Optimal Shrinkage-Based Portfolio Selection in High Dimensions

Bodnar, Taras; Okhrin, Yarema ; Parolya, N.

DOI
10.1080/07350015.2021.2004897

Publication date
2021

Document Version
Accepted author manuscript

Published in
Journal of Business and Economic Statistics

Optimal shrinkage-based portfolio selection in high dimensions[1]

Taras Bodnar[a], Yarema Okhrin[b] and Nestor Parolya[c,2]

[a] *Department of Mathematics, Stockholm University, Roslagsvägen 101, SE-10691 Stockholm, Sweden*

[b] *Department of Statistics, University of Augsburg, Universitätsstr. 16, D-86159 Augsburg, Germany*

[c] *Department of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands*

Abstract

In this paper we estimate the mean-variance portfolio in the high-dimensional case using the recent results from the theory of random matrices. We construct a linear shrinkage estimator which is distribution-free and is optimal in the sense of maximizing with probability 1 the asymptotic out-of-sample expected utility, i.e., mean-variance objective function for different values of risk aversion coefficient which in particular leads to the maximization of the out-of-sample expected utility and to the minimization of the out-of-sample variance. One of the main features of our estimator is the inclusion of the estimation risk related to the sample mean vector into the high-dimensional portfolio optimization. The asymptotic properties of the new estimator are investigated when the number of assets p and the sample size n tend simultaneously to infinity such that $p/n \to c \in (0, +\infty)$. The results are obtained under weak assumptions imposed on the distribution of the asset returns, namely the existence of the $4 + \varepsilon$ moments is only required. Thereafter we perform numerical and empirical studies where the small- and large-sample behavior of the derived estimator is investigated. The suggested estimator shows significant improvements over the existent approaches including the nonlinear shrinkage estimator and the three-fund portfolio rule, especially when the portfolio dimension is larger than the sample size. Moreover, it is robust to deviations from normality.

JEL Classification: G11, C13, C14, C58, C65

Keywords: expected utility portfolio, large-dimensional asymptotics, covariance matrix estimation, random matrix theory.

[1]Published in Journal of Business and Economic Statistics https://doi.org/10.1080/07350015.2021.2004897

[2]Corresponding author. E-mail address: n.parolya@tudelft.nl

1 Introduction

In the seminal paper of Markowitz (1952) the author suggests to determine the optimal composition of a portfolio of financial assets by minimizing the portfolio variance assuming that the expected portfolio return attains some prespecified fixed value. By varying this value we obtain the whole efficient frontier in the mean-standard deviation space. Despite of its simplicity, this approach justifies the advantages of diversification and is a standard technique and benchmark in asset management. Equivalently (see, Tobin (1958), Bodnar et al. (2013)) we can obtain the same portfolios by maximizing the expected quadratic utility (EU) with the optimization problem given by

$$\mathbf{w}'\boldsymbol{\mu}_n - \frac{\gamma}{2}\mathbf{w}'\boldsymbol{\Sigma}_n\mathbf{w} \to max \;\; \text{subject to} \;\; \mathbf{w}'\mathbf{1}_p = 1\,, \tag{1.1}$$

where $\mathbf{w} = (\omega_1, \ldots, \omega_p)'$ is the vector of portfolio weights, $\mathbf{1}_p$ is the p-dimensional vector of ones, $\boldsymbol{\mu}_n$ and $\boldsymbol{\Sigma}_n$ are the p-dimensional mean vector and the $p \times p$ covariance matrix of asset returns, respectively. The quantity $\gamma > 0$ determines the investor's behavior towards risk. It must be noted that the maximization of the mean-variance objective function (1.1) is equivalent to the maximization of the exponential utility (CARA) function under the assumption of normality of the asset returns. In this case γ equals the investor's absolute risk aversion coefficient (see, e.g., Pratt (1964)).

The solution of the optimization problem (1.1) is well known and it is given by

$$\mathbf{w}_{EU} = \mathbf{w}_{GMV} + \gamma^{-1}\mathbf{Q}_n\boldsymbol{\mu}_n\,, \tag{1.2}$$

where

$$\mathbf{Q}_n = \boldsymbol{\Sigma}_n^{-1} - \frac{\boldsymbol{\Sigma}_n^{-1}\mathbf{1}_p\mathbf{1}_p'\boldsymbol{\Sigma}_n^{-1}}{\mathbf{1}_p'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}_p} \tag{1.3}$$

and

$$\mathbf{w}_{GMV} = \frac{\boldsymbol{\Sigma}_n^{-1}\mathbf{1}_p}{\mathbf{1}_p'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}_p} \tag{1.4}$$

is the vector of the weights of the global minimum variance (GMV) portfolio. By changing the risk-aversion coefficient $\gamma \in (0, \infty)$ we obtain the set of optimal portfolios. Merton (1972) proved that this set is a parabola in the mean-variance (R-V) space (cf. Bodnar and Schmid (2009)) given by

$$(R - R_{GMV})^2 = s(V - V_{GMV}), \tag{1.5}$$

where

$$R_{GMV} = \frac{\boldsymbol{\mu}_n' \boldsymbol{\Sigma}_n^{-1} \mathbf{1}_p}{\mathbf{1}_p' \boldsymbol{\Sigma}_n^{-1} \mathbf{1}_p} \quad \text{and} \quad V_{GMV} = \frac{1}{\mathbf{1}_p' \boldsymbol{\Sigma}_n^{-1} \mathbf{1}_p} \tag{1.6}$$

are the expected return and the variance of the GMV portfolio, and

$$s = \boldsymbol{\mu}_n' \mathbf{Q}_n \boldsymbol{\mu}_n \tag{1.7}$$

is its slope parameter. The quantity s is always non-negative since $\mathbf{Q}_n$ is a positive semidefinite matrix. Moreover, when s is equal to zero, then the efficient frontier degenerates into a straight line with the GMV portfolio being the only optimal portfolio.

In practice, however, the above mentioned approach of constructing an optimal portfolio frequently shows poor out-of-sample performance in terms of various performance measures. Even naive portfolio strategies, e.g., equally weighted portfolio (see, DeMiguel et al. (2009)), often outperform the mean-variance strategy. One of the reasons is the estimation risk. The unknown parameters $\boldsymbol{\mu}_n$ and $\boldsymbol{\Sigma}_n$ have to be estimated using historical data on asset returns. This results in the "plug-in" estimator of the EU portfolio (1.2) which is a traditional and simple way to evaluate the portfolio in practice. This estimator is constructed by replacing the mean vector $\boldsymbol{\mu}_n$ and the covariance matrix $\boldsymbol{\Sigma}_n$ with their sample counterparts in (1.2). Okhrin and Schmid (2006) derive the expectation and the variance of the sample portfolio weights under the assumption that the asset returns follow a multivariate normal distribution, whereas Bodnar and Schmid (2011) obtain the exact finite-sample distribution. Recently, Bodnar et al. (2016) extended these results to the case $n < p$.

The estimation of the parameters has a negative impact on the performance of the asset

allocation strategy. This is noted in a series of papers with Merton (1980), Best and Grauer (1991), Chopra and Ziemba (1993) among others. Several approaches have arisen to reduce the consequences of the estimation risk. One strand of research opts for the Bayesian framework and using appropriate priors takes the estimation risk into account already while building the portfolio. The second strand relies on the shrinkage techniques and is related to the method exploited in this paper. A straightforward way to improve the properties of the estimators for $\boldsymbol{\mu}_n$ and $\boldsymbol{\Sigma}_n$ is to use the shrinkage approach (see, Jorion (1986), Ledoit and Wolf (2004)). Alternatively, one may apply the shrinkage estimation to the portfolio weights directly. Golosnoy and Okhrin (2007) consider the multivariate shrinkage estimator by shrinking the portfolios with and without the riskless asset to an arbitrary static portfolio. A similar technique is used by Frahm and Memmel (2010), who construct a feasible shrinkage estimator for the GMV portfolio which dominates the traditional one. At last, Bodnar et al. (2018) suggest a shrinkage estimator for the GMV portfolio which is feasible even for the singular sample covariance matrix.

An important issue nowadays is, however, the asset allocation for large portfolios. The sample estimators work well only in the case when the number of assets p is fixed and substantially smaller than the sample size n. This case is known as the standard asymptotics in statistics (see, Le Cam and Lo Yang (2000)). Under this asymptotics the traditional sample estimator is a consistent estimator for the EU portfolio. But what happens when the dimension p and the sample size n are comparable of size, say $p = 900$ and $n = 1000$? Technically, here we are in the situation when both the number of assets p and the sample size n tend to infinity. In the case when p/n tends to some concentration ratio $c > 0$ this asymptotics is known as high-dimensional asymptotics or "Kolmogorov" asymptotics (see, e.g., Bai and Silverstein (2010)). If c is close to one the sample covariance matrix tends to be close to a singular one and when $c > 1$ it becomes singular. Thus it is very unstable and tends to under- or overestimate the true parameters for c smaller but close to 1 (see, Bai and Shi (2011)). As a result, the sample estimator of the EU portfolio behaves badly in this case both from the theoretical and practical points of view (see, e.g., El Karoui (2010); Rubio et al. (2012)). For $c > 1$ the inverse sample covariance matrix does not exist and the portfolio cannot be constructed in the traditional way.

Taking the above mentioned information into account the aim of the paper is to construct a

feasible and simple shrinkage estimator of the EU portfolio which is optimal in an asymptotic sense and is additionally distribution-free. The estimator is developed using the fast growing branch of probability theory, namely random matrix theory. The main result of this theory is proved by Marčenko and Pastur (1967) and further extended under very general conditions by Silverstein (1995). Now it is called Marčenko-Pastur equation. Its importance arises in many areas of science because it shows how the true covariance matrix and its sample estimator are connected asymptotically. Knowing this we can build suitable estimators for high-dimensional quantities which depend on $\mathbf{\Sigma}_n$. In our case this refers to the shrinkage intensities. Note however, that the optimal shrinkage intensity depends again on the unknown characteristics of the asset returns. To overcome this problem we derive consistent estimators for specific functions (quadratic and bilinear forms) of the inverse sample covariance matrix and mean vector. Furthermore, we succeed to provide consistent estimators for the optimal shrinkage intensities too. Additional advantage of our approach is the simultaneous treatment of estimation risks of both the covariance matrix and the mean vector. In particular we contribute to the existent literature (see, Ledoit and Wolf (2017a)) by weakening the assumption imposed on the mean vector of the asset returns.

It is worth mentioning that there are clear links between the subject of the paper and classical methods in statistical signal processing. The data generating process considered in the paper encompasses a broad range of system configurations described by the general vector channel model. Moreover, as for the aforementioned mean-variance portfolio optimization problem, usual linear filtering schemes solving typical signal waveform estimation and detection problems in signal array processing and wireless communications are based on the estimation of the unknown population covariance matrix. Famous example is the equivalence of the GMV portfolio to the so-called Capon or minimum variance distortionless response (MVDR) beamformer (see, Verdú (1998); Van Trees (2002)).

The rest of paper is organized as follows. In the next section, we construct a shrinkage estimator for the optimal portfolio weights obtained by shrinking the EU portfolio weights to an arbitrary target portfolio. The oracle shrinkage intensity and the corresponding feasible bona-fide estimators for $c < 1$ and $c > 1$ are established as well. The derived results are evaluated in Section 3 in extensive simulation and empirical studies. All proofs are moved to

the Appendix presented in the supplementary material.

2 Optimal shrinkage estimator of mean-variance portfolio

Let $\mathbf{Y}_n = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n)$ be the $p \times n$ data matrix which consists of n vectors of the returns on $p \equiv p(n)$ assets. Let $E(\mathbf{y}_i) = \boldsymbol{\mu}_n$ and $Cov(\mathbf{y}_i) = \boldsymbol{\Sigma}_n$ for $i \in 1, ..., n$. We assume that $p/n \to c \in (0, +\infty)$ as $n \to \infty$. This type of limiting behavior is known as "the large dimensional asymptotics" or "Kolmogorov asymptotics". In this case the traditional sample estimators perform poorly or even very poorly and tend to over/underestimate the unknown parameters of the asset returns, e.g., the mean vector and the covariance matrix.

Throughout the paper it is assumed that there exists a $p \times n$ random matrix $\mathbf{X}_n$ which consists of independent and identically distributed (i.i.d.) real random variables with zero mean and unit variance such that

$$\mathbf{Y}_n = \boldsymbol{\mu}_n \mathbf{1}_n' + \boldsymbol{\Sigma}_n^{\frac{1}{2}} \mathbf{X}_n \,. \tag{2.1}$$

It must be noted that the observation matrix $\mathbf{Y}_n$ has dependent rows but independent columns. Broadly speaking, this means that we allow arbitrary cross-sectional correlations of the asset returns but assume their independence over time. Although this assumption looks quite restrictive for financial applications, there exist stronger results from random matrix theory which show that the model can be extended to (weakly) dependent variables by demanding more complicated conditions on the elements of $\mathbf{Y}_n$ (see, Bai and Zhou (2008)) or by controlling the number of dependent entries as dimension increases (see, Hui and Pan (2010), Friesen et al. (2013), Wei et al. (2016)). Although our findings can still be used when weak serial dependence structure is present between the observation vectors, like in the case of uncorrelated GARCH (generalized autoregressive conditional heteroscedastic) processes or similar ones (see, e.g., the simulation study in Bodnar et al. (2021a)), we suspect substantial changes in the analytical expressions stated in the theorems for strongly correlated observation vectors, like in the case of VAR (vector autoregressive) processes. In such situations, the estimator will depend on the

autocorrelation matrices of the underlying stochastic model and the theoretical results of the paper must be adjusted correspondingly. This interesting and important topic is not treated in the paper and is left for future research.

Nevertheless, if the entries of matrix $\mathbf{Y}_n$ are weakly dependent or so called m-dependent, this will only make the proofs more technical, but leave the results unchanged. For that reason we assume independent in time asset returns only to simplify the proofs of the main theorems and make them as transparent as possible. The three assumptions which are used throughout the paper are the following:

(A1) The covariance matrix of the asset returns $\mathbf{\Sigma}_n$ is a nonrandom p-dimensional positive definite matrix.

(A2) The elements of the matrix $\mathbf{X}_n$ have uniformly bounded $4+\varepsilon$ moments for some $\varepsilon > 0$.

(A3) The efficient frontier is asymptotically a non-degenerate object, i.e. for its slope parameter it holds that $s = \boldsymbol{\mu}_n' \mathbf{Q}_n \boldsymbol{\mu}_n > 0$ uniformly in p.

All of these regularity assumptions are general enough to fit many real world situations. The assumption (A1) together with (2.1) are usual for financial and statistical problems and they impose no strong restrictions. The assumption (A2) is a technical one. Although we demand the existence of moments of order a bit higher than four, this is solely due to the fact that the almost sure convergence is employed in the formulation of the theoretical results. In case of the convergence in probability the existence of exactly the fourth moment is sufficient. Indeed, it can be easily shown that this extra ε follows from the Borel-Cantelli lemma (see Rubio and Mestre (2011)[Proof of Lemma 4]). The assumption (A3) has an important financial interpretation. It ensures that the efficient frontier is a parabola in the mean-variance space as defined in (1.5) and it does not degenerate into a line parallel to the variance axis (cf., Bodnar and Bodnar (2010)). In the latter case, the only optimal portfolio is the GMV portfolio (1.4), a special case of the EU portfolio (1.2) with $\gamma = \infty$, and its shrinkage estimators have already been developed in Frahm and Memmel (2010) and Bodnar et al. (2018). The assumption (A3) can be tested in practice by using Theorem 1 of Bodnar et al. (2021c).

The sample covariance matrix is given by

$$\mathbf{S}_n = \frac{1}{n}\mathbf{Y}_n(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')\mathbf{Y}_n' = \frac{1}{n}\mathbf{\Sigma}_n^{\frac{1}{2}}\mathbf{X}_n(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')\mathbf{X}_n'\mathbf{\Sigma}_n^{\frac{1}{2}}\,, \tag{2.2}$$

where the symbol $\mathbf{I}_n$ stands for the n-dimensional identity matrix. The sample mean vector becomes

$$\bar{\mathbf{y}}_n = \frac{1}{n}\mathbf{Y}_n\mathbf{1}_n = \boldsymbol{\mu}_n + \mathbf{\Sigma}_n^{\frac{1}{2}}\bar{\mathbf{x}}_n \;\text{ with } \bar{\mathbf{x}}_n = \frac{1}{n}\mathbf{X}_n\mathbf{1}_n\,. \tag{2.3}$$

2.1 Oracle estimator. Case $c < 1$

In this section we consider the optimal shrinkage estimator for the EU portfolio weights presented in the introduction by finding the shrinkage parameter α and fixing some target portfolio $\mathbf{b}$.

The resulting estimator for $c < 1$ is given by

$$\hat{\mathbf{w}}_{GSE} = \alpha_n\hat{\mathbf{w}}_S + (1-\alpha_n)\mathbf{b} \;\text{ with } \mathbf{b}'\mathbf{1}_p = 1\,, \tag{2.4}$$

where the vector $\hat{\mathbf{w}}_S$ is the sample estimator of the EU portfolio given in (1.2), namely

$$\hat{\mathbf{w}}_S = \frac{\mathbf{S}_n^{-1}\mathbf{1}_p}{\mathbf{1}_p'\mathbf{S}_n^{-1}\mathbf{1}_p} + \gamma^{-1}\hat{\mathbf{Q}}_n\bar{\mathbf{y}}_n \tag{2.5}$$

with

$$\hat{\mathbf{Q}}_n = \mathbf{S}_n^{-1} - \frac{\mathbf{S}_n^{-1}\mathbf{1}_p\mathbf{1}_p'\mathbf{S}_n^{-1}}{\mathbf{1}_p'\mathbf{S}_n^{-1}\mathbf{1}_p}\,. \tag{2.6}$$

The target portfolio $\mathbf{b} \in \mathbb{R}^p$ is a given nonrandom (or random, but independent of $\mathbf{Y}_n$) vector with $\mathbf{b}'\mathbf{1}_p = 1$.No assumption is imposed on the shrinkage intensity α_n which is the object of our interest.

The aim is now to find the optimal shrinkage intensity for a given nonrandom target portfolio $\mathbf{b}$. For that reason we introduce a unified mean-variance objective function in order to calibrate

the shrinkage intensity α_n. Consider the following optimization problem

$$U(\beta) = \hat{\mathbf{w}}'_{GSE}(\alpha_n)\boldsymbol{\mu}_n - \frac{\beta}{2}\hat{\mathbf{w}}'_{GSE}(\alpha_n)\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_{GSE}(\alpha_n) \longrightarrow max \quad \text{with respect to} \quad \alpha_n\,. \tag{2.7}$$

Obviously, the mean-variance objectives (1.1) and (2.7) coincide if $\beta = \gamma$. Other special values of β which lead to widely used out-of-sample performance measures we summarize in the following proposition

Proposition 2.1 (Calibration criteria)**.** *The optimization problem (2.7) is equivalent to*

(i) maximization of the mean-variance objective (1.1) if $\beta = \gamma$*,*

(ii) minimization of the out-of-sample variance $\hat{\mathbf{w}}'_{GSE}(\alpha_n)\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_{GSE}(\alpha_n)$ *if* $\beta \to \infty$*,*

The proof of Proposition 2.1 follows from the fact that all optimal mean-variance portfolios can be obtained by maximizing the expected quadratic utility function with a specific risk aversion coefficient. As a result, the global minimum variance portfolio is a partial solution of the optimization problem (1.1). The presentation of the calibration criterion (2.7) provides an elegant way how to find the optimal shrinkage intensity $\alpha_n = \alpha_n(\beta)$ in a unified manner for several popular out-of-sample loss functions and compare them just by changing the parameter β. In Section 2.3, we provide consistent estimates of these quantities under high-dimensional asymptotic regime $p/n \to c > 0$ for $(p, n) \to \infty$.

It is worth mentioning that the coefficient β has an interesting interpretation from statistical point of view. While coefficient γ controls for investor attitude towards financial risk ("in-sample risk"), the parameter β stays for controlling the estimation risk ("out-of-sample risk"). This implies that even the mean-variance investor with arbitrary $\gamma > 0$ could choose $\beta \to \infty$ if she/he is interested, for example, in the minimization of the out-of-sample variance of the estimated portfolio.

The unified calibration criterion (2.7) can be rewritten as

$$U(\beta) = \alpha_n\hat{\mathbf{w}}'_S\boldsymbol{\mu}_n + (1-\alpha_n)\mathbf{b}'\boldsymbol{\mu}_n - \frac{\beta}{2}\left(\alpha_n^2\hat{\mathbf{w}}'_S\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_S + 2\alpha_n(1-\alpha_n)\mathbf{b}'\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_S + (1-\alpha_n)^2\mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}\right) \to max$$
$$\text{with respect to} \quad \alpha_n\,. \tag{2.8}$$

Next, taking the derivative of U with respect to α_n and setting it equal to zero we get

$$\frac{\partial U}{\partial \alpha_n} = (\hat{\mathbf{w}}_S - \mathbf{b})'\boldsymbol{\mu}_n - \beta\left(\alpha_n \hat{\mathbf{w}}_S' \boldsymbol{\Sigma}_n \hat{\mathbf{w}}_S + (1-2\alpha_n)\mathbf{b}'\boldsymbol{\Sigma}_n \hat{\mathbf{w}}_S - (1-\alpha_n)\mathbf{b}'\boldsymbol{\Sigma}_n \mathbf{b}\right) \stackrel{!}{=} 0\,.$$

From the last equation it is easy to find the optimal shrinkage intensity α_n^* given by

$$\alpha_n^* = \beta^{-1}\frac{(\hat{\mathbf{w}}_S - \mathbf{b})'(\boldsymbol{\mu}_n - \beta\boldsymbol{\Sigma}_n \mathbf{b})}{(\hat{\mathbf{w}}_S - \mathbf{b})'\boldsymbol{\Sigma}_n(\hat{\mathbf{w}}_S - \mathbf{b})}\,. \tag{2.9}$$

To ensure that α_n^* is the unique maximizer of (2.7) the second derivative of U must be negative, which is always fulfilled. Indeed, it follows from the positive definitiveness of the matrix $\boldsymbol{\Sigma}_n$, namely

$$\frac{\partial^2 U}{\partial \alpha_n^2} = -\beta(\hat{\mathbf{w}}_S - \mathbf{b})'\boldsymbol{\Sigma}_n(\hat{\mathbf{w}}_S - \mathbf{b}) < 0\,. \tag{2.10}$$

In the next theorem we derive the asymptotic properties of the optimal shrinkage intensity α_n^* under large-dimensional asymptotics.

Theorem 2.1. *Assume (A1)-(A3). Then it holds that*

$$|\alpha_n^* - \alpha^*| \xrightarrow{a.s.} 0 \quad \textit{for} \quad \frac{p}{n} \to c \in (0,1) \quad \textit{as} \quad n \to \infty$$

with

$$\alpha^* = \beta^{-1}\frac{(R_{GMV} - R_b)\left(1 + \frac{\beta/\gamma}{1-c}\right) + \beta(V_b - V_{GMV}) + \frac{\gamma^{-1}}{1-c}s}{\frac{1}{1-c}V_{GMV} - 2\left(V_{GMV} + \frac{\gamma^{-1}}{1-c}(R_b - R_{GMV})\right) + \gamma^{-2}\left(\frac{s}{(1-c)^3} + \frac{c}{(1-c)^3}\right) + V_b}, \tag{2.11}$$

where the parameters of the efficient frontier R_{GMV}, V_{GMV} and s are given in (1.6) *and* (1.7)*, respectively. The quantities $R_b = \mathbf{b}'\boldsymbol{\mu}_n$ and $V_b = \mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}$ denote the expected return and the variance of the target portfolio $\mathbf{b}$.*

Next, we assess the performance of the classical estimator of the portfolio weights $\hat{\mathbf{w}}_S$ and the optimal shrinkage weights $\hat{\mathbf{w}}_{GSE}$. As a measure of performance we consider the relative

increase in the utility of the portfolio return compared to the portfolio based on true parameters of asset returns. The results are summarized in the following corollary.

Corollary 2.1. *(a) Let U_{EU} and U_S be the mean-variance objectives in (1.1) for the true EU portfolio and its traditional estimator. Then under the assumptions of Theorem 2.1, the relative loss of the traditional estimator of the EU portfolio is given by*

$$L_S = \frac{U_{EU} - U_S}{U_{EU}} \xrightarrow{a.s.} \frac{\frac{\gamma}{2}\left(\frac{1}{1-c} - 1\right) \cdot V_{GMV} + \gamma^{-1}\left(\frac{1}{2} - \frac{1}{(1-c)} + \frac{1}{2(1-c)^3}\right) \cdot s + \frac{1}{2\gamma} \cdot \frac{c}{(1-c)^3}}{R_{GMV} + \frac{1}{2}\gamma^{-1} \cdot s - \frac{\gamma}{2} V_{GMV}} \quad (2.12)$$

for $\frac{p}{n} \to c \in (0,1)$ as $n \to \infty$.

(b) Let U_{GSE} be the expected quadratic utility for optimal shrinkage estimator of the EU portfolio. Under the assumptions of Theorem 2.1, the relative loss of the optimal shrinkage estimator is given by

$$L_{GSE} = \frac{U_{EU} - U_{GSE}}{U_{EU}} \xrightarrow{a.s.} (\alpha^*)^2 L_S + (1-\alpha^*)^2 L_{\mathbf{b}} + \alpha^*(1-\alpha^*)\frac{c}{1-c}\frac{R_b - R_{GMV} - \gamma^{-1}s}{U_{EU}} \quad (2.13)$$

for $\frac{p}{n} \to c \in (0,1)$ as $n \to \infty$ with $L_{\mathbf{b}} = (U_{EU} - U_{\mathbf{b}})/U_{EU}$ is the relative loss in the expected utility $U_{\mathbf{b}}$ of the target portfolio $\mathbf{b}$.

2.2 Oracle estimator. Case $c > 1$.

Here, similarly as in Bodnar et al. (2018), we will use the generalized inverse of the sample covariance matrix $\mathbf{S}_n$. Particularly, we use the following generalized inverse of the sample covariance matrix $\mathbf{S}_n$

$$\mathbf{S}_n^* = \mathbf{\Sigma}_n^{-1/2}\left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n' - \bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\right)^+ \mathbf{\Sigma}_n^{-1/2}, \quad (2.14)$$

where $'+'$ denotes the Moore-Penrose inverse. It can be shown that $\mathbf{S}_n^*$ is a generalized inverse of $\mathbf{S}_n$ satisfying $\mathbf{S}_n^*\mathbf{S}_n\mathbf{S}_n^* = \mathbf{S}_n^*$ and $\mathbf{S}_n\mathbf{S}_n^*\mathbf{S}_n = \mathbf{S}_n$. However, $\mathbf{S}_n^*$ is not exactly equal to the Moore-Penrose inverse because it does not satisfy the conditions $(\mathbf{S}_n^*\mathbf{S}_n)' = \mathbf{S}_n^*\mathbf{S}_n$ and $(\mathbf{S}_n\mathbf{S}_n^*)' = \mathbf{S}_n\mathbf{S}_n^*$. In case $c < 1$ the generalized inverse $\mathbf{S}_n^*$ coincides with the usual inverse $\mathbf{S}_n^{-1}$. Moreover, if $\mathbf{\Sigma}_n$ is a multiple of identity matrix, then $\mathbf{S}_n^*$ is equal to the Moore-Penrose inverse $\mathbf{S}_n^+$. In this

section, $\mathbf{S}_n^*$ is used only to determine an oracle estimator for the weights of the EU portfolio. The bona fide estimator is constructed in the next section.

Thus, the oracle estimator for $c > 1$ is given by

$$\hat{\mathbf{w}}_{GSE}^* = \alpha_n^+ \hat{\mathbf{w}}_{S^*} + (1 - \alpha_n^+)\mathbf{b} \;\text{ with } \mathbf{b}'\mathbf{1}_p = 1, \tag{2.15}$$

where the vector $\hat{\mathbf{w}}_{S^*}$ is the sample estimator of the EU portfolio given in (1.2), namely

$$\hat{\mathbf{w}}_{S^*} = \frac{\mathbf{S}_n^* \mathbf{1}_p}{\mathbf{1}_p' \mathbf{S}_n^* \mathbf{1}_p} + \gamma^{-1} \hat{\mathbf{Q}}_n^* \bar{\mathbf{y}}_n \tag{2.16}$$

with

$$\hat{\mathbf{Q}}_n^* = \mathbf{S}_n^* - \frac{\mathbf{S}_n^* \mathbf{1}_p \mathbf{1}_p' \mathbf{S}_n^*}{\mathbf{1}_p' \mathbf{S}_n^* \mathbf{1}_p}. \tag{2.17}$$

Again, the shrinkage intensity α_n^+ is the object of our interest. In order to save place we skip the optimization procedure for α_n^+ as it is only slightly different from the case $c < 1$. The optimal shrinkage intensity α_n^+ in case $c > 1$ is given by

$$\alpha_n^+ = \beta^{-1} \frac{(\hat{\mathbf{w}}_{S^*} - \mathbf{b})'(\boldsymbol{\mu}_n - \beta \boldsymbol{\Sigma}_n \mathbf{b})}{(\hat{\mathbf{w}}_{S^*} - \mathbf{b})' \boldsymbol{\Sigma}_n (\hat{\mathbf{w}}_{S^*} - \mathbf{b})}. \tag{2.18}$$

In the next theorem we find the asymptotic equivalent quantity for α_n^+ in the case $p/n \to c \in (1, +\infty)$ as $n \to \infty$.

Theorem 2.2. *Assume (A1)-(A3). Then it holds that*

$$\left| \alpha_n^+ - \alpha^+ \right| \xrightarrow{a.s.} 0 \;\; \textit{for} \;\; \frac{p}{n} \to c \in (1, +\infty) \;\; \textit{as} \;\; n \to \infty$$

with

$$\alpha^+ = \beta^{-1} \frac{(R_{GMV} - R_b)\left(1 + \dfrac{\beta/\gamma}{c(c-1)}\right) + \beta(V_b - V_{GMV}) + \dfrac{\gamma^{-1}}{c(c-1)} s}{\dfrac{c^2}{(c-1)} V_{GMV} - 2\left(V_{GMV} + \frac{\gamma^{-1}}{c(c-1)}(R_b - R_{GMV})\right) + \dfrac{\gamma^{-2}}{(c-1)^3}(s + c^2) + V_b}, \tag{2.19}$$

where R_{GMV}, V_{GMV}, R_b, V_b, and s are defined in Theorem 2.1.

Next, as for the case $c < 1$, we provide here the expression for the relative losses.

Corollary 2.2. *(a) Let U_{EU} and U_S be the mean-variance objectives in (1.1) for the true EU portfolio and its traditional estimator. Then under the assumptions of Theorem 2.2, the relative loss of the traditional estimator of the EU portfolio is given by*

$$L_S = \frac{U_{EU} - U_S}{U_{EU}} \xrightarrow{a.s.} \frac{\frac{\gamma}{2}\left(\frac{c^2}{c-1} - 1\right) \cdot V_{GMV} + \gamma^{-1}\left(\frac{1}{2} - \frac{1}{c(c-1)} + \frac{1}{2(c-1)^3}\right) \cdot s + \frac{1}{2\gamma} \cdot \frac{c^2}{(c-1)^3}}{R_{GMV} + \frac{1}{2}\gamma^{-1} \cdot s - \frac{\gamma}{2} V_{GMV}} \tag{2.20}$$

for $\frac{p}{n} \to c \in (1, +\infty)$ as $n \to \infty$.

(b) Let U_{GSE} be the expected quadratic utility for the optimal shrinkage estimator of the EU portfolio. Under the assumptions of Theorem 2.2, the relative loss of the optimal shrinkage estimator is given by

$$L_{GSE} = \frac{U_{EU} - U_{GSE}}{U_{EU}} \xrightarrow{a.s.} (\alpha^+)^2 L_S + (1-\alpha^+)^2 L_{\mathbf{b}} + \alpha^+(1-\alpha^+)\frac{1 + c - c^2}{c(c-1)} \frac{R_b - R_{GMV} - \gamma^{-1} s}{U_{EU}} \tag{2.21}$$

for $\frac{p}{n} \to c \in (1, +\infty)$ as $n \to \infty$ with $L_{\mathbf{b}} = (U_{EU} - U_{\mathbf{b}})/U_{EU}$ is the relative loss in the expected utility $U_{\mathbf{b}}$ of the target portfolio $\mathbf{b}$.

2.3 Estimation of unknown parameters. Bona fide estimator

The limiting shrinkage intensities α^* and α^+ are not feasible in practice, because they depend on R_{GMV}, V_{GMV}, s, R_b, and V_b which are unknown quantities. In this subsection we derive consistent estimators for α^* and α^+. These results are summarized in two propositions dealing with the cases $c \in (0, 1)$ and $c \in (1, \infty)$, respectively. The statements follow directly from the proofs of Theorems 2.1 and 2.2 that are provided in the supplement of the paper.

Proposition 2.2. *The consistent estimator for the limiting optimal shrinkage intensity α^* under large dimensional asymptotics $p/n \to c < 1$ as $n \to \infty$ is given by*

$$\widehat{\alpha}^* = \beta^{-1} \frac{(\hat{R}_c - \hat{R}_b)\left(1 + \frac{\beta/\gamma}{1 - p/n}\right) + \beta(\hat{V}_b - \hat{V}_c) + \frac{\gamma^{-1}}{1-p/n}\hat{s}_c}{\frac{1}{1-p/n}\hat{V}_c - 2\left(\hat{V}_c + \frac{\gamma^{-1}}{1-p/n}(\hat{R}_b - \hat{R}_c)\right) + \gamma^{-2}\left(\frac{\hat{s}_c}{(1-p/n)^3} + \frac{p/n}{(1-p/n)^3}\right) + \hat{V}_b} \tag{2.22}$$

where $\hat{R}_c$, $\hat{V}_c$, $\hat{s}_c$, $\hat{R}_b$ *and* $\hat{V}_b$ *are given by*

$$\begin{aligned}
\hat{R}_c &= \hat{R}_{GMV} && (2.23)\\
\hat{V}_c &= \frac{1}{1-p/n}\hat{V}_{GMV} && (2.24)\\
\hat{s}_c &= (1-p/n)\hat{s} - p/n && (2.25)\\
\hat{R}_b &= \mathbf{b}'\bar{\mathbf{y}}_n && (2.26)\\
\hat{V}_b &= \mathbf{b}'\mathbf{S}_n\mathbf{b}\,, && (2.27)
\end{aligned}$$

which are also ratio consistent estimators for R_{GMV}, V_{GMV}, s, R_b, *and* V_b, *respectively, while* $\hat{R}_{GMV}$, $\hat{V}_{GMV}$ *and* $\hat{s}$ *are traditional plug-in estimators.*

Using Proposition 2.2 we can immediately construct a bona-fide estimator for the expected utility portfolio weights in case $c < 1$. It holds that

$$\hat{\mathbf{w}}_{BFGSE} = \widehat{\alpha}^* \left(\frac{\mathbf{S}_n^{-1}\mathbf{1}_p}{\mathbf{1}_p'\mathbf{S}_n^{-1}\mathbf{1}_p} + \gamma^{-1}\hat{\mathbf{Q}}_n\bar{\mathbf{y}}_n \right) + (1-\widehat{\alpha}^*)\mathbf{b} \tag{2.28}$$

with $\widehat{\alpha}^*$ given in Proposition 2.2. The expression (2.28) is the optimal shrinkage estimator for a given target portfolio $\mathbf{b}$ in the sense that the shrinkage intensity $\widehat{\alpha}^*$ tends almost surely to its optimal value α^* for $p/n \to c \in (0,1)$ as $n \to \infty$.

The situation is more complex in case $c > 1$. Here we can present only oracle estimators for the unknown quantities R_{GMV}, V_{GMV} and s.

Proposition 2.3. *The consistent estimator for the limiting optimal shrinkage intensity* α^* *under large dimensional asymptotics* $p/n \to c > 1$ *as* $n \to \infty$ *is given by*

$$\widehat{\alpha}^o = \beta^{-1} \frac{(\hat{R}_c^o - \hat{R}_b)\left(1 + \frac{\beta/\gamma}{p/n(p/n-1)}\right) + \beta(\hat{V}_b - \hat{V}_c^o) + \frac{\gamma^{-1}}{p/n(p/n-1)}\hat{s}_c^o}{\frac{(p/n)^2}{p/n-1}\hat{V}_c^o - 2\left(\hat{V}_c^o + \frac{\gamma^{-1}}{p/n(p/n-1)}(\hat{R}_b - \hat{R}_c^o)\right) + \frac{\gamma^{-2}}{(p/n-1)^3}\left(\hat{s}_c^o + (p/n)^2\right) + \hat{V}_b}\,, \tag{2.29}$$

where $\hat{R}_c^o$, $\hat{V}_c^o$, $\hat{s}_c^o$ *are given by*

$$\begin{aligned}\hat{R}_c^o &= \hat{R}_{GMV}\\ \hat{V}_c^o &= \frac{1}{p/n(p/n-1)}\hat{V}_{GMV}\\ \hat{s}_c^o &= p/n[(p/n-1)\hat{s}-1]\,,\end{aligned}$$

where $\hat{R}_{GMV}$, $\hat{V}_{GMV}$ *and* $\hat{s}$ *are the traditional plug-in estimators based on the generalized inverse* $\mathbf{S}_n^*$ *from* (2.14) *and* $\hat{R}_b$ *and* $\hat{V}_b$ *are given in (2.26) and (2.27), respectively.*

Note that $\widehat{\alpha}^o$ from Proposition 2.3 is not the bona fide estimator for the unknown shrinkage intensity α^+, since the matrix $\mathbf{S}_n^*$ depends on the unknown quantities. Thus, we propose a reasonable approximation using the application of the Moore-Penrose inverse $\mathbf{S}_n^+$. As a result, the bona fide estimators of the quantities R_{GMV}, V_{GMV} and s in case $c>1$ are approximated by

$$\hat{R}_c^+ \approx \frac{\bar{\mathbf{y}}_n'\mathbf{S}_n^+\mathbf{1}_p}{\mathbf{1}_p'\mathbf{S}_n^+\mathbf{1}_p}, \quad \hat{V}_c^+ \approx \frac{1}{p/n(p/n-1)}\frac{1}{\mathbf{1}_p'\mathbf{S}_n^+\mathbf{1}_p}, \quad \hat{s}_c^+ \approx p/n[(p/n-1)\bar{\mathbf{y}}_n'\mathbf{Q}_n^+\bar{\mathbf{y}}_n-1]\,, \tag{2.30}$$

respectively. The application of (2.30) leads to the bona fide optimal shrinkage estimator of the EU portfolio in case $c>1$ expressed as

$$\hat{\mathbf{w}}_{BFGSE}^+ = \widehat{\alpha}^+\left(\frac{\mathbf{S}_n^+\mathbf{1}_p}{\mathbf{1}_p'\mathbf{S}_n^+\mathbf{1}_p}+\gamma^{-1}\hat{\mathbf{Q}}_n^+\bar{\mathbf{y}}_n\right)+(1-\widehat{\alpha}^+)\mathbf{b}_n\,, \tag{2.31}$$

with

$$\widehat{\alpha}^+ = \beta^{-1}\frac{(\hat{R}_c^+-\hat{R}_b)\left(1+\frac{\beta/\gamma}{p/n(p/n-1)}\right)+\beta(\hat{V}_b-\hat{V}_c^+)+\frac{\gamma^{-1}}{p/n(p/n-1)}\hat{s}_c^+}{\frac{(p/n)^2}{p/n-1}\hat{V}_c^+-2\left(\hat{V}_c^++\frac{\gamma^{-1}}{p/n(p/n-1)}(\hat{R}_b-\hat{R}_c^+)\right)+\frac{\gamma^{-2}}{(p/n-1)^3}\left(\hat{s}_c^++(p/n)^2\right)+\hat{V}_b}\,, \tag{2.32}$$

where $\hat{R}_b$ and $\hat{V}_b$ are given in (2.26) and (2.27), respectively; $\mathbf{Q}_n^+ = \mathbf{S}_n^+ - \dfrac{\mathbf{S}_n^+\mathbf{1}\mathbf{1}'\mathbf{S}_n^+}{\mathbf{1}'\mathbf{S}_n^+\mathbf{1}}$ and $\mathbf{S}_n^+$ is the Moore-Penrose pseudo-inverse of the sample covariance matrix $\mathbf{S}_n$.

Remark 1. It is easy to verify that if $\mathbf{\Sigma}_n = \sigma^2\mathbf{I}_p$ for any $\sigma>0$ the considered approximations in (2.30) become the exact ones. Next, we investigate the quality of this approximation in

general case without imposing restrictions on $\mathbf{\Sigma}_n$. This issue was studied for other quantities involving $\mathbf{S}_n^*$ and $\mathbf{S}_n^+$ in detail by Bodnar and Parolya (2020), who compare the limiting spectral distributions of $\mathbf{S}_n^*$ and $\mathbf{S}_n^+$ by deriving the limits for their corresponding Stieltjes transforms. It is concluded that the two inverses behave completely different in general. However, when the concentration ratio c approaches 1, then the limiting spectral distributions of both inverses $\mathbf{S}_n^*$ and $\mathbf{S}_n^+$ coincide independently of the structure of $\mathbf{\Sigma}_n$. This in turn means that one should expect a good approximation quality when c is not far away from 1.

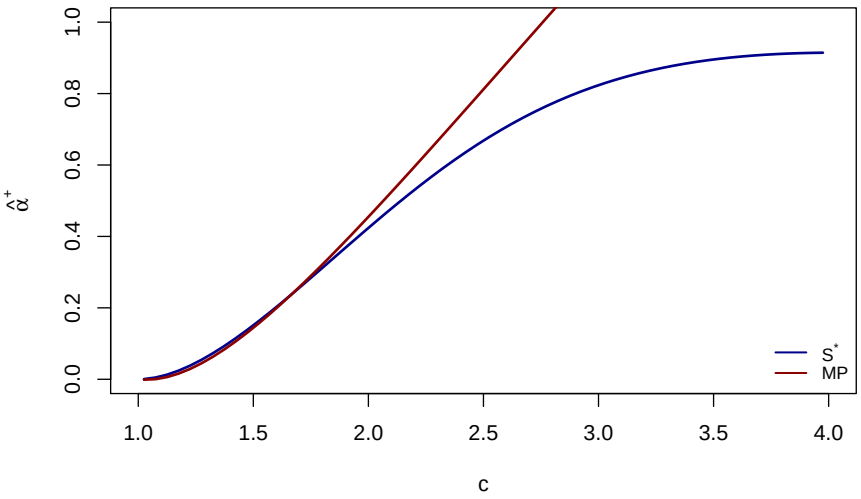

Figure 1: Estimated optimal shrinkage intensities for $\mathbf{S}_n^*$ and $\mathbf{S}_n^+$(MP) as function of concentration ratio $c > 1$ and dimension $p = 300$.

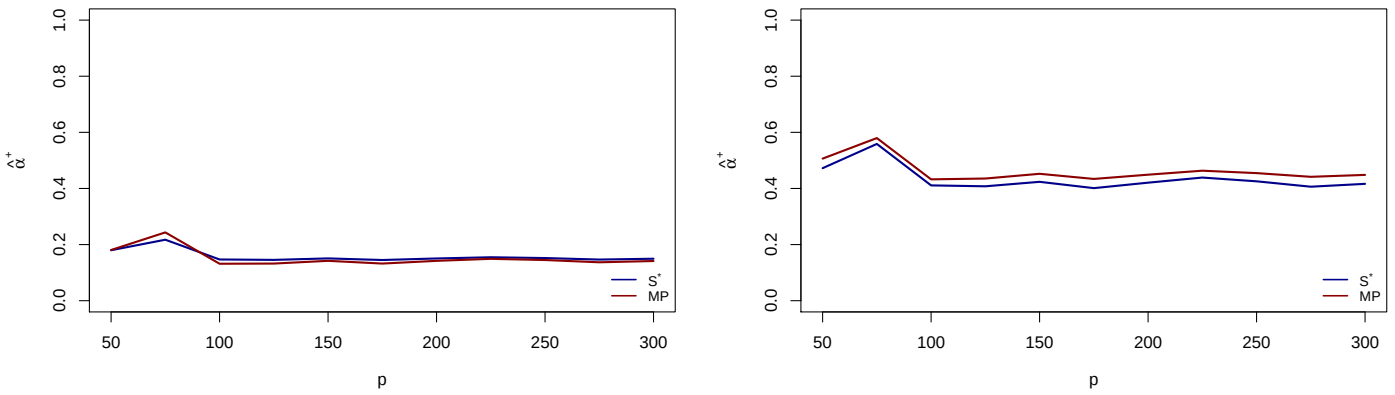

Figure 2: Estimated optimal shrinkage intensities for $\mathbf{S}_n^*$ and $\mathbf{S}_n^+$(MP) as function of dimension p for $c = 1.5$ (left) and $c = 2$ (right).

In Figure 1 we provide a comparison between the optimal shrinkage intensities computed using different types of generalized inverses, namely Moore-Penrose inverse $\mathbf{S}_n^+$ and the reflexive inverse $\mathbf{S}_n^*$ from (2.14). The optimal shrinkage intensity is calculated by (2.32) in the case of $\mathbf{S}_n^+$

and by using (2.32) where $\hat{R}_c^+$, $\hat{V}_c^+$, and $\hat{s}_c^+$ are replaced by $\hat{R}_c^o$, $\hat{V}_c^o$, and $\hat{s}_c^o$ from Proposition 2.3 in the case of $\mathbf{S}_n^*$. The design of the simulation study is exactly the same as the one described in Section 3.1.

We observe that the two optimal shrinkage intensities are quite identical till the breaking point $c = 2$. For $c > 2$ the Moore-Penrose approximation is not reliable anymore. We observe a similar behaviour for other structures of the covariance matric $\mathbf{\Sigma}_n$ and the mean vector $\boldsymbol{\mu}_n$, which indicates that the results are robust and justifies the theoretical findings of Bodnar and Parolya (2020) for the optimal shrinkage intensity given in (2.32). Figure 2 provides further numerical results related to the comparison of the two optimal shrinkage intensities. Here, we set $c = 1.5$ and $c = 2$ and study the robustness of the results in Figure 1 to changes in the dimension p from 50 to 300. We see that the shrinkage intensities are very similar uniformly over p, independently of the chosen value of c.

Summarizing the above findings, we can recommend the application of the Moore-Penrose approximation for $c \leq 2$, but there is no guarantee for a good performance for $c > 2$. Also, we observe that the Moore-Penrose inverse gets closer to $\mathbf{S}_n^*$ when the covariance matrix $\mathbf{\Sigma}_n$ is sparse. In the empirical study of Section 3.2 we consider the values of the concentration ratio c bounded by 2. The empirical results are in line with the discussion provided in this remark and, thus, $c \approx 2$ seems to be a breaking point for the approximation indeed. More theoretical treatment of this interesting phenomenon is of an independent research interest and is not within the scope of this paper.

Remark 2. Seemingly, we have handled two cases $c < 1$ and $c > 1$ (for $c \leq 2$), but not $c = 1$. The case $c = 1$ is not easy to manage because the sample covariance matrix is theoretically invertible for c equal or close to one but computationally very unstable. The reason is the smallest eigenvalue of $\mathbf{S}_n$ which is numerically very close to zero. Indeed, it is well-known that the smallest eigenvalue of $\mathbf{S}_n$ is of order $(1 - \sqrt{p/n})^2$, which converges to zero if $p/n \to 1$ and all the estimators explode (see, e.g., Bai and Yin (1993)).

In order to overcome the difficulty in a small neighborhood of $c = 1$ one has a few options to proceed:

Tikhonov (ridge) One of the possibilities, which has also been used in the simulation and empirical studies of Section 3, is the Tikhonov (ridge) approximation of the Moore-Penrose

inverse. Indeed, one can show by the eigenvalue decomposition that

$$\lim_{\delta \to 0^+} (\mathbf{S}_n + \delta \mathbf{I})^{-1} \mathbf{S}_n (\mathbf{S}_n + \delta \mathbf{I})^{-1} = \mathbf{S}_n^+ \tag{2.33}$$

if $c > 1$. For $c < 1$ the limit in (2.33) trivially exists and equals $\mathbf{S}_n^{-1}$. The advantage of the representation (2.33) is twofold. First, one has an elegant formula which incorporates both cases $c < 1$ and $c > 1$, and, secondly, one stabilizes the behaviour of the inverse matrix near singularity, i.e., near $c = 1$. The only question arises how to choose $\delta = o(1)$ in practice, but it seems that taking $\delta = 1/p$ works well in many applications. Thus, we will employ this adjustment in the empirical study in order to have a balanced and stable estimator when c is close to 1 from both sides. Although this procedure smoothes out the estimator of the precision matrix, it does not resolve the issue when c is large, i.e., $c > 2$.

Moore-Penrose Yet another option would be to derive the explicit limit of (2.18) when the Moore-Penrose inverse matrix $\mathbf{S}_n^+$ is directly used in (2.15)-(2.17). This procedure is highly nontrivial because $\mathbf{S}_n^+$ depends in a nonlinear way on the matrix $\mathbf{\Sigma}_n$ and this leads to nonlinear integral equations in the high-dimensional setting. The problem becomes even more involved when we consider quadratic and bilinear forms involving the Moore-Penrose inverse. Moreover, the case of centered sample covariance matrix (the sample mean is subtracted) makes the expressions tedious and confusing (see, e.g., Pan (2014)). Thus, the fact that the optimal shrinkage intensity depends on the mean vector $\boldsymbol{\mu}_n$ and the covariance matrix $\mathbf{\Sigma}_n$ only via the three parameters of the efficient frontier will be lost and no closed-form formulas can be derived for $\mathbf{\Sigma}_n \neq \sigma^2 \mathbf{I}$. Nevertheless, we are working on this problem in a separate project and are studying the properties of the pseudo-inverse $\mathbf{S}_n^+$ in detail, especially the limiting behaviour of its eigenvectors.

Double-shrinkage A more intuitive option is to apply a double-shrinkage approach, which incorporates both the shrinkage of the estimated portfolio weights and the regularization of the sample covariance matrix. Namely, first we shrink (regularize) the sample covariance matrix and then we shrink the estimated portfolio weights built upon the regularized sample covariance matrix. This could be done by taking either the matrix $\mathbf{S}_n + \delta \mathbf{I}$ for

some $\delta > 0$ or the one from (2.33) instead of $\mathbf{S}_n$. Then the limiting shrinkage intensity α will depend on δ and one needs to optimize the objective function (2.7) over δ as well. Unfortunately, no closed-form expression for the optimal shrinkage intensity is available in this case and the theoretical results presented in Section 2 must be changed accordingly. Actually, this procedure would generalize all of the above mentioned ideas (including the ones presented in this paper) and would provide a data-driven regularization parameter δ. This interesting and important topic is left for future research.

3 Simulation and empirical studies

In this section we illustrate the performance and the advantages of the derived results using simulated and real data. Particularly we address the estimation precision of the shrinkage coefficient and compare the bona-fide estimator with the existent approaches.

3.1 Simulation study

For simulation purposes we select the structure of the spectrum of the covariance matrix and of the mean vector to make it consistent with the characteristics of the empirical data. Particularly, for each dimension p we select the expected returns equally spread on the interval -0.3 to 0.3, capturing a typical spectrum of daily returns measured in percent. The covariance matrix has a strong impact on the properties of the shrinkage intensity and for this reason we consider several structures of its spectra. Replicating the properties of empirical data we generate covariance matrices with eigenvalues satisfying the equation $\lambda_i = 0.1e^{\delta c \cdot (i-1)/p}$ for $i = 1, ..., p$ (see, e.g., Bodnar et al. (2021b) for implementation). Thus the smallest eigenvalue is 0.1 and by selecting appropriate values for c we control the largest eigenvalue and thus the condition index of the covariance matrix. Large condition indices imply ill-conditioned covariance matrices, with the eigenvalues very sensitive to changes of the elements. We choose δ to attain the condition indices of 150, 1000 and 8000. The target portfolio weights are set equal to the weights of the equally weighted portfolio, i.e. $b_i = 1/p$ for $i = 1, ..., p$. The calibration criteria used to determine the optimal shrinkage intensities are selected as defined in Proposition 2.1.

First, we assess the general behavior of the oracle shrinkage intensities as functions of c.

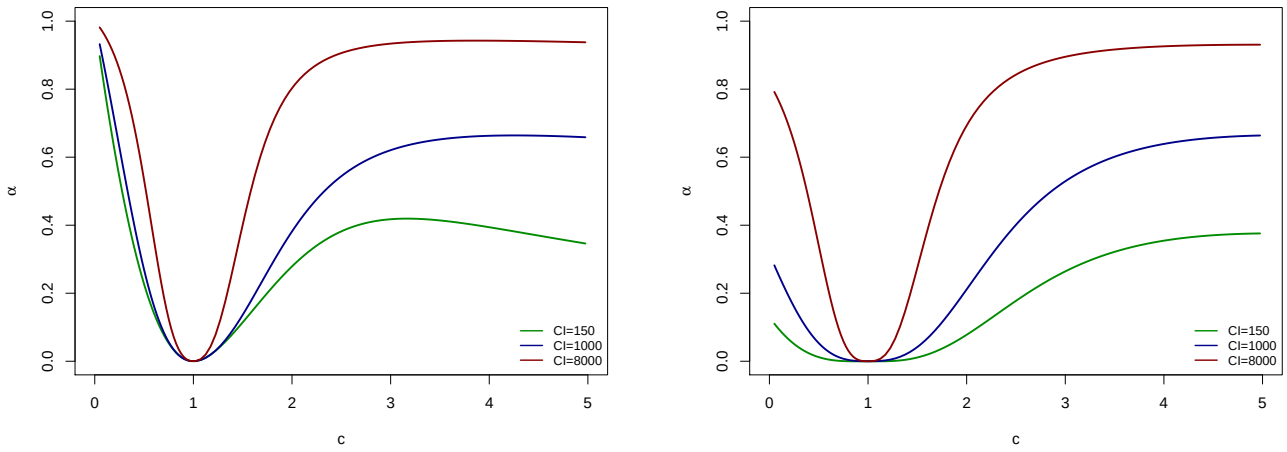

Figure 3: The asymptotic optimal shrinkage intensity as a function of c for the calibration criteria (i)-(ii) from Proposition 2.1 (left to right).

The oracle shrinkage intensities are computed using expressions in (2.11) and (2.19) for the cases $c < 1$ and $c > 1$, respectively. The parameters are computed using the true mean vector and the true covariance matrix. The results are illustrated for different condition indices and different calibration criteria in Figure 3. We observe that in all cases the shrinkage intensity falls to zero as $c \to 1_-$ and increases with c for $c > 1$. Thus if c is small the shrinkage estimator puts higher weight on the traditional estimator of the portfolio weights, due to lower estimation risk. If c tends to 1 the system becomes unstable because of nearly zero eigenvalues. In this case the portfolio weights collapse to the target portfolio weights. With c further increasing the shrinkage intensity increases too, implying that the pseudo-inverse covariance matrix can be evaluated in a proper way. The fraction of the sample EU portfolio increases with c in this case. It is worth mentioning that at some high level of c the information content in the data becomes less relevant and the shrinkage intensity starts to decrease. Note, however, that even for p much larger than n, there is still valuable information in the sample covariance matrix leading to relatively high values of α^+.

Regarding the calibration criteria we observe that if the calibration criteria coincides with the expected quadratic utility (i.e. $\beta = \gamma$), then the limit shrinkage intensities are naturally higher, compared to those minimizing the out-of-sample variance. The variance of portfolio return for the equally weighted portfolio tends to be lower than that of the sample EU portfolio. Thus the shrinkage intensity weights the equally weighted portfolio more heavily. It is important to stress that the latter calibration criterion is more sensitive to the condition index.

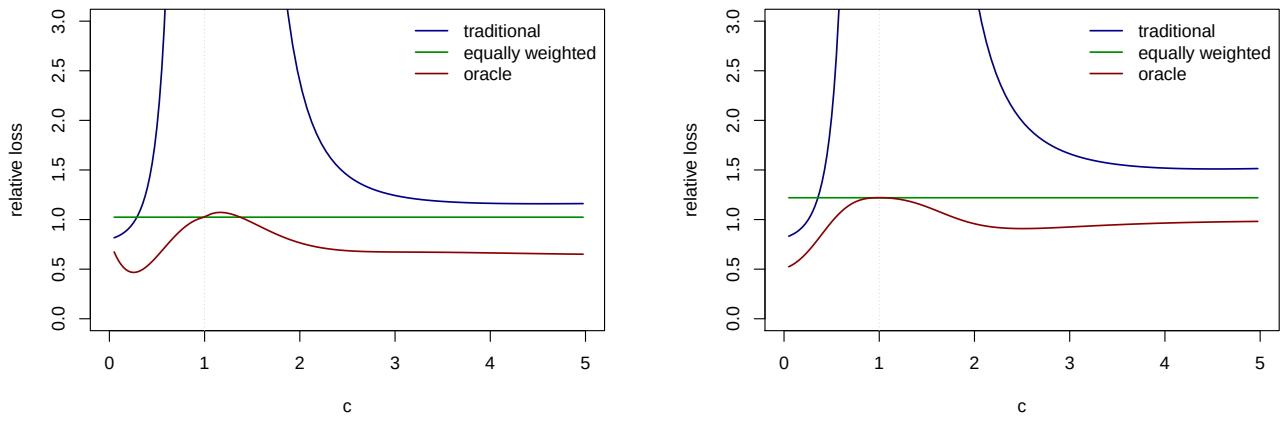

Figure 4: The relative losses for the portfolios based on the optimal shrinkage estimator, the traditional estimator and the equally weighted portfolio as a function of c for the calibration criteria (i)-(ii) from Proposition 2.1 (left to right). The dimension is set to $p = 100$ and the condition index is set to 1000.

In a similar fashion we analyze the relative losses of portfolios based on the traditional estimator and the oracle shrinkage estimator. As a benchmark, we take the equally weighted portfolio which is also the target portfolio of the shrinkage estimator. The relative losses as functions of c for fixed $p = 100$ are plotted in Figure 4. For $c < 1$ the losses of the traditional estimator show explosive behavior and are comparable to the shrinkage-based estimators only for very small values of c. Thus the traditional estimator is reliable only if the sample size is considerably larger than the portfolio dimension. The performance of the shrinkage-based estimator is relatively stable over the whole range of c and it clearly dominates both the traditional and the equally weighted benchmark in almost all of the considered cases. The losses are increasing for $c < 1$ and attain the loss of the equally weighted portfolio around $c = 1$. This is consistent with the results in Figure 3. For $c > 1$ the losses decrease and remain stable for $c > 3$.

The behavior of losses as functions of the dimension p is illustrated in Figure 5. For space reasons we provide here only the results for $\beta = \gamma$, i.e., for the first calibration criterion. The fraction c is set to 0.2 (top left), 0.5 (top right), 0.8 and 2, while the condition index equals 1000. From financial perspective it is important to note that the traditional estimator outperforms the equally weighted portfolio only for small values of c (in the particular setup for $c = 0.2$), thus when the classical estimators are stable and robust. This is consistent with the evidence from Figure 4. For $c = 0.8$ the losses of the traditional estimator increase dramatically and

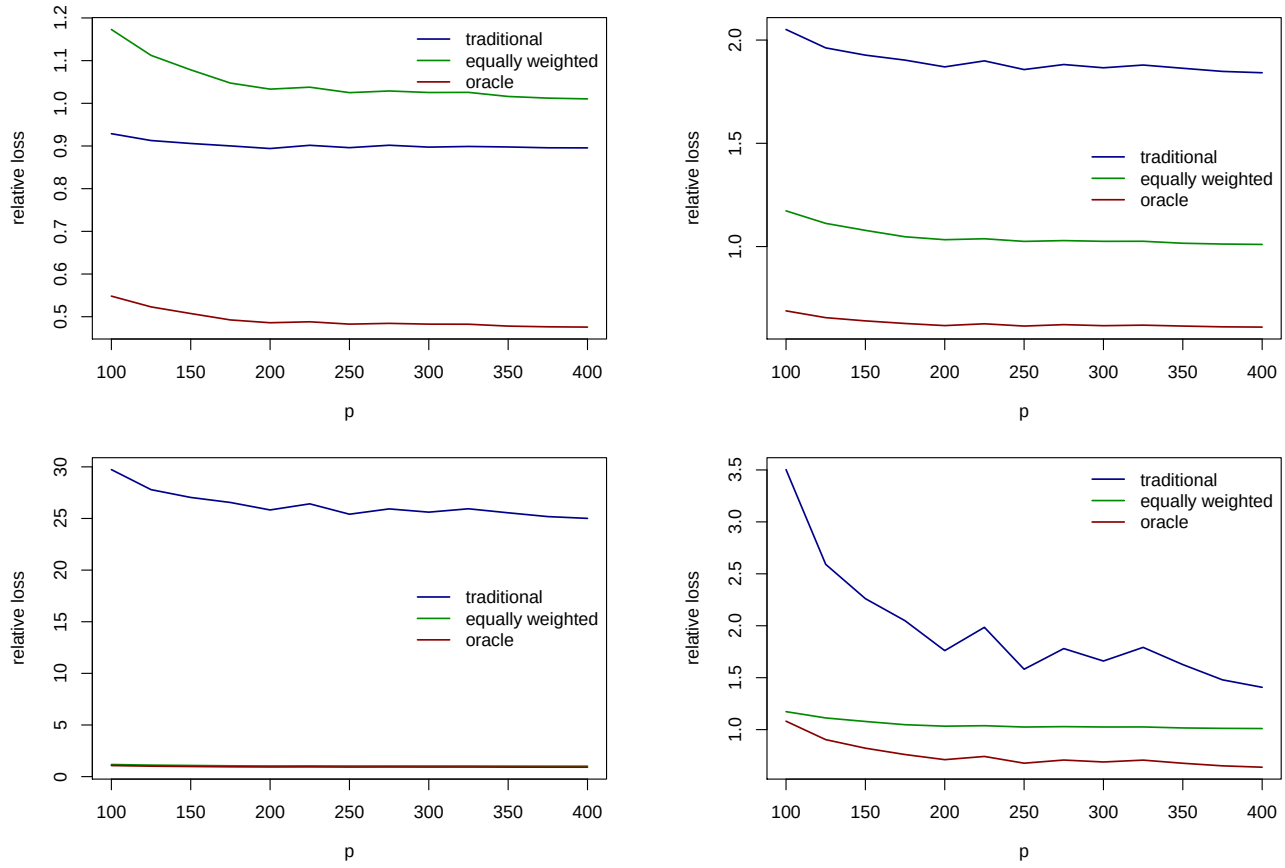

Figure 5: The relative losses for the portfolios based in the optimal shrinkage estimator, the traditional estimator and the equally weighted portfolio as a function of the dimension p for $c = 0.2$ (top left)), 0.5 (top right), 0.8 (bottom left), 2 (bottom right). The condition index is set to 1000 and the mean-variance calibration criteria is used.

they are always considerably larger than the losses of two other considered trading strategies. As before the shrinkage-based estimator clearly beats both the traditional EU portfolio and the benchmark portfolio for all c values. Furthermore, the performance of the shrinkage-based portfolio is stable for a wide range of dimensions, particularly for large values of p.

3.2 Empirical study

The data used in this study cover daily returns on 395 S&P500 constituents available for the whole period from 01.01.2000 till 23.03.2018. The investor allocates his/her wealth to the constituents with daily reallocation. We address several issues in this empirical study. First, we wish to verify the robustness of the established theoretical results for empirical data. Thus our aim is to go beyond the common practice of considering a single portfolio, but to generate a large set of different portfolios from the universe of the S&P 500 constituents. Second, we

assess the economic performance of the dynamic portfolio strategies stemming from the generalized shrinkage estimator for portfolio weights. Thus we consider several popular performance measures and test the significance in the differences between the alternative strategies. Third, the choice of the target portfolio can clearly have a substantial impact on the empirical results. For this reason, we consider several popular choices of the target. Finally, we wish to assess the dynamics of the estimated shrinkage intensities and relate their behavior to the market conditions. Next, we provide details on the setup of the empirical study.

To address the applicability of the suggested estimator in high dimensions we set $p = 300$ which is larger than a typical portfolio size in the literature. For each parameter constellation we draw randomly 1000 sets of assets from the available constituents of the S&P500 index. This guarantees a robust assessment of the empirical results. For every set of the assets we build portfolios on each of the last 1000 trading days and compute the corresponding realized returns. Afterwards, we compute the certainty equivalent (CE), Sharpe ratio (SR), Value-at-Risk (VaR) and Expected shortfall (ES) as performance measures for each path of returns and every random portfolio. To avoid potentially skewed inferences due to outliers or asymmetries we report the 10%-trimmed means and the medians of the CE and SR over the 1000 random portfolios. The VaR and ES are computed as lower empirical quantiles at 5% and 1% significance levels and are averaged over the portfolios either. For simplicity we neglect the transaction costs in the below discussion.

Target portfolios

The target portfolio weights are the key component of the shrinkage estimator. We consider three different targets: the equally weighted portfolio and two modified global minimum-variance portfolios. The equally weighted portfolio arises if we assume that all asset returns have equal expectations, equal variances and equal correlations. The covariance matrix for the first global-minimum variance portfolio assumes different variances, but equal correlations. Thus allows for more heterogeneity compared to the equally weighted portfolio. The single correlation is computed as the average correlation for all asset pairs. The corresponding target is computed by $\mathbf{b}_{ec} = \hat{\mathbf{\Sigma}}_{ec}^{-1}\mathbf{1}/\mathbf{1}'\hat{\mathbf{\Sigma}}_{ec}^{-1}\mathbf{1}$, where $\hat{\mathbf{\Sigma}}_{ec} = diag\{\hat{\mathbf{\Sigma}}\}^{1/2}\mathbf{R}_{ec}diag\{\hat{\mathbf{\Sigma}}\}^{1/2}$ with $diag\{\hat{\mathbf{\Sigma}}\}$ being the diagonal of the sample covariance matrix of returns. For the second global minimum-variance

portfolio we compute the covariance matrix of returns using the three-factor Fama-French model, i.e.

$$\hat{\mathbf{\Sigma}}_{ff} = \hat{\mathbf{B}}\hat{\mathbf{\Sigma}}_f\hat{\mathbf{B}}' + diag\{\hat{\sigma}^2_{\boldsymbol{\varepsilon}_i}\}_{i=1,\dots,p},$$

where $\hat{\mathbf{B}}$ is the matrix of estimated parameters, $\hat{\mathbf{\Sigma}}_f$ is the covariance matrix of the factors and $diag\{\hat{\sigma}^2_{\boldsymbol{\varepsilon},i}\}_{i=1,\dots,p}$ is the diagonal matrix of residual variances. The resulting target vector of weights is defined as $\mathbf{b}_{ff} = \hat{\mathbf{\Sigma}}^{-1}_{ff}\mathbf{1}/\mathbf{1}'\hat{\mathbf{\Sigma}}^{-1}_{ff}\mathbf{1}$. The latter two portfolios reduce the variation of portfolios by looking at the variance but not the mean of the underlying assets. Note, that $\mathbf{b}_{ec}$ and $\mathbf{b}_{ff}$ are stochastic by construction, since they are computed using sample characteristics of the asset returns. Thus the theoretical results in these cases are valid only conditionally on the target vector.

Benchmark models

To guarantee a fair assessment of the suggested estimator we consider two popular approaches as benchmarks. The first approach is based on the non-linear shrinkage estimator of the covariance matrix suggested by Ledoit and Wolf (2012, 2017a). The estimator relies on the spectral decomposition of the sample covariance matrix $\hat{\mathbf{\Sigma}} = \mathbf{UDU}'$, but replaces the original eigenvalues by eigenvalues $\mathbf{D}^*$ that minimize the Frobenius norm $\mathbf{D}^* = \text{argmin}_{\mathbf{D}}||\mathbf{\Sigma} - \mathbf{UDU}'||$. The solution can be approximated using a generalized version of the Marčenko-Pastur equation. First, we consider a numerical implementation of this method proposed in Ledoit and Wolf (2017b) (called *LWQuEST*). The direct numerical computation of the eigenvalues appears to be very demanding. Recently, Ledoit and Wolf (2020) suggest an analytic expression of the non-linear shrinkage estimator that uses a nonparametric estimator of the spectral density (called *LWAnalytic*). In order to determine the optimal portfolio weights we use these two approaches as a plug-in estimator of the covariance matrix. Note that in contrary to our approach, these methods shrink a parameter of the distribution of asset returns and not the portfolio weights, which are the key object of interest in asset allocation problems. The second benchmark is an extension of the three-fund portfolio of Kan and Zhou (2007). The optimal portfolio is a linear combination of the target portfolio, the sample global minimum-variance portfolio and

the portfolio that maximizes the Sharpe ratio[3]. The weights of this linear combination are determined by maximizing the expected out-of-sample utility. Note that Kan and Zhou (2007) derive the optimal portfolios assuming a finite portfolio with $p < n$. The extension of these results to the case $p/n \to c$ goes beyond the scope of this paper. For this reason we apply this estimator only for the case $p < n$. For both benchmarks we estimate the expected returns by the historical mean returns, as it is frequently done in practice.

Empirical results

The results of the empirical study are summarized in Tables 1 and 2 containing the results for mean-variance ($\beta = \gamma$) and minimum variance ($\beta = \infty$) calibration criteria, respectively. The top blocks of the table provides results for $c = 0.2$, while the following blocks correspond to c =0.5, 0.8, and 2. At the beginning of each block we summarize the performance measures for the traditional estimator and the estimator based on the nonlinear-shrinkage of Ledoit and Wolf (2012). These estimators do not depend on the target portfolio weights. Further we provide the results for the strategies involving the target, i.e. the suggested bona-fide shrinkage technique, the target portfolio itself and the extension of the estimator of Kan and Zhou (2007). This is done for equally weighted portfolio, equal correlation portfolio and Fama-French portfolio as targets. Furthermore, we include results for the traditional portfolio and for the bona-fide shrinkage portfolio, where the sample covariance matrix is replaced by its regularized version based on the Tikhonov (ridge) approximation (2.33) with $\delta = 1/p$. The corresponding strategies are called *trad ridge* and *bona fide ridge*, respectively.

First, we consider the results for the mean-variance calibration in Table 1. If the dimension is low relatively to the sample size, namely $c = 0.2$, then the traditional estimator shows a good and robust performance. According to virtually all performance measures it is better than any estimator based on equal correlation and Fama-French targets. The Ledoit-Wolf estimator is better only in terms of the Sharpe ratio. The dominant strategy for this value of c is the Kan-Zhou estimator with the equally weighted target which is closely followed by the suggested bona-fide shrinkage portfolio. The ranking of the targets and the estimators slightly changes if we increase c to 0.5. As expected the traditional estimators becomes worse and is dominated by

[3]We thank Raymond Kan for providing these results in personal communication.

		CE		SR		VaR		ES	
		average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
c = 0.2	trad	-0.84840	-0.84795	0.10033	0.10041	-0.88181	-1.43362	-1.27802	-2.01774
	trad ridge	-0.85096	-0.85028	0.10066	0.1006	-0.88193	-1.43891	-1.28130	-2.02313
	LWQuEST	-0.96539	-0.96509	**0.10310**	**0.10288**	-0.93884	-1.52905	-1.35071	-2.10094
	LWAnalytical	-0.79142	-0.79179	0.10141	0.10151	-0.8544	-1.37885	-1.24301	-1.97136
	equally weighted target								
	bona fide	-0.80701	-0.80578	0.10125	0.10137	-0.85637	-1.41141	-1.25832	-2.00392
	bona fide ridge	-0.80303	-0.80305	0.10124	0.10136	-0.85485	-1.39997	-1.25366	-1.99908
	target	-1.47973	-1.47978	0.05178	0.05183	-1.35819	-2.17993	-1.90944	-2.91308
	KZ	**-0.73868**	**-0.73831**	0.10204	0.10195	**-0.81524**	**-1.33615**	**-1.20824**	**-1.94859**
	equal correlation target								
	bona fide	-0.86149	-0.86029	0.09785	0.09803	-0.90268	-1.48278	-1.30966	-2.06512
	bona fide ridge	-0.86011	-0.85989	0.09773	0.09788	-0.90141	-1.48209	-1.30729	-2.06240
	target	-2.13972	-2.14348	0.05303	0.05291	-1.46652	-2.71956	-2.21071	-3.41674
	KZ	-0.93578	-0.93437	0.09170	0.09159	-0.94118	-1.63936	-1.40500	-2.14364
	Fama-French target								
	bona fide	-0.85595	-0.85551	0.09504	0.09433	-0.89107	-1.45683	-1.29601	-2.04557
	bona fide ridge	-0.8524	-0.85234	0.09552	0.09556	-0.88859	-1.45891	-1.29374	-2.04100
	target	-2.27656	-2.27951	0.03522	0.03529	-1.44902	-2.83263	-2.21467	-3.40873
	KZ	-0.89766	-0.89748	0.08611	0.08562	-0.89600	-1.58344	-1.34739	-2.09784
c = 0.5	trad	-1.56794	-1.56702	0.04112	0.04158	-1.25759	-1.97710	-1.74511	-2.60700
	trad ridge	-1.47446	-1.47309	0.04468	0.04461	-1.2158	-1.90656	-1.68947	-2.53159
	LWQuEST	-3.28708	-3.28730	0.02557	0.02559	-1.85906	-2.93786	-2.56534	-3.68434
	LWAnalytical	**-0.86718**	**-0.86687**	**0.06531**	**0.06548**	**-0.89736**	-1.50835	**-1.31493**	**-2.09438**
	equally weighted target								
	bona fide	-1.02898	-1.02829	0.05067	0.05098	-1.00189	-1.67912	-1.458300	-2.30347
	bona fide ridge	-0.99158	-0.9911	0.05342	0.05347	-0.98209	-1.65375	-1.43523	-2.27809
	target	-1.47854	-1.47831	0.05167	0.05170	-1.35780	-2.18133	-1.90904	-2.91312
	KZ	-0.93245	-0.93159	0.05903	0.05973	-0.95118	**-1.50578**	-1.34866	-2.10714
	equal correlation target								
	bona fide	-1.18004	-1.17898	0.05198	0.05215	-1.10187	-1.82427	-1.59237	-2.44481
	bona fide ridge	-1.15576	-1.15526	0.05312	0.05308	-1.09212	-1.81848	-1.58300	-2.43487
	target	-1.81699	-1.81663	0.04036	0.04029	-1.40816	-2.42265	-2.10357	-3.33301
	KZ	-1.01288	-1.01141	0.05760	0.05770	-1.01531	-1.57685	-1.41783	-2.14790
	Fama-French target								
	bona fide	-1.22574	-1.22594	0.04963	0.04973	-1.11090	-1.75940	-1.57373	-2.42489
	bona fide ridge	-1.19113	-1.19039	0.05083	0.05064	-1.09567	-1.73454	-1.55387	-2.40117
	target	-1.93351	-1.9336	0.04278	0.04265	-1.40284	-2.48907	-1.99845	-2.92676
	KZ	-1.01496	-1.01416	0.05761	0.05782	-1.00263	-1.55246	-1.40187	-2.16018
c = 0.8	trad	-11.10781	-11.07733	0.04031	0.03992	-3.35917	-5.40548	-4.66218	-6.70938
	trad ridge	-5.60256	-5.59981	0.04942	0.04921	-2.40138	-3.8217	-3.29866	-4.71741
	LWQuEST	-18.80892	-18.81717	0.03520	0.03521	-4.51683	-6.92530	-6.05810	-8.35046
	LWAnalytical	**-0.98453**	**-0.98438**	**0.08988**	**0.08999**	**-0.98071**	**-1.61591**	**-1.42389**	**-2.24346**
	equally weighted target								
	bona fide	-1.40577	-1.40584	0.05118	0.05134	-1.29895	-2.07527	1.83362	-2.79312
	bona fide ridge	-1.31635	-1.31596	0.05238	0.05235	-1.28390	-2.02482	-1.79171	-2.74021
	target	-1.48023	-1.47918	0.05170	0.05171	-1.35865	-2.18186	-1.91045	-2.9148
	KZ	-2.03472	-2.03185	0.06359	0.06288	-1.40796	-2.25191	-1.97686	-2.94533
	equal correlation target								
	bona fide	-1.60399	-1.60241	0.04978	0.04984	-1.29237	-2.39802	-1.98752	-3.17541
	bona fide ridge	-1.55829	-1.55768	0.05013	0.05005	-1.26800	-2.35739	-1.96495	-3.16411
	target	-1.68942	-1.68949	0.04453	0.04452	-1.3124	-2.45543	-2.05351	-3.35801
	KZ	-1.96623	-1.96299	0.06550	0.06522	-1.37667	-2.20379	-1.93809	-2.91761
	Fama-French target								
	bona fide	-1.79593	-1.79528	0.04740	0.04760	-1.39084	-2.26699	-1.94795	-2.84830
	bona fide ridge	-1.69800	-1.6972	0.04796	0.04793	-1.34454	-2.23851	-1.90270	-2.78836
	target	-1.83876	-1.83798	0.04355	0.04365	-1.37985	-2.41411	-1.98305	-2.94513
	KZ	-2.06935	-2.07129	0.06506	0.06515	-1.41329	-2.27824	-1.98427	-2.95639
c = 2.0	trad	-2.12939	-2.13080	0.06414	0.06427	-1.44683	-2.38186	-2.07522	-3.18270
	trad ridge	-2.05179	-2.05106	0.06509	0.06514	-1.42082	-2.34175	-2.03791	-3.12989
	LWQuEST	-59.03298	-58.99561	0.00629	0.00618	-7.9549	-12.68863	-10.94189	-15.31889
	LWAnalytical	-1.50503	-1.50439	0.06355	0.06375	-1.26581	-1.95316	-1.7726	-2.69539
	equally weighted target								
	bona fide	**-1.22550**	**-1.22474**	0.05904	0.05907	-1.18681	-1.99190	-1.72914	-2.75302
	bona fide ridge	-1.30872	-1.30891	0.06790	0.06782	-1.14901	-1.94775	-1.70477	-2.73827
	target	-1.47869	-1.47779	0.05175	0.05176	-1.35801	-2.17908	-1.90898	-2.91238
	equal correlation target								
	bona fide	-1.22592	-1.22761	0.07282	0.07288	**-1.09681**	-2.09756	-1.74639	-2.91502
	bona fide ridge	-1.24075	-1.23914	**0.07872**	**0.07873**	-1.10084	**-1.93814**	**-1.67112**	-2.70638
	target	-1.32850	-1.32805	0.07020	0.07020	-1.13496	-2.16850	-1.83174	-3.08545
	Fama-French target								
	bona fide	-1.44490	-1.44438	0.05890	0.05867	-1.22790	-2.09709	-1.75667	**-2.65269**
	bona fide ridge	-1.42024	-1.41913	0.06947	0.06922	-1.19613	-1.96825	-1.72476	-2.66945
	target	-1.62084	-1.62093	0.05377	0.05360	-1.29152	-2.34075	-1.86625	-2.83814

Table 1: Performance of traditional, bona-fide, the benchmark portfolios (LWQuEST - Ledoit and Wolf (2017b), LWAnalytical - Ledoit and Wolf (2020), KZ - Kan and Zhou (2007)) and the target portfolios for the *mean-variance calibration* criteria. The performance measures are averaged over 1000 random portfolios of size 300. The trading period consists of 1000 days preceding 23.03.2018 and the risk aversion is set to 5. The average values are based on trimmed mean with 10% of extreme values being dropped. The best strategies for every criteria and every values of c are highlighted in bold.

		CE		SR		VaR		ES	
		average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
c = 0.2	trad	-0.85309	-0.85231	0.10035	0.10057	-0.88250	-1.44210	-1.28383	-2.02584
	trad ridge	-0.84923	-0.84889	0.09923	0.09909	-0.88174	-1.43823	-1.27883	-2.01864
	LWQuEST	-0.96554	-0.96474	**0.10343**	**0.10359**	-0.93948	-1.52539	-1.35090	-2.10343
	LWAnalytical	-0.79114	-0.79048	0.1014	0.10161	-0.85381	-1.38047	-1.24188	-1.96952
	equally weighted target								
	bona fide	-0.81308	-0.81284	0.10123	0.10150	-0.85908	-1.41182	-1.26178	-2.0085
	bona fide ridge	-0.80998	-0.81034	0.10012	0.10001	-0.85828	-1.40687	-1.25735	-2.00243
	target	-1.47805	-1.47784	0.05169	0.05168	-1.35704	-2.17938	-1.90829	-2.91097
	KZ	**-0.73828**	**-0.73822**	**0.10187**	**0.10195**	**-0.81400**	**-1.33608**	**-1.20828**	**-1.95301**
	equal correlation target								
	bona fide	-0.85773	-0.85717	0.09806	0.09785	-0.89740	-1.47224	-1.30236	-2.05662
	bona fide ridge	-0.85262	-0.85171	0.09803	0.0981	-0.89563	-1.46898	-1.29994	-2.05272
	target	-2.13255	-2.12898	0.05319	0.05312	-1.46343	-2.71591	-2.20532	-3.40780
	KZ	-0.93731	-0.93588	0.09141	0.09120	-0.94052	-1.63821	-1.40403	-2.14399
	Fama-French target								
	bona fide	-0.85149	-0.85077	0.09663	0.09677	-0.88750	-1.45256	-1.29064	-2.03126
	bona fide ridge	-0.84949	-0.84947	0.09629	0.09625	-0.88760	-1.44700	-1.28814	-2.0294
	target	-2.27063	-2.26895	0.03538	0.03543	-1.44480	-2.82390	-2.21059	-3.40190
	KZ	-0.89663	-0.89478	0.08661	0.08659	-0.89534	-1.58062	-1.34627	-2.09320
c = 0.5	trad	-1.56971	-1.57001	0.04173	0.04161	-1.26072	-1.97665	-1.74582	-2.60683
	trad ridge	-1.47585	-1.4756	0.04465	0.0444	-1.21598	-1.90785	-1.69053	-2.54353
	LWQuEST	-3.28833	-3.29027	0.02581	0.02608	-1.86558	-2.93777	-2.56735	-3.67757
	LWAnalytical	**-0.86699**	**-0.86626**	**0.0654**	**0.06526**	**-0.89816**	-1.50838	**-1.31504**	**-2.0954**
	equally weighted target								
	bona fide	-1.0532	-1.05298	0.05066	0.05048	-1.01036	-1.69687	-1.46966	-2.30812
	bona fide ridge	-1.14637	-1.14571	0.05275	0.05322	-1.08067	-1.78718	-1.56281	-2.41015
	target	-1.48082	-1.48066	0.05171	0.05172	-1.35847	-2.18138	-1.91024	-2.91449
	KZ	-0.93323	-0.93338	0.05939	0.05941	-0.95222	**-1.50541**	-1.34938	-2.10436
	equal correlation target								
	bona fide	-1.17591	-1.17430	0.05196	0.05184	-1.09485	-1.80109	-1.57693	-2.42242
	bona fide ridge	-1.14637	-1.14571	0.05275	0.05322	-1.08067	-1.78718	-1.56281	-2.41015
	target	-1.80886	-1.80920	0.04010	0.04001	-1.40596	-2.41696	-2.09826	-3.32139
	KZ	-1.01397	-1.01394	0.05762	0.05790	-1.01524	-1.58062	-1.41661	-2.14494
	Fama-French target								
	bona fide	-1.22924	-1.22845	0.04980	0.05017	-1.11019	-1.75950	-1.56925	-2.41742
	bona fide ridge	-1.18967	-1.18982	0.05057	0.05064	-1.09012	-1.72857	-1.54528	-2.39013
	target	-1.93623	-1.93641	0.04259	0.04257	-1.40372	-2.49587	-2.00108	-2.93538
	KZ	-1.01377	-1.01328	0.05759	0.05765	-1.00204	-1.55270	-1.40064	-2.15907
c = 0.8	trad	-11.1587	-11.12847	0.03821	0.03725	-3.3777	-5.43327	-4.67808	-6.73814
	trad ridge	-5.59307	-5.5935	0.05126	0.05138	-2.39789	-3.80057	-3.28431	-4.69414
	LWQuEST	-18.84221	-18.84773	0.03577	0.03624	-4.50978	-6.96099	-6.06441	-8.39818
	LWAnalytical	**-0.98496**	**-0.98436**	**0.08942**	**0.08955**	**-0.98112**	**-1.61597**	**-1.42344**	**-2.2402**
	equally weighted target								
	bona fide	-1.43449	-1.43529	0.05084	0.0503	-1.30108	-2.08832	-1.83907	-2.79121
	bona fide ridge	-1.30591	-1.30557	0.05487	0.05487	-1.26708	-1.99403	-1.77592	-2.71366
	target	-1.47821	-1.47811	0.05168	0.05169	-1.35667	-2.18041	-1.90876	-2.91209
	KZ	-2.03415	-2.03269	0.06235	0.06256	-1.41138	-2.25474	-1.97862	-2.94248
	equal correlation target								
	bona fide	-1.60982	-1.60906	0.05073	0.05031	-1.29875	-2.39866	-1.98344	-3.15048
	bona fide ridge	-1.53174	-1.52924	0.05174	0.05171	-1.26259	-2.35322	-1.94643	-3.11905
	target	-1.68850	-1.68914	0.04481	0.04491	-1.31139	-2.45712	-2.05434	-3.36065
	KZ	-1.96089	-1.95878	0.06499	0.06450	-1.37460	-2.19955	-1.93337	-2.90941
	Fama-French target								
	bona fide	-1.82002	-1.82289	0.04739	0.04775	-1.40332	-2.25988	-1.95652	-2.86186
	bona fide ridge	-1.6811	-1.67987	0.04983	0.04971	-1.34377	-2.20183	-1.89093	-2.77691
	target	-1.83460	-1.83658	0.04329	0.0432	-1.37616	-2.40806	-1.98023	-2.93814
	KZ	-2.07064	-2.06718	0.06476	0.06399	-1.41264	-2.27641	-1.98342	-2.96439
c = 2.0	trad	-2.12727	-2.12819	0.06405	0.0641	-1.44373	-2.38377	-2.07218	-3.17711
	trad ridge	-2.06068	-2.05893	0.06518	0.06514	-1.4236	-2.34051	-2.04197	-3.13703
	LWQuEST	-58.92062	-58.87816	0.00574	0.00587	-7.96587	-12.64654	-10.9261	-15.29078
	LWAnalytical	-1.50775	-1.50743	0.064	0.06383	-1.26662	-1.9523	-1.77553	-2.70126
	equally weighted target								
	bona fide	-1.45575	-1.45568	0.05220	0.05226	-1.34997	-2.15551	-1.89422	-2.89468
	bona fide ridge	-1.2937	-1.29318	0.06757	0.06793	-1.14349	-1.94082	-1.69787	-2.73005
	target	-1.4788	-1.47899	0.05174	0.05177	-1.35821	-2.17967	-1.90906	-2.91205
	equal correlation target								
	bona fide	-1.31924	-1.32062	0.07073	0.07077	-1.13066	-2.15773	-1.82514	-3.07313
	bona fide ridge	**-1.22331**	**-1.22256**	**0.07853**	**0.07837**	**-1.09087**	**-1.93172**	**-1.66227**	-2.69818
	target	-1.32790	-1.32922	0.07046	0.07044	-1.13449	-2.16647	-1.83165	-3.08673
	Fama-French target								
	bona fide	-1.60447	-1.60370	0.05413	0.05423	-1.28581	-2.32282	-1.85587	-2.82089
	bona fide ridge	-1.41559	-1.41478	0.06915	0.06884	-1.19404	-1.97205	-1.72228	**-2.66687**
	target	-1.61993	-1.62003	0.05373	0.05383	-1.29209	-2.33933	-1.86513	-2.83576

Table 2: Performance of traditional, bona-fide, the benchmark portfolios (LWQuEST - Ledoit and Wolf (2017b), LWAnalytical - Ledoit and Wolf (2020), KZ - Kan and Zhou (2007)) and the target portfolios for the *minimum variance calibration* criteria. The performance measures are averaged over 1000 random portfolios of size 300. The trading period consists of 1000 days preceding 23.03.2018 and the risk aversion is set to 5. The average values are based on trimmed mean with 10% of extreme values being dropped. The best strategies for every criteria and every values of c are highlighted in bold.

every target-based portfolio. The leading estimator for this value of c becomes the analytical Ledoit-Wolf estimator followed by Kan-Zhou with the equally weighted portfolio as the target. Finally, we note that the application of the regularized sample covariance matrix based on the Tikhonov (ridge) approximation (2.33) leads to minor improvements in the performance of both the traditional and the bona-fide shrinkage estimators when $c = 0.2$ and $c = 0.5$.

With $c = 0.8$ we attain the ratio of dimension to sample size where the high-dimensional asymptotics becomes relevant and simpler estimators heavily suffer from estimation risk. The traditional estimator shows extremely poor performance, which is similar to that of the numerical Ledoit-Wolf estimator. Since the Kan-Zhou estimator does not take the increasing dimension into account, it becomes worse than the target portfolios and the bona-fide shrinkage portfolios. At the same time, the analytical Ledoit-Wolf estimator becomes dominant with bona-fide ridge estimator slightly behind. Large improvements are observed in the performance of the traditional estimator when the ridge regularization is employed in its construction. In contrast, the application of the ridge regularization to the bona-fide shrinkage portfolios leads to minor improvements.

Finally, if we increase c to 2 the sample covariance matrix is singular and we use the generalized inverse for the traditional estimator. Now the bona-fide estimator becomes clearly dominant, while the use of the equally weighted target and of the equally correlation target leads to similar results. As mentioned above the Kan-Zhou estimator is not feasible for $c > 1$. Also, the application of the regularized sample covariance matrix based on the Tikhonov (ridge) approximation improves the performance of the traditional estimator, while it leads to slightly worse performance of the bona-fide shrinkage estimator. To this end, we note a surprisingly poor performance of both Ledoit-Wolf estimators, which appear to be worse than most of the target portfolios. Noteworthy, the LWQuEST estimator is probably suffering from some numerical instabilities, while LWAnalytical behaves still very stable. If we switch to the minimum variance calibration criteria in Table 2, then the ranking of the estimators remains unchanged.

Summarizing, the suggested bona-fide shrinkage estimator is comparable to the analytical Ledoit-Wolf nonlinear shrinkage estimator for $c < 1$ and becomes superior starting with $c > 1$. It is dominant with respect to all performance measures. For smaller values of c the Kan-Zhou estimator outperforms the bona-fide estimator and both Ledoit-Wolf estimators, and tends to

show the best performance when it is used with the equally weighted target. For intermediate values of $c < 1$ the analytical Ledoit-Wolf estimator is dominant, while its poor performance for $c > 1$ is potentially due to the fact, that it does not take into account the estimation risk related to the sample mean vector, which is accounted for in the bona-fide shrinkage estimator and in the Kan-Zhou estimator. The numerical Ledoit-Wolf estimator seems to show huge numerical instabities in comparison to the analytic one (see, also Remark 3 below for further discussion of this finding).

The choice of the target is an important issue and relying on the results we can make general recommendations regarding its choice. For $c < 1$ the equally weighted portfolio is the best performing standalone strategy among the three alternatives. This target also leads to the shrinkage portfolio with the best overall performance. The equally correlated target is in all cases the second best choice. If $c = 2$ then the order changes and the equally correlated target becomes slightly better both as a standalone strategy and as the target for the shrinkage-based approach. Since the analysis is based on 1000 random portfolios, we can, therefore, recommend using the equally weighted target for $c < 1$ and equally correlated target for $c > 1$. Furthermore, we can conclude that taking the best standalone target strategy shall lead to the best performing shrinkage-based approach.

The time series of estimated shrinkage coefficients are depicted in Figure 6. For space reasons we provide the coefficients only for the equally weighted target and the mean-variance calibration. For other parameter constellations the results are similar. The portfolio is constructed using the first alphabetically sorted assets. We observe that for small values of c and thus a low estimation risk the shrinkage intensities are close to one. The behavior is very stable, but mimics the periods of high and low volatility of financial markets. Thus high volatility on financial markets causes higher shrinkage coefficients and a larger fraction of the sample EU portfolio. This can be justified by stronger effects of diversification during turmoil periods. With larger c the confidence in the classical portfolio diminishes leading to a stronger preference for the equally weighted portfolio. This results in lower and more volatile shrinkage intensities. However, we observe the reverse behaviour of the estimated shrinkage intensity when $c = 2$. Here, the impact of the traditional estimator in the portfolio structure increases and becomes comparable to the case of $c = 0.5$. Such results are in line with our findings of the simulation study

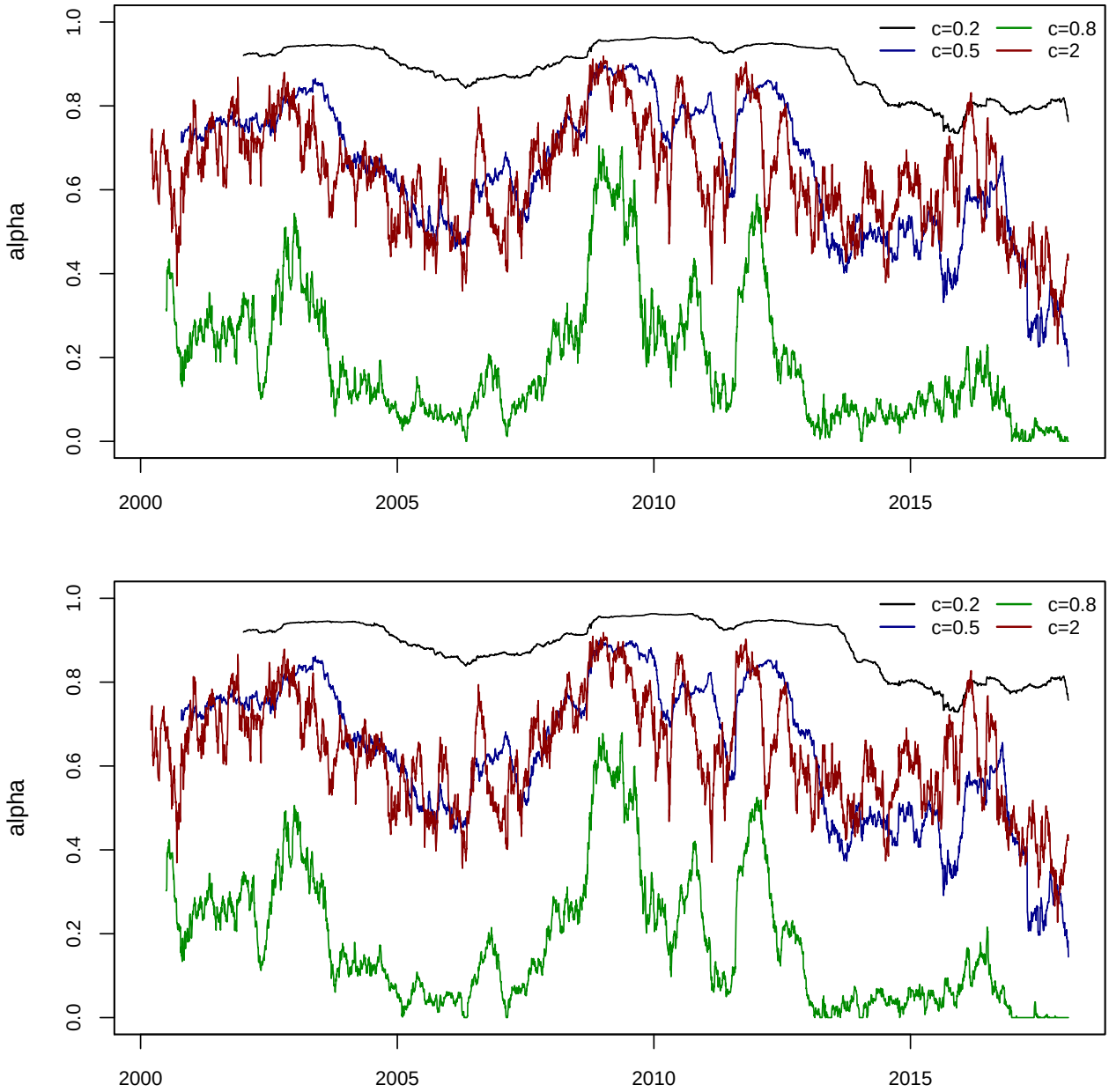

Figure 6: The bona-fide shrinkage intensities for the first 100 assets (alphabetic order) using the equally weighted target portfolio and the mean-variance calibration for $c = 0.2$, 0.5, 0.8 and 2. Above - bona fide, below - bona fide ridge (see formula (2.33)).

presented in Figure 3, where the shrinkage intensity is close to zero around $c = 1$. Finally, the results obtained by employing the regularized sample covariance matrix based on the Tikhonov (ridge) approximation leads to similar values of the shrinkage intensities independently of c. These are shown in the lower plot in Figure 6. This finding is in line with the values presented in Tables 1 and 2.

Remark 3. The numerical Ledoit-Wolf estimator shows in our empirical study surprisingly

poor performance, which is probably because of some numerical issues. That is why we recommend to use its new analytic version. Nevertheless, the analytical nonlinear shrinkage Ledoit-Wolf estimator still shows in our empirical study a poor performance[4] in case $c > 1$, but we believe there is a specific reason for that. Indeed, in Ledoit and Wolf (2017a, Assumption 5) the authors assume that the sample mean vector is independently distributed of the sample covariance matrix and its distribution is rotation invariant. This assumption appears to be a characteristic property of multivariate normal distribution following Lukacs (1979). Moreover, the assumption that the distribution of the sample mean vector is rotation invariant, imposes further restrictions on the data-generating model. It requires the population mean vector to be a zero vector and the population covariance matrix to be proportional to the identity matrix. As such, it is not clear whether the Ledoit-Wolf estimator is optimal in the case of a non-zero population mean vector, i.e., in the mean-variance framework. This is also justified by the authors themselves in Ledoit and Wolf (2017a, see Remark 5). It seems that this "sample mean" effect becomes more prominent in the case of the singular sample covariance matrix ($c > 1$) and/or small risk aversion coefficient γ, i.e. when optimal portfolios lie further away from the global minimum variance portfolio on the efficient frontier. Thus, the Ledoit-Wolf estimator should be adjusted to this type of optimal portfolios before it can be efficiently used in practice when portfolio dimension is larger than the sample size, whereas the suggested bona-fide estimator for the optimal portfolio weights incorporates both the high-dimensional effects from the sample covariance matrix and the sample mean vector simultaneously.

4 Summary

In this paper we consider the portfolio selection in the high-dimensional framework. Particularly, we assume that the number of assets p and the sample size n tend to infinity such that their ratio p/n tends to constant c where c can also be larger than one, implying that we have more assets than observations. Because of the large estimation risk we suggest a shrinkage-

[4]For $c > 1$, the analytical Ledoit-Wolf estimator was initially also very unstable because the $(p - n + 1)$th smallest eigenvalue was too close to zero (of order 10^{-12}). We have corrected this issue by treating it as "zero" and replacing it by a specific constant. To the rest of eigenvalues the optimal nonlinear shrinkage formula was applied (see, formulas (C.4) and (C.5) in Ledoit and Wolf (2020, Supplement)). The numerical implementation of Ledoit-Wolf estimator is provided in R-package *HDShOP* (see, Bodnar et al. (2021b)).

based estimator of the portfolio weights, which shrinks the mean-variance portfolio to several target portfolios, such as the equally weighted portfolio, minimum-variance portfolio, etc. For the established shrinkage intensity we derive the limiting value which depends on c and on the characteristics of the efficient frontier only. On the other side, the derived limiting expression of the shrinkage uncertainty is only an oracle value and is not feasible in practice, since it depends on unknown quantities. In order to overcome the problem, we construct a bona-fide shrinkage estimator of the optimal portfolio weights by deducing consistent estimators of the parameters of the efficient frontier under the high-dimensional setting. As a result, a fully data-driving approach is established for constructing a practically feasible estimator for the weights of the optimal mean-variance portfolios. From the technical point of view, we rely on random matrix theory and work with the asymptotic behavior of linear and quadratic forms in the sample mean vector and in the (pseudo)-inverse sample covariance matrix. In extensive simulation and empirical studies, we evaluate the performance of established results with artificial and real data. Only if the sample size is much larger than the portfolio dimension, the traditional portfolio or the benchmark portfolio dominates the portfolio suggested in the paper.

Acknowledgment

The authors thank Professor Christian Hansen and two anonymous Reviewers for their comments and suggestions which have improved the presentation of the paper. We also thank a Reviewer for the comment which prompted us to reconsider the proofs of the theorems and corollaries and, as a result, allowed us to simplify considerably the conditions needed for their validity. We gratefully acknowledge the comments from the participants at the conference "Modern Stochastics: Theory and Applications" (Kyiv, Ukraine), the 9th Annual SoFiE Conference 2016 (Hong Kong), the Meeting of the German Statistical Society "Statistical Week 2015" (Hamburg, Germany) and the 7th International Conference 2014 of the ERCIM WG on Computational and Methodological Statistics (Pisa, Italy). The authors are also thankful to Prof. Raymond Kan and Prof. Michael Wolf for fruitful discussion. Taras Bodnar was partially supported by the Swedish Research Council (VR) via the project *Bayesian Analysis of Optimal Portfolios and Their Risk Measures*.

References

Bai, J. and Shi, S. (2011). Estimating high dimensional covariance matrices and its applications. Annals of Economics and Finance, 12:199–215.

Bai, Z. and Zhou, W. (2008). Large sample covariance matrices without independence structures in columns. Statistica Sinica, 18(2):425–442.

Bai, Z. D. and Silverstein, J. W. (2010). Spectral Analysis of Large Dimensional Random Matrices. Springer, New York.

Bai, Z. D. and Yin, Y. Q. (1993). Limit of the Smallest Eigenvalue of a Large Dimensional Sample Covariance Matrix. The Annals of Probability, 21(3):1275 – 1294.

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

Bodnar, O. and Bodnar, T. (2010). On the unbiased estimator of the efficient frontier. International Journal of Theoretical and Applied Finance, 13(07):1065–1073.

Bodnar, T., Dette, H., Parolya, N., and Thorsén, E. (2021a). Sampling distributions of optimal portfolio weights and characteristics in low and large dimensions. Random Matrices: Theory and Applications, page 2250008.

Bodnar, T., Dmytriv, S., Okhrin, Y., Otryakhin, D., and Parolya, N. (2021b). HDShOP: High-Dimensional Shrinkage Optimal Portfolios. R package version 0.1.2.

Bodnar, T., Dmytriv, S., Okhrin, Y., Parolya, N., and Schmid, W. (2021c). Statistical inference for the expected utility portfolio in high dimensions. IEEE Transactions on Signal Processing, 69:1–14.

Bodnar, T., Mazur, S., and Podgórski, K. (2016). Singular inverse Wishart distribution with application to portfolio theory. Journal of Multivariate Analysis, 143:314–326.

Bodnar, T. and Parolya, N. (2020). Spectral analysis of large reflexive generalized inverse and moore-penrose inverse matrices. In Holgersson, T. and Singull, M., editors, Recent Developments in Multivariate and Random Matrix Analysis, pages 1–16. In: Holgersson T., Singull M. (eds) Recent Developments in Multivariate and Random Matrix Analysis. Springer, Cham.

Bodnar, T., Parolya, N., and Schmid, W. (2013). On the equivalence of quadratic optimization problems commonly used in portfolio theory. European Journal of Operational Research, 229:637–

644.

Bodnar, T., Parolya, N., and Schmid, W. (2018). Estimation of the global minimum variance portfolio in high dimensions. European Journal of Operational Research, 266:371–390.

Bodnar, T. and Schmid, W. (2009). Econometrical analysis of the sample efficient frontier. European Journal of Finance, 15:317–335.

Bodnar, T. and Schmid, W. (2011). On the exact distribution of the estimated expected utility portfolio weights: Theory and applications. Statistics & Risk Modeling, 28:319–342.

Chopra, V. K. and Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. The Review of Financial Studies, 19:6–11.

DeMiguel, V., L., G., and U., R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? The Review of Financial Studies, 22:1915–1953.

El Karoui, N. (2010). High-dimensionality effects in the Markowitz problem and other quadratic programs with linear constraints: Risk underestimation. Annals of Statistics, 38:3487–3566.

Frahm, G. and Memmel, C. (2010). Dominating estimators for minimum-variance portfolios. Journal of Econometrics, 159:289–302.

Friesen, O., Löwe, M., and Stolz, M. (2013). Gaussian fluctuations for sample covariance matrices with dependent data. Journal of Multivariate Analysis, 114:270–287.

Golosnoy, V. and Okhrin, Y. (2007). Multivariate shrinkage for optimal portfolio weights. European Journal of Finance, 13:441–458.

Hui, J. and Pan, G. (2010). Limiting spectral distribution for large sample covariance matrices with m-dependent elements. Communications in Statistics - Theory and Methods, 39(6):935–941.

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. The Journal of Financial and Quantitative Analysis, 21:279–292.

Kan, R. and Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. The Journal of Financial and Quantitative Analysis, 42:621–656.

Le Cam, L. and Lo Yang, G. (2000). Asymptotics in statistics: some basic concepts. Springer, New York.

Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. Journal of Multivariate Analysis, 88:365–411.

Ledoit, O. and Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. Annals of Statistics, 40:1024–1060.

Ledoit, O. and Wolf, M. (2017a). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. The Review of Financial Studies, 30(12):4349–4388.

Ledoit, O. and Wolf, M. (2017b). Numerical implementation of the quest function. Computational Statistics & Data Analysis, 115:199–223.

Ledoit, O. and Wolf, M. (2020). Analytical nonlinear shrinkage of large-dimensional covariance matrices. The Annals of Statistics, 48(5):3043 – 3065.

Lukacs, E. (1979). Some multivariate statistical characterization theorems. Journal of Multivariate Analysis, 9(2):278–287.

Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. Sbornik: Mathematics, 1:457–483.

Markowitz, H. (1952). Portfolio selection. The Journal of Finance, 7:77–91.

Merton, C. R. (1980). On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics, 8:323–361.

Merton, R. C. (1972). An analytic derivation of the efficient portfolio frontier. The Journal of Financial and Quantitive Analysis, 7(4):1851–1872.

Okhrin, Y. and Schmid, W. (2006). Distributional properties of portfolio weights. Journal of Econometrics, 134:235–256.

Pan, G. M. (2014). Comparison between two types of large sample covariance matrices. Ann. Inst. H. Poincaré Probab. Statist., 50:655–677.

Pratt, J. W. (1964). Risk aversion in the small and in the large. Econometrica, 32:122–136.

Rubio, F. and Mestre, X. (2011). Spectral convergence for a general class of random matrices. Statistics & Probability Letters, 81:592–602.

Rubio, F., Mestre, X., and Palomar, D. P. (2012). Performance analysis and optimal selection of large minimum variance portfolios under estimation risk. IEEE Journal of Selected Topics in Signal Processing, 6(4):337–350.

Silverstein, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. Journal of Multivariate Analysis, 55:331–339.

Tobin, J. (1958). Liquidity preference as behavior towards risk. Review of Economic Studies, 25:65–86.

Van Trees, H. L. (2002). Optimum Array Processing. New York: Wiley.

Verdú, S. (1998). Multiuser Detection. New York: Cambridge Univ. Press.

Wei, M., Yang, G., and Yang, L. (2016). The limiting spectral distribution for large sample covariance matrices with unbounded m-dependent entries. Communications in Statistics - Theory and Methods, 45(22):6651–6662.