0.2751) | 2.6587 (0.2882) | 2.6631 (0.2867) |
| | Exp return | 0.0772 (0.1092) | 0.0816 (0.1112) | 0.0925 (0.1148) | 0.0962 (0.1215) | 0.0917 (0.1199) |
| \mathcal{P}_4 | Variance | 1.1223 (0.0769) | 1.1632 (0.0769) | 1.1830 (0.0744) | 1.2094 (0.0738) | 1.2274 (0.0705) |
| | CVaR | 2.3334 (0.2407) | 2.4151 (0.2570) | 2.4395 (0.2625) | 2.4956 (0.2770) | 2.5304 (0.2750) |
| | Exp return | 0.08499 (0.1026) | 0.0852 (0.1071) | 0.0917 (0.1092) | 0.0956 (0.1142) | 0.0976 (0.1145) |

generally, bigger η results in a more conservative portfolio with relative low-risk index and expected return in out-of-sample analysis. The values in \mathcal{P}_4 and \mathcal{P}_5 show this trend well. But others fail, which may attribute to the instable out-of-sample performance.

To analyze the stability of out-of-sample performance in \mathcal{P}_1 – \mathcal{P}_5 easily, we draw the scatter diagram for each optimization problem at $\eta = 0.5$. For the convenience of comparison, we give the following graphical design. First, \mathcal{P}_1 , \mathcal{P}_2 , and \mathcal{P}_3 share the same coordinate system (see Figure 1), and the diagrams of \mathcal{P}_2 and \mathcal{P}_3 are identical, so we draw them in one diagram. Second, to give a vivid comparison, we zoom in the coordinate scale and draw the diagrams of \mathcal{P}_3 and \mathcal{P}_4 ; see Figure 2. The coordinate scale is further amplified in Figure 3 for the diagrams of \mathcal{P}_4 and \mathcal{P}_5 .

From Figure 1, we can find that the points in (b) are in lower risk positions and more concentrative, which means that the portfolio generated from \mathcal{P}_2 (or \mathcal{P}_3) has a better risk-controlling ability and more similar performance in 100 groups of out-of-sample data sets. Thus, the well-conditioned and sparse estimator $\hat{\Sigma}_w$ helps to stabilize the out-of-sample performance and control the risk better. This is consistent with result (2) from Table 2.

From Figures 2 and 3, we can conclude that penalty items also help to stabilize the out-of-sample performance

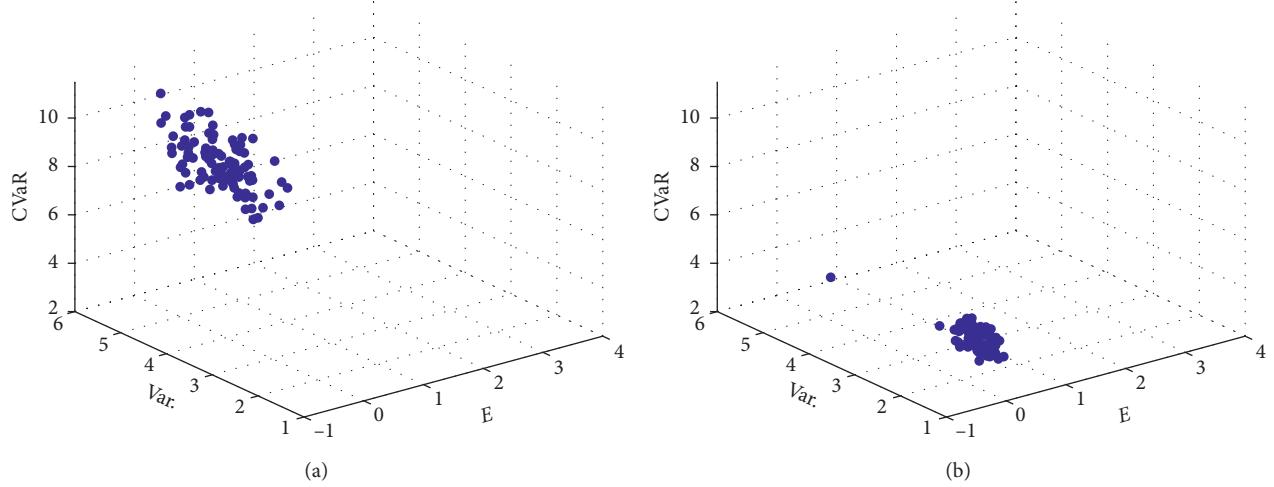


FIGURE 1: Out-of-sample performance of \mathcal{P}_1 - \mathcal{P}_3 ($\eta = 0.5$). (a) Performance of \mathcal{P}_1 ; (b) performance of \mathcal{P}_2 or \mathcal{P}_3 .

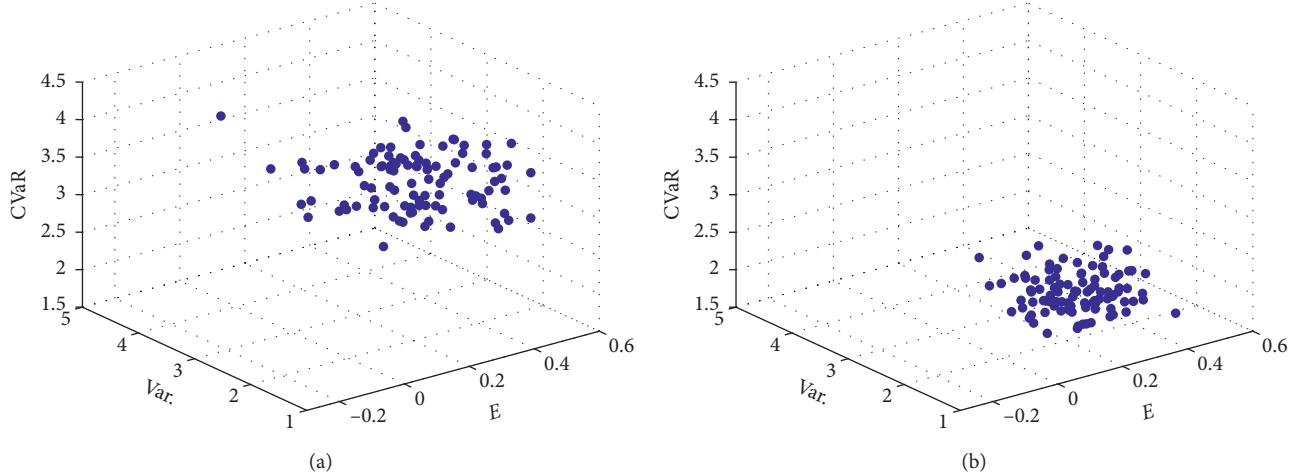


FIGURE 2: Out-of-sample performance of \mathcal{P}_3 - \mathcal{P}_4 ($\eta = 0.5$). (a) Performance of \mathcal{P}_3 ; (b) performance of \mathcal{P}_4 .

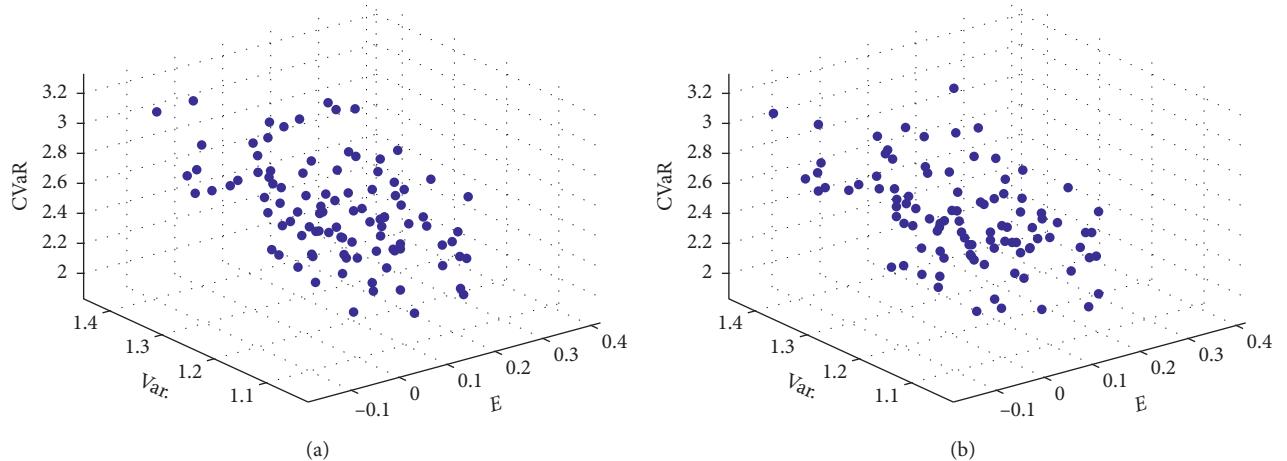


FIGURE 3: Out-of-sample performance of \mathcal{P}_4 - \mathcal{P}_5 ($\eta = 0.5$). (a) Performance of \mathcal{P}_4 ; (b) performance of \mathcal{P}_5 .

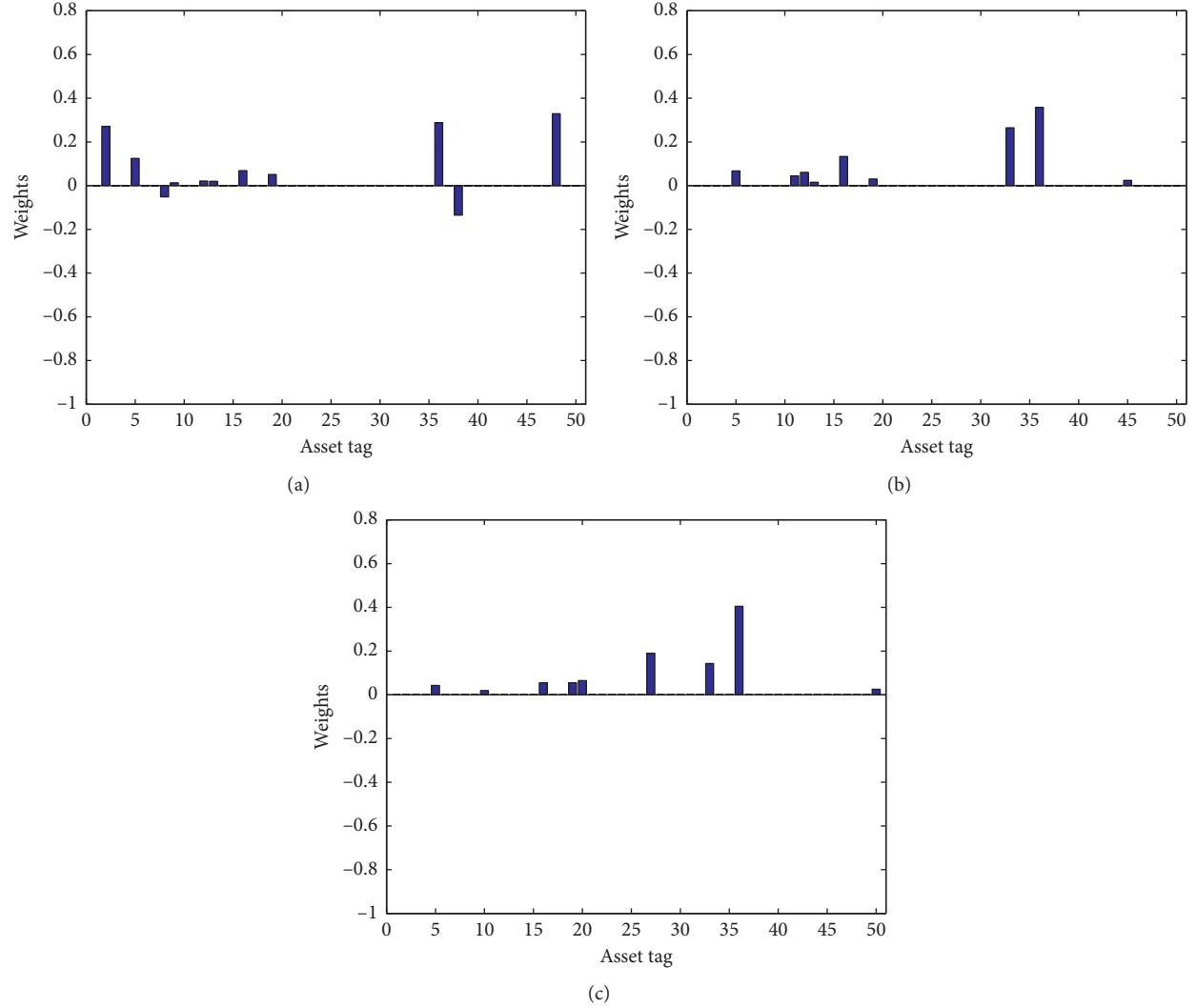


FIGURE 4: Portfolio strategy of \mathcal{P}_5 under different η . (a) Portfolio of \mathcal{P}_5 ($\eta = 0$), (b) portfolio of \mathcal{P}_5 ($\eta = 0.5$), and (c) portfolio of \mathcal{P}_5 ($\eta = 1$).

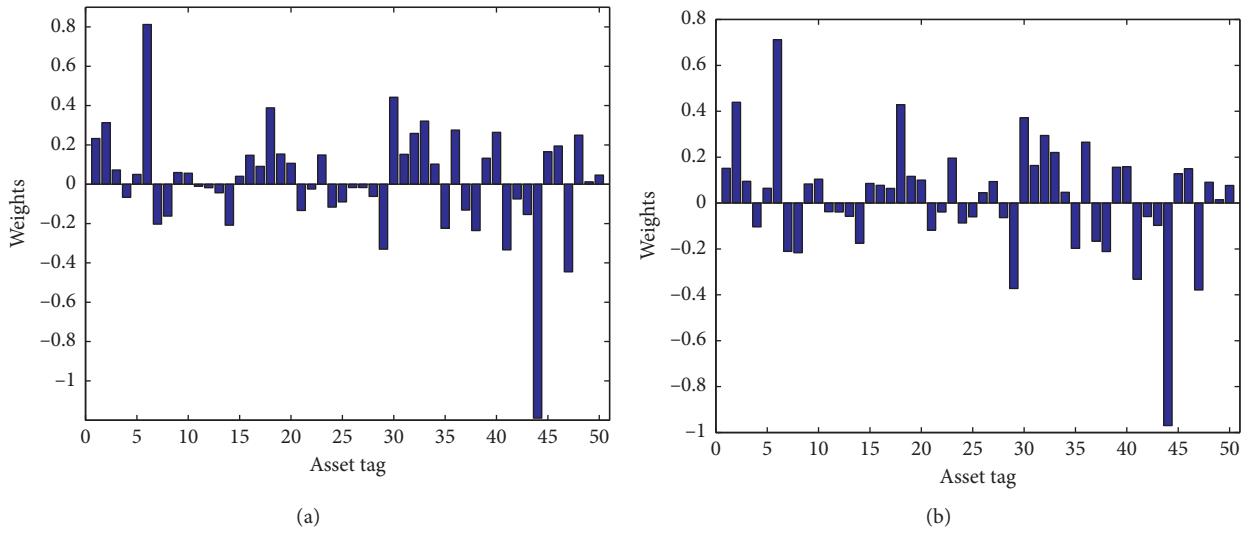


FIGURE 5: Continued.

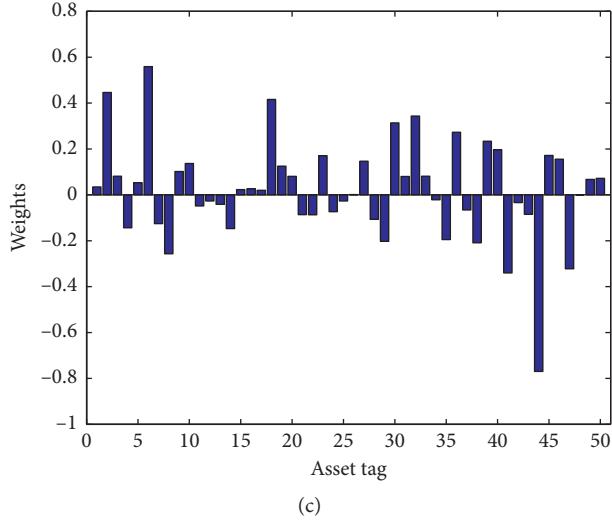


FIGURE 5: Portfolio strategy of \mathcal{P}_1 under different η . (a) Portfolio of \mathcal{P}_1 ($\eta = 0$), (b) portfolio of \mathcal{P}_1 ($\eta = 0.5$), and (c) portfolio of \mathcal{P}_1 ($\eta = 1$).

TABLE 3: Out-of-sample performance of portfolios.

| | $\eta = 0$ | $\eta = 0.5$ | $\eta = 1$ | EW | | | |
|------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-------|
| | \mathcal{P}_1 | \mathcal{P}_5 | \mathcal{P}_1 | \mathcal{P}_5 | \mathcal{P}_1 | \mathcal{P}_5 | |
| Variance | 3.609 | 1.674 | 3.326 | 1.937 | 3.158 | 1.986 | 2.787 |
| CVaR | 3.836 | 2.998 | 3.998 | 3.210 | 3.827 | 3.187 | 3.985 |
| Avg return | 0.006 | 0.061 | -0.059 | 0.075 | -0.065 | 0.0380 | 0.017 |

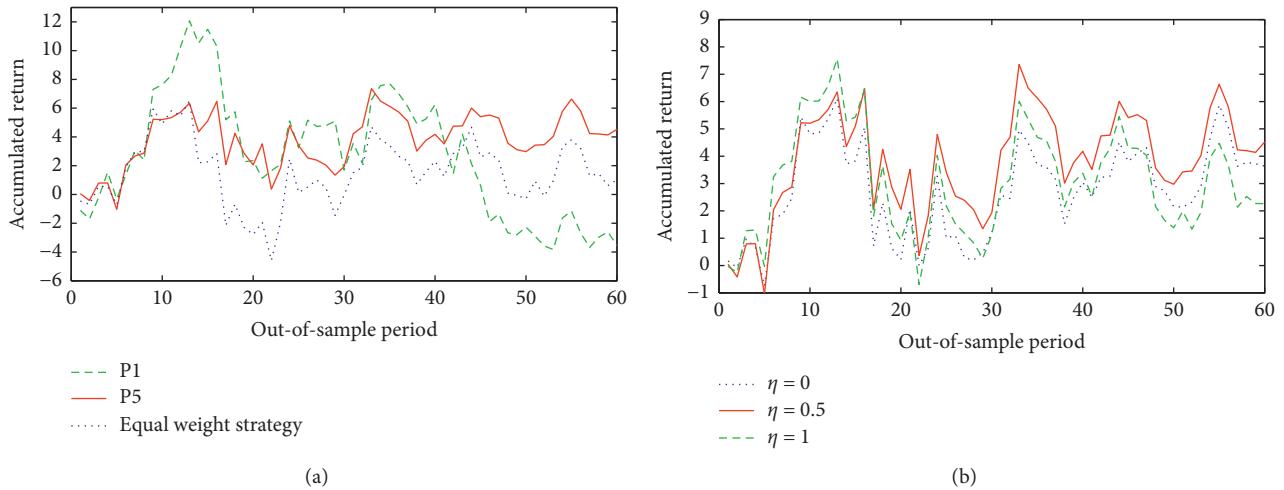


FIGURE 6: Accumulated return in out-of-sample period. (a) Different model with $\eta = 0.5$. (b) Model \mathcal{P}_5 with different η .

significantly and the reweighted L_1 norm penalty performs a little bit better than SCAD penalty under our settings. This is consistent with results (4) and (5) from Table 2.

5. Empirical Study

Suppose that an investor intends to track a stock index in a financial market allowing short selling. Through the Join-quant Data platform, we select daily data of the constituent

stocks in Shanghai Stock Exchange 50 (SSE 50) Index from Jun 11, 2018, to Dec 11, 2018. The reason why we choose data in a period of half a year elaborates as follows. The constituent stocks in SSE 50 change every half year or so, and keeping them as a group in an index apparently influences the dependency structure of assets, which further affects the optimal investment strategy. Therefore, when an investor intends to track SSE 50, it is reasonable to get the data in which the same 50 stocks always stay in SSE 50.

Then, we tag the 50 stocks from No. 1 to No. 50. After calculating, we obtain 124 observations of daily logarithm yield return, including 64 in-sample data spanning from Jun 11, 2018, to Sept 10, 2018, and 60 out-of-sample data spanning from Sept 11, 2018, to Dec 11, 2018. Here, we can find that the dimensionality of in-sample data meets the requirement of high-dimensional scenarios because $p = 50$ and $n = 64$ are comparable.

Since \mathcal{P}_5 outperforms \mathcal{P}_1 – \mathcal{P}_4 from Section 4, we calculate the optimal portfolios based on \mathcal{P}_5 in formula (13) and show the results in Figure 4. The parameter η is supposed to be 0, 0.5, and 1, respectively, and the target return μ_0 is set to be the average value of the return data of 50 stocks. Considering the idea in this paper is illuminated by the pioneering work of Roman et al. [12], we further conduct the corresponding analysis from \mathcal{P}_1 for comparison and the results are shown in Figure 5. It can be easily found that portfolios from \mathcal{P}_5 are much more sparse and can avoid extreme investment action, which will be more applicable to the practice of finance.

Next, we observe the out-of-sample performance of the portfolios. There is evidence that equal weight (EW) strategy cannot be defeated by many portfolio methods consistently (DeMiguel et al. [17]). We also add EW strategy into comparison here. The variance, CVaR_{0.95}, and average return of the portfolios are calculated, respectively; see Table 3. The accumulated returns of the portfolios in out-of-sample period are displayed in Figure 6.

From Table 3 and Figure 6(a), we can conclude that the portfolio generated from \mathcal{P}_5 has the highest terminal return and can control the risk robustly over the period, since it has a relatively slight fluctuation and avoids extreme losses and outperforms the EW strategy basically. However, the portfolio generated from \mathcal{P}_1 is unstable even compared with EW strategy and ends up with the lowest return in spite of being dominate at the beginning. Figure 6(b) also tells us that $\eta = 0.5$ in \mathcal{P}_5 will have a more promising terminal return compared with the case only considering CVaR ($\eta = 0$) or variance ($\eta = 1$).

6. Conclusion

In this paper, we have constructed five progressive optimization problems to figure out a relatively better way to conduct mean-variance-CVaR portfolio selection model in high-dimension scenarios. Specifically, two covariance matrix estimators (classical estimator and well-conditioned and sparse estimator), two estimation methods for CVaR (Rockafellar–Uryasev’s approximation and quantile regression), and two penalty functions (SCAD and reweighted L_1 norm) have been considered and different combinations of them have been used for the formulation of optimization problems. The data for simulation is generated from a hybrid distribution with multivariate normal distribution and exponential distribution. Moreover, the correlation matrix and mean vector of the multivariate normal distribution are estimated from real data. The simulation studies show that the well-conditioned and sparse estimator and penalty item devote to alleviate poor out-of-sample performance significantly

and the reweighted L_1 norm penalty has a slight superiority. Additionally, we find that the classical optimization methods for mean-variance-CVaR model in Roman et al. [12] may result in poor out-of-sample performance under high-dimension scenarios. An empirical study has been conducted based on the data of constituent stocks in SSE 50. The results solidify the foregoing findings. Therefore, we suggest to combine well-conditioned and sparse covariance estimator, quantile regression computational method for CVaR, and reweighted L_1 norm for portfolio optimization under mean-variance-CVaR criterion, since the fragile out-of-sample performance could be effectively mitigated.

The study in this paper focuses on the high-dimensional setting, where n is just a little bit greater than p . An extension of this study under a more extremely high-dimensional scenario in which $p > n$ would also be interesting. Moreover, it is noteworthy to further stabilize the out-of-sample performance by controlling the fragile brought by mean item. We will explore these problems in the following study.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by National Natural Science Foundation of China under Grant nos. 71671104 and 11971301, Key Project of National Social Science of China under Grant no. 16AZD019, and Incubation Group Project of Financial Statistics and Risk Management of SDUFE.

References

- [1] H. Markowitz, “Portfolio selection,” *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.
- [2] P. N. Kolm, R. Tütüncü, and F. J. Fabozzi, “60 years of portfolio optimization: practical challenges and current trends,” *European Journal of Operational Research*, vol. 234, no. 2, pp. 356–371, 2014.
- [3] H. Konno and H. Yamazaki, “Mean-absolute deviation portfolio optimization model and its applications to tokyo stock market,” *Management Science*, vol. 37, no. 5, pp. 519–531, 1991.
- [4] W. Ogryczak and A. Ruszczyński, “From stochastic dominance to mean-risk models: semideviations as risk measures,” *European Journal of Operational Research*, vol. 116, no. 1, pp. 33–50, 1999.
- [5] G. Consigli, “Tail estimation and mean-VaR portfolio selection in markets subject to financial instability,” *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1355–1382, 2002.
- [6] G. J. Alexander and A. M. Baptista, “Economic implications of using a mean-var model for portfolio selection: a comparison

- with mean-variance analysis,” *Journal of Economic Dynamics and Control*, vol. 26, no. 7-8, pp. 1159–1193, 2002.
- [7] Q. Xu, Y. Zhou, C. Jiang, K. Yu, and X. Niu, “A large CVaR-based portfolio selection model with weight constraints,” *Economic Modelling*, vol. 59, pp. 436–447, 2016.
- [8] A. G. Quaranta and A. Zaffaroni, “Robust optimization of conditional value at risk and portfolio selection,” *Journal of Banking & Finance*, vol. 32, no. 10, pp. 2046–2056, 2008.
- [9] H. Konno, H. Shirakawa, and H. Yamazaki, “A mean-absolute deviation-skewness portfolio optimization model,” *Annals of Operations Research*, vol. 45, no. 1, pp. 205–220, 1993.
- [10] H. Konno and K.-i. Suzuki, “A mean-variance-skewness portfolio optimization model,” *Journal of the Operations Research Society of Japan*, vol. 38, no. 2, pp. 173–187, 1995.
- [11] C. S. Robert and A. H. Philip, “On the direction of preference for moments of higher order than the variance,” *The Journal of Finance*, vol. 35, no. 4, pp. 915–919, 1980.
- [12] D. Roman, K. Darby-Dowman, and G. Mitra, “Mean-risk models using two risk measures: a multi-objective approach,” *Quantitative Finance*, vol. 7, no. 4, pp. 443–458, 2007.
- [13] J. Gao, Y. Xiong, and D. Li, “Dynamic mean-risk portfolio selection with multiple risk measures in continuous-time,” *European Journal of Operational Research*, vol. 249, no. 2, pp. 647–656, 2016.
- [14] Y. Shi, X. Zhao, and X. Yan, “Optimal asset allocation for a mean-variance-CVaR insurer under regulatory constraints,” *American Journal of Industrial and Business Management*, vol. 09, no. 07, pp. 1568–1580, 2019.
- [15] R. C. Green and B. Hollifield, “When will mean-variance efficient portfolios be well diversified?” *The Journal of Finance*, vol. 47, no. 5, pp. 1785–1809, 1992.
- [16] V. K. Chopra and W. T. Ziemba, “The effect of errors in means, variances, and covariances on optimal portfolio choice,” *The Journal of Portfolio Management*, vol. 19, no. 2, pp. 6–11, 1993.
- [17] V. DeMiguel, L. Garlappi, and R. Uppal, “Optimal versus naive diversification: how inefficient is the 1/NPortfolio strategy?” *Review of Financial Studies*, vol. 22, no. 5, pp. 1915–1953, 2009.
- [18] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.
- [19] P. J. Bickel and E. Levina, “Covariance regularization by thresholding,” *The Annals of Statistics*, vol. 36, no. 6, pp. 2577–2604, 2008.
- [20] T. T. Cai, M. Yuan, D. Giannone, and I. Loris, “Adaptive covariance matrix estimation through block thresholding,” *The Annals of Statistics*, vol. 40, no. 4, pp. 2014–2042, 2012.
- [21] A. J. Rothman, “Positive definite estimators of large covariance matrices,” *Biometrika*, vol. 99, no. 3, pp. 733–740, 2012.
- [22] O. Ledoit and M. Wolf, “A well-conditioned estimator for large-dimensional covariance matrices,” *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365–411, 2004.
- [23] A. Maurya, “A well conditioned and sparse estimate of covariance and inverse covariance matrix using a joint penalty,” *The Journal of Machine Learning Research*, vol. 17, no. 130, pp. 1–28, 2016.
- [24] J. Friedman, T. Hastie, and R. Tibshirani, “Sparse inverse covariance estimation with the graphical lasso,” *Biostatistics*, vol. 9, no. 3, pp. 432–441, 2008.
- [25] R. T. Rockafellar and S. Uryasev, “Conditional Value-at-Risk for general loss distributions,” *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443–1471, 2002.
- [26] R. T. Rockafellar and S. Uryasev, “Optimization of conditional value-at-risk,” *The Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.
- [27] G. W. Bassett, R. Koenker, and G. Kordas, “Pessimistic portfolio allocation and Choquet expected utility,” *Journal of Financial Econometrics*, vol. 2, no. 4, pp. 477–492, 2004.
- [28] A. E. B. Lim, J. G. Shanthikumar, and G.-Y. Vahn, “Conditional value-at-risk in portfolio optimization: coherent but fragile,” *Operations Research Letters*, vol. 39, no. 3, pp. 163–171, 2011.
- [29] A. Takeda and T. Kanamori, “A robust approach based on conditional value-at-risk measure to statistical learning problems,” *European Journal of Operational Research*, vol. 198, no. 1, pp. 287–296, 2009.
- [30] I. Kondor, S. Pafka, and G. Nagy, “Noise sensitivity of portfolio selection under various risk measures,” *Journal of Banking & Finance*, vol. 31, no. 5, pp. 1545–1573, 2007.
- [31] J. J. Gao and W. P. Wu, “Sparse and multiple risk measures approach for data driven mean-CVaR portfolio optimization model,” in *Optimization and Control for Systems in the Big-Data Era: Theory and Applications*, Springer, Berlin, Germany, 2017.
- [32] J. Brodie, I. Daubechies, C. De Mol, D. Giannone, and I. Loris, “Sparse and stable Markowitz portfolios,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 30, pp. 12267–12272, 2009.
- [33] J.-Y. Gotoh and A. Takeda, “On the role of norm constraints in portfolio selection,” *Computational Management Science*, vol. 8, no. 4, pp. 323–353, 2011.
- [34] B. Fastrich, S. Paterlini, and P. Winker, “Constructing optimal sparse portfolios using regularization methods,” *Computational Management Science*, vol. 12, no. 3, pp. 417–434, 2015.
- [35] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, “Coherent measures of risk,” *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.
- [36] J. Fan, J. Zhang, and K. Yu, “Vast portfolio selection with gross-exposure constraints,” *Journal of the American Statistical Association*, vol. 107, no. 498, pp. 592–606, 2012.
- [37] J. C. Emmanuel, B. W. Michael, and P. B. Stephen, “Enhancing sparsity by reweighted l_1 minimization,” *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877–905, 2008.
- [38] E. R. Lee, H. Noh, and B. U. Park, “Model selection via Bayesian information criterion for quantile regression models,” *Journal of the American Statistical Association*, vol. 109, no. 505, pp. 216–229, 2014.
- [39] L. Wang, Y. Kim, and R. Li, “Calibrating nonconvex penalized regression in ultra-high dimension,” *The Annals of Statistics*, vol. 41, no. 5, pp. 2505–2536, 2013.
- [40] Y. C. Wu and Y. F. Liu, “Variable selection in quantile regression,” *Statistica Sinica*, vol. 19, pp. 801–817, 2009.