

# Mean-Semivariance Portfolio Optimization Using Minimum Average Partial

Andrea Rigamonti  
University of Liechtenstein,  
`andrea.rigamonti@uni.li`

Katarína Lučivjanská  
Pavol Jozef Šafárik University,  
`katarina.lucivjanska@upjs.sk`

October 11, 2021

## Abstract

Mean-semivariance and minimum semivariance portfolios are a preferable alternative to mean-variance and minimum variance portfolios whenever the asset returns are not symmetrically distributed. However, similarly to other portfolios based on downside risk measures, they are particularly affected by parameter uncertainty because the estimates of the necessary inputs are less reliable than the estimates of the full covariance matrix. We address this problem by performing PCA using Minimum Average Partial on the downside correlation matrix in order to reduce the dimension of the problem and, with it, the estimation errors. We apply our strategy to several datasets and show that it greatly improves the performance of mean-semivariance optimization, largely closing the gap in out-of-sample performance with the strategies based on the covariance matrix.

*Keywords:* semivariance, principal component analysis, minimum average partial, parameter uncertainty, portfolio optimization, downside risk.

*JEL Classification Code:* C38, G11

## Acknowledgements

The authors are grateful to Alex Weissensteiner for his suggestions and comments.

# 1 Introduction

The poor out-of-sample performance of the mean-variance portfolio optimization procedure introduced by Markowitz (1952) is very well known and documented in the literature (see, e.g., Michaud, 1989; Broadie, 1993; Chopra and Ziemba, 1993). Given a set of risky assets with mean returns  $\mu$  and covariance matrix  $\Sigma$ , and a certain risk aversion coefficient  $\gamma$ , the optimal mean-variance weights in presence of a risk-free asset are given by the closed form solution

$$w_{mv} = \frac{1}{\gamma} \Sigma^{-1} \mu. \quad (1)$$

Analogously, if the investor specifies his preferences through a desired mean portfolio return  $Re$  instead of a level of risk aversion, the weights that minimize the portfolio variance given the desired mean return are

$$w_{mv} = \frac{Re}{\mu' \Sigma^{-1} \mu} \Sigma^{-1} \mu. \quad (2)$$

Since  $\mu$  and  $\Sigma$  are unknown in practice, they have to be estimated. The classical plug-in approach consists in computing sample estimates of the parameters and in treating them as if they were the true values. However, the plug-in approach leads to disappointing outcomes because of estimation errors (Kan and Zhou, 2007). The problem gets even worse if an investor optimizes a portfolio using downside risk measures (e.g. semivariance or CVaR), which are likely more in line with his real preferences than the variance but also more affected by estimation errors (Rigamonti, 2020). While studies on real empirical data have shown the importance of targeting the semivariance (see e.g. Pla-Santamaria and Bravo, 2013), up to now limited research has been done to address parameter uncertainty in this framework. Sharma et al. (2017) propose an Omega-CVaR optimization approach to avoid some of the difficulties related to the estimation of the return distribution shape or moments. Recently, fuzzy portfolio selection models have been developed as a novel approach to target various risk measures, including those used in a downside risk framework (see e.g. Chen et al., 2018; Yue et al., 2019).

In this paper, we address the problem by taking as a starting point one of the solutions devised to mitigate the impact of estimation errors in a mean-variance framework. We build on the strategy introduced by Chen and Yuan (2016), who propose to estimate a subspace mean-variance portfolio instead of the theoretically optimal mean-variance portfolio. This is done by considering only the first  $d$  eigenvectors of the sample covariance matrix  $\hat{\Sigma}$  instead of the full matrix. If  $d$  is carefully chosen, the systematic error caused by restricting the investment options is outweighed by the lower estimation error achieved thanks to the reduced dimension of the problem. Inspired by their setup, we apply a similar principle to optimize mean-semivariance

and minimum semivariance portfolios constructed using the approximation of the semicovariance matrix with the technique introduced by Rigamonti et al. (2021). For the choice of  $d$ , Chen and Yuan (2016) rely on a criterion introduced by Bai and Ng (2002) that depends on the length of the estimation window. However, when computing the semicovariance matrix not all information available from the estimation window is incorporated. Therefore, the Bai and Ng (2002) criterion may not be suitable. To overcome this obstacle, we transform the semicovariance matrix in a downside correlation matrix, and select  $d$  using the Minimum Average Partial (MAP) rule introduced by Velicer (1976), which does not depend on the length of the estimation window. We apply our portfolio rule to several empirical datasets with different size and return frequency. The results show that our strategy outperforms the portfolios based on the full semicovariance matrix, and performs well even compared to portfolios that target the variance. Moreover, contrary to existing approaches in the literature, our strategy does not involve complicated algorithms to compute the optimal weights, which are instead obtained through simple closed form solutions.

This paper provides three main contributions. First, we address the problem of parameter uncertainty for portfolios based on the semicovariance matrix. Second, we perform for the first time a principal component analysis (PCA) in a downside variability framework. Finally, we show the effectiveness of MAP for asset allocation decisions.

The remainder of the paper is organized as follows. Section 2 describes the dimension reduction methodology and the asset allocation procedure. Section 3 discusses details of principal component analysis when applied to the semicovariance matrix. Section 4 provides an empirical application to real data. Section 5 concludes.

## 2 Methodology

### 2.1 Subspace efficient portfolios

The idea which our approach is based on is analogous to that of Chen and Yuan (2016). For an investor with a certain risk aversion  $\gamma$ , given the sample mean of individual returns  $\hat{\mu}$  and the sample covariance matrix  $\hat{\Sigma}$ , they compute a subspace mean-variance portfolio whose weights are

$$\hat{w}_{mv,cov} = \frac{1}{\gamma} \sum_{k=1}^d \hat{\theta}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \hat{\mu}, \quad (3)$$

where  $\hat{\theta}_k$  is an eigenvalue of  $\hat{\Sigma}$  and  $\hat{\nu}_k$  the corresponding eigenvector, with  $\hat{\theta}_1 \geq \hat{\theta}_2 \geq \dots \geq \hat{\theta}_d$ . If we consider an investor with a desired mean portfolio

return  $Re$  instead, the portfolio weights are

$$\hat{w}_{mv,cov} = \frac{Re}{\hat{\mu}' \sum_{k=1}^d \hat{\theta}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \hat{\mu}} \sum_{k=1}^d \hat{\theta}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \hat{\mu}. \quad (4)$$

For the choice of  $d$ , Chen and Yuan (2016) rely on a criterion introduced by Bai and Ng (2002), and estimate it as

$$\hat{d} = \underset{1 \leq k \leq k_{max}}{\operatorname{argmin}} \left\{ \log \left( \sum_{i>k}^d \hat{\theta}_k^2 \right) + \frac{k(N+T)}{NT} \log \left( \frac{NT}{N+T} \right) \right\}, \quad (5)$$

where  $N$  is the number of assets,  $T$  is the estimation window and  $k_{max}$  is prespecified.

An alternative decomposition based on the eigenvalues of the correlation matrix is proposed by Peralta and Zareei (2016). Inspired by the network eigenvector centrality they decompose the weights in the following way:

$$\hat{w}_{mv}^* = \frac{Re}{\hat{\mu}' \hat{\Sigma}^{-1} \hat{\mu}} \hat{\mu}^* + \frac{Re}{\hat{\mu}' \hat{\Sigma}^{-1} \hat{\mu}} \sum_{k=1}^N (\hat{\lambda}_k^{-1} - 1) \hat{\nu}_k \hat{\nu}_k' \hat{\mu}^*, \quad (6)$$

where  $\hat{\lambda}_k$  is an eigenvalue of the correlation matrix  $\Omega$  and  $\hat{\nu}_k$  the corresponding eigenvector, and  $\hat{\mu}^*$  for every asset  $i$  corresponds to  $\hat{\mu}_i / \hat{\sigma}_i$ , with  $\hat{\mu}_i$  and  $\hat{\sigma}_i$  the estimated mean and standard deviation of that asset respectively. The optimal mean-variance weight for each asset  $i$  is then given by  $\hat{w}_{mv,i} = \hat{w}_{mv,i}^* / \hat{\sigma}_i$ .<sup>1</sup>

Peralta and Zareei (2016) consider assets as a network and concentrate their attention on the first eigenvector of the correlation matrix  $\Omega$ . They set the elements on the diagonal to zero (to discard loops). Thus, they compute the eigenvector centrality values of the network of assets, and build an asset allocation rule based on it. We take a different approach and rewrite equation (6) as<sup>2</sup>

$$\hat{w}_{mv,cor}^* = \frac{Re}{\hat{\mu}'^* \sum_{k=1}^d \hat{\lambda}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \hat{\mu}^*} \sum_{k=1}^d \hat{\lambda}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \hat{\mu}^*, \quad (7)$$

so that the weight for each asset  $i$  can then be obtained as  $\hat{w}_{mv,cor,i} = \hat{w}_{mv,cor,i}^* / \hat{\sigma}_i$ . In this way we can apply the same dimension reduction technique proposed by Chen and Yuan (2016), but using the correlation matrix

<sup>1</sup>We adopt this notation throughout the paper: a  $w$  with a  $*$  always indicates values that have to be divided by the standard deviation or downside deviation to obtain the portfolio weights.

<sup>2</sup>See Equation (7) in Peralta and Zareei (2016):  $w_{mv}^* = \frac{Re}{\mu' \Sigma^{-1} \mu} \sum_{k=1}^d \lambda_k^{-1} \nu_k \nu_k' \mu^*$ , obtained from Equation (2) by replacing  $\Sigma^{-1}$  and  $\mu$  at the numerator with the inverse of the correlation matrix  $\Omega^{-1} = \lambda_k^{-1} \nu_k \nu_k'$  and  $\mu^*$  respectively.

instead of the covariance matrix. Equations (4) and (7) give the same result (the mean-variance weights obtained with the plug-in approach) when  $d = N$ . However, when  $d < N$ , weights  $\hat{w}_{cov}$  and  $\hat{w}_{cor}$  might differ, as the eigenvalues and eigenvectors of the correlation matrix are different from those of the covariance matrix, and not directly related to them (Jolliffe and Cadima, 2016).

The estimation error in the sample mean is typically so big that minimizing the variance leads to portfolios with a higher Sharpe ratio than those computed via mean-variance optimization (see e.g. Jagannathan and Ma, 2003). Therefore, we also consider the minimum-variance portfolio, whose weights are given by

$$w_v = \frac{1}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}\Sigma^{-1}\mathbf{1}, \quad (8)$$

where  $\mathbf{1}$  is a vector of ones with length equal to the number of risky assets. The true value of  $\Sigma$  is unknown and has to be replaced by the sample estimate or the shrinkage estimate proposed by Ledoit and Wolf (2004). The strategy proposed by Chen and Yuan (2016) can be modified to compute a subspace minimum-variance portfolio, with weights given by

$$\hat{w}_{v,cov} = \frac{1}{\mathbf{1}'\sum_{k=1}^d \hat{\theta}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \mathbf{1}} \sum_{k=1}^d \hat{\theta}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \mathbf{1}. \quad (9)$$

Finally, in the framework of our strategy based on the correlation matrix, the subspace minimum-variance portfolio is computed using

$$\hat{w}_{v,cor}^* = \frac{1}{\epsilon' \sum_{k=1}^d \hat{\lambda}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \epsilon} \sum_{k=1}^d \hat{\lambda}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \epsilon, \quad (10)$$

where  $\epsilon$  for every asset  $i$  corresponds to  $1/\hat{\sigma}_i$ , and the weight for each asset  $i$  can then be obtained as  $\hat{w}_{v,cor,i} = \hat{w}_{v,cor,i}^* / \hat{\sigma}_i$ .

We now need to decide how many eigenvalues should be retained, and this is where the advantage of focusing on the correlation matrix instead of the covariance matrix becomes apparent. Warne and Larsen (2014) test several rules and find that the best performing one for factor analysis appears to be the Minimum Average Partial (MAP) introduced by Velicer (1976). This method requires to perform a PCA on the correlation matrix. Before describing how MAP works, we quickly recall the concept of partial correlation. Partial correlation measures, on a scale from -1 to 1, the linear association between two variables  $X$  and  $Y$ , while controlling for the effect of a set of additional variables  $\mathbf{Z}$ .<sup>3</sup> More formally, the partial correlation

---

<sup>3</sup>The partial correlation coincides with the conditional correlation only if the two variables are jointly distributed following a multivariate normal, elliptical, multivariate hypergeometric, multivariate negative hypergeometric, multinomial or Dirichlet distribution (Baba et al., 2004).

$\rho_{X,Y \cdot Z}$  between  $X$  and  $Y$  controlling for  $\mathbf{Z}$  is given by the correlation coefficient between the residuals obtained from the linear regression of  $X$  on  $\mathbf{Z}$  and  $Y$  on  $\mathbf{Z}$ . If one variable is partialled out of correlation we have a first order partial correlation; if two variables are partialled out we get a second order partial correlation, and so on. First order partial correlations can also be computed, as in Velicer (1976), using the formula

$$\rho_{i,j \cdot \nu_k} = \frac{\rho_{i,j} - \rho_{i,\nu_k} \rho_{j,\nu_k}}{((1 - \rho_{i,\nu_k}^2)(1 - \rho_{j,\nu_k}^2))^{1/2}}, \quad (11)$$

where  $i$  and  $j$  are any pair of the  $N$  variables and  $\nu_k$  is one of the  $N$  principal components. When we partial out a component we are removing both systematic and individual variability from the data. Removing the former decreases the value of the numerator, while removing the latter decreases the value of the denominator. As  $k$  goes from 1 to  $N$ , less systematic and more individual variability gets removed when the  $\nu_k$  component is partialled out, which means that the numerator decreases slower and the denominator faster. There is a value of  $k$  at which more individual than systematic variability gets removed from the data, which makes the denominator drop more than the numerator. This results in  $\rho_{i,j \cdot \nu_k}$  stopping going down and starting going up. The MAP criterion involves computing, for every pair of variables, the value of  $\rho_{i,j \cdot \nu_k}$ , and from these values the average squared partial correlation:

$$MAP_k = \sum_{i=1}^N \sum_{j=1}^N \frac{\rho_{i,j \cdot \nu_k}^2}{N(N-1)}, \quad (12)$$

with  $i \neq j$ . As  $k$  goes from 1 to  $N$ , the value of  $MAP_k$  decreases until it reaches a minimum before starting to increase, and that  $k$  for which  $MAP_k$  is minimum is the optimal number of components/factors to retain.

Despite the good performance found by Warne and Larsen (2014), MAP is not commonly employed in finance, and its application is mostly confined to psychometrics. One notable exception is given by Claeys and Vašíček (2004), who use MAP in their study on spillover and contagion effects in the European sovereign bond markets. To the best of our knowledge, we are the first to apply MAP in the context of financial asset allocation.

MAP does not depend on the sample size, and can therefore be directly applied even to mean-semivariance and minimum semivariance optimization problems, that require as input a semicovariance matrix. To understand why this is problematic in this setting, consider the formula for the semivariance of a single asset, provided by Markowitz (1959):

$$\sigma_B^2 = \frac{1}{T} \sum_{t=1}^T [\text{Min}(r_t - B, 0)]^2, \quad (13)$$

where  $T$  is the estimation window,  $r_t$  is the return of the asset in period  $t$ , and the benchmark  $B$  is the value below which the investor considers variability

to be downside variability. In Equation (13) we have the full sample size  $T$ , but to compute the sum we only consider the periods in which the return of the asset is lower than  $B$ . The computation of the semicovariance matrix  $\Sigma_{B,p}$ , described in Markowitz (1959), is based on an analogous principle. This makes the Bai and Ng (2002) criterion unsuitable to be used, since we are not taking advantage of all the information included in the estimation window. Using  $T$  as parameter in Equation (5) would be inappropriate, and so would be using the number of periods used to compute the numerator. Relying on MAP solves this issue.

There is one issue that needs to be carefully considered when implementing MAP. The method tends to overestimate the number of components to retain when factors are correlated (a frequent situation in real datasets) and when the sample size is not big enough. In order to correct for this effect we take advantage of the Kaiser-Guttman criterion, first proposed by Guttman (1954) and then popularized by Kaiser (1960). Applied to a correlation matrix, this rule provides an upper bound for the number of components to be retained by stating that only those with an eigenvalue greater than 1 should be retained.<sup>4</sup> Like MAP, this rule does not rely on the sample size used for the estimation, and can therefore be applied in our setting. In short, we use MAP to determine the exact number of components to be retained, but restrict this choice to the components associated to eigenvalues greater than 1 in order to prevent overfactoring.

## 2.2 Semicovariance matrix estimation

As optimization problems that use  $\Sigma_{B,p}$  are intractable (see Estrada, 2008), we need to use as input some matrix that approximates it. One such strategy is proposed by Estrada (2008), who devises a heuristic that involves computing the elements of the semicovariance matrix with respect to when the single assets, and not the portfolio as a whole, underperform the benchmark. This procedure yields an exogenous matrix, contrary to the exact semicovariance matrix  $\Sigma_{B,p}$ , which is endogenous, i.e. its terms change every time the portfolio weights change. However, the approximation error in this matrix can be substantial, and this can seriously hurt the portfolio performance (see Cheremushkin, 2009; Rigamonti, 2020). We therefore resort to the smoothed semivariance (SSV) estimator introduced by Rigamonti et al. (2021), which provides an extremely precise approximation of the sample semicovariance matrix using a simple re-weighting scheme that simultaneously estimates the optimal portfolio weights and the corresponding semicovariance matrix  $\hat{\Sigma}_B$ . Depending on the value of a single parameter

---

<sup>4</sup>While many scholars in the past have used it to determine the exact number of factors to retain, the Kaiser-Guttman rule is not meant (and does not work well) for this purpose, but only to set an upper bound (Gorsuch, 2015). This can limit overfactoring when applying some factor retention method, which in our case is MAP.

$\theta$  the algorithm can perform a smoothing of different intensity of the estimation. This leads to the estimation of the sample semicovariance matrix when  $\theta \rightarrow 0$ , of the sample covariance matrix when  $\theta \rightarrow \infty$ , and a mixture of the two with an intermediate value. Since we are interested in the sample semicovariance matrix, we always employ the SSV estimator with  $\theta = 1e^{-5}$ . While Rigamonti et al. (2021) define the SSV estimator for the minimum semivariance portfolio, we can use it to estimate also the semicovariance matrix of a mean-semivariance portfolio simply by computing the mean-semivariance weights in the re-weighting scheme.

### 2.3 Optimal portfolios computation and evaluation

As we argued, using the criterion introduced by Bai and Ng (2002) to select a value for  $d$  as done in Chen and Yuan (2016) does not seem appropriate. But if we transform  $\hat{\Sigma}_B$  into a downside correlation matrix we can use MAP (constrained by the upper bound set by the Kaiser-Guttman rule) to select an appropriate  $d$  and apply equation (7) or (10). The downside correlation matrix  $\hat{P}_B$  can be computed in a way analogous to how the correlation matrix is obtained from the covariance matrix:

$$\hat{P}_B = \hat{D}^{-1/2} \hat{\Sigma}_B \hat{D}^{-1/2}, \quad (14)$$

where  $\hat{D}^{-1/2}$  is a diagonal matrix in which every  $i$ th diagonal element is equal to the inverse of the estimated downside deviation  $\hat{\sigma}_{B,i}$  of the asset  $i$ .

We now have all the necessary inputs to apply our strategy that mitigates the impact of estimation errors in the semicovariance matrix. All the closed-form solutions available for the mean-variance and minimum-variance problems can be applied exactly as they are when using the SSV estimate of the semicovariance matrix. All we have to do is to replace the input parameter, and we obtain the solution to mean-semivariance and minimum-semivariance problems. This is a big advantage over other notable solutions proposed to address the intractability of semivariance minimization, like those devised by Harlow (1991) and Markowitz et al. (1993). The plug-in mean-semivariance and minimum-semivariance weights are therefore

$$\hat{w}_{ms} = \frac{Re}{\mu' \hat{\Sigma}_B^{-1} \mu} \hat{\Sigma}_B^{-1} \mu \quad (15)$$

and

$$\hat{w}_s = \frac{1}{\mathbf{1}' \hat{\Sigma}_B^{-1} \mathbf{1}} \hat{\Sigma}_B^{-1} \mathbf{1}, \quad (16)$$

respectively. The subspace mean-semivariance portfolio is computed using

$$\hat{w}_{ms,cor}^* = \frac{Re}{\hat{\mu}_B'^* \sum_{k=1}^d \hat{\lambda}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \hat{\mu}_B^*} \sum_{k=1}^d \hat{\lambda}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \hat{\mu}_B^*, \quad (17)$$



where  $\hat{\lambda}_k$  and  $\hat{\nu}_k$  are an eigenvalue and the corresponding eigenvector of the estimated downside correlation matrix  $\hat{P}_B$ , and the elements of  $\hat{\mu}_B^*$  correspond to  $\hat{\mu}_i/\hat{\sigma}_{B,i}$  for each asset  $i$ . Thus, the asset's weight is  $\hat{w}_{ms,cor,i} = \hat{w}_{ms,cor,i}^*/\hat{\sigma}_{B,i}$ , in which  $\hat{\sigma}_{B,i}$  is the downside deviation of the asset  $i$ . Analogously, the subspace minimum-semivariance portfolio is given by

$$\hat{w}_{s,cor}^* = \frac{1}{\epsilon_B' \sum_{k=1}^d \hat{\lambda}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \epsilon_B} \sum_{k=1}^d \hat{\lambda}_k^{-1} \hat{\nu}_k \hat{\nu}_k' \epsilon_B, \quad (18)$$

where for each asset  $i \in B$  corresponds to  $1/\hat{\sigma}_{B,i}$ , and the asset's weight can then be obtained as  $\hat{w}_{s,cor,i} = \hat{w}_{s,cor,i}^*/\hat{\sigma}_{B,i}$ .

Since the Sharpe ratio measures risk across the full distribution of returns (by the standard deviation), it is not very suitable to evaluate the out-of-sample results in the framework of our research question. Instead, we use as main performance indicator the Sortino ratio (Sortino and Prince, 1994):

$$\text{Sortino} = \frac{R - B}{\sigma_B}, \quad (19)$$

where  $R$  and  $\sigma_B$  are respectively the mean and downside deviation of the portfolio returns.

### 3 Principal component analysis on the semicovariance matrix

In this section we discuss the interpretation and implications of performing PCA on the semicovariance matrix. If the asset returns follow a symmetric distribution, focusing on the variance or on the semivariance makes no difference, and the investor should focus on the former, which is easier to estimate. However, in the presence of skewed return distributions the semivariance is a more suitable measure of risk, at least theoretically (see Rigamonti, 2020, for a more thorough discussion). The same reasoning holds when we perform PCA. When PCA is performed on the covariance matrix, the principal components explain the variance in the data. If instead we perform PCA on the semicovariance matrix, the principal components explain the semivariance in the data, that is, the variability only below a certain threshold.

To illustrate this point, we generate 10 series of 500000 returns using the moment-matching algorithm devised by Høyland et al. (2003), first with skewness equal to 0 for all the assets, and then equal to 1.3. The mean, covariance matrix and kurtosis are set in both cases equal to the respective sample moments of the returns of the Fama-French 10 industry portfolios from July 1969 to June 2021. Figure 1a and Figure 1b show the loading plots obtained by performing PCA on the covariance and semicovariance matrix (with  $B = 0$ ) estimated on the series with symmetrically distributed

returns. As we can see, the two graphs look very similar, with only minor differences. As a result, analyzing the variability or the downside variability in the data is theoretically equivalent.

Let us now focus on Figure 2a and Figure 2b. They show the loading plots obtained when performing PCA on the covariance and semicovariance matrix of the positively skewed returns. The two plots look completely different from each other, as the loadings in the first two loading vectors change significantly. This means that the information captured by the principal components is very different, which in turn shows that the information contained in the covariance and semicovariance matrix is different due to the skewness in the data. Moreover, Figure 1a and Figure 2a are extremely similar (as the covariance matrix cannot appreciate the different skewness), while Figure 1b and Figure 2b look entirely different.

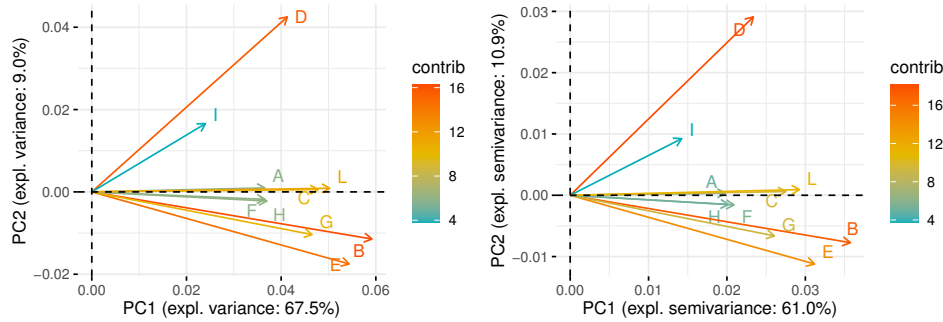


Figure 1: Loading plot with PCA performed on a) the covariance matrix (left) and b) the semicovariance matrix (right). For all variables skew = 0

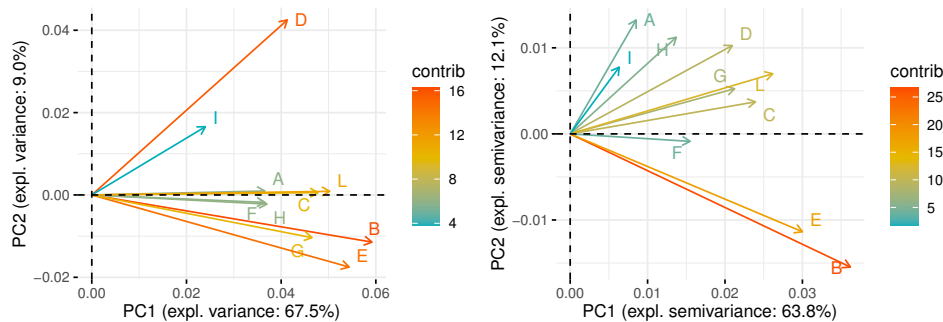


Figure 2: Loading plot with PCA performed on a) the covariance matrix (left) and b) the semicovariance matrix (right). For all variables skew = 1.3

## 4 Empirical application

We now provide an out-of-sample empirical application of our strategy. Firstly, we want to check that using MAP improves over using the Bai and Ng (2002) criterion when applied to the semicovariance matrix as if we were working with the covariance matrix like in Chen and Yuan (2016). We then compare our strategy to various competing optimal portfolios to see if we can actually improve the effectiveness of downside risk optimization. We always work with excess returns and set  $B = 0$ , so that we can have the benchmark equal to the risk-free rate, which has some desirable theoretical properties (see Hoechner et al., 2017). The estimation of the covariance matrix is always performed using the shrinkage estimator from Ledoit and Wolf (2004).

### 4.1 MAP vs Bai and Ng (2002) with downside risk

In order to compare the accuracy of MAP and the Bai and Ng (2002) criterion in selecting the optimal number of factors for the downside correlation and semicovariance matrix respectively, we compute the subspace minimum semivariance portfolio with both methods.<sup>5</sup> We also consider the unrestricted minimum semivariance portfolio to check how much we improve over it. We consider two different datasets of monthly returns: the 30 and 48 industry portfolios (value weighted) from the Kenneth R. French Data Library spanning from July 1969 (previous dates are not considered because of missing data in the 48 industry portfolios) to June 2021. We use a 180-month rolling window, so we obtain 444 out-of-sample returns.

The downside deviation achieved by the three strategies is reported in Table 1. With the 30 industry portfolios, using the Bai and Ng (2002) criterion does not lead to any improvement over the unrestricted portfolio, while our strategy that uses MAP achieves a lower downside deviation. With the 48 industry portfolios the Bai and Ng (2002) criterion improves over the unrestricted portfolio, but much less than our strategy. Overall, as expected, these results show that MAP is better suited in selecting the optimal number of factors for the subspace minimum semivariance portfolio compared to the Bai and Ng (2002) criterion.

### 4.2 Comparison between portfolio optimization strategies

We now look at how our strategy performs compared to alternative portfolio optimization strategies. In addition to the downside deviation and the Sortino ratio, we also consider two other popular downside risk measures: the Conditional Value at Risk (CVaR) and the Omega-Sharpe ratio. The

---

<sup>5</sup>To apply the Bai and Ng (2002) criterion we use the R package “nowcasting” (see de Valk et al., 2019), following all the specifications employed by Chen and Yuan (2016).

Dataset	1	2	3
30 Ind	0.0270	0.0270	0.0250
48 Ind	0.0331	0.0307	0.0259

Table 1: Downside deviation based on a 180-month rolling window. Strategies: 1 = minimum variance; 2 = subspace minimum semivariance using Bai and Ng (2002); 3 = subspace minimum semivariance using MAP

CVaR is based on the Value at Risk (VaR), to which is considered preferable, as unlike the VaR it is a coherent risk measure (see e.g. Sarykalin et al., 2008; Yamai and Yoshida, 2002). The Omega-Sharpe ratio is a simple modification of the classical Omega ratio, which leads to the same portfolio ranking but has a more familiar form (Bacon, 2008). Moreover, we look at the mean excess return and the Sharpe ratio. Finally, we compute the turnover rate at each period  $t$  as in DeMiguel et al. (2009):

$$Turnover_t = \sum_{i=1}^n |w_{i,t} - w_{i,t-1}|, \quad (20)$$

where  $i$  indicates a risky asset. We then report the average rate computed between all the periods (indicated as “TO” in the tables). An average turnover of 1 means that at each period, on average, an amount of assets equal to 100% of the value of the portfolio has to be traded.

We aim to determine which strategy works best for a real investor. Therefore, it is important to work with individual assets in order to preserve the coskewness structure, that is even more important than the idiosyncratic skewness (Albuquerque, 2012). As long series of returns are not available for a very large number of individual stocks, we work with weekly data instead of monthly data, that allow for more out-of-sample returns within a given time frame. We obtain our data from the Alpha Vantage API, which provides historical adjusted prices of single stocks starting from November 1999.<sup>6</sup> The risk-free rates were downloaded from the Kenneth R. French Data Library.<sup>7</sup> We consider three different datasets: the first includes 28 of the 30 companies that make up the Dow Jones Industrial Average, the second consists of 66 stocks included in the NASDAQ 100, and the third includes 82 components of the S&P 100.<sup>8</sup> All series span from November 19, 1999, to June 25, 2021 (the last period for which the risk-free rate is

<sup>6</sup><https://www.alphavantage.co/>

<sup>7</sup>[http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)

<sup>8</sup>The sets of stocks included in each dataset were obtained by looking at the current components of the given stock index, and selecting those for which the prices were already available at the first time period. In this way we obtain the longest possible series for the largest possible dataset.

available on the Kenneth R. French Data Library), for a total of 1128 observations. We use a 260-period rolling window for the estimation, obtaining 868 out-of-sample returns. To check the robustness of our results, we also consider an expanding window that starts with 260 periods, obtaining once again 868 out-of-sample returns.

We first compute the subspace mean-semivariance portfolio and compare it with the mean-variance and mean-semivariance portfolios, and with the mean-CVaR portfolio with the confidence interval set to 95%. To compute the mean-CVaR portfolio, we rely on the historical return distribution and use the formulation proposed by Rockafellar and Uryasev (2000). We set the target return at 0.1%. The results are shown in Table 2. While our strategy leads to a higher downside deviation compared to the competing strategies, it achieves a much higher mean return and therefore it always significantly outperforms the other strategies in terms of Sortino ratio, which is the measure that we are targeting. Moreover, we also achieve a much higher Sharpe and Omega-Sharpe ratio. The results obtained with the expanding window, shown in Table 3, are completely in line with those obtained with the rolling window, confirming the superiority of our approach.

Strategy	Mean	DwDev	Sharpe	CVaR	Omega	Sortino	TO
Panel A: DJIA (28 assets)							
Mean-V	0.0004	0.0026	0.0501	-0.0083	0.1399	0.0704	0.0570
Mean-CVaR	0.0004	0.0031	0.0376	-0.0099	0.1033	0.0528	0.1539
Mean-SV	0.0004	0.0029	0.0276	-0.0092	0.0755	0.0381	0.0734
Mean-SV MAP	0.0012	0.0094	0.0677	-0.0294	0.2890	<i>0.1017</i>	0.2285
Panel B: NDX (66 assets)							
Mean-V	0.0004	0.0020	0.0530	-0.0063	0.1494	0.0751	0.0481
Mean-CVaR	0.0004	0.0024	0.0475	-0.0079	0.1336	0.0673	0.1575
Mean-SV	0.0004	0.0022	0.0607	-0.0072	0.1737	0.0872	0.0767
Mean-SV MAP	0.0008	0.0037	0.1075	-0.0123	0.3444	<i>0.1549</i>	0.1493
Panel C: S&P 100 (82 assets)							
Mean-V	0.0004	0.0022	0.0588	-0.0072	0.1798	0.0805	0.0666
Mean-CVaR	0.0003	0.0031	0.0151	-0.0098	0.0448	0.0197	0.2427
Mean-SV	0.0003	0.0029	0.0211	-0.0091	0.0643	0.0275	0.1268
Mean-SV MAP	0.0007	0.0039	0.0883	-0.0129	0.2811	<i>0.1195</i>	0.2283

Table 2: Performance comparison; 260-week rolling window

We now consider the subspace minimum semivariance portfolio, whose competitors are the minimum variance and minimum semivariance portfolios, both with and without short selling, and the long-only minimum CVaR portfolio. The motivation for computing the minimum semivariance portfolio when the goal is to achieve a high Sortino ratio is the same for computing

Strategy	Mean	DwDev	Sharpe	CVaR	Omega	Sortino	TO
Panel A: DJIA (28 assets)							
Mean-V	0.0005	0.0038	0.0450	-0.0119	0.1298	0.0641	0.0287
Mean-CVaR	0.0005	0.0041	0.0400	-0.0132	0.1150	0.0580	0.0744
Mean-SV	0.0004	0.0039	0.0358	-0.0124	0.1019	0.0513	0.0337
Mean-SV MAP	0.0020	0.0160	0.0673	-0.0507	0.3263	<i>0.1081</i>	0.0528
Panel B: NDX (66 assets)							
Mean-V	0.0004	0.0022	0.0553	-0.0071	0.1641	0.0777	0.0152
Mean-CVaR	0.0003	0.0024	0.0052	-0.0077	0.0143	0.0072	0.0607
Mean-SV	0.0003	0.0022	0.0280	-0.0071	0.0798	0.0385	0.0185
Mean-SV MAP	0.0010	0.0040	0.1243	-0.0135	0.4052	<i>0.1776</i>	0.0553
Panel C: S&P 100 (82 assets)							
Mean-V	0.0005	0.0026	0.0751	-0.0086	0.2462	0.1052	0.0232
Mean-CVaR	0.0005	0.0034	0.0468	-0.0101	0.1518	0.0624	0.0957
Mean-SV	0.0004	0.0032	0.0457	-0.0099	0.1525	0.0598	0.0338
Mean-SV MAP	0.0010	0.0049	0.1070	-0.0167	0.3537	<i>0.1547</i>	0.1532

Table 3: Performance comparison; 260-week expanding window

the minimum variance portfolio when the goal is to achieve a high Sharpe ratio: the estimation error in the sample mean is so large that ignoring the mean altogether improves the results (Jagannathan and Ma, 2003). The additional constraint of excluding short selling can dramatically improve the out-of-sample performance when minimizing risk, as shown by Jagannathan and Ma (2003). Kremer et al. (2018) analyze the performance of several investment strategies and they find that the long-only minimum variance is the best risk minimization strategy. On the same grounds we impose the long-only constraint on the minimum CVaR portfolio, as targeting the CVaR without imposing additional constraints is known to perform poorly (Lim et al., 2011).

The results are reported in Table 4. Our strategy leads to a modest improvement over the unrestricted minimum semivariance portfolio with the Dow Jones Industrial Average, and to a large improvement with the other two datasets. Unsurprisingly, reducing the dimension of the problem is especially beneficial with large asset menus. We also improve in terms of the other two downside risk-based measures and achieve a higher Sharpe ratio, except, for the latter measure, with the Dow Jones Industrial Average. However, compared to the other four strategies, we only achieve a good Sortino and Omega-Sharpe ratio with the NASDAQ 100 dataset, while otherwise we always perform equally or worse in terms of every measure. Table 5 shows the results obtained with the expanding window. Our strategy now leads to a portfolio with a worse downside deviation and CVaR even compared to the

unrestricted minimum semivariance portfolio with the two smaller datasets, but we achieve the best results in terms of both measures with the S&P 100. We always achieve a higher Sharpe, Omega-Sharpe and Sortino ratio than the unrestricted minimum variance and minimum semivariance portfolios, but we are not competitive compared to the other three strategies.

Overall, our strategy is very competitive in a setting where the investor with a mean-semivariance utility sets a target return. However, when  $\mu$  is ignored and the investor only minimizes risk, the simple solution of excluding short positions appears to be superior in both cases where the variance and the semivariance are minimized.

Strategy	Mean	DwDev	Sharpe	CVaR	Omega	Sortino	TO
Panel A: DJIA (28 assets)							
Min V	0.0019	0.0141	0.0872	-0.0465	0.2894	0.1174	0.0762
Min V lo	0.0021	0.0137	0.0977	-0.0450	0.3264	0.1339	0.0362
Min CVaR lo	0.0019	0.0137	0.0890	-0.0458	0.2844	0.1240	0.0490
Min SV	0.0017	0.0153	0.0702	-0.0501	0.2171	0.0961	0.1281
Min SV lo	0.0020	<i>0.0134</i>	0.0953	-0.0442	0.3109	0.1334	0.0541
Min SV MAP	0.0016	0.0144	0.0678	-0.0472	0.2179	0.0939	0.0862
Panel B: NDX (66 assets)							
Min V	0.0021	0.0142	0.0930	-0.0447	0.3001	0.1278	0.1388
Min V lo	0.0027	<i>0.0141</i>	0.1257	-0.0447	0.4418	0.1770	0.0394
Min CVaR lo	0.0026	0.0154	0.1111	-0.0477	0.3811	0.1518	0.0581
Min SV	0.0021	0.0169	0.0799	-0.0557	0.2456	0.1114	0.2904
Min SV lo	0.0029	0.0143	0.1291	-0.0450	0.4473	0.1827	0.0888
Min SV MAP	0.0030	0.0152	0.1261	-0.0504	0.4271	0.1814	0.1724
Panel C: S&P 100 (82 assets)							
Min V	0.0019	0.0142	0.0893	-0.0465	0.2996	0.1161	0.1845
Min V lo	0.0022	0.0138	0.1068	-0.0440	0.3726	0.1436	0.0522
Min CVaR lo	0.0022	0.0141	0.1006	-0.0442	0.3338	0.1357	0.0619
Min SV	0.0021	0.0204	0.0648	-0.0672	0.2072	0.0884	0.5767
Min SV lo	0.0021	<i>0.0137</i>	0.0987	-0.0434	0.3408	0.1361	0.1359
Min SV MAP	0.0019	0.0142	0.0856	-0.0453	0.2846	0.1144	0.2444

Table 4: Performance comparison; 260-week rolling window

We conclude our analysis by computing the dynamics of the value of the different portfolios. Given that our strategy has a higher average turnover rate compared to most of the competitor strategies we considered, we compute the dynamics of wealth net of transaction costs. Following Frazzini et al. (2012), we set transactions costs equal to 10 basis points. Figure 3 shows the results for an investor that sets a target return. Our strategy greatly outperforms all the alternatives, especially when using an expanding

Strategy	Mean	DwDev	Sharpe	CVaR	Omega	Sortino	TO
Panel A: DJIA (28 assets)							
Min V	0.0018	<i>0.0141</i>	0.0807	-0.0460	0.2614	0.1103	0.0127
Min V lo	0.0021	0.0142	0.0928	-0.0468	0.3109	0.1279	0.0071
Min CVaR lo	0.0019	0.0150	0.0810	-0.0491	0.2624	0.1122	0.0155
Min SV	0.0017	0.0147	0.0714	-0.0483	0.2268	0.0960	0.0164
Min SV lo	0.0021	0.0142	0.0923	-0.0465	0.3044	0.1272	0.0096
Min SV MAP	0.0020	0.0150	0.0852	-0.0501	0.2812	0.1188	0.0127
Panel B: NDX (66 assets)							
Min V	0.0024	0.0150	0.1030	-0.0483	0.3477	0.1424	0.0203
Min V lo	0.0028	0.0148	0.1231	-0.0484	0.4381	0.1745	0.0062
Min CVaR lo	0.0031	0.0153	0.1288	-0.0498	0.4638	0.1839	0.0140
Min SV	0.0021	0.0157	0.0873	-0.0506	0.2851	0.1191	0.0319
Min SV lo	0.0029	<i>0.0147</i>	0.1292	-0.0474	0.4639	0.1819	0.0151
Min SV MAP	0.0029	0.0161	0.1139	-0.0531	0.3945	0.1660	0.0284
Panel C: S&P 100 (82 assets)							
Min V	0.0020	0.0139	0.0962	-0.0461	0.3191	0.1283	0.0325
Min V lo	0.0024	0.0139	0.1136	-0.0462	0.4015	0.1576	0.0102
Min CVaR lo	0.0023	0.0142	0.1068	-0.0465	0.3688	0.1484	0.0220
Min SV	0.0017	0.0160	0.0677	-0.0543	0.2145	0.0883	0.0639
Min SV lo	0.0023	0.0141	0.1053	-0.0466	0.3695	0.1455	0.0255
Min SV MAP	0.0022	<i>0.0136</i>	0.1041	-0.0447	0.3495	0.1455	0.0509

Table 5: Performance comparison; 260-week expanding window

estimation window. Figure 4 reports the value of the portfolios when risk is minimized without considering the mean return. Our strategy almost always significantly outperform the unrestricted minimum variance and minimum semivariance portfolios. However, as expected from the results in Table 4 and Table 5, we generally perform similarly or worse than the strategies that impose a long-only constraint.

## 5 Conclusions

Efficient asset allocations based on downside risk are hard to achieve because it is especially difficult to estimate the necessary inputs. In this paper we propose for the first time a strategy that directly addresses this problem when estimating the semicovariance matrix of a set of assets. In doing so, we provide two additional original contributions by performing PCA on the semicovariance matrix and on the downside correlation matrix, and by using MAP to select the optimal number of components for asset allocation decisions. The application of our strategy to a variety of datasets shows that



it can considerably improve the out-of-sample performance over portfolios constructed using the unrestricted semicovariance matrix. It is also very competitive in a setting where the investor sets a target return and performs mean-(semi)variance optimization. Given the relatively high turnover of our strategy compared to most alternatives considered in this study, one direction for future research could be the refinement of our strategy in order to reduce its transaction costs. This would further increase its attractiveness for the investors.

## Appendix

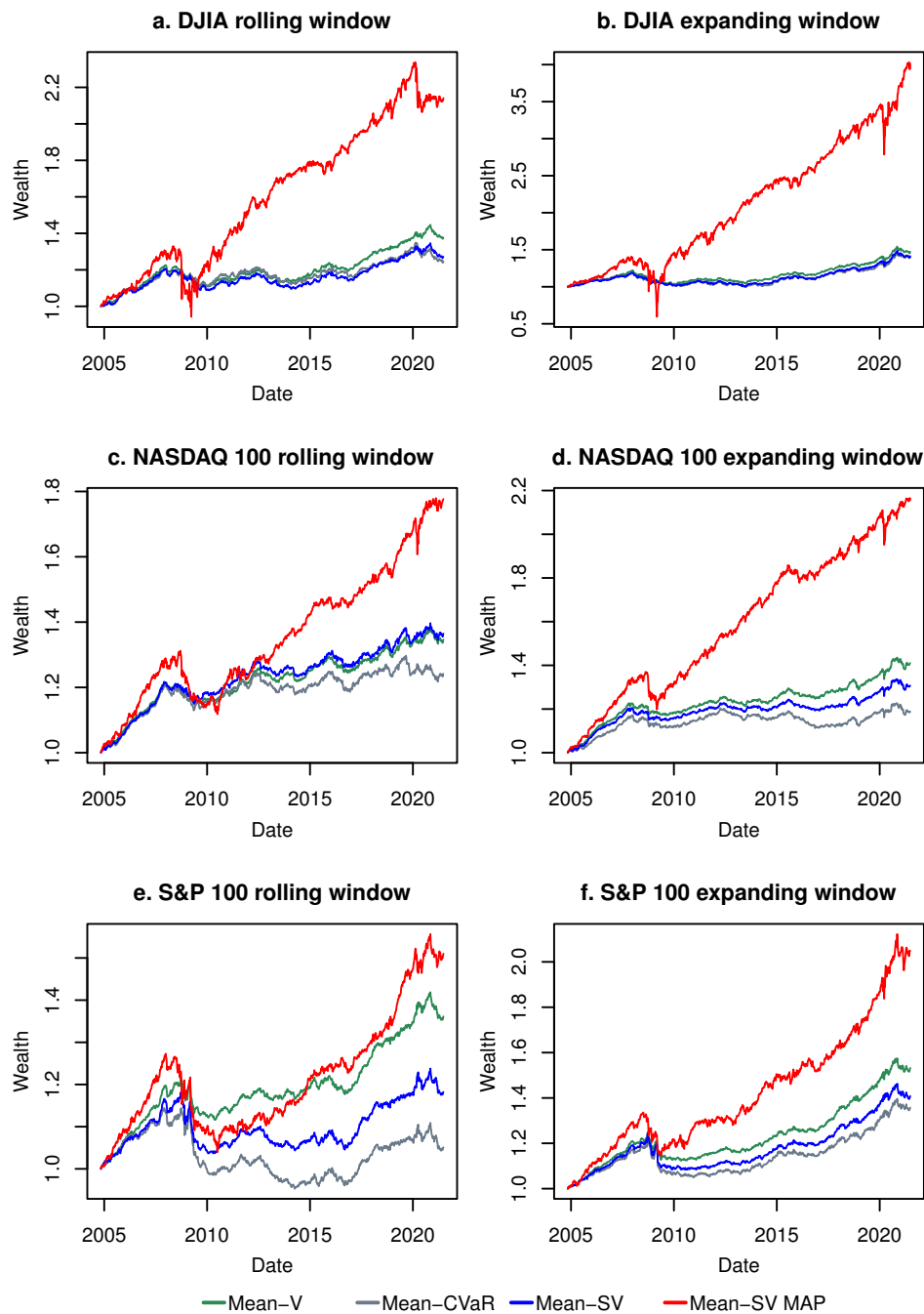


Figure 3: Portfolio value net of transaction costs

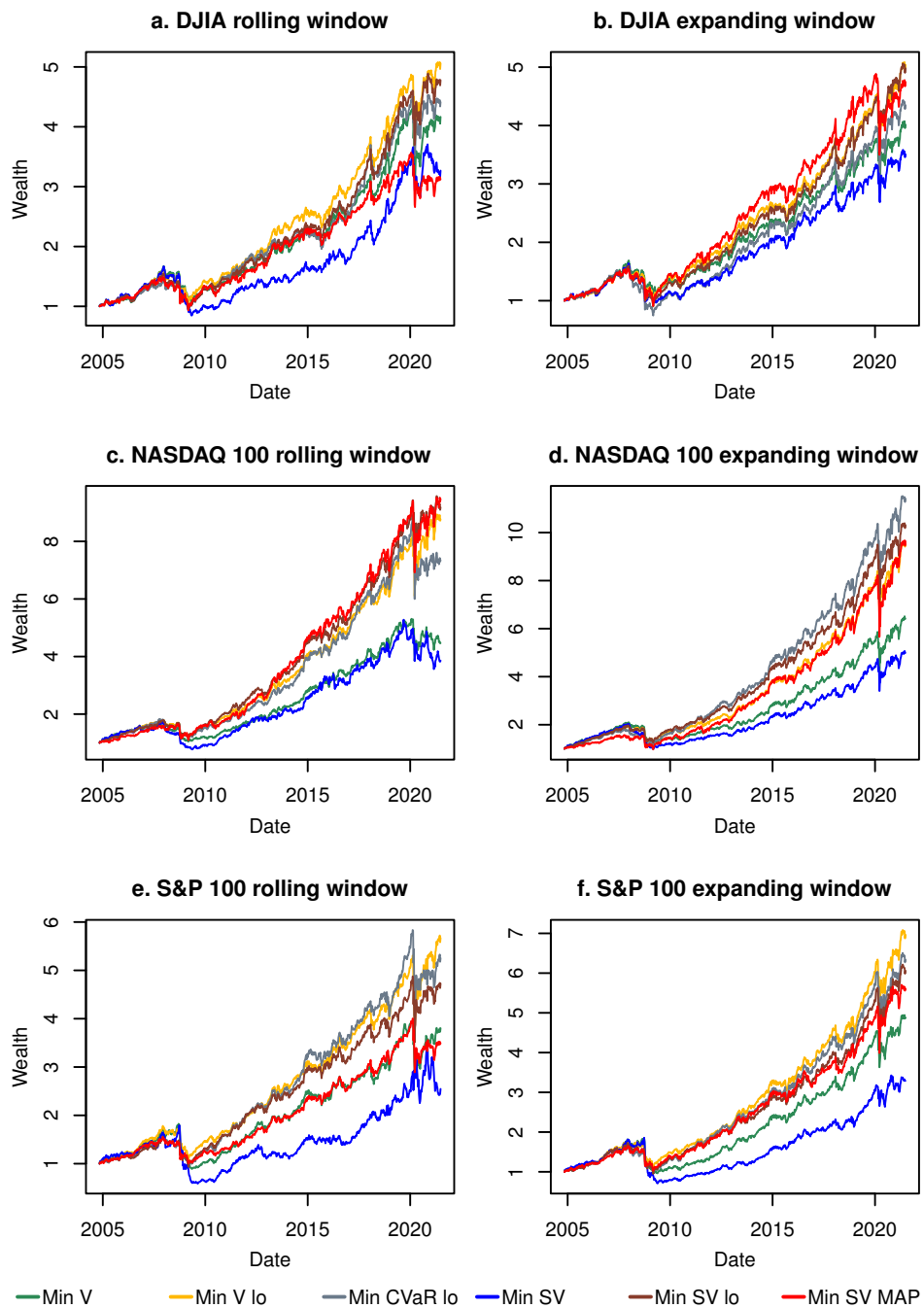


Figure 4: Portfolio value net of transaction costs

## References

- Albuquerque, R. (2012). Skewness in stock returns: Reconciling the evidence on firm versus aggregate returns. *The Review of Financial Studies*, 25(5):1630–1673.
- Baba, K., Shibata, R., and Sibuya, M. (2004). Partial correlation and conditional correlation as measures of conditional independence. *Australian & New Zealand Journal of Statistics*, 46(4):657–664.
- Bacon, C. (2008). *Practical portfolio performance measurement and attribution*. John Wiley & Sons Ltd, 2 edition.
- Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.
- Broadie, M. (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, 45:21–58.
- Chen, J. and Yuan, M. (2016). Efficient portfolio selection in a large market. *Journal of Financial Econometrics*, 14(3):496–524.
- Chen, W., Gai, Y., and Gupta, P. (2018). Efficiency evaluation of fuzzy portfolio in different risk measures via dea. *Annals of Operations Research*, 269(1-2):103–127.
- Cheremushkin, S. V. (2009). Why D-CAPM is a big mistake? The incorrectness of the cosemivariance statistics. Available at SSRN: <https://ssrn.com/abstract=1336169>.
- Chopra, V. K. and Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances in optimal portfolio choice. *Journal of Portfolio Management*, 19(2):6–11.
- Claeys, P. and Vašíček, B. (2004). Measuring bilateral spillover and testing contagion on sovereign bond markets in europe. *Journal of Banking & Finance*, 46(1):151–165.
- de Valk, S., da Mattos, D., and Ferreira, P. (2019). Nowcasting: An R package for predicting economic variables using dynamic factor models. *The R Journal*, 11(1):230–244.
- DeMiguel, V., Garlappi, L., Nogales, F. J., and 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.

- Frazzini, A., Israel, R., and Moskowitz, T. (2012). Trading costs of asset pricing anomalies. Working Paper SSRN.
- Gorsuch, R. L. (2015). *Factor Analysis*. Routledge.
- Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, 19(2):149–161.
- Harlow, W. V. (1991). Asset allocation in a downside-risk framework. *Financial Analysts Journal*, 47(5):28–40.
- Hoechner, B., Reichling, P., and Schulze, G. (2017). Pitfalls of downside performance measures with arbitrary targets. *International Review of Finance*, 17(4):597–610.
- Høyland, K., Kaut, M., and Wallace, S. W. (2003). A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24(2-3):169–185.
- Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4):1651–1684.
- Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065).
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1):141–151.
- Kan, R. and Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *The Journal of Financial and Quantitative Analysis*, 42(3):621–656.
- Kremer, P. J., Talmaciu, A., and Paterlini, S. (2018). Risk minimization in multi-factor portfolios: What is the best strategy? *Annals of operations research*, 266(1):255–291.
- Ledoit, A. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411.
- Lim, A. E., Shanthikumar, J. G., and Vahn, G. (2011). Conditional value-at-risk in portfolio optimization: Coherent but fragile. *Operations Research Letters*, 39(3):163–171.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

- Markowitz, H. (1959). *Portfolio selection*. John Wiley and Sons.
- Markowitz, H., Todd, P., Xu, G., and Y., Y. (1993). Computation of mean-semivariance efficient sets by the Critical Line Algorithm. *Annals of Operations Research*, 45:307–317.
- Michaud, O. (1989). The Markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, 45(1):31–42.
- Peralta, G. and Zareei, A. (2016). A network approach to portfolio selection. *Journal of Empirical Finance*, 38:157–180.
- Pla-Santamaria, D. and 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.
- Rigamonti, A. (2020). Mean-variance optimization is a good choice, but for other reasons than you might think. *Risks*, 8(1).
- Rigamonti, A., Weissensteiner, A., Ferrari, D., and Paterlini, S. (2021). Smoothed semicovariance estimation. Available at SSRN: <https://ssrn.com/abstract=3786023>.
- Rockafellar, R. T. and Uryasev, S. (2000). Optimization of Conditional Value-at-Risk. *Journal of risk*, 2(3):21–41.
- Sarykalin, S., Serraino, G., and Uryasev, S. (2008). Value-at-Risk vs. Conditional Value-at-Risk in risk management and optimization. *INFORMS Tutorials in Operations Research*, pages 270–294.
- Sharma, A., Utz, S., and Mehra, A. (2017). Omega-cvar portfolio optimization and its worst case analysis. *OR Spectrum*, 39:505–539.
- Sortino, F. A. and Prince, L. N. (1994). Performance measurement in a downside risk framework. *The Journal of Investing*, 3(3):59–64.
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3):321–327.
- Warne, R. T. and Larsen, R. (2014). Evaluating a proposed modification of the guttman rule for determining the number of factors in an exploratory factor analysis. *Psychological Test and Assessment Modeling*, 56(1):104–123.
- Yamai, Y. and Yoshida, T. (2002). On the validity of Value-at-Risk: Comparative analyses with expected shortfall. *Monetary and economic studies*, 20(1):57–85.

Yue, W., Wang, Y., and Xuan, H. (2019). Fuzzy multi-objective portfolio model based on semi-variance-semi-absolute deviation risk measures. *Soft Computing*, 23(17):8159–8179.