



# Smoothed semicovariance estimation for portfolio selection

Davide Ferrari<sup>1</sup> · Sandra Paterlini<sup>2</sup> · Andrea Rigamonti<sup>3</sup> ·  
Alex Weissensteiner<sup>1</sup>

Received: 28 September 2023 / Accepted: 3 May 2024  
© The Author(s) 2024

## Abstract

Downside risk measures, such as semivariance, are essential for evaluating investment risk. Focusing on semivariance allows investors to emphasize loss mitigation without considering upside volatility as risk. However, minimizing the semivariance of a portfolio is an analytically intractable and numerically challenging problem due to the endogeneity of the parameters in the semicovariance matrix. We introduce a methodology for consistent estimation of the portfolio semivariance based on a smooth approximation of the empirical semicovariance matrix. Differently from existing methods, the new estimator does not rely on biased surrogate semicovariance models and enables the treatment of large problems with many assets. The extent of smoothing is determined by a single tuning constant, which allows our method to span an entire set of optimal portfolios with limit cases represented by the minimum semivariance and the minimum variance portfolios. The methodology is implemented through an iteratively reweighted algorithm, which is computationally efficient for high-dimensional problems with many assets. Our numerical studies confirm the theoretical convergence of the smoothed semivariance estimator to the traditional sample semivariance. The resulting minimum smoothed semivariance portfolio performs well in- and out-of-sample compared to other popular selection rules.

**Keywords** Semivariance · Smoothed semicovariance · Portfolio optimization · Skewness

**JEL Classification** G11 · D81

---

✉ Sandra Paterlini  
sandra.paterlini@unitn.it

Davide Ferrari  
davferrari@unibz.it

Andrea Rigamonti  
andrea.rigamonti@econ.muni.cz

Alex Weissensteiner  
alex.weissensteiner@unibz.it

<sup>1</sup> Faculty of Economics and Management, Free University of Bozen-Bolzano, Piazza Università 1, 39100 Bozen-Bolzano, Italy

<sup>2</sup> Department of Economics and Management, University of Trento, Via Inama 5, 38122 Trento, Italy

<sup>3</sup> Department of Finance, Masaryk University, Lipová 507/41a, 602 00 Brno, Czech Republic

# 1 Introduction

A strand of literature suggests that investors do not treat gains and losses in the same way. As a consequence, a distinction between “good” and “bad” target deviations is helpful for volatility forecasts, risk management, and asset pricing (for an exhaustive overview see Bollerslev, 2021). This distinction is crucial in financial markets to explain stylized facts such as fat tails, negative asymmetry, and non-normality in asset returns, and to construct more robust portfolio selection strategies. By recognizing the empirical regularities in asset returns, especially those implying negative asymmetry and aggregational Gaussianity over longer time horizons, loss-averse investors can use alternative portfolio selection rules like minimum semivariance optimization (see Cont, 2001, 2010; Ratliff-Crain et al., 2023).

Shortly after introducing the mean-variance framework, Markowitz (1959) formally describes the concept of semivariance of a portfolio and illustrates its theoretical advantages over the variance as a measure of risk. Some decades later, Sortino and Prince (1994) introduce the Sortino ratio, which is similar to the Sharpe ratio but relies on the downside risk, i.e., the square root of the semivariance, to quantify risk.

So far, research has mostly focused on mean-variance efficient portfolios and minimum-variance portfolios (see Zopounidis et al., 2014, and the references within the special issue on Markowitz’s contributions in portfolio theory), although targeting the downside volatility allows the investor to focus on loss minimization without considering the upside volatility as risk. Downside risk measures are better suited for assessing the investment risk of investors who consider potential losses relative to their target returns. For further elucidation on downside risk measures, see Harlow (1991) and Nawrocki (1991), and more recently Pla-Santamaria and Bravo (2013) and Klebaner et al. (2016).

The standard deviation is more widely used as a risk measure because calculating optimal semivariance portfolios is more challenging from an optimization perspective compared to using quadratic programming for the mean-variance framework (see Ballester, 2005; Estrada, 2008; Cumova & Nawrocki, 2011). In order to minimize the portfolio semivariance it is necessary to compute a semicovariance matrix. Differently from the covariance matrix, the semicovariance matrix only considers periods in which the portfolio underperforms the benchmark. As a consequence, the semicovariance matrix is endogenous in the sense that changes in the portfolio weights affect the periods in which the portfolio return underperforms the benchmark, which in turn determines the value of the semivariance. This means that the semivariances and semicovariances contained in the semicovariance matrix depend not only on a given set of data but also on the chosen portfolio weights.

From a mathematical viewpoint, the semivariance relies on a non-smooth indicator function to identify periods where the portfolio is below a certain threshold. As a consequence, analytical solutions do not exist, and naïve numerical optimization procedures, such as grid search, are only feasible for a relatively small number of assets. Previous attempts to estimate the semicovariance matrix exist in the literature (see Hogan & Warren, 1972; Nawrocki, 1991; Ballester, 2005; Estrada, 2008; Cumova & Nawrocki 2011). However, all available approaches suffer from serious drawbacks, including lack of matrix symmetry, presence of simplifying assumptions on the semivariance structure or on the preferences of investors, restrictions on the choice of the benchmark value, and poor performance in terms of approximation error (see Estrada, 2008; Cheremushkin, 2009; Rigamonti, 2020).

In this work, we introduce an estimator for the portfolio semivariance, which we refer to as smoothed semivariance (SSV) estimator in the rest of the paper. The main idea is to approximate the indicator function in the classical definition of the empirical semivariance by

a continuous function with accuracy of approximation controlled by a smoothing constant. The resulting estimator has the advantage to be a smooth function in the portfolio weights, while it can be made arbitrarily close to the usual semivariance estimator. We explore the minimization of the SSV estimator as a rule for portfolio selection. Depending on the value of the smoothing constant, the resulting minimum SSV portfolios allow spanning an entire set of solutions, ranging from the minimum semivariance to the portfolio obtained by equally weighting all the observations in the estimation set (i.e., without discriminating between upside and downside volatility). When the benchmark is equal to the sample mean, the latter case coincides with the minimum variance portfolio. In order to present this additional theoretical property, in this paper we focus on the special case in which the benchmark is equal to the mean. However, our methodology can be used with any benchmark.<sup>1</sup> To compute the minimum SSV portfolios, we introduce an easy-to-implement iteratively reweighted scheme, which allows tackling large problems with many assets.

To corroborate the precision of our SSV algorithm, for a small problem we show that our results are almost identical to the one obtained via grid search, which can be considered as the global optimum. Moreover, we illustrate that the estimated semicovariance matrix might suffer from greater parameter uncertainty than the estimated covariance matrix. This suggests that an estimator that perfectly separates positive and negative data points might not be optimal in terms of out-of-sample performance, and that intermediate levels of the SSV tuning parameter might lead to better out-of-sample properties. This is confirmed on simulated and on empirical data when comparing the SSV approach with benchmarks such as the global minimum-variance portfolio and the minimum semivariance portfolios obtained using the techniques proposed by Ballesterro (2005), Estrada (2008) and Cumova and Nawrocki (2011). Furthermore, in order to address parameter uncertainty in the out-of-sample exercise, we show that the proposed SSV approach can be easily combined with well-known shrinkage techniques.

The remainder of this paper is structured as follows. Section 2 describes the SSV estimator and its theoretical properties. Section 3 addresses the issue of portfolio selection through the optimization of the SSV and describes the algorithm. Sections 4 and 5 illustrate the properties of the new method using simulated and real data, respectively. Section 6 concludes and discusses future research. Proofs of propositions are deferred to the “Appendix A”. Robustness results are provided in “Appendix B”.

## 2 Methodology

### 2.1 Background and setup

Let  $X = (X_1, \dots, X_N)^\top$  be a random vector representing the returns for  $N$  assets with mean vector  $E(X) = \eta \in \mathbb{R}^N$  and covariance matrix  $cov(X) = \Sigma \in \mathbb{R}^{N \times N}$ . The portfolio return is defined by the weighted average  $Y(w) = w^\top X$ , where  $w = (w_1, \dots, w_N)^\top$  is the vector of portfolio weights to be determined and satisfying the budget constraint  $\sum_{j=1}^N w_j = 1$ . The expected return and variance of the portfolio are given by

$$\mu(w) = E[Y(w)] = w^\top \eta, \quad (1)$$

$$\sigma^2(w) = var[Y(w)] = w^\top \Sigma w. \quad (2)$$

<sup>1</sup> As robustness test, we report also empirical results for the benchmark being equal to the median return as in Bernard et al. (2019), see “Appendix B”.

The portfolio semivariance is defined as

$$\sigma_s^2(w) = E[\min\{Y(w) - b, 0\}^2] = E[(Y(w) - b)^2 I(Y(w) < b)], \quad (3)$$

where  $I(\cdot)$  represents the indicator function, and  $b$  is a benchmark set by the investor. The portfolio semivariance is similar to the variance, but it considers only the variability of portfolio returns below  $b$ .

Let  $X_1, \dots, X_T$  be  $T$  independent observations on  $X$ , where  $X_t = (X_{1t}, \dots, X_{Nt})^\top$  and  $X_{it}$  denotes the observation on asset  $i$  at time  $t$ . Consider the empirical semivariance estimator

$$\hat{\sigma}_s^2(w) = \frac{1}{T} \sum_{t=1}^T \min\{Y_t(w) - b, 0\}^2 = \frac{1}{T} \sum_{t=1}^T (Y_t(w) - b)^2 I(Y_t(w) < b), \quad (4)$$

where  $Y_t(w) = w^\top X_t$ . Particularly, we are interested in finding the optimal portfolio  $\hat{w}$  by solving the minimum semivariance problem

$$\min_w \hat{\sigma}_s^2(w) \quad \text{s.t.} \quad \sum_{j=1}^N w_j = 1. \quad (5)$$

The optimal portfolio weights obtained by solving (5) do not coincide with those of the minimum variance portfolio if asset returns are asymmetric and/or the benchmark differs from the mean (see Klebaner et al., 2016). Solving the optimization task in (5) is challenging as it relies on the non-smooth indicator function  $I(Y_t(w) < b)$ , which impedes the use of standard optimizers, and requires the semicovariance matrix as an endogenous input.

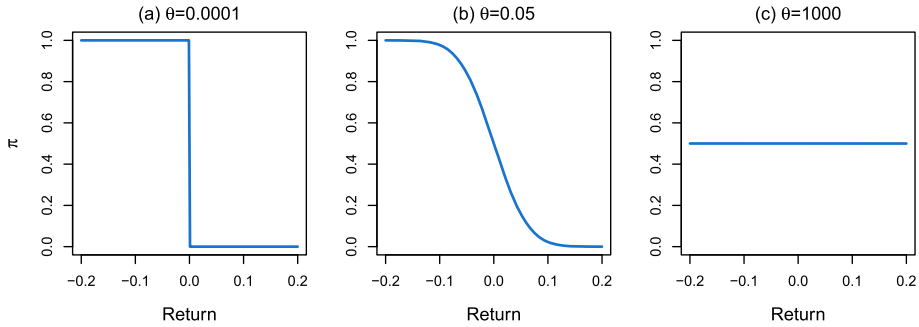
To address the above challenges, several algorithms based on linear or quadratic programming have been suggested to minimize the semivariance of a portfolio (for a review see Estrada, 2008). Taking a different approach, Estrada (2008) proposes to approximate the semicovariance matrix using a heuristic approach based on whether the single assets, and not the portfolio as a whole, underperform the benchmark. Although this yields a symmetric and positive semi-definite matrix which is easy to compute, the discrepancy between this heuristic definition and the actual semicovariance  $\sigma_s^2(w)$  can be substantial and generally increases with the number of assets. The approximation error implied by this heuristic approach cannot be neglected and implies a significant estimation bias and performance loss (see Cheremushkin, 2009; Rigamonti, 2020). Differently from existing approaches, we propose a smoothed version of the empirical semicovariance matrix. The estimator is shown to converge to the sample semicovariance estimator in (4). The proposed approximation can be conveniently optimized through a reweighting algorithm.

## 2.2 The smoothed semivariance estimator

Motivated by the above computational issues, we introduce the smoothed semivariance (SSV) estimator defined by

$$\hat{\sigma}_s^2(w, \theta) = \frac{1}{T} \sum_{t=1}^T (Y_t(w) - b)^2 \pi(Y_t(w) - b; \theta), \quad (6)$$

where  $\pi(\cdot; \theta)$  is a continuous function indexed by the smoothing parameter  $\theta$ , which we use to relax the binary weights  $I(Y_t(w) - b < 0)$  in (4). In this paper,  $\pi$  is given by the parametric



**Fig. 1** Smoothing weights  $\pi(z; \theta)$  for different values of the smoothing parameter  $\theta$ : (a)  $\theta = 0.0001$ , (b)  $\theta = 0.05$  and (c)  $\theta = 1000$

model

$$\pi(z; \theta) = \begin{cases} 1 - F(z/\theta), & \theta > 0 \\ I(z < 0), & \theta = 0, \end{cases} \tag{7}$$

where  $F(\cdot)$  is the distribution function of  $Y(w) - b$ , and the smoothing parameter  $\theta$  may be determined based on the sample. We let  $\theta = \theta_T$  be a positive sequence decreasing to zero with the sample size  $T$ . This choice, while ensuring smoothness of the estimator (6) with respect to  $w$ , it also makes the discrepancy between the smoothed estimator (6) and the traditional sample semivariance estimator (4) negligible as  $T$  diverges.

For  $\theta \rightarrow 0$ , the smoothing function  $\pi$  converges to the indicator function  $I(\cdot)$  in (4), and in the limit  $\hat{\sigma}_s^2(w; \theta)$  coincides with the traditional empirical semivariance  $\hat{\sigma}_s^2(w) = \hat{\sigma}^2(w; 0)$ , regardless of the benchmark  $b$  set by the investor; see Fig. 1a. If  $b = \bar{Y}(w) = T^{-1} \sum_{t=1}^T Y_t(w)$  and  $\theta \rightarrow \infty$ , the SSV weighs equally all the observations, and converges to the sample covariance estimator; see Fig. 1c. Therefore, we focus on the special case where the SSV is defined as

$$\hat{\sigma}_s^2(w, \theta) = \frac{1}{T} \sum_{t=1}^T [Y_t(w) - \bar{Y}(w)]^2 \pi(Y_t(w) - \bar{Y}(w); \theta), \tag{8}$$

with smooth function  $\pi(\cdot; \theta)$  as defined in (7). In Sect. 4, we show that, due to parameter uncertainty, the lowest downside deviation in an out-of-sample exercise might be achieved with an intermediate value of  $\theta$ . This corresponds to the situation in Fig. 1b.

The following proposition shows that the SSV estimator  $\hat{\sigma}^2(w, \theta)$  is asymptotically equivalent to the non-smoothed semivariance  $\hat{\sigma}_s^2(w)$ , and uniformly in  $w$ , as the sample size  $T$  increases. The proposition also provides information on the convergence rate.

**Proposition 2.1** *Let  $\theta = \theta_T$  be a sequence such that  $\theta_T \rightarrow 0$  as  $T \rightarrow \infty$ . Assume: (i)  $\{X_t, t \geq 1\}$  are iid vectors with finite moment up to second moment; and (ii) the portfolio  $Y_t = w^\top X_t$  has continuous and bounded density  $f(\cdot)$  such that  $yf(ay) \rightarrow 0$  as  $|y| \rightarrow \infty$  for all  $a \in \mathbb{R}^+$  and  $w \in \mathbb{R}^N$ . Then for any  $w$  such that  $\|w\|_2 < C, C > 0$ , we have*

$$|\hat{\sigma}_s^2(w, \theta_T) - \hat{\sigma}_s^2(w)| = O_p \left( \theta_T^{-1} f(k\theta_T^{-1}) \right), \tag{9}$$

where  $k > 0$  is some positive constant.

The main requirement is that the portfolio  $Y_t(w) = w^\top X_t$  has density with sub-linear tail behavior, meaning its tails go to zero quicker than  $y$  as  $y \rightarrow \infty$ . This is a relatively mild

condition satisfied by a large number of distributions used in financial modelling. Then, the rate of convergence between the smoothed and non-smoothed estimators is governed by the tail behavior. For example, when  $X_t$  follows the  $N$ -variate normal distribution  $\mathcal{N}_N(\eta, \Sigma)$ , then  $(Y_t(w) - b)$  follows the univariate normal distribution  $\mathcal{N}_1(w^\top \eta - b, w^\top \Sigma w)$ . In this case, for any  $a > 0$  and  $\|w\|_2 < C$  we have

$$yf(ay) = y \frac{1}{\sqrt{2\pi w^\top \Sigma w}} \exp \left\{ -\frac{a(y - w^\top \eta + b)^2}{2w^\top \Sigma w} \right\} \rightarrow 0, \quad \text{as } |y| \rightarrow \infty.$$

The tail of the normal distribution decreases exponentially fast in  $y^2$  and clearly satisfies the requirement of Proposition 2.1.

Proposition 2.1 contains information on the convergence rate of the smooth estimator, thus providing us with some guidance on how to set  $\theta_T$  in applications. For the case of normally distributed assets, the SSV estimator approximates the non-smoothed estimator up to a vanishing term of order  $\exp\{-k\theta_T^{-2}\}$ ,  $k > 0$ , in probability, because of the dominating effect of the exponential term in the normal pdf. This means that setting  $\theta_T = T^{-1/2}$  ensures an approximation of order  $O_p(\exp\{-kT\})$  for sufficiently large  $T$ . In real data applications, we recommend setting  $\theta_T = T^{-1/2}$  or even smaller, such as  $\theta = T^{-1}$ , as these choices are expected to have negligible impact on the statistical accuracy of either smoothed or non-smoothed estimators. Particularly, the standard deviation of such estimators is expected to be of order  $T^{-1/2}$  by the Central Limit Theorem.

### 3 Optimal portfolio selection

Consider a random sample  $X_1, \dots, X_T$  from  $X \sim \mathcal{N}_N(\eta, \Sigma)$  and define the sample covariance matrix  $\hat{\Sigma} = T^{-1} \sum_{t=1}^T (X_t - \bar{X})(X_t - \bar{X})^\top$ . The global minimum variance (GMV) portfolio  $\hat{w}_{mv}$  is found by minimizing the portfolio variance  $\hat{\sigma}^2(w) = w^\top \hat{\Sigma} w$  subject to the constraint  $\sum_{j=1}^N w_j = 1$ . The GMV portfolio has the closed-form expression  $\hat{w}_{mv} = \hat{\Sigma}^{-1} \mathbf{1} / (\mathbf{1}^\top \hat{\Sigma}^{-1} \mathbf{1})$ , where  $\mathbf{1}$  is the  $N \times 1$  vector of ones.

The semivariance minimization problem (5) can be similarly formulated. Particularly, note that the portfolio semivariance is given by  $\hat{\sigma}_s^2(w) = w^\top \hat{\Sigma}_s(w) w$ , where  $\hat{\Sigma}_s(w)$  is the  $N \times N$  semicovariance matrix

$$\hat{\Sigma}_s(w) = \frac{1}{T} \sum_{t=1}^T (X_t - \bar{X})(X_t - \bar{X})^\top I\{w^\top (X_t - \bar{X}) < 0\}. \quad (10)$$

The minimum semivariance portfolio  $\hat{w}_0$  is found by solving

$$\min_w w^\top \hat{\Sigma}_s(w) w \quad \text{s.t.} \quad \sum_{j=1}^N w_j = 1. \quad (11)$$

This is a challenging optimization task that cannot be tackled using gradient-based optimization techniques due to the non-smoothness of  $\hat{\Sigma}_s(w)$ . To address this issue, we replace  $\hat{\Sigma}_s(w)$  in (11) by the smoothed semicovariance matrix

$$\hat{\Sigma}_s(w; \theta) = \frac{1}{T} \sum_{t=1}^T (X_t - \bar{X})(X_t - \bar{X})^\top \pi \left( w^\top (X_t - \bar{X}); \theta \right), \quad (12)$$

where the function  $\pi(\cdot; \theta)$  is defined in (7). Thus, given  $\theta > 0$ , we compute the smoothed minimum semivariance portfolio by solving the approximate optimization task

$$\min_w w^\top \hat{\Sigma}_s(w; \theta) w \quad \text{s.t.} \quad \sum_{j=1}^N w_j = 1. \tag{13}$$

The solution to the above problem clearly depends on the choice of the smoothing parameter  $\theta$ , and we denote it by  $\hat{w}_\theta$  in the rest of the paper.

Note that the minimum semivariance portfolio  $\hat{w}_0$  and the minimum variance portfolio  $\hat{w}_{mv}$  are recovered as limit cases from the optimization (13). In particular, for  $\theta \rightarrow 0$ , we have  $\hat{\Sigma}_s(w, \theta) \rightarrow \hat{\Sigma}_s(w)$  for all  $w \in \mathbb{R}^N$ , meaning that the minimum SSV portfolio  $\hat{w}_\theta$  approximates the target (non-smooth) minimum semivariance portfolio  $\hat{w}_0$  for a sufficiently small  $\theta$ . For certain special cases, the minimum semivariance portfolio  $\hat{w}_0$  may be obtained by brute-force computation, for example using grid search. However, this is typically unfeasible in realistic applications where the number of assets  $N$  is moderate or large. On the other hand, our smoothing approach with  $\theta > 0$  can be applied to situations with arbitrary, and possibly very large, number of assets  $N$ . For  $\theta \rightarrow \infty$ , the SSV estimator converges to the sample minimum variance portfolio  $\hat{w}_{mv}$ .

The next proposition shows that the minimum SSV portfolio  $\hat{w}_\theta$  converges to the population minimum semivariance portfolio as long as  $\theta$  decreases to 0 with the sample size  $T$ . For the following analysis, we define the Lagrangian function  $L_s(v; \theta) = \sigma_s^2(w; \theta) + \lambda \sum_{j=1}^N w_j$  with  $v = (w^\top, \lambda)^\top$  and let  $v^* = (w^*, \lambda^*)^\top$  be the minimizer of  $L_s(v; 0)$ . Here  $w^*$  is the optimal portfolio that minimizes the theoretical non-smooth population semivariance (3) in the absence of sampling variability. Note that (13) can be written in terms of the empirical objective

$$\hat{L}_s(v; \theta) = \frac{1}{T} \sum_{t=1}^T \left\{ w^\top (X_t - \bar{X})(X_t - \bar{X})^\top \pi(w^\top (X_t - \bar{X}); \theta) w + \lambda \sum_{j=1}^N w_j \right\}.$$

Particularly,  $\hat{L}_s(v; \theta)$  can be regarded as an M-estimating function, with  $w$  representing a statistical parameter. From this viewpoint, the minimizer  $\hat{v}_\theta = (\hat{w}_\theta^\top, \hat{\lambda}_\theta)^\top$  of  $\hat{L}(w; \theta)$  can be viewed as an M-estimator, which enables us to exploit existing theory to analyze its properties.

**Proposition 3.1** *Assume: (i) for every  $\epsilon > 0$ ,  $\inf_{v: \|v-v^*\|_2 > \epsilon} L_s(v; 0) > L_s(v^*; 0)$ , and (ii)  $X_t$  has finite moments up to fourth moment. Then, under the conditions given Proposition 2.1,  $\hat{w}_{\theta_T} \rightarrow w^*$  in probability as  $T \rightarrow \infty$ .*

Consistency for  $\hat{w}_\theta$  with  $\theta$  decreasing to zero is obtained through standard arguments for M-estimators (e.g. see Theorem 5.7 in van der Vaart, 2000). The main assumptions for consistency of  $\hat{w}_\theta$  are the uniform convergence of the semivariance estimator shown in Proposition 2.1 and the uniqueness of the minimizer  $w^*$  of  $\sigma_s^2(w; 0)$  in a neighborhood of  $w^*$ .

### 3.1 Iteratively reweighted algorithm

The optimization task in (11) cannot be solved by a gradient-based approach due to the non-smoothness of the objective function. Solving (11) using a generic optimization algorithm may still be challenging when the number of elements in  $w$  is moderate or large. To address this

issue, we propose a reweighting optimization approach whereby, starting from some initial portfolio weights, we alternate the estimation of the weighted covariance matrix  $\hat{\Sigma}_s(w; \theta)$  and the solution of (13) to update the weights. Note that, given  $\hat{\Sigma}_s(w; \theta)$ , the optimization task in (13) is a standard minimum-variance problem which admits a closed-form solution for  $T > N$ . This reweighting strategy leads to a fast and easy-to-implement algorithm, which we describe next.

Let  $x_1, \dots, x_T$  be  $T$  realizations of the vector of asset returns with  $x_t = (x_{1t}, \dots, x_{Nt})^\top$  and  $x_{it}$  denotes the realization of the return for asset  $i$  at time  $t$ . Moreover, define  $z_t(w) = w^\top (x_t - \bar{x})$ , with  $\bar{x} = T^{-1} \sum_{t=1}^T x_t$ . We use the superscript  $[k]$  to denote the  $k$ -th step of our algorithm, i.e.,  $w^{[k]} = (w_1^{[k]}, \dots, w_N^{[k]})'$  are the portfolio weights obtained at Step  $k$ . For a given  $\theta$ , we carry out the following steps:

Algorithm 1: Iteratively reweighted algorithm for minimum SSV portfolio

0. *Initialization.* Set  $k = 0$  and  $\pi_t^{[0]} = I(z_t(w^{[0]} < 0))$ , for  $t = 1, \dots, T$ , where  $w^{[0]}$  are initial weights. For example, set  $w_1^{[0]} = \dots = w_N^{[0]} = 1/N$ .
1. *Parametric smoothing update.* Compute  $\pi_t^{[k]} = \pi(z_t(w^{[k]}; \theta); \theta)$ , for  $1 \leq t \leq T$ , where  $\pi(\cdot; \theta)$  is the parametric smoother in (7).
2. *Update of the optimal portfolio.* Given  $\pi_t^{[k]}$ ,  $t = 1, \dots, T$ , update the portfolio weights  $\hat{w}^{[k+1]}$  by solving

$$\min_w \hat{\sigma}_s^{2[k]}(w) = w^\top \hat{\Sigma}_s^{[k]} w \quad \text{s.t.} \quad \sum_{j=1}^N w_j = 1, \quad (14)$$

where  $\hat{\Sigma}_s^{[k]} = T^{-1} \sum_{t=1}^T (x_t - \bar{x})(x_t - \bar{x})^\top \pi_t^{[k]}$  is the sample covariance matrix with observations reweighted using the smoothing constants  $\pi_t^{[k]}$ ,  $t = 1, \dots, T$ . The solution of (14) is

$$\hat{w}^{[k+1]} = \frac{\left(\hat{\Sigma}_s^{[k]}\right)^{-1} \mathbf{1}}{\mathbf{1}^\top \left(\hat{\Sigma}_s^{[k]}\right)^{-1} \mathbf{1}}.$$

3. Set  $k \leftarrow k + 1$  and repeat Steps 1 and 2 until a convergence criterion is met; e.g.,  $\|\hat{w}^{[k+1]} - \hat{w}^{[k]}\|_2 / \|\hat{w}^{[k]}\|_2 < \tau$ , for some tolerance  $\tau > 0$ .

A greater degree of flexibility is achieved by replacing Step 1 with non-parametric weights computed by kernel smoothing.

- 1'. *Non-parametric smoothing update.* Compute  $\pi_t^{[k]} = 1 - \hat{F}(z_t(w^{[k]})/\theta)$ , for  $1 \leq t \leq T$ , where  $\hat{F}$  is a non-parametric estimator for the cdf of  $Z_t(w^{[k]}) = Y_t(w^{[k]}) - \bar{Y}(w^{[k]})$  with observations  $z_1(w^{[k]}), \dots, z_t(w^{[k]})$ .

To find the final solution, the above algorithm solves a sequence of convex optimization tasks. In particular, note that Step 2 corresponds to a minimum variance portfolio optimization task where the usual sample covariance matrix is replaced by a weighted covariance matrix. Since the solution is available in closed-form, the computational time for this step is usually negligible. Based on our numerical experiments, the algorithm converges very quickly within a few iterations, even when the number of assets is large. Finally, note that our procedure is not limited to the constraint  $\sum_{j=1}^N w_j = 1$ . A different set of constraints on the portfolio weights may be included in Step 2 of the algorithm, depending on the application at hand. Then a quadratic programming solver may be used to compute the update  $\hat{w}^{[k+1]}$  instead of the explicit formula given above.



## 4 Simulation study

In this section, we study the performance of the SSV estimator compared to the global minimum variance portfolio (GMV) and the following competing semivariance estimation algorithms: Estrada (2008), Cumova and Nawrocki (2011), Ballester (2005), both in-sample and out-of-sample. First, we compare our method with the minimum semivariance portfolio obtained via grid search, which can be considered as the global optimum for computationally feasible cases. To this end, we focus on a simple setting with two assets only, and we consider both assets having either negatively skewed or positively skewed returns. We generate 100 000 returns from a multivariate skew-normal with the skewness of both assets equal to  $-0.5$  or  $0.5$  (see Azzalini and Dalla Valle, 1996).<sup>2</sup> The variance of the two assets is set equal to 0.0050 and 0.0034 respectively, while their covariance is set equal to 0.0019.<sup>3</sup> We set the benchmark  $b = \bar{Y}(w)$ , such that the SSV estimator converges to the sample covariance matrix estimator when  $\theta \rightarrow \infty$ , and we evaluate the results based on the portfolio downside deviation.

We use rolling windows of different lengths to compare our SSV approach with various other proposed in the literature. We rely on a parametric smoothing based on a normal cdf, testing three different values for the smoothing parameter  $\theta = 1/T$ , which gives a portfolio very close to the one obtained with grid search,  $\theta = 100$ , a high value that results in a portfolio close to the sample minimum variance portfolio, and an intermediate value  $\theta = 1/\sqrt{T}$ . These choices for  $\theta$  are corroborated by our theoretical analysis in Proposition 2.1. We consider also other portfolio selection rules: the minimum semivariance portfolio obtained via grid search, the sample global minimum variance portfolio, and the minimum semivariance portfolios obtained with the techniques proposed by Ballester (2005), Estrada (2008) and Cumova and Nawrocki (2011).

Figure 2 shows the in-sample and the out-of-sample results with negatively skewed (Panels a and b) and with positively skewed (Panels c and d) returns. In the in-sample analysis, all strategies except for Cumova and Nawrocki (2011) achieve a lower downside deviation with shorter rather than longer estimation windows. This is due to the fact that, instead of the true (unknown) inputs, sample estimates are used in the calculations. In other words, overfitting creates the appearance of better results in-sample when the sample size is not very large. In case of multivariate normal distributed returns the size of the effect could be derived analytically.<sup>4</sup> For longer estimation windows, as the sample estimates converge to the true values, this effect wanes out.

In the out-of-sample evaluation, the total downside deviation is larger for shorter estimation windows due to the unknown input parameters.<sup>5</sup> As the estimation window increases, the downside deviation converges to the true downside deviation under known parameters. Asymptotically, the in-sample and the out-of-sample estimates of the downside deviation tend to the same values.

For all panels in Fig. 2, the dashed black line, which corresponds to the SSV estimator with  $\theta = 1/T$ , perfectly overlaps the green line, obtained by minimizing the semivariance via grid search and which gives the global optimum. This means that our estimator provides

<sup>2</sup> We randomly generated the returns using the R package “sn”, see Azzalini (2020).

<sup>3</sup> We use the sample covariance matrix of two of the 30 industry portfolios from July 1926 to November 2022.

<sup>4</sup> For the minimum-variance portfolio, the relationship between the estimated variance  $\hat{\sigma}_T^2 = \hat{w}'\hat{\Sigma}\hat{w}$  and the true variance  $\sigma^2 = w'\Sigma w$  is given by  $T\hat{\sigma}_T^2/\sigma^2 \sim \chi_{T-N}^2$ , see Kempf and Memmel (2006) and Frahm and Memmel (2010).

<sup>5</sup> To quantify the effect for the minimum-variance portfolio with multivariate-normal distributed returns, see Frahm and Memmel (2010).

**Table 1** In-sample, monthly data, benchmark equal to sample mean return

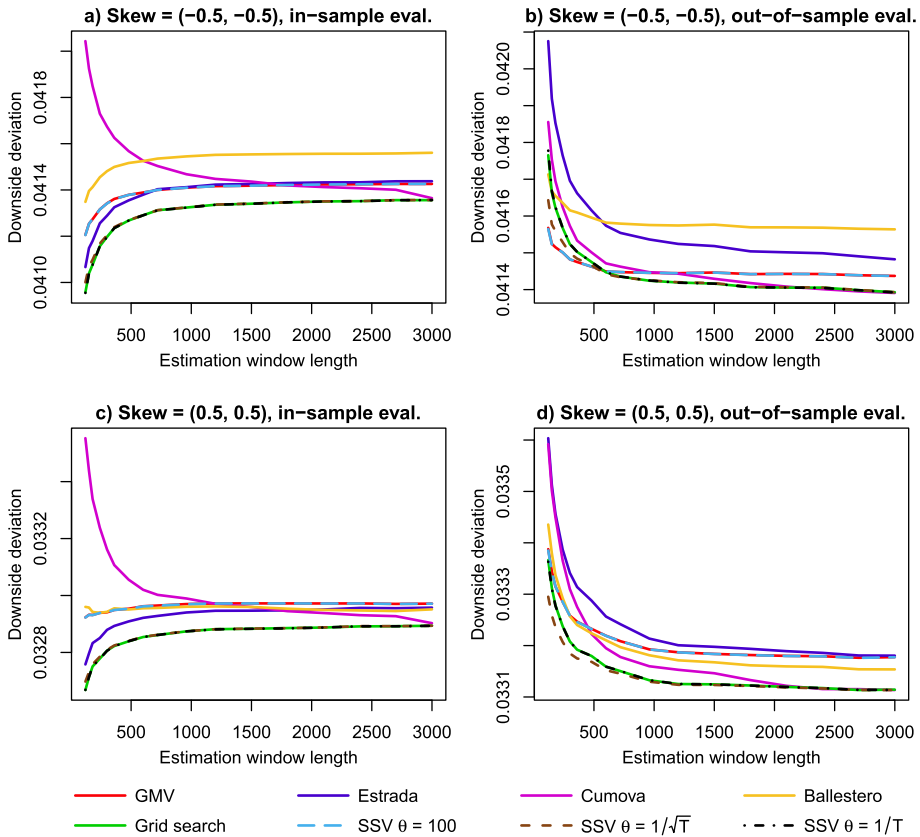
Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt	-	0.0307	0.0431	-0.4867	1.6609	0.0954	0.1466	-
17 Ind	GMV	0.0214	0.0317	-0.2299	0.8857	0.1518	0.1682	0.1600
	Ballestero	0.0219	0.0321	-0.3192	1.3341	0.1445	0.1721	0.1299
	Cumova	0.0208	0.0326	0.0674	0.6286	0.1909	0.1748	0.2145
	Estrada	0.0218	0.0344	0.1365	0.4597	0.1856	0.1582	0.2358
	SSV $\theta = 1/T$	<b>0.0205</b>	0.0326	0.1314	0.9368	0.2070	0.1861	0.1933
	SSV $\theta = 1/\sqrt{T}$	0.0206	0.0324	0.1092	1.1046	0.1952	0.1820	0.1778
30 Ind	SSV $\theta = 100$	0.0214	0.0317	-0.2297	0.8855	0.1519	0.1682	0.1600
	GMV	0.0185	0.0283	-0.1175	0.5972	0.1828	0.1895	0.2358
	Ballestero	0.0190	0.0286	-0.2059	0.9567	0.1730	0.1936	0.2002
	Cumova	0.0183	0.0301	0.3018	0.4393	0.2122	0.1739	0.3264
	Estrada	0.0202	0.0332	0.2641	0.6667	0.2167	0.1437	0.3288
	SSV $\theta = 1/T$	<b>0.0174</b>	0.0299	0.4880	0.9789	0.2440	0.1913	0.2823
	SSV $\theta = 1/\sqrt{T}$	0.0176	0.0292	0.3818	1.0604	0.2252	0.1902	0.2563
	SSV $\theta = 100$	0.0185	0.0283	-0.1172	0.5969	0.1828	0.1895	0.2358

Table 1 continued

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
48 Ind	GMV	0.0186	0.0269	-0.3446	0.2586	0.1303	0.1756	0.3575
	Ballesterro	0.0188	0.0268	-0.4470	0.5252	0.1314	0.1875	0.3135
	Cumova	0.0192	0.0300	0.2559	0.3685	0.1261	0.1493	0.5479
	Estrada	0.0231	0.0359	-0.0118	0.4511	0.1685	0.1484	0.5096
	SSV $\theta = 1/T$	<b>0.0166</b>	0.0281	0.5745	1.3036	0.1864	0.1704	0.4596
	SSV $\theta = 1/\sqrt{T}$	0.0170	0.0273	0.3731	1.1662	0.1661	0.1746	0.4009
SBM 100 (70)	SSV $\theta = 100$	0.0186	0.0269	-0.3443	0.2581	0.1303	0.1756	0.3575
	GMV	0.0187	0.0290	-0.1811	0.3679	0.2142	0.2573	0.8141
	Ballesterro	0.0193	0.0292	-0.2824	0.6040	0.1965	0.2594	0.7102
	Cumova	0.4005	0.4841	-25.4689	747.4059	-0.0158	-0.0091	39.2733
	Estrada	0.0244	0.0391	0.0328	1.6265	0.2131	0.1626	1.3823
	SSV $\theta = 1/T$	<b>0.0171</b>	0.0312	0.6482	0.7032	0.2509	0.2565	1.1220
SSV $\theta = 1/\sqrt{T}$	SSV $\theta = 1/\sqrt{T}$	0.0174	0.0303	0.5043	0.8566	0.2399	0.2547	0.9442
	SSV $\theta = 100$	0.0187	0.0290	-0.1806	0.3673	0.2142	0.2573	0.8141

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballesterro (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T, 1/\sqrt{T}, 100$ . Portfolio returns spanning from April 1989 to April 2023 for the 48 industry portfolios; from July 1946 to April 2023 for the other datasets

Best DD values reported in bold

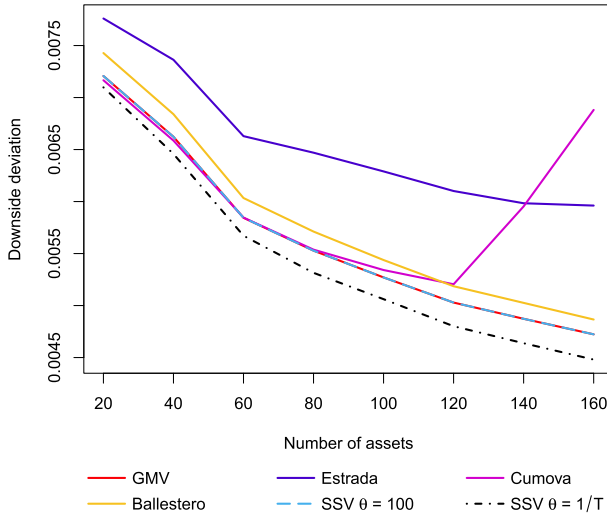


**Fig. 2** The figure illustrates the in-sample (left panels) and out-of-sample (right panels) downside deviation achieved by minimum variance and minimum semivariance portfolios with two assets and skewness of both assets equal to  $-0.5$  (top panels) and  $0.5$  (bottom panels). Competing semivariance estimation algorithms: Estrada (2008), Cumova and Nawrocki (2011), Ballestero (2005)

an extremely precise approximation of the true sample semicovariance matrix. The SSV estimator with  $\theta = 100$ , indicated as dashed light blue line, matches the red line obtained by using the sample covariance matrix. As expected, the SSV estimator corresponds to the sample covariance estimator when  $\theta$  is sufficiently large. Targeting the semivariance always gives the best results in-sample, but using the variance or an intermediate approach given by smoothing with  $\theta = 1/\sqrt{T}$  (dashed brown line) can work better out-of-sample, especially for negatively skewed returns and short estimation windows, where parameter uncertainty is high.<sup>6</sup> Overall, our strategy significantly outperforms the three competitor semicovariance matrix estimation algorithms across all considered settings in approximating the target sample covariance matrix.

Besides considering different estimation windows, we also study the in-sample effect of varying the number of assets. We generate 63,000 returns for a number of assets  $N \in \{20, 40, \dots, 160\}$  by randomly sampling with replacement daily returns of S&P 500

<sup>6</sup> A positive skewness reduces the estimation error as more observations—used in the calculation of the downside deviation—are below the mean (for a more in-depth discussion see Rigamonti, 2020).



**Fig. 3** Downside deviation for positively skewed data when varying the number of assets  $N$ . Smoothing with  $\theta = 1/\sqrt{T}$  not plotted to improve the clarity of the illustration. Competing semivariance estimation algorithms: Estrada (2008), Cumova and Nawrocki (2011), Ballesterio (2005)

constituents with data spanning from November 2, 1999, to April 28, 2023.<sup>7</sup> As we use the same sample of days for all assets, we capture the cross-correlation of stock returns. We use a rolling estimation window of 5 years (i.e., 1 260 days) to calculate the asset weights.<sup>8</sup> We then compute the downside deviation achieved by the different portfolio selection rules. Results are reported in Fig. 3.

An interesting finding is that the approximation from Cumova and Nawrocki (2011) performs well with small portfolios, but it no longer works as  $N$  gets larger as the estimated semicovariance matrix is quasi-singular. The Estrada (2008) heuristic generally performs very poorly, and the algorithm proposed by Ballesterio (2005), although much better than Estrada (2008), also fails to beat the sample covariance matrix. The SSV, on the other hand, matches the covariance matrix with  $\theta = 100$  and significantly improves over it with  $\theta = 1/T$ . Moreover, the improvement over the other strategies gets larger as  $N$  grows. For the largest portfolio, the SSV estimation time with  $\theta = 1/T$  and required convergence precision set at 0.01% (i.e., we stop the iterations when the difference between the portfolio weights estimated in the current and previous round is less than 0.01%) is only 0.27 s. This is significantly less than the 0.67 s required by Cumova and Nawrocki (2011), although Ballesterio (2005) and Estrada (2008) are by far the fastest methods, with 0.041 and 0.042 s respectively. The exercise was carried out using a computer equipped with an Intel Core i7-8565U processor and 24 GB of RAM.

<sup>7</sup> The dataset from which we sample includes 350 stocks whose returns are available for the entire time-period considered. We sample from the first  $N$  stocks sorted alphabetically according to their ticker symbol. Data were downloaded from Alpha Vantage: <https://www.alphavantage.co/>.

<sup>8</sup> This is a typical estimation window for daily data. See e.g. Lim et al. (2011), who suggests not using daily returns older than 5 years to avoid including outdated information in the estimation sample.

## 5 Real data application

We now consider real-world financial data to evaluate the in-sample and the out-of-sample performance of the minimum SSV portfolios. As in our Monte Carlo simulations, for the in-sample results of the SSV portfolios we use the normal cdf with  $\theta = 1/T$ ,  $\theta = 1/\sqrt{T}$  and  $\theta = 100$ . According to the theoretical analysis in Proposition 2.1, a choice for the smoothing parameter  $\theta$  of order  $1/\sqrt{T}$ , or smaller, provides an accurate approximation of the empirical SSV to the standard non-smoothed empirical SV assuming normally distributed portfolios. We compare these results with the global minimum variance portfolio and the minimum semivariance portfolios obtained with the methods of Ballestero (2005), Estrada (2008) and Cumova and Nawrocki (2011). We employ monthly returns of the 17, 30 and 48 industry portfolios, and those of the 100 portfolios formed on size and book-to-market, all downloaded from the Kenneth R. French data library.<sup>9</sup> The 48 industry portfolios returns start from July 1969, as the dataset contains missing data before that date. The other datasets start from July 1926, but for the 100 portfolios formed on size and book-to-market we do not consider 30 portfolios that contain missing data, and therefore for this dataset we have  $N = 70$ . All datasets span up to April 2023.

For estimation, we use a rolling window of 20 years, i.e.,  $T = 240$ , obtaining 406 portfolio returns for the 48 industry portfolios and 922 returns for the other datasets. To calculate realized portfolio returns, for the in-sample analysis we use the asset returns of the last month  $x_T$  in our estimation sample, i.e., the rolling window, while for the out-of-sample analysis we use the returns  $x_{T+1}$ . As in our simulation experiments, we set the benchmark equal to the sample mean such that for  $\theta \rightarrow \infty$  the SSV estimator perfectly converges to the sample covariance estimator. For the in- and out-of-sample calculated portfolio returns, in addition to the downside deviation (DD), we report the standard deviation (SD), skewness, excess kurtosis, the Sortino ratio and the Sharpe ratio. To better compare the last two measures, we work with excess returns over the risk-free rate from the Kenneth R. French data library. Furthermore, to consider potential trading costs of the various portfolio selection rules, we report the average turnover ( $TO$ ). The turnover at time  $t$  is  $TO_t = \sum_{i=1}^N |w_i(t) - w_i(t-1)|$ , where each weight  $w_i(t-1)$  is adjusted for the effect of the returns realized in the previous period, as in DeMiguel et al. (2009). A  $TO$  of 1 means that, on average, an amount of assets equal to 100% of the wealth has to be traded.

For the out-of-sample exercise, we present only the SSV results for minimizing the semi-variance of the portfolio with  $\theta = 1/T$ .<sup>10</sup> In addition to the comparison with the other strategies used in Table 1, we also use our method in combination with a *shrinkage*-technique to consider parameter uncertainty and to improve the out-of-sample results.<sup>11</sup> In particular, when estimating the sample covariance matrix (for the GMV strategy) and the sample semi-covariance matrix (for the SSV and the Ballestero, 2005, strategies) we shrink the implied correlation matrix  $R$  towards the identity matrix  $\mathbf{I}$ . For the minimum-variance portfolio strategy (GMV), we decompose the sample covariance matrix in a vector of sample standard deviations  $\hat{\sigma}$  and a sample correlation matrix  $\hat{R}$ . The shrunk correlation matrix is then given by  $\hat{R}\hat{S} = \delta\hat{R} + (1 - \delta)\mathbf{I}$ , with  $\mathbf{I}$  the identity matrix and  $0 < \delta < 1$ . For the calculation of the portfolio weights  $\hat{w}$ , we use the *shrunk* covariance matrix  $\hat{\Sigma} := \text{diag}(\hat{\sigma})\hat{R}\hat{S}\text{diag}(\hat{\sigma})$ . We

<sup>9</sup> [https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

<sup>10</sup> Results for  $\theta = 1/\sqrt{T}$  and  $\theta = 100$  are available upon request. Again, the SSV results for  $\theta = 100$  correspond to those of the minimum variance portfolios.

<sup>11</sup> While using an intermediate  $\theta$  value can also improve results out-of-sample, as shown in the simulation, choosing the best value is not straightforward and left for future research.

**Table 2** Out-of-sample, monthly data, benchmark equal to sample mean return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt 17 Ind	-	0.0308	0.0431	-0.4867	1.6609	0.0950	0.1466	-
	GMV	0.0238	0.0354	-0.0812	1.5337	0.1425	0.1546	0.1569
	Ballestero	0.0240	0.0351	-0.2052	1.8874	0.1365	0.1600	0.1270
	Cumova	0.0245	0.0364	-0.1637	0.8738	0.1532	0.1504	0.2115
	Estrada	0.0252	0.0378	-0.0825	0.7429	0.1521	0.1381	0.2321
	SSV $\theta = 1/T$	0.0241	0.0361	-0.1220	1.1625	0.1673	0.1622	0.1905
	SSV $\theta = 1/\sqrt{T}$	0.0240	0.0359	-0.1137	1.2547	0.1612	0.1598	0.1750
	SSV $\theta = 100$	0.0238	0.0354	-0.0813	1.5333	0.1425	0.1547	0.1569
	GMV (shrink)	0.0236	0.0344	-0.2421	1.8802	0.1400	0.1762	0.0686
	Ballestero (shrink)	0.0243	0.0351	-0.3344	2.1697	0.1315	0.1735	0.0607
30 Ind	SSV $\theta = 1/T$ (shrink)	<b>0.0233</b>	0.0343	-0.2587	1.5559	0.1544	0.1830	0.0756
	GMV	0.0225	0.0342	0.0088	1.5192	0.1589	0.1623	0.2316
	Ballestero	0.0225	0.0337	-0.0875	1.7887	0.1528	0.1686	0.1961
	Cumova	0.0236	0.0359	-0.0150	0.5595	0.1551	0.1394	0.3224
	Estrada	0.0246	0.0381	-0.0221	0.8005	0.1836	0.1289	0.3237
	SSV $\theta = 1/T$	0.0228	0.0351	0.0555	1.0823	0.1772	0.1565	0.2788
	SSV $\theta = 1/\sqrt{T}$	0.0225	0.0345	0.0227	1.0992	0.1713	0.1577	0.2523
	SSV $\theta = 100$	0.0225	0.0342	0.0088	1.5189	0.1589	0.1623	0.2316
	GMV (shrink)	0.0218	0.0327	-0.1177	1.6893	0.1632	0.1827	0.1004
	Ballestero (shrink)	0.0225	0.0332	-0.2356	2.0551	0.1516	0.1819	0.0908
SSV $\theta = 1/T$ (shrink)	<b>0.0216</b>	0.0326	-0.0861	1.4924	0.1722	0.1803	0.1071	

Table 2 continued

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
48 Ind	GMV	0.0262	0.0366	-0.3829	0.7546	0.0780	0.1184	0.3558
	Ballesterro	0.0255	0.0355	-0.4569	1.0690	0.0831	0.1318	0.3112
	Cumova	0.0287	0.0411	-0.1547	0.5777	0.0675	0.0972	0.5523
	Estrada	0.0297	0.0437	-0.2450	0.2872	0.1429	0.1296	0.5074
	SSV $\theta = 1/T$	0.0270	0.0384	-0.2170	1.5171	0.0905	0.1076	0.4590
	SSV $\theta = 1/\sqrt{T}$	0.0263	0.0372	-0.2460	1.0632	0.0869	0.1133	0.3988
	SSV $\theta = 100$	0.0262	0.0366	-0.3828	0.7546	0.0780	0.1184	0.3558
	GMV (shrink)	0.0234	0.0329	-0.4978	1.5713	0.1118	0.1855	0.1474
	Ballesterro (shrink)	0.0238	0.0332	-0.5929	1.8271	0.1100	0.1864	0.1351
	SSV $\theta = 1/T$ (shrink)	<b>0.0231</b>	0.0330	-0.3134	1.5526	0.1124	0.1793	0.1586
SBM 100 (70)	GMV	0.0286	0.0420	-0.2515	1.2314	0.1386	0.1769	0.8115
	Ballesterro	0.0281	0.0407	-0.3137	1.4723	0.1337	0.1852	0.7071
	Cumova	0.4060	0.8720	17.5083	539.1981	0.0407	0.0212	27.4596
	Estrada	0.0335	0.0511	-0.0012	1.3866	0.1691	0.1335	1.3789
	SSV $\theta = 1/T$	0.0310	0.0452	-0.2824	1.0656	0.1251	0.1684	1.1202
	SSV $\theta = 1/\sqrt{T}$	0.0297	0.0436	-0.2435	0.8673	0.1318	0.1710	0.9442
	SSV $\theta = 100$	0.0286	0.0420	-0.2515	1.2312	0.1386	0.1769	0.8116
	GMV (shrink)	<b>0.0262</b>	0.0377	-0.2997	1.3193	0.1205	0.1746	0.2284
	Ballesterro (shrink)	0.0268	0.0380	-0.3794	1.6702	0.1121	0.1801	0.2154
	SSV $\theta = 1/T$ (shrink)	<b>0.0262</b>	0.0379	-0.2644	1.0810	0.1186	0.1731	0.2376

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballesterro (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T$ , GMV with shrunk covariance matrix, GMV (shrink), Ballesterro with shrunk covariance matrix, Ballesterro (shrink), and SSV  $\theta = 1/T$  with shrunk semicovariance matrix, SSV (shrink), are based on a shrunk correlation matrix  $R^s$ , with  $R^s = 0.8R + (1 - 0.8)\mathbf{I}$ ,  $R$  being the sample correlation matrix and  $\mathbf{I}$  the  $N \times N$  identity matrix. Portfolio returns spanning from April 1989 to April 2023 for the 48 industry portfolios; from July 1946 to April 2023 for the other datasets. Best DD values reported in bold



apply the same shrinkage procedure for the estimated semicovariance matrix of Ballestero (2005), and iteratively in our SSV approach (Algorithm 1). For each loop  $k$ , after estimating the sample semicovariance matrix  $\hat{\Sigma}_s^{[k]} = T^{-1} \sum_{t=1}^T (x_t - \bar{x})(x_t - \bar{x})^\top \pi_t^{[k]}$ , we decompose  $\hat{\Sigma}_s^{[k]}$  in  $\hat{\sigma}_s^{[k]}$  and  $\hat{R}_s^{[k]}$ , we shrink the semi-correlation matrix  $\widehat{RS}_s^{[k]} = \delta \hat{R}_s^{[k]} + (1 - \delta)\mathbf{I}$ , and we use the semicovariance matrix  $\hat{\Sigma}_s^{[k]} := \text{diag}(\hat{\sigma}_s^{[k]}) \widehat{RS}_s^{[k]} \text{diag}(\hat{\sigma}_s^{[k]})$  to calculate the portfolio weights.

The aim of this exercise is to show that our SSV approach can be easily combined with a shrinkage approach, with greater benefit compared to the Ballestero (2005) approach, which is a difficult benchmark to beat out-of-sample. An exhaustive comparison of the proposed minimum-semivariance strategy (SSV) under different shrinkage approaches is beyond the scope of the present paper, and it is left for future research.

Tables 1 and 2 report the in-sample and out-of-sample results, respectively, with the lowest downside deviation for each data set in bold. The first row of each table reports values for the broad US market as defined in the Kenneth R. French dataset, which uses all NYSE, AMEX and NASDAQ firms to compute its returns. Almost all portfolio selection rules improve out-of-sample over the broad market portfolio: They show lower downside deviation, standard deviation, and excess kurtosis, and higher skewness, Sortino and Sharpe Ratio.

In Table 1, when considering the in-sample downside deviation (DD), parameter uncertainty is avoided, and the SSV estimator is a very close approximation to the traditional sample semicovariance matrix. The SSV with  $\theta = 1/T$  shows the best results in terms of DD. On the other hand, the other three semicovariance matrix estimation methods suffer because of the worse approximation of the semicovariance matrix compared to the proposed SSV method. In line with intuition, GMV shows the lowest standard deviation among the competitors. SSV with  $\theta = 100$  matches, as expected, the GMV portfolio, showing the flexibility of the proposed approach. In terms of turnover, the SSV strategy is competitive with the other portfolio selection rules. Despite the estimation window of 240 observations is larger than the portfolio size, for an increasing number of assets the estimated semicovariance matrix using Cumova and Nawrocki (2011) is quasi-singular. As a consequence, the extreme portfolio weights imply that this method strongly underperforms for SBM 100.

The positive performance of the SSV approach also largely holds out-of-sample, see Table 2. While the SSV performs worse than the GMV and the Ballestero (2005) approaches without shrinkage, it still significantly improves over Estrada (2008) and Cumova and Nawrocki (2011). As mentioned before, the performance of the latter strategy quickly degrades as  $N$  increases, leading to very poor results with  $N = 70$ . The good performance of GMV reflects the trade-off between minimizing downside deviation and minimizing the estimation error in a out-of-sample exercise where the data generating process changes over time. Although GMV uses another objective function, i.e., optimizes over variance rather than semivariance, it benefits from using all the observations for estimation. As in our data sets returns are not heavily skewed, the benefit from using all observations and having lower estimation errors might outweigh the drawback of using variance instead of semivariance. When combined with shrinking, the SSV approach outperforms for all datasets, except for the largest dataset, where it achieves the same downside deviation as the GMV. Furthermore, the turnover more than halves when using shrinkage, with basically the same values for SSV  $\theta = 1/N$  (shrink) and GMV (shrink).

As a robustness analysis, in ‘‘Appendix B’’ we report additional in- and out-of-sample results for daily and quarterly data. For daily data we use an estimation window of 1260

observations, i.e., approximately 5 years.<sup>12</sup> For quarterly data we use 80 observations in the estimation window, i.e., 20 years (as with monthly data). Furthermore, for all data frequencies we also report results for the benchmark being equal to sample median return (see Bernard et al., 2019). Our findings remain unaltered for these additional settings. The problem affecting the Cumova and Nawrocki (2011) algorithm disappears with daily data (due to the much larger sample size used for estimation), while it becomes more serious with quarterly data (as fewer observations are used for the estimation). Again, out-of-sample most portfolio selection rules improve over the broad US market in terms of downside deviation and the other risk-return measures, and—across the different portfolio selection rules—the SSV results are the most promising one. As optimal shrinkage is not the focus of our analysis, for the out-of-sample results we simply set the shrinkage factor equal to 0.9 for daily data, to 0.8 for monthly data, and to 0.7 for quarterly data. This corresponds to a stronger shrinkage effect with lower frequency data, justified by the fact that less observations are used for estimation, which is therefore less precise. Selecting an optimal, time-varying, value for the shrinkage parameter would likely improve our results, but it is beyond the scope of this work.

## 6 Conclusions

The SSV matrix estimator introduced in this paper provides an extremely precise approximation of the sample semicovariance matrix of a set of assets. By changing the single tuning parameter  $\theta$ , the SSV can span the entire set of portfolios included between the minimum semivariance and the minimum variance portfolio. The low computational intensity of this procedure makes it suitable for portfolio optimization problems with many assets, contrary to the grid search algorithm, which despite being accurate, is only feasible for relatively small portfolios. Compared to other approaches considered in this paper, the proposed method is unbiased in large samples since it targets the actual portfolio semivariance instead of a heuristic approximation. Our theoretical derivation and numerical findings support these claims. Although we illustrate our algorithm on the relatively simple case of minimum semivariance optimization, our procedure is very general. With minor modifications, our algorithm can accommodate more sophisticated constraints on the portfolio weights, depending on the specific application at hand. Furthermore, other types of objective functions involving the semivariance, beyond the quadratic objective function studied in this paper, may be considered in the future using a similar framework.

For real-world data, we show that in-sample the proposed SSV approach outperforms the competing portfolio selection rules in terms of downside deviation. Out-of-sample, while the SSV approach underperforms some of the considered portfolio selection rules, it outperforms all of them when combined with shrinkage estimation of the semivariance. This indicates that the higher parameter uncertainty implied by the semicovariance estimation can offset the benefits from targeting a more realistic objective function. Overall, our SSV approach provides a robust and flexible framework for semivariance portfolio minimization. Out of sample, almost all of our considered portfolio selection rules improve over the broad market portfolio. They show lower downside deviation, standard deviation, and excess kurtosis, and higher skewness, Sortino ratio and Shape ratio. Across the different settings, our

---

<sup>12</sup> This choice allows for portfolio return series that cover a larger time-span compared to monthly data, where 20 years are used in the estimation window. However, for consistency and better comparability, we only consider the portfolio returns during the same time-span for which they are available when using monthly data.

SSV approach results as the most promising portfolio selection technique, especially when combined out-of-sample with a shrinkage approach.

Looking forward, several avenues for further research present themselves. Firstly, extending the application of the smoothed semivariance (SSV) estimator to periods where asset returns present particularly heavy-tailed distributions or skewed distributions could provide insights into its performance under different market conditions and risk profiles. Although we have considered both  $b$  equal to the mean and the median and different time frequencies (i.e., daily, monthly and quarterly), further analysis is required to provide additional insights on the applicability and performance of the SSV method. While the performance of smooth weights depending on the normal cdf worked reasonably well in our numerical studies, a more detailed understanding on the choice of the smoothing function in a wider range of scenarios would be useful. Secondly, investigating adaptive approaches for selecting the smoothing parameter  $\theta$  in the SSV estimator based on changing market conditions or asset characteristics could enhance its flexibility and robustness. For instance,  $\theta$  might be considered as a tuning parameter in an appropriate data-fitting procedure. Thirdly, further analyses should focus on combining SSV with shrinkage techniques as well as the development of other procedures to mitigate the effects of parameter uncertainty in the semicovariance matrix estimation. Fourthly, analyzing the risk-return tradeoff in portfolio optimization using the SSV approach under different risk preferences and investment objectives could offer valuable guidance for investors. Lastly, extending the application of the SSV estimator beyond traditional asset classes to alternative investments like cryptocurrencies, commodities, or real estate could broaden its applicability across different investment domains.

## Appendix A: Proof of Propositions

### Proof of Proposition 1

For fixed  $w$  and  $b$ , define  $Z_t = Y_t(w) - b$ . Let  $\theta_T = \xi s_T^{-1}$  where  $s_T \rightarrow \infty$  is an arbitrary sequence diverging with  $T$ . Consider the following first-order Taylor expansion of  $\widehat{\sigma}_s^2(w, \xi s_T^{-1})$  around  $\xi = 0$

$$\begin{aligned} \widehat{\sigma}_s^2\left(w, \frac{\xi}{s_T}\right) &= \sigma_s^2(w, 0) + \xi \left( \left. \frac{\partial}{\partial y} \widehat{\sigma}_s^2\left(w, \frac{\xi}{s_T}\right) \right|_{\xi=\bar{\xi}} \right) \\ &= \sigma_s^2(w, 0) + \frac{\xi}{\bar{\xi}} \frac{1}{T} \sum_{t=1}^T \left( -\frac{Z_t s_T}{\bar{\xi}} \right) f\left(\frac{Z_t s_T}{\bar{\xi}}\right), \end{aligned} \tag{A1}$$

where  $\bar{\xi}$  is some constant such that  $0 < \bar{\xi} < \xi$ . By our assumptions we have

$$E[(Z_t s_T / \bar{\xi})^2 f(Z_t s_T / \bar{\xi})^2] < \infty,$$

for all  $T \geq 1$ . Therefore, the weak Law of Large Numbers for triangular arrays implies

$$\left| \frac{1}{T} \sum_{t=1}^T \frac{Z_t s_T}{\bar{\xi}} f\left(\frac{Z_t s_T}{\bar{\xi}}\right) - E\left[\frac{Z_t s_T}{\bar{\xi}} f\left(\frac{Z_t s_T}{\bar{\xi}}\right)\right] \right| = o_p(1). \tag{A2}$$

Moreover, by Cauchy–Schwartz inequality we can write

$$\left| E \left[ \frac{Z_t s_T}{\xi} f \left( \frac{Z_t s_T}{\xi} \right) \right] \right| \leq \frac{s_T}{\xi} (E[Z_t^2])^{1/2} E \left[ f \left( \frac{Z_t s_T}{\xi} \right)^2 \right]^{1/2}. \quad (\text{A3})$$

By our assumptions,  $E[Z_t^2] < \infty$  for any  $\|w\|_2 < C$ . Moreover, for sufficiently large  $s_T$  there is a constant  $k_1$  such that  $f(Z_t s_T / \xi) \leq f(k_1 s_T / \xi)$ . This means that right-hand side of (A3) is bounded by  $|k_2 \theta_T| f(k_1 \theta_T)$  for some constants  $k_1, k_2 > 0$ . This shows that the reminder term in (A1) is  $O_p(\theta_T f(k_1 \theta_T))$ .

## Proof of Proposition 2

First, we show  $\left| \hat{L}_s^2(v; \theta_T) - L_s(v; 0) \right|$  converges to zero in probability uniformly in  $v$ . By the triangle inequality, we have

$$\sup_{(w^\top, \lambda): \|w\|_2 < C} \left| \hat{L}_s(w; \theta_T) - L_s(w; 0) \right| = \sup_{w: \|w\|_2 < C} \left| \hat{\sigma}_s^2(w; \theta_T) - \sigma_s^2(w; 0) \right| \quad (\text{A4})$$

$$\begin{aligned} &\leq \sup_{w: \|w\|_2 < C} \left| \hat{\sigma}_s^2(w; \theta_T) - \hat{\sigma}_s^2(w; 0) \right| \\ &\quad + \sup_{w: \|w\|_2 < C} \left| \hat{\sigma}_s^2(w; 0) - \sigma_s^2(w; 0) \right| \quad (\text{A5}) \end{aligned}$$

The first term converges in probability to zero by Proposition 2.1. For the second term, Markov's inequality implies, for any  $\varepsilon > 0$

$$\begin{aligned} P \left( \sup_{w: \|w\|_2 < C} \left| \hat{\sigma}_s^2(w; 0) - \sigma_s^2(w; 0) \right| > \varepsilon \right) &\leq \frac{1}{\varepsilon^2 T} E \left[ \sup_{w: \|w\|_2 < C} (w^\top X_t)^4 I(w^\top X_t < b)^2 \right] \\ &\leq \frac{C^4 E[\|X_t\|_2^4]}{\varepsilon^2 T}. \end{aligned}$$

The right-hand side converges to zero as  $T \rightarrow \infty$ . This shows that the second term in (A4) converges to zero in probability. Hence, by Theorem 5.7 in van der Vaart (2000) we have  $\hat{v}_{\theta_T} \xrightarrow{P} v^*$  and  $\hat{w}_{\theta_T} \xrightarrow{P} w^*$ .

## Appendix B: Tables

See Tables 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.

**Table 3** In-sample, daily data, benchmark equal to sample mean return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt 17 Ind	-	0.00697	0.00961	-0.60420	16.44055	0.01923	0.03148	-
	GMV	0.00443	0.00622	-0.40392	13.60668	0.04322	0.04285	0.02434
	Balletero	0.00455	0.00633	-0.50012	16.03263	0.04116	0.04355	0.01969
	Cumova	0.00435	0.00629	0.14118	13.87184	0.05202	0.04394	0.02904
	Estrada	0.00449	0.00650	0.18269	12.81047	0.05145	0.04050	0.03156
30 Ind	SSV $\theta = 1/T$	<b>0.00431</b>	0.00623	0.09924	13.81104	0.05249	0.04490	0.02692
	SSV $\theta = 1/\sqrt{T}$	0.00434	0.00622	0.05900	14.39285	0.04775	0.04400	0.02490
	SSV $\theta = 100$	0.00443	0.00622	-0.40366	13.60535	0.04322	0.04285	0.02434
	GMV	0.00414	0.00582	-0.46245	11.08304	0.04314	0.04062	0.03313
	Balletero	0.00426	0.00594	-0.58396	14.09969	0.04136	0.04194	0.02756
48 Ind	Cumova	0.00404	0.00591	0.23403	10.66211	0.05353	0.04302	0.03951
	Estrada	0.00421	0.00619	0.30939	9.83836	0.05318	0.03741	0.04255
	SSV $\theta = 1/T$	<b>0.00399</b>	0.00586	0.25352	10.93668	0.05482	0.04456	0.03684
	SSV $\theta = 1/\sqrt{T}$	0.00403	0.00583	0.18114	11.86282	0.04974	0.04334	0.03387
	SSV $\theta = 100$	0.00414	0.00582	-0.46215	11.08024	0.04315	0.04062	0.03313
48 Ind	GMV	0.00479	0.00662	-0.42861	5.10347	0.02212	0.03343	0.05310
	Balletero	0.00492	0.00676	-0.51630	6.79729	0.02180	0.03605	0.04467
	Cumova	0.00469	0.00677	0.37781	9.68503	0.03055	0.03804	0.06299
	Estrada	0.00510	0.00738	0.37828	9.36389	0.03226	0.03291	0.06779
	SSV $\theta = 1/T$	<b>0.00459</b>	0.00671	0.46785	9.73502	0.03867	0.04314	0.06088
48 Ind	SSV $\theta = 1/\sqrt{T}$	0.00464	0.00668	0.35941	10.01436	0.03128	0.03882	0.05484
	SSV $\theta = 100$	0.00479	0.00662	-0.42842	5.10264	0.02212	0.03344	0.05310

Table 3 continued

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
SBM 100 (70)	GMV	0.00372	0.00516	-0.59122	5.39060	0.04947	0.08009	0.07768
	Ballesterro	0.00384	0.00527	-0.72494	6.65195	0.04661	0.07749	0.06773
	Cumova	0.00365	0.00540	0.39521	10.63873	0.06912	0.08658	0.09895
	Estrada	0.00376	0.00554	0.02362	6.83350	0.07260	0.08406	0.10321
	SSV $\theta = 1/T$	<b>0.00356</b>	0.00526	0.10803	4.71152	0.06848	0.08939	0.09245
	SSV $\theta = 1/\sqrt{T}$	0.00363	0.00518	-0.11311	5.32431	0.05705	0.08343	0.08084
	SSV $\theta = 100$	0.00372	0.00516	-0.59105	5.38932	0.04947	0.08009	0.07768

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballesterro (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T, 1/\sqrt{T}, 100$ . Portfolio returns spanning from April 28, 1989 to April 28, 2023 for the 48 industry portfolios; from July 31, 1946 to April 28, 2023 for the other datasets

Best DD values reported in bold

**Table 4** Out-of-sample, daily data, benchmark equal to sample mean return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt	-	0.00698	0.00961	-0.60420	16.44055	0.01921	0.03148	-
17 Ind	GMV	0.00463	0.00646	-0.55577	21.36361	0.04086	0.04087	0.02398
	Ballesterro	0.00472	0.00654	-0.72572	24.56278	0.03876	0.04156	0.01940
	Cumova	0.00464	0.00653	-0.45761	21.37757	0.04409	0.03897	0.02871
	Estrada	0.00476	0.00673	-0.36740	19.19959	0.04409	0.03604	0.03116
	SSV $\theta = 1/T$	<b>0.00461</b>	0.00647	-0.51536	21.53820	0.04301	0.03891	0.02660
	SSV $\theta = 1/\sqrt{T}$	<b>0.00461</b>	0.00645	-0.51010	21.50115	0.04099	0.03956	0.02456
30 Ind	SSV $\theta = 100$	0.00463	0.00646	-0.55575	21.36336	0.04086	0.04087	0.02398
	GMV (shrink)	0.00466	0.00647	-0.62519	25.02453	0.03963	0.04331	0.01577
	Ballesterro (shrink)	0.00481	0.00664	-0.76796	26.70502	0.03734	0.04283	0.01354
	SSV $\theta = 1/T$ (shrink)	<b>0.00461</b>	0.00644	-0.61087	25.05277	0.04207	0.04193	0.01712
	GMV	0.00445	0.00621	-0.77965	24.37934	0.03959	0.03770	0.03260
	Ballesterro	0.00453	0.00627	-0.94283	28.10805	0.03800	0.03909	0.02711
	Cumova	0.00448	0.00629	-0.72492	26.75549	0.04091	0.03515	0.03899
	Estrada	0.00464	0.00655	-0.65076	24.50253	0.04189	0.03083	0.04202
	SSV $\theta = 1/T$	0.00444	0.00623	-0.69907	24.68835	0.03957	0.03502	0.03636
	SSV $\theta = 1/\sqrt{T}$	<b>0.00443</b>	0.00620	-0.70005	24.70880	0.03938	0.03662	0.03338
30 Ind	SSV $\theta = 100$	0.00445	0.00621	-0.77961	24.37859	0.03960	0.03770	0.03260
	GMV (shrink)	0.00446	0.00621	-0.80312	27.84516	0.03929	0.04029	0.02246
	Ballesterro (shrink)	0.00460	0.00635	-0.94338	30.01735	0.03732	0.04062	0.01958
	SSV $\theta = 1/T$ (shrink)	<b>0.00443</b>	0.00619	-0.75146	28.41310	0.04060	0.03862	0.02440

Table 4 continued

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
48 Ind	GMV	<b>0.00526</b>	0.00725	-0.42094	10.72827	0.01836	0.02924	0.05213
	Ballesterro	0.00532	0.00730	-0.53769	12.23434	0.01815	0.03193	0.04386
	Cumova	0.00528	0.00735	-0.01109	13.54152	0.01623	0.02721	0.06216
	Estrada	0.00564	0.00789	-0.04726	10.65344	0.02159	0.02536	0.06698
	SSV $\theta = 1/T$	0.00527	0.00730	-0.15651	12.19704	0.01757	0.02805	0.06000
	SSV $\theta = 1/\sqrt{T}$	0.00523	0.00724	-0.20091	11.82388	0.01765	0.02848	0.05391
	SSV $\theta = 100$	<b>0.00526</b>	0.00725	-0.42086	10.72793	0.01836	0.02924	0.05213
	GMV (shrink)	0.00543	0.00745	-0.44211	12.36652	0.01306	0.03308	0.03509
	Ballesterro (shrink)	0.00557	0.00761	-0.54660	13.15421	0.01317	0.03510	0.03061
	SSV $\theta = 1/T$ (shrink)	0.00537	0.00740	-0.22758	13.30367	0.01098	0.03056	0.03855
SBM 100 (70)	GMV	0.00372	0.00516	-0.59122	5.39060	0.04947	0.08009	0.07768
	Ballesterro	0.00384	0.00527	-0.72494	6.65195	0.04661	0.07749	0.06773
	Cumova	0.00365	0.00540	0.39521	10.63873	0.06912	0.08658	0.09895
	Estrada	0.00376	0.00554	0.02362	6.83350	0.07260	0.08406	0.10321
	SSV $\theta = 1/T$	<b>0.00356</b>	0.00526	0.10803	4.71152	0.06848	0.08939	0.09245
	SSV $\theta = 1/\sqrt{T}$	0.00427	0.00581	-1.02554	14.62417	0.03900	0.06746	0.07997
	SSV $\theta = 100$	0.00424	0.00578	-1.07680	15.46975	0.04074	0.06961	0.07684
	GMV (shrink)	0.00433	0.00583	-1.24093	19.64921	0.03693	0.06806	0.04926
	Ballesterro (shrink)	0.00445	0.00597	-1.30440	20.00388	0.03517	0.06572	0.04488
	SSV $\theta = 1/T$ (shrink)	0.00432	0.00585	-1.07239	17.37217	0.03730	0.06677	0.05541

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballesterro (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T$ , GMV with shrunk covariance matrix, GMV (shrink), Ballesterro with shrunk covariance matrix, Ballesterro (shrink), and SSV  $\theta = 1/T$  with shrunk semicovariance matrix, SSV (shrink), are based on a shrunk correlation matrix  $R^s$ , with  $R^s = 0.9R + (1 - 0.9)\mathbf{I}$ ,  $R$  being the sample correlation matrix and  $\mathbf{I}$  the  $N \times N$  identity matrix. Portfolio returns spanning from April 28, 1989 to April 28, 2023 for the 48 industry portfolios; from July 31, 1946 to April 28, 2023 for the other datasets. Best DD values reported in bold



**Table 5** In-sample, quarterly data, benchmark equal to sample mean return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt	-	0.0575	0.0818	-0.5892	0.9519	0.1492	0.2395	-
17 Ind	GMV	0.0327	0.0523	-0.1573	0.3686	0.3008	0.2927	0.5339
	Ballestero	0.0339	0.0530	-0.2777	0.6406	0.2803	0.3017	0.4530
	Cumova	0.0398	0.0643	-0.2359	2.0870	0.2559	0.2177	0.9273
	Estrada	0.0401	0.0681	0.1108	0.8469	0.2981	0.1988	0.8418
	SSV $\theta = 1/T$	<b>0.0296</b>	0.0548	0.4689	0.8706	0.3851	0.2967	0.6988
30 Ind	SSV $\theta = 1/\sqrt{T}$	0.0301	0.0535	0.3692	0.9092	0.3680	0.2994	0.6095
	SSV $\theta = 100$	0.0326	0.0523	-0.1566	0.3682	0.3009	0.2927	0.5340
	GMV	0.0254	0.0422	-0.2389	0.1811	0.4295	0.3786	0.9091
	Ballestero	0.0263	0.0425	-0.4019	0.4161	0.4040	0.3878	0.8182
	Cumova	0.4548	0.6407	-4.3383	64.4616	0.0061	-0.0042	77.8008
48 Ind	Estrada	0.0395	0.0659	-0.0190	1.5505	0.3309	0.1905	1.5639
	SSV $\theta = 1/T$	<b>0.0224</b>	0.0470	0.8613	0.9157	0.4814	0.2878	1.3290
	SSV $\theta = 1/\sqrt{T}$	0.0228	0.0441	0.6461	1.3879	0.4911	0.3369	1.0548
	SSV $\theta = 100$	0.0254	0.0422	-0.2381	0.1806	0.4296	0.3786	0.9091
	GMV	0.0191	0.0337	-0.0458	0.6909	0.4050	0.3439	1.8586
48 Ind	Ballestero	0.0195	0.0334	-0.1698	0.8206	0.3979	0.3643	1.7558
	Cumova	0.7854	6.0376	10.6597	119.4090	0.8181	0.1032	217.1535
	Estrada	0.0711	0.0993	-0.5701	2.3427	-0.0247	-0.0077	3.5904
	SSV $\theta = 1/T$	<b>0.0153</b>	0.0484	3.7306	20.0256	0.3168	0.1766	2.9083
	SSV $\theta = 1/\sqrt{T}$	0.0167	0.0461	3.2576	18.5857	0.5315	0.2689	2.1141
SSV $\theta = 100$	0.0191	0.0337	-0.0449	0.6904	0.4050	0.3439	1.8587	

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballestero (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T, 1/\sqrt{T}, 100$ . Portfolio returns spanning from September 1989 to March 2023 for the 48 industry portfolios; from September 1946 to March 2023 for the other datasets

Best DD values reported in bold

**Table 6** Out-of-sample, quarterly data, benchmark equal to sample mean return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt	-	0.0583	0.0818	-0.5892	0.9519	0.1469	0.2395	-
17 Ind	GMV	0.0470	0.0707	-0.2664	1.0492	0.2139	0.2195	0.5389
	Ballesterro	0.0464	0.0689	-0.3434	1.2312	0.2071	0.2335	0.4557
	Cumova	0.0602	0.0852	-0.6981	2.1668	0.1565	0.1547	0.9457
	Estrada	0.0563	0.0842	-0.3276	1.4826	0.2133	0.1603	0.8468
	SSV $\theta = 1/T$	0.0493	0.0742	-0.4196	1.6117	0.2370	0.2230	0.6971
	SSV $\theta = 1/\sqrt{T}$	0.0478	0.0722	-0.3563	1.3471	0.2364	0.2252	0.6096
	SSV $\theta = 100$	0.0470	0.0707	-0.2665	1.0496	0.2139	0.2195	0.5389
	GMV (shrink)	0.0445	0.0645	-0.5803	1.3649	0.2252	0.2893	0.1377
	Ballesterro (shrink)	0.0459	0.0659	-0.6200	1.4556	0.2094	0.2853	0.1255
	SSV $\theta = 1/T$ (shrink)	<b>0.0439</b>	0.0642	-0.5883	1.3957	0.2469	0.3006	0.1617
30 Ind	GMV	0.0464	0.0720	-0.2605	0.8589	0.2816	0.2501	0.9157
	Ballesterro	0.0454	0.0697	-0.3616	1.0128	0.2764	0.2630	0.8231
	Cumova	0.5914	1.5756	10.5015	158.4353	0.1810	0.0644	49.5986
	Estrada	0.0618	0.0938	-0.4271	1.7202	0.2766	0.1758	1.5647
	SSV $\theta = 1/T$	0.0525	0.0795	-0.2663	0.8777	0.2268	0.1841	1.3487
	SSV $\theta = 1/\sqrt{T}$	0.0488	0.0752	-0.2801	1.2012	0.2748	0.2265	1.0681
	SSV $\theta = 100$	0.0464	0.0720	-0.2604	0.8593	0.2816	0.2501	0.9158
	GMV (shrink)	0.0412	0.0614	-0.6309	1.2984	0.2947	0.3131	0.1999
	Ballesterro (shrink)	0.0425	0.0625	-0.6908	1.4869	0.2701	0.3099	0.1878
	SSV $\theta = 1/T$ (shrink)	<b>0.0411</b>	0.0613	-0.6218	1.1864	0.2970	0.3138	0.2224

Table 6 continued

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
48 Ind	GMV	0.0605	0.0966	0.0582	2.0123	0.1996	0.1652	1.9319
	Ballestero	0.0589	0.0935	0.0202	2.1335	0.2010	0.1742	1.8203
	Cumova	1.8289	3.0593	2.2359	43.5775	0.0359	0.0148	284.1181
	Estrada	0.0957	0.1415	-0.2452	2.0928	0.1173	0.0857	3.7369
	SSV $\theta = 1/T$	0.0761	0.1112	-0.0029	2.9260	0.0669	0.0792	3.0766
	SSV $\theta = 1/\sqrt{T}$	0.0662	0.1014	-0.0037	2.0735	0.1381	0.1246	2.2398
	SSV $\theta = 100$	0.0605	0.0967	0.0581	2.0126	0.1996	0.1652	1.9320
	GMV (shrink)	0.0426	0.0600	-0.7631	0.8610	0.2024	0.3177	0.2927
	Ballestero (shrink)	0.0436	0.0606	-0.8514	1.1261	0.1957	0.3170	0.2806
	SSV $\theta = 1/T$ (shrink)	<b>0.0422</b>	0.0599	-0.7534	1.0711	0.2138	0.3211	0.3153

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballestero (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T$ ; GMV with shrunk covariance matrix, GMV (shrink), Ballestero with shrunk covariance matrix, Ballestero (shrink), and SSV  $\theta = 1/T$  with shrunk semicovariance matrix, SSV (shrink), are based on a shrunk correlation matrix  $R^s$ , with  $R^s = 0.7R + (1 - 0.7)\mathbf{I}$ ,  $R$  being the sample correlation matrix and  $\mathbf{I}$  the  $N \times N$  identity matrix. Portfolio returns spanning from September 1989 to March 2023 for the 48 industry portfolios; from September 1946 to March 2023 for the other datasets. Best DD values reported in bold

Table 7 In-sample, daily data, benchmark equal to sample median return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt	–	0.00711	0.00961	–0.60420	16.44055	–0.02005	0.03148	–
17 Ind	GMV	0.00451	0.00622	–0.40392	13.60668	0.00438	0.04285	0.02434
	Ballesterro	0.00462	0.00633	–0.50012	16.03263	0.00538	0.04355	0.01969
	Cumova	0.00441	0.00629	0.14118	13.87184	0.02542	0.04394	0.02904
	Estrada	0.00452	0.00652	0.19199	12.74172	0.03796	0.04171	0.03916
30 Ind	SSV $\theta = 1/T$	<b>0.00432</b>	0.00630	0.13737	13.43809	0.06173	0.04670	0.04252
	SSV $\theta = 1/\sqrt{T}$	0.00439	0.00622	0.06277	14.38308	0.02183	0.04454	0.02666
	SSV $\theta = 100$	0.00451	0.00622	–0.40209	13.60228	0.00569	0.04291	0.02451
	GMV	0.00420	0.00582	–0.46245	11.08304	0.01579	0.04062	0.03313
	Ballesterro	0.00431	0.00594	–0.58396	14.09969	0.01695	0.04194	0.02756
	Cumova	0.00408	0.00591	0.23403	10.66211	0.03966	0.04302	0.03951
48 Ind	Estrada	0.00422	0.00620	0.34641	10.07960	0.05027	0.03940	0.05151
	SSV $\theta = 1/T$	<b>0.00395</b>	0.00594	0.32678	10.98059	0.09564	0.04836	0.05738
	SSV $\theta = 1/\sqrt{T}$	0.00405	0.00583	0.19563	12.03907	0.03702	0.04426	0.03583
	SSV $\theta = 100$	0.00419	0.00582	–0.46023	11.07867	0.01710	0.04072	0.03327
	GMV	0.00484	0.00662	–0.42861	5.10347	0.00474	0.03343	0.05310
	Ballesterro	0.00496	0.00676	–0.51630	6.79729	0.00373	0.03605	0.04467
48 Ind	Cumova	0.00472	0.00677	0.37781	9.68503	0.02314	0.03804	0.06299
	Estrada	0.00509	0.00739	0.46780	10.48953	0.03095	0.03833	0.08334
	SSV $\theta = 1/T$	<b>0.00448</b>	0.00684	0.66547	11.62950	0.10313	0.05245	0.09805
	SSV $\theta = 1/\sqrt{T}$	0.00464	0.00668	0.41457	10.73511	0.03189	0.04090	0.05864
SSV $\theta = 100$	0.00483	0.00662	–0.42567	5.10963	0.00591	0.03363	0.05326	

Table 7 continued

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
SBM 100 (70)	GMV	0.00383	0.00516	-0.59122	5.39060	-0.00145	0.08009	0.07768
	Ballesterro	0.00395	0.00527	-0.72494	6.65195	-0.00220	0.07749	0.06773
	Cumova	0.00373	0.00540	0.39521	10.63873	0.03783	0.08658	0.09895
	Estrada	0.00377	0.00559	0.13232	6.13062	0.08362	0.08717	0.12546
	SSV $\theta = 1/T$	<b>0.00355</b>	0.00553	0.27973	4.97429	0.14778	0.09420	0.16444
	SSV $\theta = 1/\sqrt{T}$	0.00369	0.00521	-0.07869	5.51197	0.03281	0.08462	0.08731
	SSV $\theta = 100$	0.00382	0.00516	-0.58512	5.37000	0.00426	0.08049	0.07885

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballesterro (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T, 1/\sqrt{T}, 100$ . Portfolio returns spanning from April 28, 1989 to April 28, 2023 for the 48 industry portfolios; from July 31, 1946 to April 28, 2023 for the other datasets  
 Best DD values reported in bold

**Table 8** Out-of-sample, daily data, benchmark equal to sample median return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt	–	0.00712	0.00961	–0.60420	16.44055	–0.02004	0.03148	–
17 Ind	GMV	0.00470	0.00646	–0.55577	21.36361	0.00365	0.04087	0.02398
	Ballesterro	0.00479	0.00654	–0.72572	24.56278	0.00434	0.04156	0.01940
	Cumova	0.00469	0.00653	–0.45761	21.37757	0.01920	0.03897	0.02871
	Estrada	0.00481	0.00675	–0.43842	20.33155	0.02890	0.03548	0.03880
	SSV $\theta = 1/T$	0.00465	0.00654	–0.57006	22.22356	0.04749	0.03794	0.04241
	SSV $\theta = 1/\sqrt{T}$	0.00466	0.00646	–0.52239	21.56328	0.01575	0.03945	0.02641
	SSV $\theta = 100$	0.00470	0.00646	–0.55689	21.34212	0.00478	0.04084	0.02416
	GMV (shrink)	0.00473	0.00647	–0.62519	25.02453	0.00445	0.04331	0.01577
	Ballesterro (shrink)	0.00488	0.00664	–0.76796	26.70502	0.00446	0.04283	0.01354
	SSV $\theta = 1/T$ (shrink)	<b>0.00464</b>	0.00647	–0.64816	26.13895	0.03365	0.04131	0.02473
30 Ind	GMV	0.00450	0.00621	–0.77965	24.37934	0.01418	0.03770	0.03260
	Ballesterro	0.00458	0.00627	–0.94283	28.10805	0.01510	0.03909	0.02711
	Cumova	0.00451	0.00629	–0.72492	26.75549	0.02852	0.03515	0.03899
	Estrada	0.00467	0.00656	–0.73532	27.03890	0.03510	0.02989	0.05104
	SSV $\theta = 1/T$	0.00445	0.00631	–0.74985	25.71043	0.06833	0.03382	0.05717
	SSV $\theta = 1/\sqrt{T}$	0.00446	0.00619	–0.71272	24.94223	0.02640	0.03646	0.03544
	SSV $\theta = 100$	0.00450	0.00621	–0.78138	24.39409	0.01522	0.03767	0.03275
	GMV (shrink)	0.00451	0.00621	–0.80312	27.84516	0.01715	0.04029	0.02246
	Ballesterro (shrink)	0.00465	0.00635	–0.94338	30.01735	0.01664	0.04062	0.01958
	SSV $\theta = 1/T$ (shrink)	<b>0.00443</b>	0.00622	–0.81151	29.51667	0.05726	0.03801	0.03527

Table 8 continued

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
48 Ind	GMV	0.00530	0.00725	-0.42094	10.72827	0.00254	0.02924	0.05213
	Ballester	0.00537	0.00730	-0.53769	12.23434	0.00148	0.03193	0.04386
	Cumova	0.00530	0.00735	-0.01109	13.54152	0.00971	0.02721	0.06216
	Estrada	0.00566	0.00789	-0.01151	11.03544	0.01293	0.02521	0.08239
	SSV $\theta = 1/T$	0.00523	0.00742	-0.07461	12.46332	0.05842	0.02735	0.09737
	SSV $\theta = 1/\sqrt{T}$	0.00523	0.00724	-0.18531	12.10324	0.01526	0.02835	0.05778
	SSV $\theta = 100$	0.00530	0.00725	-0.41887	10.76776	0.00341	0.02927	0.05229
	GMV (shrink)	0.00547	0.00745	-0.44211	12.36652	0.00015	0.03308	0.03509
	Ballester (shrink)	0.00561	0.00761	-0.54660	13.15421	-0.00122	0.03510	0.03061
	SSV $\theta = 1/T$ (shrink)	<b>0.00533</b>	0.00742	-0.21095	13.37312	0.03339	0.03043	0.05683
SBM 100 (70)	GMV	0.00435	0.00578	-1.07683	15.47024	-0.00385	0.06961	0.07684
	Ballester	0.00442	0.00583	-1.20987	17.18805	-0.00464	0.06804	0.06704
	Cumova	0.00445	0.00601	-0.91948	14.36501	0.01489	0.06529	0.09818
	Estrada	0.00444	0.00617	-0.77457	12.72306	0.05029	0.06417	0.12470
	SSV $\theta = 1/T$	0.00435	0.00615	-0.70748	10.21029	0.09054	0.06344	0.16406
	SSV $\theta = 1/\sqrt{T}$	<b>0.00434</b>	0.00584	-0.99719	14.25320	0.01610	0.06677	0.08666
	SSV $\theta = 100$	<b>0.00434</b>	0.00578	-1.07586	15.44742	0.00049	0.06946	0.07807
	GMV (shrink)	0.00444	0.00583	-1.24093	19.64921	-0.00828	0.06806	0.04926
	Ballester (shrink)	0.00456	0.00597	-1.30440	20.00388	-0.00784	0.06572	0.04488
	SSV $\theta = 1/T$ (shrink)	0.00437	0.00596	-0.93694	15.23862	0.04868	0.06511	0.08573

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballester (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T$ , GMV with shrunk covariance matrix, GMV (shrink), Ballester with shrunk covariance matrix, Ballester (shrink), and SSV  $\theta = 1/T$  with shrunk semicovariance matrix, SSV (shrink), are based on a shrunk correlation matrix  $R^s$ , with  $R^s = 0.9R + (1 - 0.9)\mathbf{I}$ ,  $R$  being the sample correlation matrix and  $\mathbf{I}$  the  $N \times N$  identity matrix. Portfolio returns spanning from April 28, 1989 to April 28, 2023 for the 48 industry portfolios; from July 31, 1946 to April 28, 2023 for the other datasets. Best DD values reported in bold

**Table 9** In-sample, monthly data, benchmark equal to sample median return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt 17 Ind	-	0.0325	0.0431	-0.4867	1.6609	-0.0142	0.1466	-
	GMV	0.0219	0.0317	-0.2299	0.8857	0.1002	0.1682	0.1600
	Ballesterio	0.0225	0.0321	-0.3192	1.3341	0.0954	0.1721	0.1299
	Cumova	0.0211	0.0326	0.0674	0.6286	0.1626	0.1748	0.2145
	Estrada	0.0209	0.0342	0.0959	0.3007	0.2600	0.1603	0.2876
	SSV $\theta = 1/T$	<b>0.0199</b>	0.0348	0.2185	0.9622	0.3806	0.1793	0.3360
30 Ind	SSV $\theta = 1/\sqrt{T}$	0.0205	0.0327	0.1217	1.1429	0.2285	0.1810	0.2104
	SSV $\theta = 100$	0.0219	0.0317	-0.2283	0.8835	0.1060	0.1678	0.1610
	GMV	0.0187	0.0283	-0.1175	0.5972	0.1710	0.1895	0.2358
	Ballesterio	0.0192	0.0286	-0.2059	0.9567	0.1595	0.1936	0.2002
	Cumova	0.0180	0.0301	0.3018	0.4393	0.2594	0.1739	0.3264
	Estrada	0.0173	0.0325	0.3346	0.4321	0.4900	0.1747	0.3949
	SSV $\theta = 1/T$	<b>0.0158</b>	0.0340	0.5291	0.6971	0.7816	0.1953	0.5357
	SSV $\theta = 1/\sqrt{T}$	0.0167	0.0298	0.4357	1.1811	0.3836	0.1908	0.2938
	SSV $\theta = 100$	0.0187	0.0284	-0.1176	0.6003	0.1687	0.1872	0.2368



Table 9 continued

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
48 Ind	GMV	0.0184	0.0269	-0.3446	0.2586	0.1557	0.1756	0.3575
	Ballesterro	0.0187	0.0268	-0.4470	0.5252	0.1463	0.1875	0.3135
	Cumova	0.0181	0.0300	0.2559	0.3685	0.2346	0.1493	0.5479
	Estrada	0.0187	0.0353	0.1973	0.3723	0.5406	0.1802	0.5901
	SSV $\theta = 1/T$	<b>0.0132</b>	0.0370	0.7434	0.2984	1.2898	0.1504	1.0258
SBM 100 (70)	SSV $\theta = 1/\sqrt{T}$	0.0153	0.0283	0.6343	2.1561	0.4201	0.1632	0.4542
	SSV $\theta = 100$	0.0186	0.0269	-0.3374	0.2627	0.1329	0.1696	0.3566
	GMV	0.0195	0.0290	-0.1811	0.3679	0.1353	0.2573	0.8141
	Ballesterro	0.0200	0.0292	-0.2824	0.6040	0.1257	0.2594	0.7102
	Cumova	0.3545	0.4841	-25.4689	747.4059	-0.0164	-0.0091	39.2733
48 Ind	Estrada	0.0182	0.0375	0.4345	0.8926	0.6723	0.2122	1.5116
	SSV $\theta = 1/T$	<b>0.0110</b>	0.0477	0.8773	0.5606	2.6934	0.2087	3.7860
	SSV $\theta = 1/\sqrt{T}$	0.0152	0.0315	0.6859	1.5205	0.5842	0.2560	1.1480
	SSV $\theta = 100$	0.0191	0.0289	-0.1729	0.3623	0.1703	0.2607	0.8200

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballesterro (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T, 1/\sqrt{T}, 100$ . Portfolio returns spanning from April 1989 to April 2023 for the 48 industry portfolios; from July 1946 to April 2023 for the other datasets

Best DD values reported in bold

Table 10 Out-of-sample, monthly data, benchmark equal to sample median return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt	-	0.0326	0.0431	-0.4867	1.6609	-0.0142	0.1466	-
17 Ind	GMV	0.0242	0.0354	-0.0812	1.5337	0.0962	0.1546	0.1569
	Ballestero	0.0244	0.0351	-0.2052	1.8874	0.0918	0.1600	0.1270
	Cumova	0.0247	0.0364	-0.1637	0.8738	0.1299	0.1504	0.2115
	Estrada	0.0245	0.0374	-0.1651	0.5285	0.2003	0.1322	0.2840
	SSV $\theta = 1/T$	0.0244	0.0385	-0.1829	1.1446	0.2818	0.1440	0.3352
	SSV $\theta = 1/\sqrt{T}$	0.0240	0.0361	-0.1589	1.2551	0.1813	0.1548	0.2085
	SSV $\theta = 100$	0.0242	0.0354	-0.0789	1.5378	0.1014	0.1543	0.1578
	GMV (shrink)	0.0240	0.0344	-0.2421	1.8802	0.1059	0.1762	0.0686
	Ballestero (shrink)	0.0247	0.0351	-0.3344	2.1697	0.0964	0.1735	0.0607
	SSV $\theta = 1/T$ (shrink)	<b>0.0233</b>	0.0345	-0.3156	1.4014	0.1875	0.1799	0.1075
30 Ind	GMV	0.0225	0.0342	0.0088	1.5192	0.1499	0.1623	0.2316
	Ballestero	0.0226	0.0337	-0.0875	1.7887	0.1420	0.1686	0.1961
	Cumova	0.0231	0.0359	-0.0150	0.5595	0.1919	0.1394	0.3224
	Estrada	0.0221	0.0371	-0.0401	0.5217	0.3691	0.1441	0.3920
	SSV $\theta = 1/T$	0.0224	0.0393	-0.0472	0.8359	0.5160	0.1488	0.5351
	SSV $\theta = 1/\sqrt{T}$	0.0218	0.0349	-0.0067	1.0572	0.2816	0.1547	0.2911
	SSV $\theta = 100$	0.0226	0.0342	0.0104	1.5278	0.1476	0.1603	0.2323
	GMV (shrink)	0.0220	0.0327	-0.1177	1.6893	0.1451	0.1827	0.1004
	Ballestero (shrink)	0.0227	0.0332	-0.2356	2.0551	0.1336	0.1819	0.0908
	SSV $\theta = 1/T$ (shrink)	<b>0.0208</b>	0.0328	-0.1398	1.5114	0.2845	0.1817	0.1553

Table 10 continued

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
48 Ind	GMV	0.0258	0.0366	-0.3829	0.7546	0.0962	0.1184	0.3558
	Balletero	0.0252	0.0355	-0.4569	1.0690	0.0945	0.1318	0.3112
	Cumova	0.0274	0.0411	-0.1547	0.5777	0.1373	0.0972	0.5523
	Estrada	0.0257	0.0429	-0.1864	0.3109	0.3969	0.1507	0.5877
	SSV $\theta = 1/T$	0.0266	0.0487	-0.1082	0.5568	0.5975	0.0926	1.0386
	SSV $\theta = 1/\sqrt{T}$	0.0249	0.0381	-0.1777	1.0561	0.2325	0.1043	0.4553
	SSV $\theta = 100$	0.0260	0.0366	-0.3746	0.7714	0.0793	0.1135	0.3553
	GMV (shrink)	0.0236	0.0329	-0.4978	1.5713	0.0871	0.1855	0.1474
	Balletero (shrink)	0.0241	0.0332	-0.5929	1.8271	0.0834	0.1864	0.1351
	SSV $\theta = 1/T$ (shrink)	<b>0.0217</b>	0.0339	-0.2074	1.4073	0.2829	0.1762	0.2312
SBM 100 (70)	GMV	0.0293	0.0420	-0.2515	1.2314	0.0884	0.1769	0.8115
	Balletero	0.0287	0.0407	-0.3137	1.4723	0.0860	0.1852	0.7071
	Cumova	0.4177	0.8720	17.5083	539.1981	0.0407	0.0212	27.4596
	Estrada	0.0288	0.0487	-0.0909	0.9207	0.4266	0.1638	1.5185
	SSV $\theta = 1/T$	0.0339	0.0661	-0.2380	0.9996	0.8562	0.1416	3.8234
	SSV $\theta = 1/\sqrt{T}$	0.0283	0.0449	-0.2381	0.9910	0.3081	0.1759	1.1540
	SSV $\theta = 100$	0.0289	0.0419	-0.2471	1.2434	0.1112	0.1796	0.8182
	GMV (shrink)	0.0270	0.0377	-0.2997	1.3193	0.0651	0.1746	0.2284
	Balletero (shrink)	0.0275	0.0380	-0.3794	1.6702	0.0635	0.1801	0.2154
	SSV $\theta = 1/T$ (shrink)	<b>0.0246</b>	0.0384	-0.2984	1.0359	0.2992	0.1881	0.3788

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Balletero (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T$ , GMV with shrunk covariance matrix, GMV (shrink), Balletero with shrunk covariance matrix, Balletero (shrink), and SSV  $\theta = 1/T$  with shrunk semicovariance matrix, SSV (shrink), are based on a shrunk correlation matrix  $R^s$ , with  $R^s = 0.8R + (1 - 0.8)\mathbf{I}$ ,  $R$  being the sample correlation matrix and  $\mathbf{I}$  the  $N \times N$  identity matrix. Portfolio returns spanning from April 1989 to April 2023 for the 48 industry portfolios; from July 1946 to April 2023 for the other datasets. Best DD values reported in bold

**Table 11** In-sample, quarterly data, benchmark equal to sample median return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt	-	0.0617	0.0818	-0.5892	0.9519	0.0114	0.2395	-
17 Ind	GMV	0.0316	0.0523	-0.1573	0.3686	0.4126	0.2927	0.5339
	Ballesterro	0.0332	0.0530	-0.2777	0.6406	0.3624	0.3017	0.4530
	Cumova	0.0362	0.0643	-0.2359	2.0870	0.5240	0.2177	0.9273
	Estrada	0.0283	0.0677	0.5713	0.5551	1.0561	0.2289	0.9394
	SSV $\theta = 1/T$	<b>0.0211</b>	0.0789	1.0304	2.1402	2.2384	0.1990	1.8538
30 Ind	SSV $\theta = 1/\sqrt{T}$	0.0250	0.0555	0.5381	1.1673	0.8960	0.2894	0.7494
	SSV $\theta = 100$	0.0321	0.0523	-0.1747	0.3712	0.3642	0.2937	0.5354
	GMV	0.0247	0.0422	-0.2389	0.1811	0.5361	0.3786	0.9091
	Ballesterro	0.0257	0.0425	-0.4019	0.4161	0.4935	0.3878	0.8182
	Cumova	0.4281	0.6407	-4.3383	64.4616	0.0496	-0.0042	77.8008
48 Ind	Estrada	0.0284	0.0660	0.3707	0.4236	1.0625	0.1942	1.6574
	SSV $\theta = 1/T$	<b>0.0089</b>	0.1427	2.6698	12.2311	10.3799	0.2296	7.1309
	SSV $\theta = 1/\sqrt{T}$	0.0186	0.0456	0.7360	1.3974	1.1139	0.3399	1.2053
	SSV $\theta = 100$	0.0250	0.0424	-0.2521	0.1811	0.4954	0.3860	0.9128
	GMV	0.0183	0.0337	-0.0458	0.6909	0.5755	0.3439	1.8586
48 Ind	Ballesterro	0.0189	0.0334	-0.1698	0.8206	0.5413	0.3643	1.7558
	Cumova	0.7544	6.0376	10.6597	119.4090	0.8682	0.1032	217.1535
	Estrada	0.0616	0.1016	-0.3123	2.4629	0.2038	-0.0460	3.6642
	SSV $\theta = 1/T$	<b>0.0005</b>	0.0782	2.2920	7.2194	81.0715	-0.0026	11.4988
	SSV $\theta = 1/\sqrt{T}$	0.0133	0.0476	3.8013	24.9795	1.1998	0.2559	2.3896
SSV $\theta = 100$	0.0187	0.0338	-0.0631	0.5836	0.4685	0.3583	1.8912	

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballesterro (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T, 1/\sqrt{T}, 100$ . Portfolio returns spanning from September 1989 to March 2023 for the 48 industry portfolios; from September 1946 to March 2023 for the other datasets

Best DD values reported in bold

**Table 12** Out-of-sample, quarterly data, benchmark equal to sample median return

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
Mkt 17 Ind	-	0.0623	0.0818	-0.5892	0.9519	0.0110	0.2395	-
	GMV	0.0456	0.0707	-0.2664	1.0492	0.2911	0.2195	0.5389
	Ballesterro	0.0453	0.0689	-0.3434	1.2312	0.2675	0.2335	0.4557
	Cumova	0.0561	0.0852	-0.6981	2.1668	0.3241	0.1547	0.9457
	Estrada	0.0473	0.0827	-0.2066	1.4281	0.5943	0.1661	0.9334
	SSV $\theta = 1/T$	0.0490	0.0994	-0.0447	1.3751	0.9482	0.1502	1.8664
	SSV $\theta = 1/\sqrt{T}$	0.0438	0.0738	-0.3345	1.4583	0.5140	0.2190	0.7424
	SSV $\theta = 100$	0.0461	0.0704	-0.2749	0.9567	0.2566	0.2199	0.5395
	GMV (shrink)	0.0451	0.0645	-0.5803	1.3649	0.1935	0.2893	0.1377
	Ballesterro (shrink)	0.0465	0.0659	-0.6200	1.4556	0.1832	0.2853	0.1255
30 Ind	SSV $\theta = 1/T$ (shrink)	<b>0.0424</b>	0.0655	-0.5394	1.5906	0.3685	0.2872	0.2296
	GMV	0.0454	0.0720	-0.2605	0.8589	0.3396	0.2501	0.9157
	Ballesterro	0.0446	0.0697	-0.3616	1.0128	0.3287	0.2630	0.8231
	Cumova	0.5665	1.5756	10.5015	158.4353	0.2215	0.0644	49.5986
	Estrada	0.0530	0.0895	-0.4880	1.4912	0.5601	0.1370	1.6780
	SSV $\theta = 1/T$	0.1018	0.1855	-0.8163	5.3771	0.7279	0.0752	8.2505
	SSV $\theta = 1/\sqrt{T}$	0.0459	0.0778	-0.2454	1.1454	0.4899	0.2219	1.2235
	SSV $\theta = 100$	0.0456	0.0715	-0.2891	0.8692	0.3180	0.2562	0.9189
	GMV (shrink)	0.0424	0.0614	-0.6309	1.2984	0.2267	0.3131	0.1999
	Ballesterro (shrink)	0.0436	0.0625	-0.6908	1.4869	0.2171	0.3099	0.1878
SSV $\theta = 1/T$ (shrink)	<b>0.0395</b>	0.0627	-0.5481	1.4192	0.4406	0.2890	0.3045	

Table 12 continued

Dataset (N)	Strategy	DD	SD	Skew	Kurt	Sortino	Sharpe	TO
48 Ind	GMV	0.0594	0.0966	0.0582	2.0123	0.2509	0.1652	1.9319
	Ballestero	0.0580	0.0935	0.0202	2.1335	0.2476	0.1742	1.8203
	Cumova	2.1086	3.0593	2.2359	43.5775	0.0374	0.0148	284.1181
	Estrada	0.0739	0.1248	-0.1539	1.4785	0.2962	0.0343	3.7174
	SSV $\theta = 1/T$	0.0979	0.1790	0.3991	1.1511	0.4482	0.0037	11.9830
	SSV $\theta = 1/\sqrt{T}$	0.0640	0.1017	-0.0220	1.9674	0.2340	0.1100	2.5161
	SSV $\theta = 100$	0.0605	0.0968	0.0441	1.8273	0.2138	0.1682	1.9533
	GMV (shrink)	0.0449	0.0600	-0.7631	0.8610	0.0900	0.3177	0.2927
	Ballestero (shrink)	0.0457	0.0606	-0.8514	1.1261	0.0921	0.3170	0.2806
	SSV $\theta = 1/T$ (shrink)	<b>0.0389</b>	0.0601	-0.7298	1.0010	0.4788	0.3108	0.4411

Downside deviation (DD), standard deviation (SD), skewness (Skew), excess kurtosis (Kurt), Sortino ratio (Sortino), Sharpe ratio (Sharpe), and average turnover rate (TO). Strategies: global minimum-variance portfolio (GMV); minimum semivariance portfolio using Ballestero (2005), Estrada (2008) and Cumova and Nawrocki (2011); minimum SSV portfolios with  $\theta = 1/T$ ; GMV with shrink covariance matrix, GMV (shrink), Ballestero with shrink covariance matrix, Ballestero (shrink), and SSV  $\theta = 1/T$  with shrink semicovariance matrix, SSV (shrink), are based on a shrink correlation matrix  $R^s$ , with  $R^s = 0.7R + (1 - 0.7)\mathbf{I}$ ,  $R$  being the sample correlation matrix and  $\mathbf{I}$  the  $N \times N$  identity matrix. Portfolio returns spanning from September 1989 to March 2023 for the 48 industry portfolios; from September 1946 to March 2023 for the other datasets. Best DD values reported in bold

**Acknowledgements** The authors wish to thank the associate editor and the referees for their insightful feedback and valuable contributions to improving the quality of the manuscript.

**Funding** Open access funding provided by Università degli Studi di Trento within the CRUI-CARE Agreement.

**Data availability** The data used are freely downloadable from the sources indicated in the manuscript.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Code availability** Custom R code.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Azzalini, A. (2020). *The R package sn: The skew-normal and related distributions such as the skew-t (version 1.6-1)*. Università di Padova, Padua. <http://azzalini.stat.unipd.it/SN>
- Azzalini, A., & Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, 83(4), 715–726.
- Ballesterio, E. (2005). Mean-semivariance efficient frontier: A downside risk model for portfolio selection. *Applied Mathematical Finance*, 12(12), 1–15.
- Bernard, C., Bondarenko, O., & Vanduffel, S. (2019). Option-implied dependence and correlation risk premium. Canadian Derivatives Institute
- Bollerslev, T. (2021). Realized semi(co)variation: Signs that all volatilities are not created equal. *Journal of Financial Econometrics*, 20(2), 219–252.
- Cheremushkin, S. V. (2009). Why D-CAPM is a big mistake? The incorrectness of the cosemivariance statistics. <https://ssrn.com/abstract=1336169>
- Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236.
- Cont, R. (2010). Stylized properties of asset returns. In *Encyclopedia of quantitative finance*.
- Cumova, D., & Nawrocki, D. (2011). A symmetric LPM model for heuristic mean-semivariance analysis. *Journal of Economics and Business*, 63(3), 217–236.
- DeMiguel, V., Garlappi, L., Nogales, F. J., & Uppal, R. (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55(5), 798–812.
- Estrada, J. (2008). Mean-semivariance optimization: A heuristic approach. *Journal of Applied Finance*, 18(1), 57–72.
- Frahm, G., & Memmel, C. (2010). Dominating estimators for minimum-variance portfolios. *Journal of Econometrics*, 159(2), 289–302.
- Harlow, W. V. (1991). Asset allocation in a downside-risk framework. *Financial Analysts Journal*, 47(5), 28–40.
- Hogan, W. W., & Warren, J. M. (1972). Computation of the efficient boundary in the E-S portfolio selection model. *Journal of Financial and Quantitative Analysis*, 7(4), 1881–1896.
- Kempf, A., & Memmel, C. (2006). Estimating the global minimum variance portfolio. *Schmalenbach Business Review*, 58(4), 332–348.
- Klebaner, F., Landsman, Z., Makov, U., & Yao, J. (2016). Optimal portfolios with downside risk. *Quantitative Finance*, 17(3), 315–325.

- Lim, A. E., Shanthikumar, J. G., & Vahn, G. (2011). Conditional value-at-risk in portfolio optimization: Coherent but fragile. *Operations Research Letters*, 39(3), 163–171.
- Markowitz, H. (1959). *Portfolio selection*. Wiley.
- Nawrocki, D. N. (1991). Optimal algorithms and lower partial moment: Ex-post results. *Applied Economics*, 23(3), 465–470.
- Pla-Santamaria, D., & Bravo, M. (2013). Portfolio optimization based on downside risk: A mean-semivariance efficient frontier from dow jones blue chips. *Annals of Operations Research*, 205(1), 189–201.
- Ratliff-Crain, E., Oort, C. M. V., Bagrow, J., Koehler, M. T. K., & Tivnan, B. F. (2023). Revisiting stylized facts for modern stock markets.
- Rigamonti, A. (2020). Mean-variance optimization is a good choice, but for other reasons than you might think. *Risks*, 8(1), 29.
- Sortino, F. A., & Prince, L. N. (1994). Performance measurement in a downside risk framework. *The Journal of Investing*, 3(3), 59–64.
- van der Vaart, A. W. (2000). *Asymptotic statistics* (Vol. 3). Cambridge University Press.
- Zopounidis, C., Doumpos, M., & Fabozzi, F. J. (2014). Preface to the special issue: 60 years following Harry Markowitz's contributions in portfolio theory and operations research. *European Journal of Operational Research*, 234(2), 343–345.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.