



# A risk perspective of estimating portfolio weights of the global minimum-variance portfolio

Thomas Holgersson<sup>1</sup> · Peter Karlsson<sup>1</sup> · Andreas Stephan<sup>2</sup> 

Received: 28 March 2017 / Accepted: 22 December 2018 / Published online: 4 February 2019  
© The Author(s) 2019

## Abstract

The problem of how to determine portfolio weights so that the variance of portfolio returns is minimized has been given considerable attention in the literature, and several methods have been proposed. Some properties of these estimators, however, remain unknown, and many of their relative strengths and weaknesses are therefore difficult to assess for users. This paper contributes to the field by comparing and contrasting the risk functions used to derive efficient portfolio weight estimators. It is argued that risk functions commonly used to derive and evaluate estimators may be inadequate and that alternative quality criteria should be considered instead. The theoretical discussions are supported by a Monte Carlo simulation and two empirical applications where particular focus is set on cases where the number of assets ( $p$ ) is close to the number of observations ( $n$ ).

**Keywords** Global minimum-variance portfolio · Portfolio theory · High dimensional · Risk functions

**JEL Classification** C13 · C18 · C44 · G11

---

We thank two anonymous reviewers for their insightful comments and helpful suggestions and the session participants at the German Statistical Society conference in Augsburg 2016 and those attending the seminar at the European University Viadrina in January 2017 for useful suggestions. The usual disclaimer applies.

---

✉ Andreas Stephan  
andreas.stephan@ju.se

Thomas Holgersson  
thomas.holgersson@lnu.se

Peter Karlsson  
peter.karlsson@lnu.se

<sup>1</sup> Present Address: Department of Economics and Statistics, Linnæus University, 351 95 Växjö, Sweden

<sup>2</sup> Jönköping International Business School, Jönköping, Sweden

## 1 Introduction

Markowitz (1952, 1959) developed the theoretical foundation for the modern portfolio theory<sup>1</sup> providing investors with a tool to solve a key issue of how to distribute their wealth among a set of available assets. The problem was postulated as a choice of a portfolio mean return and variance of portfolio returns. This led to two principles where an investor for a given level of portfolio variance maximizes the portfolio return, or likewise for a given portfolio return minimizes the portfolio variance. Hence, according to Markowitz, an investor under such constraints needs only to be concerned about two moments of the assets multivariate distribution: the mean vector and the covariance matrix. In practical situations, these two quantities are unknown and must be estimated in order to perform the optimization. Many authors have shown that this optimization procedure often fails in practice (Bai et al. 2009; Best and Grauer 1991; Kempf et al. 2002; Merton 1980). Some authors go so far as to argue that the estimation error (sampling variance) dominates the procedure to the extent that the equally weighted non-random portfolio performs better than those optimized from data (Frankfurter et al. 1971; DeMiguel et al. 2009; Michaud 1989).

The estimation problem becomes particularly serious when the numbers of assets (say  $p$ ) are close to the number of observations ( $n$ ). This is mainly so because the sample covariance matrix becomes stochastically unstable and may not even be invertible. It then comes natural that the standard “plug-in” estimators, defined by simply replacing the unknown mean vector and covariance matrix by the standard text book estimators, should be replaced by some improved covariance estimator. A significant number of improvements over the plug-in estimator have been developed over the last decades (Frost and Savarino 1986; Ledoit and Wolf 2003, 2004). The vast majority of these fall into two categories or families of estimators. The first category is based on the fact that the sample covariance matrix is a poor approximation of the true covariance matrix and, therefore, the estimation problem is concerned with developing improved estimators of the covariance. These improved estimators are then simply substituted for the unknown parameter within Markowitz’s optimal weight function. The other approach, which appears to be the more common one during recent years, relies on principles developed by Stein (1956) and James and Stein (1961). With such estimators, the standard estimator is weighted (“shrunk”) toward a non-random target quantity. This type of estimators seems to have found a new renaissance in portfolio estimation theory, for which they appear to be particularly well suited. A sample of some recent developments includes Bodnar et al. (2018), Frahm and Memmel (2010), Golosnoy and Okhrin (2007), Kempf and Memmel (2006), Okhrin and Schmid (2007). Each of these methods naturally has its own merits and uses.

Investors want to know the basic properties and the relative risk of favoring one estimator over another before applying any specific method. There is, however, no consensus about how the concept of “risk” should be defined. From a statistical point of view, risk refers to some measure of the difference between a quantity of interest (which could be either random or fixed) and our inference target. Risk is usually

---

<sup>1</sup> Roy (1952) wrote a paper around the same time as Markowitz in which he developed a similar theoretical foundation.

expressed through moments of differences, such as the mean squared error (MSE), but could also involve forecast bias, or angles between true and estimated vectors. It is obvious that the optimality properties of any estimator depend on the specific risk measure, or quality criteria, being used to describe it. Indeed, it is well known that the ranking of estimators' performance, such as estimators of the inverse covariance matrix, may change or even reverse when evaluated with alternative loss functions (Haff 1979; Muirhead 1982).

While most recent papers in portfolio optimization theory have been concerned with the extremely important problem of developing efficient estimators of portfolio weights and related quantities, this paper focuses on the risks associated with these estimators. Specifically, the purpose of this paper is to compare and contrast a number of risk measures of the GMVP estimator to give investors and developers of statistical methods a fair understanding of their differences and similarities and, hence, a foundation for determining the weight estimator that is best suited for a given specific problem.

The paper proceeds as follows. In Sect. 2, the problem of minimum-variance portfolio estimation is stated, Sect. 3 introduces the concept of risk function, and Sect. 4 classifies different GMVP estimators. Section 5 describes the Monte Carlo study design and provides a discussion of the derived results. Section 6 outlines two empirical applications, and Sect. 7 summarizes the findings and concludes.

## 2 Preliminaries

We consider the problem of constructing an investment portfolio  $\Pi$ , defined as a weighted sum of risky assets,  $\mathbf{R} = (R_1, \dots, R_p)'$ . In order to construct a portfolio of assets, we define a vector of weights  $\mathbf{v} \in \mathbb{R}^p$  under the common constraint  $\mathbf{v}'\mathbf{1} = 1$ , where  $\mathbf{1}: p \times 1$  is a vector of ones. The mean and variance of  $\mathbf{R}$  are defined by  $\boldsymbol{\mu} := E[\mathbf{R}]$  and  $\boldsymbol{\Sigma} := \text{Cov}[\mathbf{R}]$ , respectively, where  $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$  is positive definite by assumption. The variance of the portfolio excess return is uniquely minimized by the *global minimum-variance portfolio* (GMVP) which is given by solving the following minimization problem:

$$\underset{\substack{\mathbf{w} \\ (p \times 1)}}{\text{arg min}} (\mathbf{v}'\boldsymbol{\Sigma}\mathbf{v}), \tag{1}$$

and the well-known solution to (1) is given by:

$$\underset{\substack{\mathbf{w} \\ (p \times 1)}}{\text{arg min}} = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}. \tag{2}$$

The expected return and the return variance of the global minimum-variance (MV) portfolio are given by

$$\mu_{MV} = \boldsymbol{\mu}'\mathbf{w} = \frac{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}$$

and

$$\sigma_{\text{MV}}^2 = \mathbf{w}' \boldsymbol{\Sigma} \mathbf{w} = \frac{\mathbf{1}}{\mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1}}.$$

The weight vector  $\mathbf{w}$  depends on the unknown parameter  $\boldsymbol{\Sigma}^{-1}$  which needs to be estimated from data. The classical, or plug-in estimator, of the weight vector  $\mathbf{w}$  is obtained by replacing  $\boldsymbol{\Sigma}$  in (2) by the inverse of the sample covariance matrix  $\mathbf{S} := n^{-1} \sum_{i=1}^n (\mathbf{R}_i - \bar{\mathbf{R}}) (\mathbf{R}_i - \bar{\mathbf{R}})'$ , where  $\bar{\mathbf{R}} := n^{-1} \sum_{i=1}^n \mathbf{R}_i$ . We define this estimator as

$$\hat{\mathbf{w}}_I = \frac{\mathbf{S}^{-1} \mathbf{1}}{\mathbf{1}' \mathbf{S}^{-1} \mathbf{1}}. \quad (3)$$

The distribution of  $\hat{\mathbf{w}}_I$  when sampling from a Gaussian distribution is well established (Okhrin and Schmid 2006; Bodnar and Zabolotsky 2017). In particular, the estimator  $\mathbf{S}^{-1}$  is, by the law of large numbers, a consistent estimator of  $\boldsymbol{\Sigma}^{-1}$  and the consistency of (3) follows directly. However, it is well known that  $\mathbf{S}^{-1}$  is a poor approximation of  $\boldsymbol{\Sigma}^{-1}$  when the number of assets  $p$  is large relative to the number of observations  $n$ , and as a consequence  $\hat{\mathbf{w}}_I$  will not adequately approximate  $\mathbf{w}$ . Because of this, a number of authors have suggested improved estimators of the GMVP weights. An obvious solution is to simply replace  $\mathbf{S}^{-1}$  with a more efficient estimator of  $\boldsymbol{\Sigma}^{-1}$ . Another important family of improved estimators is given by Stein-type estimators, which, in terms of our portfolio estimation problem, are of the form  $\hat{\mathbf{w}}_S = a \hat{\mathbf{w}} + b \mathbf{w}_0$ , where  $a$  and  $b$  are constants, and  $\mathbf{w}_0$  is a pre-determined reference portfolio, which is usually defined non-random. If  $(\hat{\mathbf{w}} - \mathbf{w}_0)$  is small, the improvement in terms of the mean squared error  $E \left[ (\hat{\mathbf{w}} - \mathbf{w})' (\hat{\mathbf{w}} - \mathbf{w}) \right]$  may be considerable.

### 3 The risk of portfolio estimators

Generally speaking, there are several ways to view the weight vector  $\mathbf{w}$  and the formulation of the inference problem. When deriving statistical estimators and inference procedures for portfolio weights, there is no consensus regarding which quantity to optimize. For example, since  $\mathbf{w} \in \mathbb{R}^p$ , it is natural to think of estimators in terms of estimating a parameter vector. Alternatively, upon noting that  $\mathbf{w} = (\mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1})^{-1} \boldsymbol{\Sigma}^{-1} \mathbf{1}$  only depends on the unknown parameter  $\boldsymbol{\Sigma}^{-1} \in \mathbb{R}^{p \times p}$ , the inference problem may be thought of as one concerned with estimating a matrix, a problem rather different from that of estimating a vector. Yet another view of the inference problem is the out-of-sample prediction variance (Frahm and Memmel 2010). It is obvious that the properties of any estimator  $\hat{\mathbf{w}}$  will depend on the quality criteria, or risk function, being used to judge it and that no single estimator can optimize all relevant properties simultaneously. In fact, the performance ranking of estimators of  $\mathbf{w}$  may be changed or even reversed when evaluated on alternative loss functions (Haff 1979; Muirhead 1982). An investor searching the literature for “the best” estimator of the GMVP is

likely to end up with a battery of proposed estimators, each being “optimal” in some sense. In this section, we will present and discuss similarities and differences between a number of risk functions for the GMVP problem, some of which are commonly used while others appears to be new in the GMVP context.

The  $L_2$ -norm risk is commonly used to assess estimators of  $\Sigma^{-1}$  alone, i.e., without respect to the bigger problem in which it is an ingredient. It is defined as follows:

$$\mathfrak{R}_0(\hat{\Sigma}^{-1}) := p^{-1}trE\left[\left(\hat{\Sigma}^{-1} - \Sigma^{-1}\right)^2\right].$$

The  $\mathfrak{R}_0(\hat{\Sigma}^{-1})$  risk is clearly naïve, in the sense that it is only indirectly related to the actual problem of estimating the optimal weight vector  $\mathbf{w}$ . In other words, an estimator of  $\Sigma^{-1}$  which is optimal with respect to  $\mathfrak{R}_0$  need not perform very well when substituted into Eq. (2).

A risk function that is more adequate for the portfolio weight problem may be derived as follows: The denominator of  $\mathbf{w} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}$  is merely a scaling factor used to impose the length-one condition. The quantity  $\Sigma^{-1}\mathbf{1}$  contains all necessary information for determining  $\mathbf{w}$  and the task of estimating  $\mathbf{w}$  reduces to that of estimating  $\Sigma^{-1}\mathbf{1}$ . A risk function for the minimum-variance portfolio estimator may accordingly be defined by

$$\begin{aligned} \mathfrak{R}_1(\hat{\Sigma}, \Omega) &:= p^{-1}E\left[\left(\hat{\Sigma}^{-1}\mathbf{1} - \Sigma\mathbf{1}\right)' \Omega \left(\hat{\Sigma}^{-1}\mathbf{1} - \Sigma^{-1}\mathbf{1}\right)\right] \\ &= p^{-1}\mathbf{1}'\left\{E\left[\left(\hat{\Sigma}^{-1} - \Sigma^{-1}\right) \Omega \left(\hat{\Sigma}^{-1} - \Sigma^{-1}\right)\right]\right\}\mathbf{1} \end{aligned}$$

where  $\Omega$  is a positive semi-definite non-random matrix. Two important special cases are given by

$$\mathfrak{R}_1(\hat{\Sigma}^{-1}, \mathbf{I}) = p^{-1}\mathbf{1}'\left\{E\left[\left(\hat{\Sigma}^{-1} - \Sigma^{-1}\right)^2\right]\right\}\mathbf{1}$$

and

$$\mathfrak{R}_1(\hat{\Sigma}^{-1}, \mathbf{1}\mathbf{1}') = p^{-1}E\left[\left(\mathbf{1}'\left(\hat{\Sigma}^{-1} - \Sigma^{-1}\right)\mathbf{1}\right)^2\right].$$

Frahm and Memmel (2010) utilize a somewhat different risk function defined by

$$\mathfrak{R}_2(\hat{\mathbf{w}}, \Omega) := p^{-1}E\left[(\hat{\mathbf{w}} - \mathbf{w})' \Omega (\hat{\mathbf{w}} - \mathbf{w})\right].$$

The  $\mathfrak{R}_2$  risk function has been used frequently in the context of estimating mean value vectors and in regression analysis (Anderson 2003; Efron and Morris 1976; James and Stein 1961; Muirhead 1982; Serdobolskii 2000; Srivastava 2002). This risk function explicitly evaluates the second-order moment properties (variance plus squared bias)

of an estimator  $\hat{\mathbf{w}}$ . It allows us to conveniently split mean square errors in different directions to  $\mathbf{w}$  by appropriate choice of  $\mathbf{\Omega}$ .

Yet another risk criterion, the out-of-sample variance of  $\hat{\mathbf{w}}'\mathbf{R}_m$ , advocated by Frahm and Memmel (2010), is somewhat different from  $\mathfrak{R}_1$  and  $\mathfrak{R}_2$ . For some  $\mathbf{R}_m$  not included in the estimate of  $\hat{\mathbf{w}}$ , this variance is determined by

$$\mathfrak{R}_3(\hat{\mathbf{w}}) = \text{Var}(\hat{\mathbf{w}}'\mathbf{R}_m). \tag{4}$$

Following Frahm and Memmel (2010), this variance may be decomposed as follows:

$$\begin{aligned} \mathfrak{R}_3(\hat{\mathbf{w}}, \mathbf{R}_m) &= \text{Var}[\hat{\mathbf{w}}'\mathbf{R}_m] = E_{\hat{\mathbf{w}}}[\text{Var}[\hat{\mathbf{w}}'\mathbf{R}_m | \hat{\mathbf{w}}]] + \text{Var}_{\hat{\mathbf{w}}}[E[\hat{\mathbf{w}}'\mathbf{R}_m | \hat{\mathbf{w}}]] \\ &= E_{\hat{\mathbf{w}}}[\hat{\mathbf{w}}'\mathbf{\Sigma}_R\hat{\mathbf{w}}] + \text{Var}_{\hat{\mathbf{w}}}[\hat{\mathbf{w}}'\boldsymbol{\mu}_R] = E_{\hat{\mathbf{w}}}[\hat{\mathbf{w}}'\mathbf{\Sigma}_R\hat{\mathbf{w}}] + \boldsymbol{\mu}'_R \text{Cov}[\hat{\mathbf{w}}]\boldsymbol{\mu}_R \\ &= \frac{1}{\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}} + E[(\hat{\mathbf{w}} - \mathbf{w})'\mathbf{\Sigma}_R(\hat{\mathbf{w}} - \mathbf{w})] + \boldsymbol{\mu}'_R \text{Cov}[\hat{\mathbf{w}}]\boldsymbol{\mu}_R. \end{aligned}$$

The out-of-sample variance may accordingly be decomposed into three terms. The first one,  $\frac{1}{\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}}$ , is the risk due to randomness of assets and hence not subject to estimation issues. Frahm and Memmel (2010) argue that the third term,  $\boldsymbol{\mu}'_R \text{Cov}[\hat{\mathbf{w}}]\boldsymbol{\mu}_R$ , is negligible in most practical situations, and hence that the term  $E_{\hat{\mathbf{w}}}[\hat{\mathbf{w}}'\mathbf{\Sigma}_R\hat{\mathbf{w}}]$  is the one of main interest to us. The decomposition of  $\mathfrak{R}_3$  specified above is, however, not necessarily the most versatile one. An important concern with  $\mathfrak{R}_3$  is that it is a measure of the variance of  $\hat{\mathbf{w}}'\mathbf{R}_m$  but does not depend on the actual value that  $\mathbf{R}_m$  assigns.

A different expression of the out-of-sample variance may be obtained by conditioning on  $\mathbf{R}_m$ . We define the conditional out-of-sample variance as follows:

$$\begin{aligned} \mathfrak{R}_4(\hat{\mathbf{w}}, \mathbf{r}_m) &:= E\left[\left(\hat{\mathbf{w}}'\mathbf{R}_m - \mathbf{w}'\mathbf{R}_m\right)^2 \mid \mathbf{R}_m = \mathbf{r}_m\right] \\ &= \text{Var}[\hat{\mathbf{w}}'\mathbf{R}_m \mid \mathbf{R}_m = \mathbf{r}_m] + \text{Bias}^2[\hat{\mathbf{w}}'\mathbf{R}_m \mid \mathbf{R}_m = \mathbf{r}_m]. \end{aligned}$$

Although  $\mathfrak{R}_3$  and  $\mathfrak{R}_4$  both describe the out-of-sample variance of  $\hat{\mathbf{w}}'\mathbf{R}_m$ , the difference between them is that the latter conditions on  $\mathbf{R}_m = \mathbf{r}_m$  and also emphasizes the bias. It therefore allows for investigation of the portfolio risk when  $\mathbf{R}_m$  assigns some specific value. The risk of a portfolio as a function of some estimator  $\hat{\mathbf{w}}$  will be different when  $\mathbf{R}_m$  assigns a value close to  $E[\mathbf{R}_m]$  compared to a scenario when  $\mathbf{R}_m$  assigns a value far from  $E[\mathbf{R}_m]$ . Hence,  $\mathfrak{R}_4$  allows us to investigate the risk when, for example, the market reacts dramatically to a specific event. This matter becomes particularly important when using an estimator which is biased because  $E[\hat{\mathbf{w}}'\mathbf{R}_m] = E_{\mathbf{r}}[E[\hat{\mathbf{w}}'\mathbf{R}_m \mid \mathbf{R}_m = \mathbf{r}_m]] = E[\hat{\mathbf{w}}'\boldsymbol{\mu}_R]$  regardless of a possible bias in  $\hat{\mathbf{w}}$ . In contrast, the forecast bias conditioned on  $E[\hat{\mathbf{w}}' - \mathbf{w}']\mathbf{r}_m$  may be considerable since it increases linearly with  $\mathbf{r}_m$ . The concept of explicitly conditioning on a specific return value  $\mathbf{r}_m$  provides us with the possibility of assessing the behavior of a weight estimator under certain scenarios. For example, investors may be particularly interested in cases when the market is turbulent ( $\|\mathbf{r}_m\|$  is large relative to  $\sigma_R^2$ ), the market reacts additively to an event (say  $\mathbf{r}_m \mapsto \mathbf{a} + \mathbf{r}_m, \mathbf{a} \in \mathbb{R}_+^p$ ) or when the market is stable ( $\mathbf{r}_m \approx E[\mathbf{R}_m]$ ).

Remark (i) (*Directional risks*) Any estimator of the GMVP may be decomposed into components orthogonal and parallel to  $\mathbf{w}$ : Let  $\mathbf{A}^+$  denote the Moore–Penrose pseudoinverse of some matrix  $\mathbf{A}$ . Then, the component of  $\hat{\mathbf{w}}$  parallel to  $\mathbf{w}$  is given by  $\mathbf{v} = (\mathbf{w}^+ \hat{\mathbf{w}}) \mathbf{w} = \left( \frac{\mathbf{w}' \hat{\mathbf{w}}}{\mathbf{w}' \mathbf{w}} \right) \mathbf{w}$  and the component orthogonal to  $\mathbf{w}$  is determined by  $\mathbf{u} = \mathbf{\Omega}_\perp \hat{\mathbf{w}}$  where  $\mathbf{\Omega}_\perp = \left( \mathbf{I} - \left( \frac{\mathbf{w} \mathbf{w}'}{\mathbf{w}' \mathbf{w}} \right) \right)$  is a projection matrix (Rao 2008, pp. 46–47). We can thus decompose  $\hat{\mathbf{w}}$  according to  $\hat{\mathbf{w}} = \mathbf{v} + \mathbf{u}$ , where  $\mathbf{v}$  is parallel to  $\mathbf{w}$  and  $\mathbf{u}$  is orthogonal to  $\mathbf{w}$ . Some special cases of the above-defined risk functions in the direction orthogonal to the GMVP, which is our direction of main interest, are then given by  $\mathfrak{R}_1(\hat{\mathbf{\Sigma}}^{-1}, \mathbf{\Omega}_\perp)$ ,  $\mathfrak{R}_2(\hat{\mathbf{\Sigma}}^{-1}, \mathbf{\Omega}_\perp)$ ,  $\mathfrak{R}_3(\mathbf{u})$  and  $\mathfrak{R}_4(\mathbf{u}, \mathbf{r}_m)$ . Although the risk in a certain direction to  $\mathbf{w}$  alone is of limited interest, it does provide some insight into the relative performance of one estimator to another. For example, it may be shown that  $E \left[ (\hat{\mathbf{w}}_I - \mathbf{w})' \mathbf{\Omega}_\perp (\hat{\mathbf{w}}_I - \mathbf{w}) \right] = \frac{1}{(n-p+1)} \frac{1}{\mathbf{1}' \mathbf{\Sigma}^{-1} \mathbf{1}} \left( \text{tr} \{ \mathbf{\Sigma}^{-1} \} - \left( \frac{\mathbf{1}' \mathbf{\Sigma}^{-3} \mathbf{1}}{\mathbf{1}' \mathbf{\Sigma}^{-2} \mathbf{1}} \right) \right)$ , whereas  $(\mathbf{w}_0 - \mathbf{w})' \mathbf{\Omega}_\perp (\mathbf{w}_0 - \mathbf{w}) = p^{-1} \left( 1 - \frac{p^{-1} (\mathbf{1}' \mathbf{\Sigma}^{-1} \mathbf{1})^2}{\mathbf{1}' \mathbf{\Sigma}^{-2} \mathbf{1}} \right)$  (see Appendix A).

Remark (ii) (*Implicit Covariance Matrix*) There always exists a p.d. diagonal matrix  $\mathbf{\Lambda}$  such that  $\mathbf{P} \mathbf{1} = \mathbf{\Lambda} \mathbf{P} \mathbf{w}_0$  or, equivalently,  $\mathbf{w}_0 = \mathbf{P} \mathbf{\Lambda}^{-1} \mathbf{P}' \mathbf{1}$ , where  $\mathbf{w}_0$  is any reference portfolio (Frahm and Memmel 2010, Theorem 8). The Stein-type estimator defined by  $\hat{\mathbf{w}}_S = (1 - \alpha) \hat{\mathbf{w}}_I + \alpha \mathbf{w}_0$  is therefore associated with an “implicit” covariance matrix estimator in the sense that there exists a matrix  $\hat{\mathbf{\Sigma}}_S^{-1}$  such that  $\hat{\mathbf{w}}_S = \frac{\hat{\mathbf{\Sigma}}_S^{-1} \mathbf{1}}{\mathbf{1}' \hat{\mathbf{\Sigma}}_S^{-1} \mathbf{1}}$ , where  $\hat{\mathbf{\Sigma}}_S^{-1} = (1 - \alpha) \mathbf{S}^{-1} + \alpha \mathbf{\Sigma}_0^{-1}$ ,  $0 \leq \alpha \leq 1$ , and  $\hat{\mathbf{\Sigma}}_S^{-1}$  may, or may not, be positive definite. If we define our implicit covariance matrix by  $\mathbf{\Sigma}_S^{-1} = (1 - \alpha) \mathbf{S}^{-1} + \alpha \mathbf{\Sigma}_0^{-1} = (1 - \alpha) \mathbf{S}^{-1} + \alpha p^{-1} \mathbf{I} (\mathbf{1}' \mathbf{S}^{-1} \mathbf{1})$ , we obtain the identity  $\hat{\mathbf{w}}_S = \frac{\hat{\mathbf{\Sigma}}_S^{-1} \mathbf{1}}{\mathbf{1}' \hat{\mathbf{\Sigma}}_S^{-1} \mathbf{1}} = \frac{\{(1-\alpha)\mathbf{S}^{-1} + \alpha p^{-1}\mathbf{I}(\mathbf{1}'\mathbf{S}^{-1}\mathbf{1})\}\mathbf{1}}{\mathbf{1}'\{(1-\alpha)\mathbf{S}^{-1} + \alpha p^{-1}\mathbf{I}(\mathbf{1}'\mathbf{S}^{-1}\mathbf{1})\}\mathbf{1}} = (1 - \alpha) \frac{\mathbf{S}^{-1} \mathbf{1}}{\mathbf{1}' \mathbf{S}^{-1} \mathbf{1}} + \alpha p^{-1} \mathbf{1} = (1 - \alpha) \hat{\mathbf{w}}_I + \alpha \mathbf{w}_0$ . This identity allows us to investigate the risk of  $\hat{\mathbf{w}}_S$  with respect to any risk function designed for estimators of the (inverse) covariance matrices. For example, although  $\mathfrak{R}_0(\hat{\mathbf{\Sigma}}^{-1})$  is a function of  $\hat{\mathbf{\Sigma}}^{-1}$  and does not involve an explicit weight estimator, it is nevertheless possible to evaluate  $\hat{\mathbf{w}}_S$  via  $\mathfrak{R}_0(\hat{\mathbf{\Sigma}}_S^{-1})$ .

#### 4 Families of GMVP weight estimators

We will consider a few estimators for further investigation. The first one is the “standard estimator” defined in Eq. (3). The next is suggested by Frahm and Memmel (2010) and is defined as

$$\hat{\mathbf{w}}_{II} = (1 - \kappa_m) \hat{\mathbf{w}}_I + \kappa_m \mathbf{w}_0, \tag{5}$$

where  $\mathbf{w}_0 := p^{-1} \mathbf{1}$ ,  $\kappa_m = \min[\kappa_s, 1]$ ,  $\kappa_s = \frac{p-3}{n-p+2} \left[ \frac{\mathbf{w}_0' \mathbf{S} \mathbf{w}_0 - \hat{\mathbf{w}}_I' \mathbf{S} \hat{\mathbf{w}}_I}{\hat{\mathbf{w}}_I' \mathbf{S} \hat{\mathbf{w}}_I} \right]^{-1}$ . The estimator in Eq. (5) may be thought of as a weighted average between the traditional estimator

and a reference portfolio, here represented by  $\mathbf{w}_0 := p^{-1}\mathbf{1}$ , although the reference portfolio could essentially be any non-random portfolio such that  $\mathbf{w}_0'\mathbf{1} = 1$ .

An alternative estimator within the same family was proposed by Bodnar et al. (2018) who suggested the estimator

$$\hat{\mathbf{w}}_{\text{III}} = (1 - \kappa_b) \hat{\mathbf{w}}_{\text{I}} + \kappa_b \mathbf{w}_0$$

where  $\kappa_b = 1 - \frac{(1-c)((1-c)\mathbf{b}'\mathbf{S}\mathbf{b}\mathbf{1}'\mathbf{S}^{-1}\mathbf{1}-1)}{c+(1-c)((1-c)\mathbf{b}'\mathbf{S}\mathbf{b}\mathbf{1}'\mathbf{S}^{-1}\mathbf{1}-1)}$  and  $c = p/n$ .

Although the estimators  $\hat{\mathbf{w}}_{\text{II}}$  and  $\hat{\mathbf{w}}_{\text{III}}$  have shown great potential in improving the standard estimator  $\hat{\mathbf{w}}_{\text{I}}$ , improved estimators can be developed from a variety of different points of view. In particular, resolvent-type estimators, defined by  $\hat{\Sigma}_k^{-1} = (\mathbf{S} + k\mathbf{I})^{-1}$ ,  $k \in \mathbb{R}_+$ , have shown great potential in estimating the precision matrix, particularly in high-dimensional settings (Holgersson and Karlsson 2012; Serdobolskii 1985).

These estimators add a small constant to the eigenvalues before inversion, thereby creating a more stable estimator. They play an important role in spectral analysis (Serdobolskii 1985) but have also proved to be efficient in more applied problems (Holgersson and Karlsson 2012). The “regularizing” coefficient  $k$  imposes a (small) bias on estimators of the precision matrix and hence offers a form of variance-bias trade-off rather different from the Stein-type estimators. Since the poor performance of the standard plug-in estimator is largely due to high sample variance in the precision matrix, the resolvent estimators are interesting candidates for improved estimation of the GMVP.

While the Stein-type estimators depend on the regularizing coefficient  $\kappa$ , the estimator  $\hat{\Sigma}_k^{-1}$  depends on the coefficient  $k$ , which usually has to be determined from data. We will use the  $\mathfrak{R}_0$  risk to derive the optimal value of  $k$ , i.e., we search for the value of  $k$  which minimize  $\mathfrak{R}_0(\hat{\Sigma}_k^{-1}) = E \left\{ p^{-1} \text{tr} \left( \hat{\Sigma}_k^{-1} - \Sigma^{-1} \right)^2 \right\}$ . A consistent estimate of  $\mathfrak{R}_0(\hat{\Sigma}_k^{-1})$  has been derived by Serdobolskii (2000), defined as follows:

$$\begin{aligned} \hat{\mathfrak{R}}_0(\hat{\Sigma}_k^{-1}) &= \hat{\Lambda}_{-2} - 2k^{-1} \left( p^{-1} - n^{-1} \right) \text{tr} \left( \mathbf{S}^{-1} \tilde{\Sigma}_k^{-1} \right) \\ &\quad + 2k^{-1} p^{-1} n^{-1} \left( \text{tr} \tilde{\Sigma}_k^{-1} \right)^2 + k^{-2} p^{-1} \text{tr} \tilde{\Sigma}_k^{-2}, \end{aligned}$$

where  $\tilde{\Sigma}_k^{-1} = (\mathbf{I} + k^{-1}\mathbf{S})^{-1}$ , and  $\hat{\Lambda}_{-2} = (1 - pn^{-1})^2 p^{-1} \text{tr}(\mathbf{S}^{-2}) + (1 - pn^{-1}) p^{-1} n^{-1} (\text{tr}(\mathbf{S}^{-1}))^2$ . The optimal value of  $k$  is defined by  $\hat{k} = \min_k \hat{\mathfrak{R}}_0(\hat{\Sigma}_k^{-1})$  which is obtained numerically. A feasible resolvent-type portfolio estimator is then defined by

$$\hat{\mathbf{w}}_{\text{IV}} = \frac{\hat{\Sigma}_{\hat{k}}^{-1} \mathbf{1}}{\mathbf{1}' \hat{\Sigma}_{\hat{k}}^{-1} \mathbf{1}}. \quad (6)$$



Defined by Eqs. (3)–(6), we have a set of different estimators of the GMVP weights. These will be used in a Monte Carlo simulation of the next section in order to compare the risk functions and the performances of estimators.

### 5 Monte Carlo study

To investigate the efficiency of GMVP estimators from a risk perspective, we conduct Monte Carlo experiments using three different data generating processes (DGP I, DGP II and DGP III). DGP I is based on a multivariate normal distribution with different covariance structures and zero mean vector, and DGP II is based on a multivariate skewed  $t$  distribution with mean vector equal to a zero vector. DGP III is based on a skewed distribution with nonzero mean vector. DGP I is specified as follows:

$$\mathbf{R}_t \sim N_p(\mathbf{0}, \Sigma),$$

where  $\mathbf{R}_t$  is a vector of  $p$  different assets returns in time period  $t$ ,  $\{\mathbf{R}_t\}_{t=1}^n$  is independent and identically distributed (IID), where  $n$  corresponds to the number of observations. The specification of  $\Sigma$  will assign different values in the simulations. We will use a Toeplitz covariance structure given by

$$\Sigma = \begin{bmatrix} 1 & \phi^1 & \phi^2 & \dots & \phi^{p-1} \\ \phi^1 & \ddots & \ddots & \ddots & \vdots \\ \phi^2 & \ddots & \ddots & \ddots & \phi^2 \\ \vdots & \ddots & \ddots & \ddots & \phi^1 \\ \phi^{p-1} & \dots & \phi^2 & \phi^1 & 1 \end{bmatrix},$$

and also a covariance matrix estimated from stock return data.

For DGP II, we first define the distribution of the vector of returns as

$$\mathbf{R}_t \sim \text{Skewed } t_p(v, \boldsymbol{\lambda}, \boldsymbol{\Omega}, \boldsymbol{\gamma}),$$

where  $\boldsymbol{\lambda} \in \mathbb{R}^p$  and  $\boldsymbol{\gamma} \in \mathbb{R}^p$  are parameter vectors,  $\boldsymbol{\Omega} \in \mathbb{R}^{p \times p}$  is positive definite and  $v > 4$ . When  $\boldsymbol{\gamma} \neq \mathbf{0}$ , this yields a  $p$  dimensional skewed  $t$  distribution. There exists a number of different multivariate distributions in the literature that all share the name multivariate skewed  $t$  distribution (Kotz and Nadarajah 2004). In this paper, we use the larger class of multivariate normal mixture distributions to get a skewed multivariate distribution which is referred to as a skewed multivariate  $t$  distribution (Demarta and McNeil 2005). This is achieved by setting

$$\mathbf{R}_t = \boldsymbol{\lambda} + W_t^{-1} \boldsymbol{\gamma} + W_t^{-0.5} \mathbf{X}_t$$

where  $\mathbf{X}_t \sim N_p(\mathbf{0}, \mathbf{\Omega})$  and  $W_t^{-1} \sim \text{InverseGamma}(v/2, v/2)$  which is independent of  $\mathbf{X}_t$ . The first moments of  $\mathbf{R}_t$  are given by

$$E[\mathbf{R}_t] = E_{W_t^{-1}} \left[ E[\mathbf{R}_t | W_t^{-1}] \right] = \boldsymbol{\lambda} + \frac{v}{v-2} \boldsymbol{\gamma},$$

$$\begin{aligned} \text{Cov}[\mathbf{R}_t] &= E_{W_t^{-1}} \left[ \text{Cov}(\mathbf{R}_t | W_t^{-1}) \right] + \text{Cov}_{W_t^{-1}} \left( E[\mathbf{R}_t | W_t^{-1}] \right) \\ &= \frac{v}{v-2} \mathbf{\Omega} + \frac{2v^2\beta^2}{(v-2)^2(v-4)} \boldsymbol{\gamma}\boldsymbol{\gamma}'. \end{aligned}$$

(For details refer to Appendix B).

We assume that all stock returns are skewed to the same extent which in turn is achieved by setting  $\boldsymbol{\gamma} = \beta \mathbf{1}_p$ . Further, it is also of interest to have the DGP centered at the zero vector and this is achieved by  $\boldsymbol{\lambda} = \beta v / (v - 2) \mathbf{1}_p$ , hence the DGP II is specified as  $\mathbf{R}_t \sim \text{Skewed } t_p(v, \boldsymbol{\lambda} = \beta v / (v - 2) \mathbf{1}_p, \mathbf{\Omega}, \boldsymbol{\gamma} = \beta \mathbf{1}_p)$ , with moment

$$E[\mathbf{R}_t] = \mathbf{0},$$

and

$$\text{Cov}(\mathbf{R}_t) = \frac{v}{v-2} \mathbf{\Omega} + \frac{2v^2\beta^2}{(v-2)^2(v-4)} \mathbf{1}\mathbf{1}'.$$

Following the procedure of Holgersson and Mansoor (2013), DGP III is specified as follows: Let  $Z_0 \sim \chi^2_{(1)}$ ,  $Q_j \sim \chi^2_{(1)}$ ,  $U_j \sim \chi^2_{(1)}$ ,  $j = 1, \dots, p$ , where all variables are mutually and individually independent. Then, each variable  $R_{jt}$  in  $\mathbf{R}_t$  is equal to  $R_{jt} = Z_{0t} + Q_{jt} + U_{jt}$ , and hence, each variable in  $\mathbf{R}_t$  has a  $\chi^2_{(3)}$  distribution and

covariance structure:  $\text{Cov}(R_t) = \begin{bmatrix} 6 & 2 & \dots & 2 \\ 2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 2 \\ 2 & \dots & 2 & 6 \end{bmatrix}$ . The specifications of DGP I, II and

III are summarized in Tables 1 and 2.

Finally, as performance measures we take the five risk functions ( $\mathfrak{R}_0 - \mathfrak{R}_4$ ), and for each estimator ( $\hat{\mathbf{w}}_{II}, \hat{\mathbf{w}}_{III}, \hat{\mathbf{w}}_{IV}$ ) we divide its risk by the corresponding risk for  $\hat{\mathbf{w}}_I$  to get their relative performance. Furthermore, for  $\mathfrak{R}_4$  we choose three different conditioned returns, specified in Table 3.

### 5.1 Results from Monte Carlo simulations

Based on the results from the Monte Carlo experiments displayed in Tables 4, 5, 6 and 7, the estimator  $\hat{\mathbf{w}}_{IV}$  performs well if  $c$  is larger than 0.1 for both DGP I with covariance structure from real data and for DGP II. But for DGP I with a Toeplitz covariance structure,  $\hat{\mathbf{w}}_{IV}$  performs well only for  $c$  close to one. On the other hand,

**Table 1** Specification of the distribution of the stochastic terms in DGP I, II III

Parameter	Specification of DGP I
$\Sigma$	i. $\Sigma_{p \times p} = n^{-1} \sum_{t=1}^n (\mathbf{R}_t - \bar{\mathbf{R}})(\mathbf{R}_t - \bar{\mathbf{R}})'$ , $\bar{\mathbf{R}} = n^{-1} \sum_{t=1}^n \mathbf{R}_t$ , where $\mathbf{R}_t$ is a vector of observed returns on $p$ stocks ii. $\Sigma_{p \times p} = \text{Toeplitz}(\phi^0, \phi^1, \dots, \phi^{p-1})$ , $\phi = 0.5$
$\mu$	$\mathbf{0}_{p \times 1}$
Parameter	Specification of DGP II
$\Omega$	$\Omega_{p \times p} = \text{Toeplitz}(\phi^0, \phi^1, \dots, \phi^{p-1})$ , $\phi = 0.5$
$\nu$	5
$\gamma$	$\gamma = \beta \mathbf{1}_p$ , $\beta = 1$
$\lambda$	$\lambda = \beta \nu / (\nu - 2) \mathbf{1}_p$
$\mu$	$\mathbf{0}_{p \times 1}$
$\Sigma$	$\Sigma = \frac{\nu}{\nu-2} \Omega + \frac{2\nu^2\beta^2}{(\nu-2)^2(\nu-4)} \mathbf{1}_p \mathbf{1}_p'$
Parameter	Specification of DGP III
$df$	1
$\Sigma$	$\Sigma_{p \times p} = \text{Toeplitz}(6, 2, \dots, 2)$ ,
$\mu$	$3 \cdot \mathbf{1}_{p \times 1}$

**Table 2** Design of the Monte Carlo experiments used for DGP I, II III

Factor	Symbol	Design
$p/n$ ratio (dimension in relation to sample size)	$c$	0.1, 0.5, 0.98
Dimension of portfolio (number of assets)	$p$	100
Number of time observations	$n$	Follows from $p$ and $c$ above

**Table 3** Specifications of the conditioned observation  $\mathbf{R}_t$  used in  $\mathfrak{N}_4$

Risk	Conditioned on observation $\mathbf{R}_t$
$\mathfrak{N}_4(\hat{w}_j   \mathbf{R}_t)$	single random draw from specified DGP and $\mathbf{R}_t$ is kept fixed over all replicates
$\mathfrak{N}_4(\hat{w}_j   \mathbf{R}_t + l)$	vector $l$ has ones in its $p/2$ upper rows and zeros elsewhere
$\mathfrak{N}_4(\hat{w}_j   2\mathbf{R}_t)$	multiplying $\mathbf{R}_t$ with 2 when $\mathbf{R}_t$ is generated as specified above

**Table 4** MC simulation results for DGP I with  $\Sigma$  according to Table 1 (i), ( $p = 100$ )

Estimator	(1) $\mathfrak{R}_0$	(2) $\mathfrak{R}_1$	(3) $\mathfrak{R}_2$	(4) $\mathfrak{R}_3$	(5) $\mathfrak{R}_4$	(6) $\mathfrak{R}_4(\mathbf{R}_t + 1)$	(7) $\mathfrak{R}_4(2\mathbf{R}_t)$
$c = 0.98$							
$\hat{\mathbf{w}}_{II}$	0.9975	0.9996	0.9007	0.9022	0.9020	0.9049	0.9043
$\hat{\mathbf{w}}_{III}$	0.8660	0.9714	0.1441	0.1971	0.1514	0.1411	0.1459
$\hat{\mathbf{w}}_{IV}$	7.43E-09	2.38E-09	0.0132	0.0270	0.0291	0.0112	0.0115
$c = 0.5$							
$\hat{\mathbf{w}}_{II}$	0.8245	0.8744	0.8975	0.9599	0.8980	0.8934	0.8953
$\hat{\mathbf{w}}_{III}$	0.6699	0.7599	0.8108	0.9484	0.8096	0.7949	0.8023
$\hat{\mathbf{w}}_{IV}$	0.6208	0.2027	1.1517	0.9054	1.5729	0.6837	0.4599
$c = 0.1$							
$\hat{\mathbf{w}}_{II}$	0.8866	0.9563	0.9786	0.9988	0.9788	0.9769	0.9774
$\hat{\mathbf{w}}_{III}$	0.8710	0.9501	0.9757	0.9988	0.9759	0.9735	0.9742
$\hat{\mathbf{w}}_{IV}$	1	1	1	1	1	1	1

Columns (1)–(5) show risk  $\mathfrak{R}_i(\hat{\mathbf{w}}_j)$  of estimator  $\hat{\mathbf{w}}_j$  relative to  $\mathfrak{R}_i(\hat{\mathbf{w}}_I)$ , where  $i \in \{0, 1, 2, 3, 4\}$ , columns (6) and (7) show  $\mathfrak{R}_4(\hat{\mathbf{w}}_j | \mathbf{R}_t + 1)$  relative to  $\mathfrak{R}_4(\hat{\mathbf{w}}_I | \mathbf{R}_t + 1)$  and  $\mathfrak{R}_4(\hat{\mathbf{w}}_j | 2\mathbf{R}_t)$  relative to  $\mathfrak{R}_4(\hat{\mathbf{w}}_I | 2\mathbf{R}_t)$ , respectively  $\hat{\mathbf{w}}_{II}$ : Frahm and Memmel (2010),  $\hat{\mathbf{w}}_{III}$ : Bodnar et al. (2018),  $\hat{\mathbf{w}}_{IV}$ : resolvent-type estimator

estimator  $\hat{\mathbf{w}}_{III}$  performs best among the four estimators. This holds for all investigated values of  $c$ . Furthermore,  $\hat{\mathbf{w}}_{II}$  is a good estimator if  $c$  is not close to one, and its performance is close to  $\hat{\mathbf{w}}_{III}$ . However, as  $c$  gets close to one,  $\hat{\mathbf{w}}_{II}$  is outperformed both by  $\hat{\mathbf{w}}_{III}$  and  $\hat{\mathbf{w}}_{IV}$ .

If we examine the estimators performance under  $\mathfrak{R}_4(R_t + 1)$  and  $\mathfrak{R}_4(2R_t)$ , both of which could reflect a shock in the financial market, the result for the estimators remains in the same internal ordering as indicated by  $\mathfrak{R}_3$ . Thus, with regard to the discussed results, the recommendation is that one should primarily consider  $\hat{\mathbf{w}}_{III}$  because it performs well regardless of the value on  $c$ , unless  $c$  is very close to one, in which case  $\hat{\mathbf{w}}_{IV}$  is dominating. It should, however, be stressed that the above performance rankings are made only on basis of point estimations. While more general inferences such as interval estimation lies outside the scope of this paper, it should be mentioned that the estimators’ performances in terms of, for example, coverage rates need not correspond to their point estimation efficiency.

## 6 Empirical study

The empirical evaluation of the investigated estimators of the weights in the GMVP is achieved through a moving window approach on two different data sets for which we use different sampling methods. In the first method (fixed sampling method), we simply apply the estimators on all available assets, and in the second approach (random sampling method), we repeatedly randomly pick a given number of assets and then evaluate the estimators performance by the one-period out-of-sample returns. The reason for applying a moving window is that the mean-variance portfolio theory was developed as a one-period model.

**Table 5** MC simulation results for DGP I with  $\Sigma$  according to Table 1 (ii), ( $p = 100$ )

Estimator	(1) $\mathfrak{R}_0$	(2) $\mathfrak{R}_1$	(3) $\mathfrak{R}_2$	(4) $\mathfrak{R}_3$	(5) $\mathfrak{R}_4$	(6) $\mathfrak{R}_4 (\mathbf{R}_t + 1)$	(7) $\mathfrak{R}_4 (2\mathbf{R}_t)$
<i>c</i> = 0.98							
$\hat{\mathbf{w}}_{II}$	0.9735	0.9987	0.4616	0.4666	0.4771	0.4354	0.4630
$\hat{\mathbf{w}}_{III}$	0.7139	0.9338	0.0147	0.0245	0.0156	0.0140	0.0131
$\hat{\mathbf{w}}_{IV}$	9.05E-07	9.57E-07	0.0365	0.0485	0.0427	0.0401	0.0391
<i>c</i> = 0.5							
$\hat{\mathbf{w}}_{II}$	0.1745	0.0599	0.0276	0.5150	0.0243	0.0247	0.0316
$\hat{\mathbf{w}}_{III}$	0.2419	0.0538	0.0137	0.5083	0.0100	0.0101	0.0176
$\hat{\mathbf{w}}_{IV}$	0.1520	0.2821	0.4407	0.7514	0.5315	0.5182	0.4667
<i>c</i> = 0.1							
$\hat{\mathbf{w}}_{II}$	6.7586	1.1462	0.0490	0.9075	0.0213	0.0230	0.0756
$\hat{\mathbf{w}}_{III}$	7.9439	1.3396	0.0508	0.9079	0.0198	0.0209	0.0820
$\hat{\mathbf{w}}_{IV}$	0.7016	0.8513	0.8792	0.9905	0.9160	0.9100	0.8903

See Table 4

**Table 6** MC simulation results for DGP II with  $\Sigma$  according to Table 1, ( $p = 100$ )

Estimator	(1) $\mathfrak{R}_0$	(2) $\mathfrak{R}_1$	(3) $\mathfrak{R}_2$	(4) $\mathfrak{R}_3$	(5) $\mathfrak{R}_4$	(6) $\mathfrak{R}_4 (\mathbf{R}_t + 1)$	(7) $\mathfrak{R}_4 (2\mathbf{R}_t)$
<i>c</i> = 0.98							
$\hat{\mathbf{w}}_{II}$	0.9987	0.9996	0.7997	0.7998	0.7962	0.7964	0.7963
$\hat{\mathbf{w}}_{III}$	0.9358	0.9658	0.0782	0.0790	0.0801	0.0802	0.0801
$\hat{\mathbf{w}}_{IV}$	1.21E-07	5.90E-07	0.0425	0.0459	0.0488	0.0501	0.0488
<i>c</i> = 0.5							
$\hat{\mathbf{w}}_{II}$	0.5633	0.8330	0.6544	0.6612	0.6572	0.6575	0.6572
$\hat{\mathbf{w}}_{III}$	0.3755	0.7484	0.4679	0.4784	0.4726	0.4730	0.4726
$\hat{\mathbf{w}}_{IV}$	0.1681	0.7059	0.5581	0.6349	0.6559	0.6685	0.6559
<i>c</i> = 0.1							
$\hat{\mathbf{w}}_{II}$	1.3380	0.9903	0.7809	0.7893	0.7800	0.7819	0.7801
$\hat{\mathbf{w}}_{III}$	1.3953	0.9891	0.7554	0.7648	0.7544	0.7565	0.7544
$\hat{\mathbf{w}}_{IV}$	1.0721	0.9980	0.9255	0.9498	0.9566	0.9606	0.9566

See Table 4

### 6.1 Fixed sample

The evaluation procedure in the fixed sampling method is as follows: For each stock listed on the stock exchange, we take  $n$  observations starting at time point  $t - n$  and ending at time point  $t$ . We then calculate monthly returns, and based on these observations, each estimator presented in this paper is used to estimate the weights of the global minimum-variance portfolio. For each estimator, the return on the GMVP is calculated for the first out-of-sample observation, i.e., the observation at time period

**Table 7** MC simulation results for DGP III with  $\Sigma$  according to Table 1, ( $p = 100$ )

Estimator	(1) $\mathfrak{R}_0$	(2) $\mathfrak{R}_1$	(3) $\mathfrak{R}_2$	(4) $\mathfrak{R}_3$	(5) $\mathfrak{R}_4$	(6) $\mathfrak{R}_4(\mathbf{R}_t + 1)$	(7) $\mathfrak{R}_4(2\mathbf{R}_t)$
$c = 0.98$							
$\hat{\mathbf{w}}_{II}$	0.9992	0.9999	0.4334	0.4390	0.4179	0.4168	0.4179
$\hat{\mathbf{w}}_{III}$	0.9672	0.9946	0.019	0.0288	0.0168	0.0121	0.0168
$\hat{\mathbf{w}}_{IV}$	2.66E-08	1.97E-08	0.00358	0.0454	0.036	0.0358	0.0360
$c = 0.5$							
$\hat{\mathbf{w}}_{II}$	0.1263	0.0371	0.0242	0.5042	0.0242	0.024	0.0242
$\hat{\mathbf{w}}_{III}$	0.1745	0.0163	0.0106	0.4972	0.0105	0.0104	0.0105
$\hat{\mathbf{w}}_{IV}$	0.3083	0.4337	0.5783	0.7857	0.5763	0.5766	0.5763
$c = 0.1$							
$\hat{\mathbf{w}}_{II}$	4.9783	0.1493	0.0124	0.9014	0.0122	0.012	0.0122
$\hat{\mathbf{w}}_{III}$	6.1863	0.1803	0.0191	0.9902	0.0192	0.019	0.0192
$\hat{\mathbf{w}}_{IV}$	1	1	1	1	1	1	1

See Table 4

$t + 1$  for each stock. We repeat this procedure, but the starting point is moved one step forward in time (starting at time point  $t - n + 1$  and ending at time point  $t + 1$ ). The procedure is repeated until 10 sample returns are generated. Thus, for each estimated portfolio weight vector, the one-period out-of-sample portfolio return is calculated as:

$$R_{t+1,j} = \hat{\mathbf{w}}'_j \mathbf{R}_{t+1}, \quad j = I, \dots, V,$$

where  $\mathbf{R}_{t+1}$  is a  $p \times 1$  vector of stock excess returns observed in time period  $t + 1$ ,  $\hat{\mathbf{w}}_j$  is a  $p \times 1$  vector of estimated weights for the GMVP and  $\hat{\mathbf{w}}_V$  corresponds to an equally weighted portfolio.

The evaluation of each estimator of the GMVP weights is based on the portfolio risk measured by the standard deviations of the portfolio returns:

$$\hat{\sigma}_{\text{portfolio},j} = \sqrt{\frac{1}{10} \sum_{t=1}^{10} (R_{t,j} - \bar{R}_j)^2}, \quad \text{where } \bar{R}_j = \frac{1}{10} \sum_{t=1}^{10} R_{t,j}, \quad j = I, \dots, V. \tag{7}$$

In addition, we also calculate the out-of-sample Sharpe ratio

$$S\hat{R}_j = \frac{\bar{R}_j}{\hat{\sigma}_{\text{portfolio},j}}, \quad j = I, \dots, V. \tag{8}$$

**Example 1** Stocks listed on the Nasdaq stock exchange

For this empirical application, 89 stocks with complete past values are selected from SP100 over the period 1997-04 to 2010-07 (159 monthly returns). Note that gross

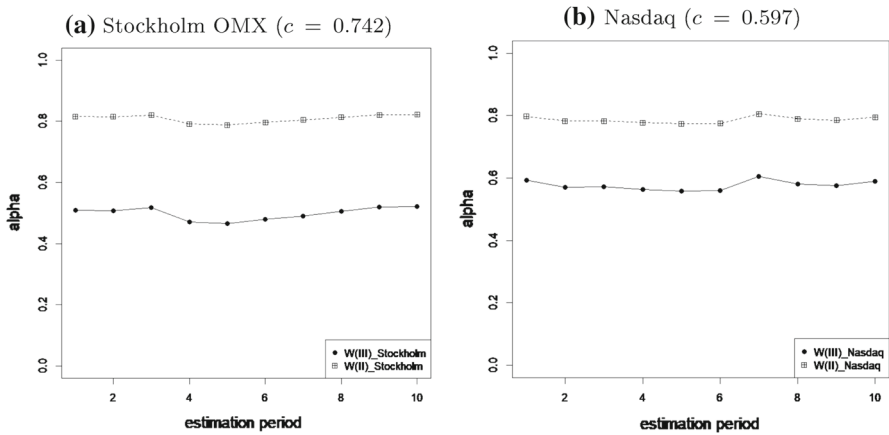


Fig. 1 Smoothing coefficients for estimators  $\hat{w}_{II}$ ,  $\hat{w}_{III}$  applied to data from Nasdaq and Stockholm OMX stock exchange

returns are used, i.e., the risk-free return is not subtracted. A moving window approach is employed with a length of 149 months, giving 10 out-of-sample returns.

**Example 2** Stocks listed on the Stockholm stock exchange

In this empirical application, 112 stocks are selected out of 283 stocks listed on the Stockholm OMX stock exchange over the period 1997-01 to 2010-06 (161 monthly returns). The same procedure as described above is used with a moving window length of 151 months.

The results of the empirical applications (Table 8) confirm what was already established in the Monte Carlo simulations of Sect. 5. That is, in the empirical application using data from the Nasdaq stock exchange in which  $c$  is around 0.6, the performances of  $\hat{\sigma}_{portfolio,j}$  for  $\hat{w}_I$ ,  $\hat{w}_{II}$ ,  $\hat{w}_{III}$  and  $\hat{w}_{IV}$  are relatively close to each other. However, for the Stockholm stock exchange in which  $c$  is around 0.74, we find that both  $\hat{w}_{III}$  and  $\hat{w}_{IV}$  yield portfolios with much lower  $\hat{\sigma}_{portfolio,j}$  compared to  $\hat{w}_I$  and  $\hat{w}_{II}$ . In both settings,  $\hat{w}_I$  is outperformed by all other estimators. On the other hand, if we shift measure from out-of-sample standard deviation to the out-of-sample Sharpe ratio, the performance is somewhat reversed in that  $\hat{w}_{II}$  performs better than  $\hat{w}_{IV}$  and the equally weighted portfolio,  $\hat{w}_V$ , is then outperforming all investigated estimators. To get a better understanding of the relative behavior of the two Stein-type estimators, i.e.,  $\hat{w}_{II}$  and  $\hat{w}_{III}$ , their smoothing coefficients are displayed in Fig. 1. The values of the smoothing coefficients are approximately the same, but the estimator  $\hat{w}_{II}$  tends to weight heavier toward the traditional estimator ( $\hat{w}_I$ ), while the estimator  $\hat{w}_{III}$  puts less weight on the traditional estimator.

**6.2 Random samples**

The evaluation procedure in the random sampling method is similar to the fixed sampling method, with the difference that we are now randomly selecting without

**Table 8** Performance of GMVP estimators applied to portfolios from Nasdaq and Stockholm OMX stock exchange, fixed sampling method

	$\hat{w}_I$	$\hat{w}_{II}$	$\hat{w}_{III}$	$\hat{w}_{IV}$	$\hat{w}_V$
Nasdaq ( $c = 0.597$ )					
$\hat{\sigma}_j$	0.0589	0.0560	0.0540	0.0443	0.0540
$\bar{R}_j$	0.0006	0.0021	0.0035	1.3E-05	0.0064
$S\hat{R}_j$	0.0098	0.0374	0.0640	0.0003	0.1189
Stockholm OMX ( $c = 0.742$ )					
$\hat{\sigma}_j$	0.0774	0.0636	0.0447	0.0479	0.0456
$\bar{R}_j$	0.0079	0.0112	0.0165	0.0085	0.0251
$S\hat{R}_j$	0.1019	0.1757	0.3695	0.1773	0.5503

$\hat{w}_I$ : standard estimator,  $\hat{w}_{II}$ : Frahm and Memmel (2010),  $\hat{w}_{III}$ : Bodnar et al. (2018),  $\hat{w}_{IV}$ : resolvent-type estimator,  $\hat{w}_V$ : equally weighted. For definitions, see Eqs. (7) and (8). Nasdaq stock exchange:  $p = 89$ ,  $n = 149$ ; Stockholm OMX stock exchange:  $p = 112$ ,  $n = 151$

replacement  $p = 20, 50, 80$  stocks and apply the moving window approach. This procedure is then repeated 100 times which results in 100 time-series of 10 one-period out-of-sample returns of the GMVP for each estimator. Note that  $c$  is kept constant when shifting from the different portfolio sizes by adjusting  $n$ . Based on Eqs. (7) and (8), we calculate the average out-of-sample variance, average out-of-sample mean return and the average out-of-sample Sharpe ratio of the 100 replications:

$$\text{Mean}(\hat{\sigma}_{\text{portfolio},j}) = \frac{1}{100} \sum_{i=1}^{100} \hat{\sigma}_{\text{portfolio},i,j}, \quad j = I, \dots, V,$$

$$\text{Mean}(\bar{R}_j) = \frac{1}{100} \sum_{i=1}^{100} \hat{\sigma}_{\text{portfolio},i,j}, \quad j = I, \dots, V.$$

In addition to the above we also calculate the mean out-of-sample Sharpe ratio as

$$\text{Mean}\left(\hat{SR}_{\text{portfolio},j}\right) = \frac{1}{100} \sum_{i=1}^{100} \frac{\bar{R}_{i,j}}{\hat{\sigma}_{\text{portfolio},i,j}}, \quad j = I, \dots, V.$$

**Example 3 Stocks listed on the Nasdaq stock exchange (random sampling method)**

For this empirical application  $p = 20, 50, 80$  stock are randomly selected out of 89 stocks with complete past values from SP100 over the period 1997-04 to 2010-07 (159 monthly returns). A moving window approach is employed with  $c = 0.537$  ( $n = 159, 103, 47$ ), giving 10 out-of-sample returns. This procedure is repeated 100 times.

**Example 4 Stocks listed on the Stockholm stock exchange (random sampling method)**

In this empirical application  $p = 20, 50, 80$  stocks are randomly selected out of 112 stocks used in Example 2. The same procedure as described before is used where



**Table 9** Performance of GMVP estimators applied to portfolios from Nasdaq stock exchange, random sampling method

	$\hat{\mathbf{w}}_I$	$\hat{\mathbf{w}}_{II}$	$\hat{\mathbf{w}}_{III}$	$\hat{\mathbf{w}}_{IV}$	$\hat{\mathbf{w}}_V$
Nasdaq ( $p = 80, c = 0.537$ )					
Mean( $\hat{\sigma}_{\text{portfolio}, j}$ )	0.0535	0.0519	0.0511	0.0432	0.0540
Mean( $\bar{R}_j$ )	0.0007	0.0021	0.0032	0.0010	0.0064
Mean( $\hat{SR}_{\text{portfolio}, j}$ )	0.0138	0.0411	0.0629	0.0229	0.1192
Nasdaq ( $p = 50, c = 0.537$ )					
Mean( $\hat{\sigma}_{\text{portfolio}, j}$ )	0.0535	0.0517	0.0508	0.0434	0.0540
Mean( $\bar{R}_j$ )	-0.0012	0.0001	0.0014	0.0023	0.0066
Mean( $\hat{SR}_{\text{portfolio}, j}$ )	-0.0219	0.0028	0.0278	0.0540	0.1213
Nasdaq ( $p = 20, c = 0.537$ )					
Mean( $\hat{\sigma}_{\text{portfolio}, j}$ )	0.0496	0.0469	0.0462	0.0409	0.0557
Mean( $\bar{R}_j$ )	0.0035	0.0039	0.0045	0.0041	0.0068
Mean( $\hat{SR}_{\text{portfolio}, j}$ )	0.0701	0.0837	0.0980	0.0996	0.1225

$\hat{\mathbf{w}}_I$ : standard estimator,  $\hat{\mathbf{w}}_{II}$ : Frahm and Memmel (2010),  $\hat{\mathbf{w}}_{III}$ : Bodnar et al. (2018),  $\hat{\mathbf{w}}_{IV}$ : resolvent-type estimator,  $\hat{\mathbf{w}}_V$ : equally weighted. For definitions, see Eqs. (7) and (8)

a moving window is employed with  $c = 0.53$  ( $n = 161, 104, 48$ ). The results of the empirical applications based on random sampling method (Tables 9 and 10) also confirm what was established in the empirical applications based on fixed sampling method. Namely, using data from the Nasdaq stock exchange in which  $c$  is around 0.5, the performances of  $\hat{\sigma}_{\text{portfolio}, j}$  for  $\hat{\mathbf{w}}_I$ ,  $\hat{\mathbf{w}}_{II}$ ,  $\hat{\mathbf{w}}_{III}$  and  $\hat{\mathbf{w}}_{IV}$  are relatively close to each other but  $\hat{\mathbf{w}}_{II}$ ,  $\hat{\mathbf{w}}_{III}$  and  $\hat{\mathbf{w}}_{IV}$  outperform  $\hat{\mathbf{w}}_I$  and  $\hat{\mathbf{w}}_V$ . Shifting the evaluation measure to the out-of-sample Sharpe ratio yields a different picture, since now  $\hat{\mathbf{w}}_V$  is outperforming the other estimators while  $\hat{\mathbf{w}}_I$  yields consistently the lowest result. But for the Stockholm stock exchange in which  $c$  is also around 0.5, we find that  $\hat{\mathbf{w}}_{III}$  yield portfolios with lower  $\hat{\sigma}_{\text{portfolio}, j}$  compared to  $\hat{\mathbf{w}}_{IV}$  and  $\hat{\mathbf{w}}_{II}$ , and all three estimators are outperforming the regular estimator  $\hat{\mathbf{w}}_I$ . Hence, in both settings,  $\hat{\mathbf{w}}_I$  is outperformed by  $\hat{\mathbf{w}}_{II}$ ,  $\hat{\mathbf{w}}_{III}$  and  $\hat{\mathbf{w}}_{IV}$ . On the other hand, if we shift measure from out-of-sample standard deviation to out-of-sample Sharpe ratio, the estimator  $\hat{\mathbf{w}}_V$  is still outperforming all the other estimators and the regular estimator  $\hat{\mathbf{w}}_I$  yields the lowest Sharpe ratio. It is also found that among  $\hat{\mathbf{w}}_{II}$ ,  $\hat{\mathbf{w}}_{III}$  and  $\hat{\mathbf{w}}_{IV}$ , the performance is similar, but  $\hat{\mathbf{w}}_{III}$  has the highest Sharpe ratio.

## 7 Summary

The global minimum-variance portfolio (GMVP) solution developed by Markowitz is considered to be a fundamental concept in portfolio theory. The early researchers investigating this matter usually applied a simple plug-in estimator for estimating the weights and paid very little attention to the distributional property of the estimator.

**Table 10** Performance of GMVP estimators applied to portfolios from Stockholm stock exchange, random sampling method

	$\hat{w}_I$	$\hat{w}_{II}$	$\hat{w}_{III}$	$\hat{w}_{IV}$	$\hat{w}_V$
Stockholm OMX ( $p = 80, c = 0.53$ )					
Mean( $\hat{\sigma}_{\text{portfolio}, j}$ )	0.0502	0.0469	0.0444	0.0445	0.0460
Mean( $\bar{R}_j$ )	0.0113	0.0132	0.0149	0.0125	0.0248
Mean( $\hat{SR}_{\text{portfolio}, j}$ )	0.2243	0.2808	0.3363	0.2798	0.5389
Stockholm OMX ( $p = 50, c = 0.53$ )					
Mean( $\hat{\sigma}_{\text{portfolio}, j}$ )	0.0757	0.0685	0.0623	0.0684	0.0472
Mean( $\bar{R}_j$ )	0.0225	0.0229	0.0232	0.0245	0.0251
Mean( $\hat{SR}_{\text{portfolio}, j}$ )	0.2977	0.3336	0.3722	0.3573	0.5325
Stockholm OMX ( $p = 20, c = 0.53$ )					
Mean( $\hat{\sigma}_{\text{portfolio}, j}$ )	0.0711	0.0614	0.0533	0.0571	0.0500
Mean( $\bar{R}_j$ )	0.0149	0.0168	0.0189	0.0172	0.0246
Mean( $\hat{SR}_{\text{portfolio}, j}$ )	0.2099	0.2743	0.3552	0.3021	0.4912

See Table 9

More recently, the full distribution of the standard estimator has been derived (Okhrin and Schmid 2006), and it is now recognized that the standard estimator offers a poor approximation of the true GMVP. Within a relatively short period of time, a variety of improvements to the standard estimator have been developed. Naturally, each of these improvements has its pros and cons, but there does not seem to be a consensus about how to evaluate the performance, or efficiency, of GMVP estimators. Perhaps this is because there are, in fact, several possible measures one can use for assessing the properties of a portfolio estimator. In this paper, we discuss a number of different risk functions for the weight estimator. These include: risk functions of covariance matrix estimators, forecast mean square errors, directional risks and conditional risks. The risk functions are labeled with an index determined by the degree to which they are specialized for portfolio estimation:  $\mathfrak{R}_2$  is generally preferred over  $\mathfrak{R}_1$  which is preferred over  $\mathfrak{R}_0$  etc. However, this ordering does not mean that  $\mathfrak{R}_2$  is uniformly better than  $\mathfrak{R}_1$  and  $\mathfrak{R}_0$ . For example,  $\mathfrak{R}_4$  does not exist in closed form for the regularized portfolio estimator used in this paper. Hence,  $\hat{w}_{IV}$  has to be optimized through  $\mathfrak{R}_0$  rather than  $\mathfrak{R}_4$ . In other words, one would typically use  $\mathfrak{R}_4$  or  $\mathfrak{R}_3$  as a tool for deriving an estimator of  $w_{GMVP}$ , but there are settings where risk functions of lower rank-order must be used because of their simpler functional form. A selection of recent GMVP estimators is used in a Monte Carlo simulation for purposes of: (i) comparing different risk measures for a given estimator and (ii) comparing different estimators for a given risk. Moreover, a new estimator, based on a resolvent estimator, is proposed. The analysis focuses on asset data where the number of observations ( $n$ ) is comparable to the number of assets ( $p$ ). This case is important because investors might be reluctant to use long data sets as the economy is not expected to be stable over long time periods, and hence, investors are encountering a high-dimensional setting. The simulations are

complemented by an analysis of two real data sets: One data set is drawn from the Nasdaq stock exchange, and the other one employs Stockholm OMX data. The general finding of the paper is that no estimator dominates uniformly over all risk functions. We can, however, establish that there are dominating tendencies, in the sense that some estimators tend to perform better with respect to most risk aspects. A Stein-type estimator developed by Frahm and Memmel (2010) is found to perform well in cases when  $n \gg p$ , whereas another Stein-type estimator proposed by Bodnar et al. (2018) dominates when  $n$  is proportional to  $p$ . A resolvent-type estimator is found to perform surprisingly well over a large number of settings. While this paper is restricted to properties of point estimators, future research could involve more general inferential aspects, such as hypotheses testing.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## A Appendix

From Okhrin and Schmid (2006), we have that

$$\hat{\mathbf{w}}_{I,p} \left( \mathbf{w}, \left( \sigma^2 \boldsymbol{\Sigma}^{-1} - \mathbf{w}\mathbf{w}' \right) / (n - p + 1), n - p + 1 \right),$$

$$\sigma^2 = \mathbf{w}' \boldsymbol{\Sigma} \mathbf{w} = \frac{1}{\mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1}}.$$

Hence,  $E \left[ (\hat{\mathbf{w}}_I - \mathbf{w}) (\hat{\mathbf{w}}_I - \mathbf{w})' \right] = \frac{1}{(n-p+1)} (\sigma^2 \boldsymbol{\Sigma}^{-1} - \mathbf{w}\mathbf{w}')$ . Therefore,

$$\begin{aligned} E \left[ (\hat{\mathbf{w}}_I - \mathbf{w})' \boldsymbol{\Omega}_\perp (\hat{\mathbf{w}}_I - \mathbf{w}) \right] &= \text{tr} \left( \left( \mathbf{I} - \mathbf{w} \left( \frac{\mathbf{w}'}{\mathbf{w}\mathbf{w}'} \right) \right) \frac{1}{(n-p+1)} (\sigma^2 \boldsymbol{\Sigma}^{-1} - \mathbf{w}\mathbf{w}') \right) = \\ &= \frac{\sigma^2}{(n-p+1)} \text{tr} \left( (\boldsymbol{\Sigma}^{-1} - \mathbf{w}\mathbf{w}') - \mathbf{w} \left( \frac{\mathbf{w}'}{\mathbf{w}\mathbf{w}'} \right) (\boldsymbol{\Sigma}^{-1} - \mathbf{w}\mathbf{w}') \right) = \\ &= \frac{\sigma^2}{(n-p+1)} \text{tr} \left( \left( \mathbf{I} - \frac{\mathbf{w}\mathbf{w}'}{\mathbf{w}\mathbf{w}'} \right) \boldsymbol{\Sigma}^{-1} \right) = \frac{\sigma^2}{(n-p+1)} \text{tr} \left( \left( \mathbf{I} - \frac{\boldsymbol{\Sigma}^{-1} \mathbf{1} \mathbf{1}' \boldsymbol{\Sigma}^{-1}}{\mathbf{1}' \boldsymbol{\Sigma}^{-2} \mathbf{1}} \right) \boldsymbol{\Sigma}^{-1} \right) = \\ &= \frac{\sigma^2}{(n-p+1)} \text{tr} \left( \text{tr} (\boldsymbol{\Sigma}^{-1}) - \frac{\mathbf{1}' \boldsymbol{\Sigma}^{-3} \mathbf{1}}{\mathbf{1}' \boldsymbol{\Sigma}^{-2} \mathbf{1}} \right). \end{aligned}$$

Also,

$$\begin{aligned} E \left[ (\mathbf{w}_0 - \mathbf{w})' \boldsymbol{\Omega}_\perp (\mathbf{w}_0 - \mathbf{w}) \right] &= \\ \text{tr} \left\{ \left( \mathbf{I} - \left( \frac{\mathbf{w}\mathbf{w}'}{\mathbf{w}\mathbf{w}'} \right) \right) (\mathbf{w}_0 \mathbf{w}_0' - 2\mathbf{w}_0 \mathbf{w}' + \mathbf{w}\mathbf{w}') \right\} &= \\ \text{tr} \left\{ \left( \mathbf{I} - \left( \frac{\mathbf{w}\mathbf{w}'}{\mathbf{w}\mathbf{w}'} \right) \right) \mathbf{w}_0 \mathbf{w}_0' \right\} - 2\text{tr} \{ (\mathbf{w}\mathbf{w}'_0 - (\mathbf{w}\mathbf{w}'_0)) \} + \text{tr} \{ (\mathbf{w}\mathbf{w}' - (\mathbf{w}\mathbf{w}')) \} &= \\ p^{-1} \left( 1 - \frac{p^{-1} (\mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1})^2}{\mathbf{1}' \boldsymbol{\Sigma}^{-2} \mathbf{1}} \right). & \end{aligned}$$

## B Appendix

Let the random vector  $\mathbf{R}_t$  be equal to  $\mathbf{R}_t = \boldsymbol{\lambda} + W_t^{-1}\boldsymbol{\gamma} + W_t^{-0.5}\mathbf{X}_t$ , where  $\mathbf{X}_t \sim N_p(\mathbf{0}, \boldsymbol{\Omega})$  and  $W_t^{-1} \sim \text{InverseGamma}(\nu/2, \nu/2)$  which is independent of  $\mathbf{X}_t$ , and let  $\boldsymbol{\gamma} = \beta\mathbf{1}_p$ . Then, the moment of this distribution is given by

$$\begin{aligned} E[\mathbf{R}_t] &= E_{W_t^{-1}} \left[ E[\mathbf{R}_t | W_t^{-1}] \right] = \boldsymbol{\lambda} + E_{W_t^{-1}} \left[ W_t^{-1} \right] \\ &\boldsymbol{\gamma} = \boldsymbol{\lambda} + \frac{\nu}{\nu-2}\boldsymbol{\gamma} = \boldsymbol{\lambda} + \frac{\nu\beta}{\nu-2}\mathbf{1}_p, \end{aligned} \quad (\text{B.1})$$

and

$$\text{Cov}[\mathbf{R}_t] = E_{W_t^{-1}} \left[ \text{Cov}(\mathbf{R}_t | W_t^{-1}) \right] + \text{Cov}_{W_t^{-1}} \left( E[\mathbf{R}_t | W_t^{-1}] \right). \quad (\text{B.2})$$

The first part of the r.h.s in (B.1) is equal to

$$\begin{aligned} E_{W_t^{-1}} \left[ \text{Cov}(\mathbf{R}_t | W_t^{-1}) \right] &= E_{W_t^{-1}} \left[ \text{Cov}(\boldsymbol{\lambda} + W_t^{-1}\boldsymbol{\gamma} + W_t^{-0.5}\mathbf{X}_t | W_t^{-1}) \right] = \\ &E_{W_t^{-1}} \left[ \left( W_t^{-0.5} \right)^2 \text{Cov}(\mathbf{X}_t | W_t^{-1}) \right] = \\ &E_{W_t^{-1}} \left[ W_t^{-1} \right] \boldsymbol{\Omega} = \frac{\nu/2}{\nu/2-1} \boldsymbol{\Omega} = \frac{\nu}{\nu-2} \boldsymbol{\Omega}. \end{aligned} \quad (\text{B.3})$$

The second part of the r.h.s in (B.2) is equal to

$$\begin{aligned} \text{Cov}_{W_t^{-1}} \left( E[\mathbf{R}_t | W_t^{-1}] \right) &= \text{Cov}_{W_t^{-1}} \left( E[\boldsymbol{\lambda} + W_t^{-1}\boldsymbol{\gamma} + W_t^{-0.5}\mathbf{X}_t | W_t^{-1}] \right) \\ &= \text{Cov}_{W_t^{-1}} \left( \boldsymbol{\lambda} + W_t^{-1}\boldsymbol{\gamma} + W_t^{-0.5} E[\mathbf{X}_t | W_t^{-1}] \right) = \text{Cov}_{W_t^{-1}} \left( \boldsymbol{\lambda} + W_t^{-1}\boldsymbol{\gamma} + \mathbf{0} \right) \\ &= \text{Cov}_{W_t^{-1}} \left( W_t^{-1}\boldsymbol{\gamma} \right) = \boldsymbol{\gamma} \text{Cov}_{W_t^{-1}} \left( W_t^{-1} \right) \boldsymbol{\gamma}' \\ &= \boldsymbol{\gamma} \frac{(\nu/2)^2}{(\nu/2-1)^2(\nu/2-2)} \boldsymbol{\gamma}' = \frac{2\nu^2\beta^2}{(\nu-2)^2(\nu-4)} \mathbf{1}\mathbf{1}'. \end{aligned} \quad (\text{B.4})$$

Then, by substituting in equations (B.3) and (B.4) into (B.2) we have (B.5)

$$\begin{aligned} \text{Cov}[\mathbf{R}_t] &= E_{W_t^{-1}} \left[ \text{Cov}(\mathbf{R}_t | W_t^{-1}) \right] + \text{Cov}_{W_t^{-1}} \left( E[\mathbf{R}_t | W_t^{-1}] \right) \\ &= \frac{\nu}{\nu-2} \boldsymbol{\Omega} + \frac{2\nu^2\beta^2}{(\nu-2)^2(\nu-4)} \mathbf{1}\mathbf{1}'. \end{aligned} \quad (\text{B.5})$$

## References

Anderson, T.W.: An Introduction to Multivariate Statistical Analysis, 3rd edn. Wiley, New York (2003)

- Bai, Z., Liu, H., Wong, W.K.: Enhancement of the applicability of Markowitz's portfolio optimization by utilizing random matrix theory. *Math. Finance* **19**(4), 639–667 (2009)
- Best, M.J., Grauer, R.R.: On the sensitivity of mean–variance-efficient portfolios to changes in asset means: some analytical and computational results. *Rev. Financ. Stud.* **4**(2), 315–342 (1991)
- Bodnar, T., Zabolotsky, T.: How risky is the optimal portfolio which maximizes the sharpe ratio? *ASTA Adv. Stat. Anal.* **101**(1), 1–28 (2017)
- Bodnar, T., Parolya, N., Schmid, W.: Estimation of the global minimum variance portfolio in high dimensions. *Eur. J. Oper. Res.* **266**(1), 371–390 (2018)
- Demarta, S., McNeil, A.J.: The t copula and related copulas. *Int. Stat. Rev.* **73**(1), 111–129 (2005)
- DeMiguel, V., Garlappi, L., Uppal, R.: Optimal versus naive diversification: how inefficient is the 1/n portfolio strategy? *Rev. Financ. Stud.* **22**(5), 1915–1953 (2009)
- Efron, B., Morris, C.: Multivariate empirical Bayes and estimation of covariance matrices. *Ann. Stat.* **4**(1), 22–32 (1976)
- Frahm, G., Memmel, C.: Dominating estimators for minimum-variance portfolios. *J. Econom.* **159**(2), 289–302 (2010)
- Frankfurter, G.M., Phillips, H.E., Seagle, J.P.: Portfolio selection: the effects of uncertain means, variances, and covariances. *J. Financ. Quant. Anal.* **6**(5), 1251–1262 (1971)
- Frost, P.A., Savarino, J.E.: Portfolio size and estimation risk. *J. Portf. Manag.* **12**(4), 60–64 (1986)
- Golosnoy, V., Okhrin, Y.: Multivariate shrinkage for optimal portfolio weights. *Eur. J. Finance* **13**(5), 441–458 (2007)
- Haff, L.R.: Estimation of the inverse covariance matrix: random mixtures of the inverse Wishart matrix and the identity. *Ann. Stat.* **7**(6), 1264–1276 (1979)
- Holgersson, T., Karlsson, P.S.: Three estimators of the Mahalanobis distance in high-dimensional data. *J. Appl. Stat.* **39**(12), 2713–2720 (2012)
- Holgersson, T., Mansoor, R.: Assessing normality of high-dimensional data. *Commun. Stat. Simul. Comput.* **42**(2), 360–369 (2013)
- James, W., Stein, C.: Estimation with quadratic loss. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pp. 361–379, University of California Press, Berkeley, CA (1961)
- Kempf, A., Memmel, C.: Estimating the global minimum variance portfolio. *Schmalenbach Bus. Rev.* **58**(4), 332–348 (2006)
- Kempf, A., Kreuzberg, K., Memmel, C.: How to incorporate estimation risk into Markowitz optimization. In: *Operations Research Proceedings 2001*, pp. 175–182. Springer (2002)
- Kotz, S., Nadarajah, S.: *Multivariate T-Distributions and Their Applications*. Cambridge University Press, Cambridge (2004)
- Ledoit, O., Wolf, M.: Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empir. Finance* **10**(5), 603–621 (2003)
- Ledoit, O., Wolf, M.: Honey, I shrunk the sample covariance matrix. *J. Portf. Manag.* **30**(4), 110–119 (2004)
- Markowitz, H.M.: Portfolio selection. *J. Finance* **7**(1), 77–91 (1952)
- Markowitz, H.M.: *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press, New Haven (1959)
- Merton, R.C.: On estimating the expected return on the market: an exploratory investigation. *J. Financ. Econ.* **8**(4), 323–361 (1980)
- Michaud, R.O.: The Markowitz optimization enigma: is optimized optimal? *ICFA Contin. Edu. Ser.* **1989**(4), 43–54 (1989)
- Muirhead, R.J.: *Aspects of Multivariate Statistical Theory*. Wiley, New York (1982)
- Okhrin, Y., Schmid, W.: Distributional properties of portfolio weights. *J. Econom.* **134**(1), 235–256 (2006)
- Okhrin, Y., Schmid, W.: Comparison of different estimation techniques for portfolio selection. *ASTA Adv. Stat. Anal.* **91**(2), 109–127 (2007)
- Rao, C.R.: *Linear Statistical Inference and Its Applications*, 2nd edn. Wiley, New York (2008)
- Roy, A.D.: Safety first and the holding of assets. *Econometrica* **20**(3), 431–449 (1952)
- Serdobolskii, V.I.: The resolvent and the spectral functions of sample covariance matrices of increasing dimension. *Russ. Math. Surv.* **40**(2), 232–233 (1985)
- Serdobolskii, V.I.: Estimation of high-dimensional inverse covariance matrices. In: *Multivariate Statistical Analysis*, pp. 87–101. Springer (2000)
- Srivastava, M.S.: *Methods of Multivariate Statistics*. Wiley, New York (2002)

---

Stein, C.: Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 197–206. University of California Press, Berkeley, CA (1956)

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