SPECTRALLY-CORRECTED ESTIMATION FOR HIGH-DIMENSIONAL MARKOWITZ MEAN-VARIANCE OPTIMIZATION[1]

ZHIDONG BAI[a], HUA LI[b], MICHAEL MCALEER[cdef] AND WING-KEUNG WONG[gh]

EI2016-20

Abstract

This paper considers the portfolio problem for high dimensional data when the dimension and size are both large. We analyze the traditional Markowitz mean-variance (MV) portfolio by large dimension matrix theory, and find the spectral distribution of the sample covariance is the main factor to make the expected return of the traditional MV portfolio overestimate the theoretical MV portfolio. A correction is suggested to the spectral construction of the sample covariances to be the sample spectrally-corrected covariance, and to improve the traditional MV portfolio to be spectrally corrected. In the expressions of the expected return and risk on the MV portfolio, the population covariance matrix is always a quadratic form, which will direct MV portfolio estimation. We provide the limiting behavior of the quadratic form with the sample spectrally-corrected covariance matrix, and explain the superior performance to the sample covariance as the dimension increases to infinity proportionally with the sample size. Moreover, this paper deduces the limiting behavior of the expected return and risk on the spectrally-corrected MV portfolio, and illustrates the superior properties of the spectrally-corrected MV portfolio. In simulations, we compare the spectrally-corrected estimates with the traditional and bootstrap-corrected estimates, and show the performance of the spectrally-corrected estimates are the best in portfolio returns and portfolio risk. We also compare the performance of the new proposed estimation with different optimal portfolio estimates for real data from S&P 500. The empirical findings are consistent with the theory developed in the paper.

KEYWORDS: Markowitz Mean-Variance Optimization, Optimal Return, Optimal Portfolio Allocation,

[1] The authors are grateful to Raymond Kan, Harry Markowitz, Michael Wolf, and participants at the 2013 Asian Meeting of the Econometric Society for their valuable comments that have significantly improved the paper. The fourth author would like to thank Robert B. Miller and Howard E. Thompson for their continuous guidance and encouragement. This research is partially supported by Northeast Normal University, National University of Singapore, Chang Chun University, Hong Kong Baptist University, Research Grants Council (RGC) of Hong Kong, Australian Research Council, and Natural Science Council, Taiwan.

[a]KLASMOE and School of Mathematics and Statistics, Northeast Normal University, China.

[b]School of Sciences, Chang Chun University, China.

[c]Department of Quantitative Finance, National Tsing Hua University, Hsinchu, Taiwan.

[d]Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam.

[e]Tinbergen Institute, The Netherlands.

[f]Department of Quantitative Economics, Complutense University of Madrid, Spain.

[g]Department of Economics, Hong Kong Baptist University, China.

[h]Research Grants Council of Hong Kong, Hong Kong.

Large Random Matrix, Bootstrap Method, Spectrally-corrected Covariance Matrix.
JEL classification: G11, C13, C61.

.

1. INTRODUCTION

Mean-Variance (MV) portfolio optimization has been one of the most important topics in finance since Markowitz (1952) developed the theory. It provides a powerful tool for investors to allocate their wealth, incorporating their preferences according to their expectations of returns and risks. According to the theory, portfolio optimizers respond to the uncertainty of an investment by selecting portfolios that maximize profit, subject to achieving a specified level of calculated risk or, equivalently, minimize variance subject to obtaining a predetermined level of expected gain (see Markowitz (1952, 1959, 1991); Kroll et al. (1984)).

More precisely, we assume that there are p branches of assets with random returns $\mathbf{r} = (r_1, \cdots, r_p)^T$, having expectation $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_p)^T$ and covariance matrix $\Sigma = (\sigma_{ij})$. For any investable capital, C, and investment plan, $\mathbf{c} = (c_1, ..., c_p)^T$, satisfying $\sum_{i=1}^p c_i = C$, the anticipated return is a random variable, $\mathbf{c}^T\mathbf{r}$, with expectation, $\mathbf{c}^T\boldsymbol{\mu}$, and variance or risk, $\mathbf{c}^T\Sigma\mathbf{c}$. For convenience, we also call $\mathbf{c} = (c_1, ..., c_p)^T$ a portfolio. Without loss of generality, we assume $C \leq 1$, in which the strict inequality infers that portfolio optimizers invest their wealth only partially. We further assume that short selling is allowed; that is, any component of $\mathbf{c}$ could be negative.

In this model, the MV portfolio optimization problem can be reformulated as:

$$\max \mathbf{c}^T\boldsymbol{\mu}, \text{ subject to } \mathbf{c}^T\mathbf{1} \leq 1 \text{ and } \mathbf{c}^T\Sigma\mathbf{c} \leq \sigma_0^2, \tag{1.1}$$

in which $\mathbf{1}$ represents the vector of ones, and σ_0^2 is a given level of risk. We call $R = \max \mathbf{c}^T\boldsymbol{\mu}$ satisfying (1.1) the **optimal expected (OE) return**, and the solution $\mathbf{c}$ to the maximization the **optimal allocation (OA) plan**. Bai et al. (2009) extend the separation theorem (Cass and Stiglitz (1970)) and the mutual fund theorem (Merton (1972)) to obtain the analytical solution of equation (1.1), as shown in the following proposition:[1]

[1]In the expression of $\mathbf{c}$, $\frac{\sigma_0\Sigma^{-1}\boldsymbol{\mu}}{\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}}$ is the solution of (1.1) only with one restriction $\mathbf{c}^T\Sigma\mathbf{c} \leq \sigma_0$, if it satisfies $\mathbf{c}^T\mathbf{1} \leq 1$, that is, $\frac{\sigma_0\mathbf{1}^T\Sigma^{-1}\boldsymbol{\mu}}{\sqrt{\boldsymbol{\mu}\Sigma^{-1}\boldsymbol{\mu}}} \leq 1$. This is the OA plan. Otherwise, $\mathbf{c}(\boldsymbol{\mu}, \Sigma) = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}} + b(\boldsymbol{\mu}, \Sigma)\left(\Sigma^{-1}\boldsymbol{\mu} - \frac{\mathbf{1}^T\Sigma^{-1}\boldsymbol{\mu}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}\Sigma^{-1}\mathbf{1}\right)$. See Bai et al. (2009) and the references therein for further information.

Proposition 1.1 *For the optimization problem shown in (1.1), the optimal allocation and the corresponding expected return are:*

$$
(1.2) \qquad \mathbf{c} := \mathbf{c}(\boldsymbol{\mu}, \Sigma) = \begin{cases} \frac{\sigma_0 \Sigma^{-1}\boldsymbol{\mu}}{\sqrt{\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}}} & \text{if} \quad \frac{\sigma_0 \mathbf{1}^T\Sigma^{-1}\boldsymbol{\mu}}{\sqrt{\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}}} \leq 1, \\ \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}} + b(\boldsymbol{\mu}, \Sigma)\left(\Sigma^{-1}\boldsymbol{\mu} - \frac{\mathbf{1}^T\Sigma^{-1}\boldsymbol{\mu}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}\Sigma^{-1}\mathbf{1}\right) & \text{if} \quad \frac{\sigma_0 \mathbf{1}^T\Sigma^{-1}\boldsymbol{\mu}}{\sqrt{\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}}} > 1, \end{cases}
$$

and

$$
(1.3) \qquad R := R(\boldsymbol{\mu}, \Sigma) = (\mathbf{c}(\boldsymbol{\mu}, \Sigma))' \ \boldsymbol{\mu} \ ,
$$

respectively, in which:

$$
b(\boldsymbol{\mu}, \Sigma) = \sqrt{\frac{\mathbf{1}^T\Sigma^{-1}\mathbf{1}\sigma_0^2 - 1}{\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}\mathbf{1}^T\Sigma^{-1}\mathbf{1} - (\mathbf{1}^T\Sigma^{-1}\boldsymbol{\mu})^2}} \ .
$$

Proposition 1.1 provides investors with the best OA plan with the corresponding OE return, and also an excellent solution to Markowitz's MV optimization procedure. However, in real-life applications, practitioners have to estimate both an unknown expectation, $\boldsymbol{\mu}$, and an unknown covariance matrix, Σ. Nevertheless, in classical estimation, the sample mean and sample covariance are not consistent estimates of their counterpart parameters in the Markowitz optimization problem. In the past five decades, there have been over 300 papers written on the estimation of $\boldsymbol{\mu}$, as mentioned in the report of Green et al. (2013), with many possible estimates of $\boldsymbol{\mu}$.

Nevertheless, a difficult task is how to provide accurate estimates of the population covariance matrix to be used in the expression of the OA plan (1.2) that will lead to a more accurate estimate of the MV optimal return. In contrast, there have been few papers written on how to estimate the covariance matrix accurately.

It is well known that the sample covariance matrix is not a good choice as the estimator of the population covariance matrix in the MV optimization. This is because the sample covariance matrix tends to be far from its population counterpart when the dimension of the sample plays an important role compared with the sample size. When the dimension of the sample and the sample size increase to infinity proportionally, it is well known that: (1) the spectral distribution of the sample covariance matrix follows the MP-Law when the population covariance is a unit matrix (see Marcenko and Pastur (1967)); or (2) follows a nonrandom distribution with the form of several implicit functions for the common population covariance when the population covariance satisfies some regularity conditions (Siverstein (1995)).

This finding gives inspiration to explore further information for the population covariance matrix, including the spectral structure (see, for example, El Karoui (2008), Rao et al. (2008), Mestre (2008), Bai et al. (2010), Li et al. (2013), Li and Yao (2013) among others), and the eigenvector matrix (Bai et al. (2007), Siverstein (1990, 1989, 1984)) among others, when both the dimension and the size of the sample are large.

In this paper, we apply the spectral theory of the population covariance to correct the spectrum of the sample covariance matrix that enables further development of the spectrally-corrected (SC) estimates for the MV portfolio optimization. We first develop some limiting properties for the SC estimates for both return and risk in the MV portfolio optimization. Thereafter, we compare the SC estimates with the corresponding traditional plug-in (PI) and bootstrap-corrected (BC) estimates (see Bai et al. (2009) and the references therein for further information).

There are many proposals to improve the population covariance matrix estimation, which can be divided into two schools. The first suggests building on the additional knowledge in the estimation process, such as sparseness, graph model or factor model (see Bickel and Levina (2008), Rohde and Tsybakov (2011), Cai et al. (2012), Ravikumar et al. (2008), Rajaratnam et al. (2008), Khare and Rajaratnam (2011), Fan et al. (2008), among others). The second recommends correcting the spectrum of the sample covariance, such as the optimal linear shrinkage estimator in Ledoit and Wolf (2004) and the nonlinear shrinkage estimator in Ledoit and Wolf (2012). The SC estimates given in this paper belong to the second school. We improve estimation about the quadratic form associated with the population covariance matrix and its inverse. The details are given in the following sections.

The organization of this paper is as follows: In Section 2, we discuss the Markowitz MV optimization enigma, and develop some properties for the limiting behavior of the classical Markowitz optimal portfolio estimator. In Section 3, BC estimation has been designed to solve the protfolio estimator but its performance in risk is even worse than the classical Markowitz optimal portfolio, that is the PI portfolio. In Sections 4 and 5, we introduce the SC method and derive properties for the limiting behavior of the SC optimal portfolio estimator. Simulation studies and empirical illustrations are provided in Sections 6 and 7. Section 8 gives some concluding remarks.

2. PLUG-IN ESTIMATION AND MARKOWITZ MEAN-VARIANCE OPTIMIZATION ENIGMA

We denote, $\hat{\boldsymbol{\mu}}$, and, $\hat{\Sigma}$, as the estimates of the population mean, $\boldsymbol{\mu}$, and covariance matrix, Σ (PCOV), respectively, for the random return vector $\mathbf{r}$. Substitution of $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$ in (1.2) gives the OA estimate and the corresponding random portfolio return as:

(2.1) $\quad \hat{\mathbf{c}} = \mathbf{c}(\hat{\boldsymbol{\mu}}, \hat{\Sigma}) \quad \text{and} \quad r_{\mathbf{c}} = \hat{\mathbf{c}}'\mathbf{r}.$

Then, for the expectation, $R_{\hat{\mathbf{c}}} = \hat{\mathbf{c}}\boldsymbol{\mu}$, and the risk (or variance), $\sigma^2_{\hat{c}} = \hat{\mathbf{c}}'\Sigma\hat{\mathbf{c}}$, we have following proposition:

PROPOSITION 2.1 *For the optimization problem shown in (1.1) and given $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$, the expectation, $R_{\hat{\mathbf{c}}}$, and risk, $\sigma^2_{\hat{\mathbf{c}}}$, of the random portfolio return, $r_{\hat{\mathbf{c}}}$, respectively are:*

$$(2.2) \qquad R_{\hat{\mathbf{c}}} = \begin{cases} \dfrac{\sigma_0\, \boldsymbol{\mu}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}{\sqrt{\hat{\boldsymbol{\mu}}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}} & \text{if} \quad \dfrac{\sigma_0\mathbf{1}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}{\sqrt{\hat{\boldsymbol{\mu}}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}} \le 1, \\ \dfrac{\boldsymbol{\mu}'\hat{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\hat{\Sigma}^{-1}\mathbf{1}} + b_{\hat{\mathbf{c}}}\left(\boldsymbol{\mu}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}} - a_{\hat{\mathbf{c}}}\,\boldsymbol{\mu}'\hat{\Sigma}^{-1}\mathbf{1}\right) & \text{if} \quad \dfrac{\sigma_0\mathbf{1}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}{\sqrt{\hat{\boldsymbol{\mu}}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}} > 1, \end{cases}$$

and

$$(2.3) \qquad \sigma^2_{\hat{\mathbf{c}}} = \begin{cases} \dfrac{\sigma_0^2\, e_{\mu,\mu}}{\hat{\boldsymbol{\mu}}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}} & \text{if} \quad \dfrac{\sigma_0\mathbf{1}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}{\sqrt{\hat{\boldsymbol{\mu}}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}} \le 1, \\ \dfrac{e_{1,1}}{(\mathbf{1}'\hat{\Sigma}^{-1}\mathbf{1})^2} + b^2_{\hat{\mathbf{c}}}(e_{\mu,\mu} - 2a_{\hat{\mathbf{c}}}e_{1,\mu} + a^2_{\hat{\mathbf{c}}}e_{1,1}) & \text{if} \quad \dfrac{\sigma_0\mathbf{1}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}{\sqrt{\hat{\boldsymbol{\mu}}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}} > 1, \end{cases}$$

in which $a_{\hat{\mathbf{c}}} = \frac{\mathbf{1}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}{\mathbf{1}'\hat{\Sigma}^{-1}\mathbf{1}}$, $b_{\hat{\mathbf{c}}} = \sqrt{\frac{\mathbf{1}'\hat{\Sigma}^{-1}\mathbf{1}\sigma_0^2-1}{\hat{\boldsymbol{\mu}}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}\mathbf{1}'\hat{\Sigma}^{-1}\mathbf{1}-(\mathbf{1}'\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}})^2}}$, $e_{\mu,\mu} = \hat{\boldsymbol{\mu}}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}$, $e_{1,1} = \mathbf{1}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mathbf{1}$, and $e_{1,\mu} = \mathbf{1}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\hat{\boldsymbol{\mu}}$.

From Proposition 2.1, $R_{\hat{\mathbf{c}}}$ is a function of the quadratic form $\mathbf{a}'\hat{\Sigma}^{-1}\mathbf{b}$, and $\sigma^2_{\hat{\mathbf{c}}}$ is a function of $\mathbf{a}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mathbf{b}$, in which $\mathbf{a}$ and $\mathbf{b}$ could be $\boldsymbol{\mu}$, $\mathbf{1}$, or $\hat{\boldsymbol{\mu}}$. In order to obtain improved estimates for the return and risk, we intend to obtain improved estimates for both $\mathbf{a}'\hat{\Sigma}^{-1}\mathbf{b}$ and $\mathbf{a}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mathbf{b}$. For purposes of obtaining a superior estimate $\hat{\Sigma}$ of Σ, so that it will provide improved performance in both (2.2) and (2.3), we develop properties for $\mathbf{a}'\hat{\Sigma}^{-1}\mathbf{b}$ and $\mathbf{a}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mathbf{b}$. Estimation of a quadratic form, $\mathbf{a}\hat{\Sigma}^{-1}\mathbf{b}$, closer to its population counterpart is more important than making $\hat{\Sigma}^{-1}$ closer to Σ^{-1} in the Euclidean distance. For simplicity, let :

(2.4) $\quad \phi_{\mathbf{a}}^{\mathbf{b}}(\mathbf{A}) = \mathbf{a}'\mathbf{A}^{-1}\mathbf{b} \quad \text{and} \quad \varphi_{\mathbf{a}}^{\mathbf{b}}(\mathbf{A}) = \mathbf{a}'\mathbf{A}^{-1}\Sigma\mathbf{A}^{-1}\mathbf{b},$

for $A = \Sigma$ or any estimate $\hat{\Sigma}$. For an estimate $\hat{\Sigma}$ of Σ, $\phi_{\mathbf{a}}^{\mathbf{b}}(\hat{\Sigma})$ is an accurate estimate of $\phi_{\mathbf{a}}^{\mathbf{b}}(\Sigma)$ and $\varphi_{\mathbf{a}}^{\mathbf{b}}(\hat{\Sigma})$ is an accurate estimate of $\varphi_{\mathbf{a}}^{\mathbf{b}}(\Sigma)$ if $\phi_{\mathbf{a}}^{\mathbf{b}}(\hat{\Sigma})$ is close to $\phi_{\mathbf{a}}^{\mathbf{b}}(\Sigma)$ and $\varphi_{\mathbf{a}}^{\mathbf{b}}(\hat{\Sigma})$ is close to $\varphi_{\mathbf{a}}^{\mathbf{b}}(\Sigma)$ for any large sample size n.

2.1. *The limiting behavior of the sample covariance matrix*

It is standard practice to use the sample covariance matrix in PCOV estimation. This practice is useful if the effect of the dimension of the sample is neglectable when compared with the sample size since, in the classical limit theory, the sample covariance matrix is a consistent estimator of the PCOV as the sample size tends to infinity for a given dimension. However, in the large dimensional setup, in which both of the sample size and dimension are large, the classical law of large numbers is not applicable because the sample covariance matrix diverges from the PCOV. In the large dimensional setup, the most interesting situation is when the sample size, n, and the dimension, p, increase to infinity proportionally, such that:

$$p/n \to y > 0 \quad \text{with} \quad p, n \to \infty \,. \tag{2.5}$$

The statement in (2.5) is the fundamental assumption in this paper. In addition, we consider $y \in (0, 1)$ and do not study the case where $y > 1$ as we have to deal with the inverse of the singular matrix in the latter case, which is not the purpose of the paper.

Under this assumption, the limiting properties of the sample covariance have been well investigated, and we will use this property to study Markowitz's MV optimization estimation.

Suppose that $\mathbf{x}_k = (x_{1k}, \cdots, x_{pk})'$ $(k = 1, 2, \cdots, n)$ are i.i.d. random vectors with mean vector, $\boldsymbol{\mu}$, and covariance matrix, Σ. Define the sample covariance matrix as:

$$\mathbf{S}_n = \frac{1}{n-1}\sum_{k=1}^{n}(\mathbf{x}_k - \overline{\mathbf{x}})(\mathbf{x}_k - \overline{\mathbf{x}})', \tag{2.6}$$

in which $\overline{\mathbf{x}} = \sum_{k=1}^{n} \mathbf{x}_k/n$ is the sample mean. For any $p \times p$ real symmetric $\mathbf{S}$, the empirical spectral distribution (ESD) $F^{\mathbf{S}}$ is defined as:

$$F^{\mathbf{S}}(x) = \frac{1}{p}\sum_{i=1}^{p}\delta_{[\lambda_i, +\infty)}(x), \tag{2.7}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$ are the eigenvalues of $\mathbf{S}$, and $\delta_A(x)$ is 1 if $x \in A$ and 0 otherwise.

For a distribution sequence $F_n = F^{\mathbf{S}_n}$, if it converges to a nonrandom distribution F as $p/n \to y > 0$ with $p, n \to \infty$, F is called the limiting spectral distribution (LSD) of the sequence of $\{\mathbf{S}_n\}$. We let $\underline{m}$ denote the Stieltjes transform[2] of $\underline{F} = yF + (1-y)\delta_0$. There is an obvious one-to-one mapping between $\underline{F}$ and $\underline{m}$, where $\underline{m}$ is the unique solution on the upper complex plane of the

[2]If $F(x)$ is a function of bounded variation on the real line, then its Stieltjes transform is defined by $m_F(z) = \int \frac{1}{\lambda - z} dF(\lambda)$, $z \in \mathbb{C}^+$, and so $\underline{m}(z) = -\frac{1-y}{z} + ym(z)$ for any $z \in \mathbb{C}^+$.

following Marčenko-Pastur equation (see Siverstein (1995)):

$$z = -\frac{1}{\underline{m}} + y \int \frac{t dH(t)}{1 + t\underline{m}} \qquad z \in \mathbb{C}^+ \;, \tag{2.8}$$

in which $H = \lim_{p\to\infty} F^{\Sigma_p}$.

According to the spectrum analysis of the sample covariance, we deduce the limiting behavior of the quadratic form $\mathbf{a}_p'\mathbf{S}_n^{-1}\mathbf{b}_p$ and $\mathbf{a}_p'\mathbf{S}_n^{-1}\Sigma\mathbf{S}_n^{-1}\mathbf{b}_p$ for any pair of sequences $\{\mathbf{a}_p\}$ and $\{\mathbf{b}_p\}$ under appreciate regularity conditions. Consider the following assumptions that will often be used below:

Assumption (I) $\mathbf{Z}_p = (\mathbf{z}_1, \cdots, \mathbf{z}_n) = (z_{i,j})_{p,n}$, *in which* $z_{i,j}$ $(i = 1, \cdots, p, j = 1, \cdots, n)$ *are i.i.d. random variables, with* $Ez_{ij} = 0$, $E|z_{ij}|^2 = 1$, $E|z_{ij}|^4 < \infty$, *and* $\mathbf{x}_k = \boldsymbol{\mu} + \Sigma_p^{1/2}\mathbf{z}_k$, *where* Σ_p *is a spectrally bounded nonsingular matrix, and* $k = 1, 2, \cdots, n$;

Assumption (II) $\Sigma_p = \mathbf{V}_p\Delta_p\mathbf{V}_p'$ *is nonrandom Hermitian and nonnegative definite with its spectral norm bounded in p, where* $\mathbf{V}_p = (\mathbf{V}_{1,p_1}, \mathbf{V}_{2,p_2}, \cdots, \mathbf{V}_{L,p_L})$ *and*

$$\Delta_p(\boldsymbol{\tau}, \mathbf{w}_p) = diag(\tau_1\mathbf{I}_{p_1}, \tau_2\mathbf{I}_{p_2}, ..., \tau_L\mathbf{I}_{p_L}) \qquad (\tau_1 > \tau_2 > \cdots > \tau_L), \tag{2.9}$$

in which $\boldsymbol{\tau} = (\tau_1, ..., \tau_L)$, $\mathbf{w}_p = (p_1, ..., p_L)/p$, $p_1 + ... + p_L = p$, *and* $\mathbf{I}_{p_i}$ *is the* p_i *dimension unit matrix* $(i = 1, ..., L)$.

Assumption (III) $\mathbf{w}_p \to \mathbf{w} = (w_1, w_2, ..., w_L)$, *as* $p \to \infty$, $(w_1 + w_2 + \cdots + w_L = 1)$.

We now present some results that form the foundation of the paper in the following theorem:

Theorem 2.1 *Under Assumption (I), if the empirical spectral distribution (see (4.3)) of* Σ_p, F^{Σ}, *converges to a given distribution function H, we have:*

$$\mathbf{a}_p'\mathbf{S}_n^{-1}\mathbf{b}_p - \frac{\mathbf{a}_p'\Sigma_p^{-1}\mathbf{b}_p}{(1-y)} \longrightarrow 0 \text{ and } \mathbf{a}_p'\mathbf{S}_n^{-1}\Sigma_p\mathbf{S}_n^{-1}\mathbf{b}_p - \frac{\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p}{(1-y)^3} \longrightarrow 0 \tag{2.10}$$

in probability for any pair of uniform bounded sequences $\{\mathbf{a}_p\}$ *and* $\{\mathbf{b}_p\}$, *where* $\mathbf{S}_n$ *is defined in* (2.6).

We conduct Monte Carlo simulations to check the assertions made in Theorem 2.1, and display the simulation results in Observations 6.1 and 6.2 in the Simulation Section. The simulation results displayed in Observation 6.1 confirm that the assertion shown in the first limiting equation (2.10) holds true, while the simulation results displayed in Observation 6.2 confirm that the second limiting equation (2.10) holds true.

Applying Theorem 2.1, the quadratic form (that is, $\mathbf{a}_p'\mathbf{S}_n^{-1}\mathbf{b}_p$) with the inverse of $\mathbf{S}_n$ is asymptotically $(1-y)^{-1}$ (> 1) times that of (that is, $\mathbf{a}_p'\Sigma_p^{-1}\mathbf{b}_p$) with Σ_p^{-1}. This property could be used to explain the Markowitz mean-variance optimization enigma, which will be discussed in detail in the next subsection.

2.2. *Markowitz mean-variance optimization enigma*

Before discussing the solution for the Markowitz mean-variance optimization enigma, we examine the performance of the sample covariance matrix in the MV optimization portfolio by assuming that the estimates of the population mean vector are fixed. In this paper, we examine the property of the portfolio[3]:

$$\mathbf{c}_p = \mathbf{c}(\boldsymbol{\mu}, \mathbf{S}_n), \tag{2.11}$$

which is constructed by plugging $\mathbf{S}_n$ into (1.2), and then the property of its expected return given by $\mathbf{c}'_p\boldsymbol{\mu}_p$ (denoted as R_p). Bai et al. (2009) refer to $\mathbf{c}(\bar{\mathbf{x}}, \mathbf{S}_n)$ and $\mathbf{c}(\bar{\mathbf{x}}, \mathbf{S}_n)'\bar{\mathbf{x}}$ as "**plug-in allocation**" and "**plug-in return**," respectively. The values of $\mathbf{c}(\bar{\mathbf{x}}, \mathbf{S}_n)$ and $\mathbf{c}(\bar{\mathbf{x}}, \mathbf{S}_n)'\bar{\mathbf{x}}$ are obtained from plugging both $\mathbf{S}_n$ and $\bar{\mathbf{x}}$ into (1.2). Define:

$$R_p = \mathbf{c}'_p\boldsymbol{\mu}_p \quad \textit{and} \quad Risk^p_{\mathbf{c}} = \mathbf{c}'_p\Sigma_p\mathbf{c}_p\ . \tag{2.12}$$

As $\bar{\mathbf{x}}$ is a consistent estimator of $\boldsymbol{\mu}$, without loss of generality, in this paper we refer to $\mathbf{c}_p$ as "**plug-in allocation**", $R_p = \mathbf{c}'_p\boldsymbol{\mu}_p$ as "**plug-in return**," and $Risk^p_{\mathbf{c}} = \mathbf{c}'_p\Sigma_p\mathbf{c}_p$ as "**plug-in risk**".

According to the classical theory of large numbers, as $n \to \infty$, $\mathbf{S}_n$ is a consistent estimator of Σ for given p, so that as $n \to \infty$ for a given p, R_p is consistent for R. However, if p tends to infinity, R_p could become an inaccurate estimate of R. Bai et al. (2009) have analyzed this situation. We extend their work by deriving the following lemma and theorem:

LEMMA 2.1 *Under Assumptions (I) to (III), supposing* $\left(\frac{\mathbf{1}}{\sqrt{p}}, \frac{\mathbf{1}}{\sqrt{p}}\right)$, $\left(\frac{\mathbf{1}}{\sqrt{p}}, \frac{\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|}\right)$ *and* $\left(\frac{\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|}, \frac{\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|}\right)$ *belong to* $\Omega = \left\{(\upsilon_1, \upsilon_2) : \upsilon_1^T U_{p_i} U_{p_i}^T \upsilon_2 = d_i \in \mathbb{R}, i = 1, \cdots, L, \max\{\|\upsilon_1\|, \|\upsilon_2\|\} \le M(>0)\right\}$, *we have:*

$$\frac{\mathbf{1}'\Sigma_p^{-1}\mathbf{1}}{p} \to \varsigma^0_{1,1}, \quad \frac{\mathbf{1}'\Sigma_p^{-1}\boldsymbol{\mu}}{\sqrt{p}\|\boldsymbol{\mu}\|} \to \varsigma^0_{1,\mu}, \quad \textit{and} \quad \frac{\boldsymbol{\mu}'\Sigma_p^{-1}\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|^2} \to \varsigma^0_{\mu,\mu}; \tag{2.13}$$

$$\frac{\mathbf{1}'\mathbf{S}_n^{-1}\mathbf{1}}{p} \to \gamma\varsigma^0_{1,1}, \quad \frac{\mathbf{1}'\mathbf{S}_n^{-1}\boldsymbol{\mu}}{\sqrt{p}\|\boldsymbol{\mu}\|} \to \gamma\varsigma^0_{1,\mu}\ , \quad \textit{and} \quad \frac{\boldsymbol{\mu}'\mathbf{S}_n^{-1}\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|^2} \to \gamma\varsigma^0_{\mu,\mu}; \tag{2.14}$$

$$\frac{\mathbf{1}'\mathbf{S}_n^{-1}\Sigma\mathbf{S}_n^{-1}\mathbf{1}}{p} \to \gamma^3\varsigma^0_{\mathbf{1},\mathbf{1}}, \quad \frac{\mathbf{1}'\mathbf{S}_n^{-1}\Sigma\mathbf{S}_n^{-1}\boldsymbol{\mu}}{\sqrt{p}\|\boldsymbol{\mu}\|} \to \gamma^3\varsigma^0_{\mathbf{1},\mu}\ , \quad \textit{and} \quad \frac{\boldsymbol{\mu}'\mathbf{S}_n^{-1}\Sigma\mathbf{S}_n^{-1}\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|^2} \to \gamma^3\varsigma^0_{\mu,\mu}; \tag{2.15}$$

in which $\gamma = 1/(1-y)$ $(0 < y < 1)$.

[3] In order to eliminate the disturbance from the estimation of $\boldsymbol{\mu}$, we consider it as a known vector. In the empirical analysis, we select $\bar{\mathbf{x}}$ as $\hat{\boldsymbol{\mu}}$, which is a consistent estimator of $\boldsymbol{\mu}$.

THEOREM 2.2 *Under Assumptions (I) to (III), if $\sigma_0 = \xi_0/\sqrt{p}$, $\|\boldsymbol{\mu}\|/\sqrt{p} = \xi_\mu + o(1)$ then, for three pairs of sequences $\left(\frac{\mathbf{1}}{\sqrt{p}}, \frac{\mathbf{1}}{\sqrt{p}}\right)$, $\left(\frac{\mathbf{1}}{\sqrt{p}}, \frac{\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|}\right)$ and $\left(\frac{\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|}, \frac{\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|}\right)$ in Ω, we have:*

a. the limit of the theoretical optimal return R exists and:

$$R \longrightarrow \begin{cases} \xi_0 \xi_\mu \sqrt{\varsigma^0_{\mu,\mu}} & \text{if} \quad \xi_0 \varsigma^0_{1,\mu}/\varsigma^0_{\mu,\mu} < 1, \\ \xi_\mu \frac{\varsigma^0_{1,\mu}}{\varsigma^0_{1,1}} + \xi_\mu \sqrt{\frac{\varsigma^0_{\mathbf{1},\mathbf{1}}\xi_0 - 1}{\varsigma^0_{\mu,\mu}\varsigma^0_{1,1} - (\varsigma^0_{1,\mu})^2}} \left(\varsigma^0_{\mu,\mu} - \frac{(\varsigma^0_{1,\mu})^2}{\varsigma^0_{1,1}}\right) & \text{if} \quad \xi_0 \varsigma^0_{1,\mu}/\varsigma^0_{\mu,\mu} > 1, \end{cases} \tag{2.16}$$

b. the limit of "the plug-in return" exists and:

$$R_p \longrightarrow \begin{cases} \xi_0 \xi_\mu \sqrt{\gamma \varsigma^0_{\mu,\mu}} & \text{if} \quad \sqrt{\gamma}\xi_0 \varsigma^0_{1,\mu}/\varsigma^0_{\mu,\mu} < 1, \\ \xi_\mu \frac{\varsigma^0_{1,\mu}}{\varsigma^0_{1,1}} + \gamma \xi_\mu \sqrt{\frac{\varsigma^0_{1,1}\xi_0^2 - \gamma^{-1}}{\varsigma^0_{\mu,\mu}\varsigma^0_{1,1} - (\varsigma^0_{1,\mu})^2}} \left(\varsigma^0_{\mu,\mu} - \frac{(\varsigma^0_{1,\mu})^2}{\varsigma^0_{1,1}}\right) & \text{if} \quad \sqrt{\gamma}\xi_0 \varsigma^0_{1,\mu}/\varsigma^0_{\mu,\mu} > 1. \end{cases} \tag{2.17}$$

c. the limit of $p \cdot Risk_p$ exists and:

$$p \cdot Risk^p_{\mathbf{c}} \rightarrow \begin{cases} \gamma^2 \xi_0^2 & \text{if} \quad \sqrt{\gamma}\xi_0 \varsigma^0_{1,\mu}/\varsigma^0_{\mu,\mu} < 1, \\ \gamma + \gamma \cdot \frac{\gamma \xi_0^2 \varsigma^0_{1,1} - 1}{\varsigma^0_{\mu,\mu}\varsigma^0_{1,1} - \varsigma^2_{1,\mu}} \left(\varsigma^0_{\mu,\mu} - \frac{\left(\varsigma^0_{1,\mu}\right)^2}{\varsigma^0_{1,1}}\right) & \text{if} \quad \sqrt{\gamma}\xi_0 \varsigma^0_{1,\mu}/\varsigma^0_{\mu,\mu} > 1, \end{cases} \tag{2.18}$$

in which $\gamma = 1/(1-y)$ $(0 < y < 1)$. In addition, $p \cdot Risk^p_{\mathbf{c}} \to \infty$, as $y \to 1$.

From Theorem 2.2, we have the following remark:

REMARK 2.1 *According to (2.16) and (2.17), we have:*

a. when $\xi_0 \varsigma^0_{1,\mu} < \sqrt{1-y}\,\varsigma^0_{\mu,\mu}$ or $\xi_0 \varsigma^0_{1,\mu} > \varsigma^0_{\mu,\mu}$, the plug-in return R_p is always asymptotically greater than the theoretical optimal return;

b. when $\xi_0 \varsigma^0_{1,\mu} < \sqrt{1-y}\,\varsigma^0_{\mu,\mu}$, the plug-in return R_p is asymptotically $1/\sqrt{1-y}$ times the theoretic optimal return R;

c. when $\sqrt{1-y}\,\varsigma^0_{\mu,\mu} < \xi_0 \varsigma^0_{1,\mu} < \varsigma^0_{\mu,\mu}$, we have:

$$R = \xi_0 \xi_\mu \sqrt{\varsigma^0_{\mu,\mu}},$$

and

$$R_p = \xi_\mu \frac{\varsigma^0_{1,\mu}}{\varsigma^0_{1,1}} + \frac{\xi_\mu}{1-y} \sqrt{\frac{\varsigma^0_{1,1}\xi_0^2 - (1-y)}{\varsigma^0_{\mu,\mu}\varsigma^0_{1,1} - (\varsigma^0_{1,\mu})^2}} \left(\varsigma^0_{\mu,\mu} - \frac{(\varsigma^0_{1,\mu})^2}{\varsigma^0_{1,1}}\right),$$

according to (1.2) and (2.17).

Theorem 2.2 and Remark 2.1 show that the findings in Bai et al. (2009) that the plug-in return, R_p, is asymptotically greater than its corresponding theoretical optimal return, R, holds only in point a. of Remark 2.1, but not in point b. Thus, one should not be surprised if the plug-in return,

R_p, is smaller than its corresponding theoretical optimal return, R. We show that it is possible that the plug-in return, R_p, is smaller than its corresponding theoretical optimal return, R, in the following example :

Example 2.1 *Considering the special case in which $\Sigma_p = I_p$ and $\xi_0 = 1$, we have $\varsigma^0_{1,1} = \varsigma^0_{\mu,\mu} = 1$, and so, when $\sqrt{1-y} < \varsigma^0_{1,\mu} < 1$:*

$$R = \xi_\mu \quad \text{and} \quad R_p = \xi_\mu \left(\varsigma^0_{1,\mu} + \frac{\sqrt{y(1-\varsigma^0_{1,u})}}{1-y} \right).$$

For a small enough y, we have $R_p < R$ as $|\varsigma^0_{1,\mu}| < 1$ in (2.13).

In order to demonstrate the assertions in Theorem 2.2, we simulate the plug-in returns, R_p, and its corresponding theoretical optimal return, R, by setting the population covariance to be a unit matrix. We display the results in Figure 1. From the figure, R_p can be larger than R, and the deviation between R_p and R increases exponentially when the number of assets increases. Bai et al. (2009) call this phenomenon “over-prediction”, which is consistent with the finding in Remark 2.1 points a and b. The result is also consistent with the finding in Theorem 2.2 that the plug-in estimator is not accurate for the return estimation in the optimal portfolio.

We also note that the plug-in estimator is inaccurate in the plug-in risk defined in (2.12) because, as y increases toward 1, the risk of the portfolio $\mathbf{c}(\boldsymbol{\mu}, \mathbf{S}_n)$ will increase dramatically. In order to demonstrate this phenomenon, we conduct simulations for the performance of the plug-in risk for different pairs of (p, n) by setting the risk level $\sigma_0 = 1$ in (1.1), and report the results in Table I. From the table, we find that all $risk^p_{\mathbf{c}} > 3$ and $risk^p_{\mathbf{c}}$ could be larger than 100 when $p/n = 0.9$. This means that the plug-in portfolio not only has the over-prediction problem for the estimated return, but also yields much higher risk than its corresponding theoretical optimal portfolio. Table IV provides further information and confirmation of the result.

3. BOOTSTRAP-CORRECTED ESTIMATION

In order to circumvent the limitation of the plug-in estimation, Bai et al. (2009) introduce a bootstrap-corrected approach to improve estimation and solve the over-prediction problem. The bootstrap-corrected method requires a draw from the resample $\chi^* = \{\mathbf{x}^*_1, \cdots, \mathbf{x}^*_n\}$ of the p-variate normal distribution with mean, $\boldsymbol{\mu}$, and covariance matrix, $\mathbf{S}_n$, as defined in equation (2.6). Thereafter, one has to compute the sample covariance matrix from the resample χ^*, denoted as

$\mathbf{S}_n^*$, and then plug $\mathbf{S}_n^*$ into (1.2) to obtain $\mathbf{c}_p^* := \mathbf{c}(\boldsymbol{\mu}, \mathbf{S}_n^*)$ and $R_p^* := R(\boldsymbol{\mu}, \mathbf{S}_n^*)$. Under suitable conditions, Bai et al. (2009) prove the following proposition to provide asymptotic properties for the bootstrap-corrected estimation :

Proposition 3.1 *Under Assumption (I) and using the bootstrapped plug-in procedure, as described above, the* **bootstrap-corrected allocation***, $\boldsymbol{c}_b$, and* **bootstrap-corrected return estimate***, R_b, are:*

$$\boldsymbol{c}_b = \boldsymbol{c}_p + \frac{1}{\sqrt{\gamma}}(\boldsymbol{c}_p - \boldsymbol{c}_p^*) \quad \textit{and} \quad R_b = R_p + \frac{1}{\sqrt{\gamma}}(R_p - R_p^*), \tag{3.1}$$

where $\gamma = 1/(1-y)$, and $\mathbf{c}_p$ and R_p are plug-in allocation and return, respectively.

The bootstrap-corrected allocation is deduced from correcting the bias of R_p, and so it is expected to circumvent the over-prediction problem. Bai et al. (2009) conduct simulations to show that the bootstrap-corrected allocation is indeed closer to the theoretical allocation than is the plug-in allocation, and the bootstrap-corrected return performs better than the plug-in return. We conduct simulations to reexamine the issue and find that, under Assumptions (I) to (III), the bootstrap-corrected allocation is indeed closer to the theoretical allocation than is the plug-in allocation, and the bootstrap-corrected return performs better than the plug-in return. However, we also find that the bootstrap-corrected return could sometimes be smaller than its theoretical optimal return, or even be negative. This shows that the bootstrap-corrected approach can be improved.

We call the risk of the bootstrap-corrected return:

$$Risk_{\mathbf{c}}^b = \mathbf{c}_b' \Sigma_p \mathbf{c}_b \tag{3.2}$$

"bootstrap-corrected risk." According to (3.1) and Part c of Theorem 2.2, we obtain the following theorem:

Theorem 3.1 *Under Assumption (I), for any given p, we have:*

$$p \cdot Risk_{\mathbf{c}}^b = p \cdot Risk_{\mathbf{c}}^p + O(\gamma^{-1/2}), \tag{3.3}$$

in which $\gamma = 1/(1-y) \to \infty$, as $y \to 1$.

In the simulation study, we find that the bootstrap-corrected risk is not stable, and is sometimes even higher than the plug-in risk, $Risk_{\mathbf{c}}^p$, defined in (2.12), implying that the bootstrap-corrected risk, $Risk_{\mathbf{c}}^b$, could perform even worse than the plug-in risk. We report $Risk_{\mathbf{c}}^p$ and $Risk_{\mathbf{c}}^b$ in Table I for the following cases: (a) fix $p/n = 0.5$, and vary p from 100 to 500; and (b) fix $p = 252$ and vary $p/n = 0.5$ to 0.9, with $\sigma_0 = 1$ in both situations. We obtain the following results: (1) the performance of both plug-in and bootstrap-corrected risks are inaccurate as all exceed 3; (2) when $p = 252$ and vary p/n from 0.5 to 0.6, both plug-in and bootstrap-corrected risks are greater than 100; and (3) the bootstrap-corrected risk is larger than the plug-in risk in all cases reported in Table I. The results in Table IV confirm that the performance of both plug-in and bootstrap-corrected risks can be inaccurate.

4. THE LIMITING BEHAVIOR OF THE SAMPLE SPECTRALLY-CORRECTED COVARIANCE MATRIX

According to the theory of large dimensional random matrix, the sample covariance matrix deviates from the population covariance matrix dramatically as $p, n \to \infty$ when its ratio is $y = p/n > 0$. In order to explain this phenomenon, we express the spectral decomposition for the sample covariance matrix as :

$$\mathbf{S}_n = \mathbf{U}_n' \Lambda_n \mathbf{U}_n \tag{4.1}$$

in which $\Lambda_n = diag(\lambda_1, ..., \lambda_p)$ $(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p)$ is the eigenvalue matrix, and $\mathbf{U}_n$ is the corresponding matrix of eigenvectors. In order to solve the problem of the large deviations, the deviation of the sample covariance from the PCOV estimation can be separated into two parts, namely: (i) the deviation of the eigenvalue matrix of the sample covariance; and (ii) the corresponding eigenvector matrix.

For data of large dimensions, it is well known that the eigenvalue matrix of the sample covariance is far from the PCOV matrix (see Bai et al. (2007)). However, it is still an open problem as how best to correct the eigenvectors for the PCOV. First, we have to correct the spectral element, and thereafter correct the corresponding eigenvector. Correcting both the spectral element and the corresponding eigenvector to improve the PCOV estimation is a useful approach. There are many papers on the spectral estimation of the PCOV for the large dimensional data (see, for example, Mestre (2008), Li and Yao (2013), Li et al. (2013), Yao et al. (2012), El Karoui (2008)). The problem is complicated as the eigenvector matrix of the PCOV is not unique when there are two or more eigenvalues of the PCOV of the same value. Therefore, we make the following conjecture:

Conjecture 4.1 *It is not possible to obtain an efficient estimate for the eigenvector.*

In order to provide a possible solution to the problem stated in Conjecture 4.1, we propose an approach to correct the eigenvalue matrix of $\mathbf{S}_n$, and thereafter obtain the sample spectral corrected covariance matrix:

(4.2) $\quad \widetilde{\mathbf{S}}_n = \mathbf{U}_n'\Delta_p\mathbf{U}_n,$

in which Δ_p is given by (2.9). We believe that this estimate will outperform the sample covariance matrix in estimating the quadratic form of the PCOV. We will develop asymptotic properties for the limiting behavior of $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n) = \mathbf{a}_p{}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}_p$, and then conduct simulations to show that $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$ performs better than $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n) = \mathbf{a}_p{}'\mathbf{S}_n^{-1}\mathbf{b}_p$.

Before making the comparison, for a given $p \times p$ symmetric matrix $\mathbf{A}$, define the **empirical spectral distribution** (ESD) as:

(4.3) $$F^{\mathbf{A}}(x) = \frac{1}{p}\sum_{i=1}^{p} I_{\{\lambda_i^{\mathbf{A}} \le x\}}, \quad x \in \mathbb{R},$$

in which $\lambda_1^{\mathbf{A}} \le \lambda_2^{\mathbf{A}} \le \cdots \le \lambda_p^{\mathbf{A}}$ are the eigenvalues of $\mathbf{A}$, and $I_{\{\cdot\}}$ denotes the indicator function.

Theorem 4.2 *Under Assumptions (I) to (III), assume $F^{\mathbf{S}_n}$ has a limit spectral distribution $F^{\boldsymbol{\tau},\mathbf{w}}$, with L splitting support, Θ.*[4] *Then for any pair of sequences $\{\mathbf{a}_p\}$ and $\{\mathbf{b}_p\}$ in Ω (as defined in Lemma 2.1), we have:*

(4.4) $$\mathbf{a}_p{}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}_p \longrightarrow \sum_{k=1}^{L}\frac{d_k}{\tau_k}\sum_{j=1}^{L}\frac{\tau_k(u_j-\tau_j)}{\tau_j(u_j-\tau_k)} \doteq \varsigma(\mathbf{a},\mathbf{b}) \;\; a.s.,$$

where u_j is the solution of $1 + y\int \frac{t}{u-t}dF^{\boldsymbol{\tau},\mathbf{w}}(t) = 0$, for any $j = 1, \cdots, L$ with $\tau_1 > u_1 > \tau_2 > \cdots > \tau_L > u_L > 0$, and $\varsigma(\mathbf{a},\mathbf{b})$ is the limit of $\mathbf{a}_p{}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}_p$.

We conduct Monte Carlo simulations to check that the assertion made in Theorem 4.2 holds true, and display the simulation results in Observation 6.1 in the Simulation Section. The simulation results displayed in Observation 6.1 confirm that the assertion shown Theorem 4.2 is correct.

As explained in Section 2, the accurvacy of the portfolio optimization depends on the accuracy of the estimates of the quadratic forms $\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p$ listed in (1.2) and (1.3). For simplicity, we let $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{A}) = \mathbf{a}_p'\mathbf{A}^{-1}\mathbf{b}_p$, as defined in (2.4), for $A = \Sigma$ or any estimate $\widehat{\Sigma}$ of Σ. The traditional estimate $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n) = \mathbf{a}_p'\mathbf{S}_n^{-1}\mathbf{b}_p$ is asymptotically equal to $\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p/(1-y)$, and answers the following question: (1) What is the characteristic of $\varsigma(\mathbf{a},\mathbf{b})$ defined in (4.4), and (2) does $\varsigma(\mathbf{a},\mathbf{b})$ approach or diverge from $\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p/(1-y)$ as compared with $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) = \mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p$? We hypothesize that $\widetilde{\mathbf{S}}_n$ defined in (4.2) will perform better than the sample covariance matrix $\mathbf{S}_n$ defined in (2.6) in the sense that it provides better estimates of $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$, as shown in the following conjecture:

[4] The L splitting support is the support of $F^{\boldsymbol{\tau},\mathbf{w}}$ that can be covered by L disjoint intervals.

Conjecture 4.3 *Under the conditions stated in Theorem 4.2, when p is large, we have:*

$$
\begin{aligned}
&\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p < \varsigma(\mathbf{a},\mathbf{b}) < \gamma\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p \quad \text{if} \quad \mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p > 0,\\
&\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p > \varsigma(\mathbf{a},\mathbf{b}) > \gamma\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p \quad \text{if} \quad \mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p < 0\,,
\end{aligned}
\tag{4.5}
$$

where $\gamma = 1/(1-y)$*, and* $\varsigma(\mathbf{a},\mathbf{b})$ *is the limit of* $\mathbf{a}_p{}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}_p$*, as defined in Theorem 4.2.*

We conduct Monte Carlo simulations to check the assertion made in Conjecture 4.3, and display the result in Observation 6.1. The simulations confirm the assertion is correct. From the above discussion and the simulations, we expect $\mathbf{c}(\boldsymbol{\mu},\widetilde{\mathbf{S}}_n)$ will perform better than $\mathbf{c}(\boldsymbol{\mu},\mathbf{S}_n)$ in estimating $\mathbf{c}(\boldsymbol{\mu},\Sigma)$ in (1.2) in portfolio optimization. Nonetheless, besides using the estimate of $\mathbf{c}(\boldsymbol{\mu},\Sigma)$, we have to check the accuracy in estimating the risk defined in (2.3). For any portfolio strategy, the corresponding risk is an important measure to evaluate the performance of the strategy. From (2.3), one could find that the risk, $\sigma^2(\boldsymbol{\mu},\hat{\Sigma})$, is determined by $\varphi_{\mathbf{a}}^{\mathbf{b}}(\hat{\Sigma}) = \mathbf{a}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mathbf{b}$ for $\mathbf{a},\mathbf{b} = \mathbf{1}$, and $\boldsymbol{\mu}$ (see (2.4)). Considering $\hat{\Sigma} = \widetilde{\mathbf{S}}_n$, we have following theorem:

Theorem 4.4 *Under Assumptions (I) and (II), for any pair of sequences* $\{\mathbf{a}_p\}$ *and* $\{\mathbf{b}_p\}$ *in* Ω*, we have:*

$$
\mathbf{a}_p{}'\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}_p \longrightarrow \sum_{k=1}^{L} d_k\lambda_k\left(\sum_{j=1}^{L}\frac{(u_j-\lambda_j)}{\lambda_j(u_j-\lambda_k)}\right)^2 \doteq \varrho(\mathbf{a},\mathbf{b}) \quad a.s.
\tag{4.6}
$$

Similar to Conjecture 4.3 to hypothesize the behavior of $\varsigma(\mathbf{a},\mathbf{b})$ defined in Theorem 4.2, we have the following conjecture to hypothesize the behavior of $\varrho(\mathbf{a},\mathbf{b})$ defined in Theorem 4.4:

Conjecture 4.5 *Under the conditions stated in Theorem 4.4, when p is large, we have:*

$$
\begin{aligned}
&\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p < \varrho(\mathbf{a},\mathbf{b}) < \gamma^3\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p \quad \text{if} \quad \mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p > 0\,,\\
&\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p > \varrho(\mathbf{a},\mathbf{b}) > \gamma^3\mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p \quad \text{if} \quad \mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p < 0\,,
\end{aligned}
\tag{4.7}
$$

in which $\gamma = 1/(1-y)$*, and* $\varrho(\mathbf{a},\mathbf{b})$ *is the limiting behavior of* $\mathbf{a}_p'\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}_p$*, as defined in (4.6).*

Conjecture 4.3 hypothesizes the behavior of the estimates of the components for the optimal return, while Conjecture 4.5 hypothesizes the behavior of the estimates of the components for risk. Therefore, we conduct simulations to check whether we could reject the assertions made for the estimates of the risk components in Conjecture 4.5, and in Theorems 2.1 and 4.2. In the simulations, we compute $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)(=\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) = \mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p)$, $\varrho(\mathbf{a},\mathbf{b})$, $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)(=\mathbf{a}_p\mathbf{S}_n^{-1}\Sigma\mathbf{S}_n^{-1}\mathbf{b}_p)$ and $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)(=\mathbf{a}_p\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}_p)$, and report the results in Table III. We find that: (1) $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) < \varrho(\mathbf{a},\mathbf{b}) < \gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$; (2) $\varrho(\mathbf{a},\mathbf{b})$ is close to $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$; (3) $\gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ is further from $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$; (4) $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n) \to \varrho(\mathbf{a},\mathbf{b})$

with small standard deviation; and (5) $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n) \to \gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ with a much higher standard deviation than for $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$. For example, when $y = 0.9$ in Panel A of Table III, $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) = 2.1266$, $\varrho(\mathbf{a},\mathbf{b}) = 4.3561$, $\gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) = 2126.6$, $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n) = 6.7951$ with standard deviation $= 2.1544$, while $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n) = 3422.9$ with standard deviation $= 7450.3$. The results support the assertions that $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$ is a more accurate estimate of $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ than is $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$.

5. THE SPECTRALLY-CORRECTED OPTIMAL PORTFOLIO

We now develop the theory of the spectrally-corrected estimation for the optimal portfolio. Suppose the expected return vector, $\boldsymbol{\mu}$, is given, and plugging the sample spectrally-corrected covariance matrix into (1.2) gives the spectrally-corrected optimal portfolio:

$$\mathbf{c}_s := \mathbf{c}(\boldsymbol{\mu}, \widetilde{\mathbf{S}}_n), \tag{5.1}$$

where $\mathbf{c}(\cdot,\cdot)$ is defined in (1.2). As the estimator $\widetilde{\mathbf{S}}_n$ is obtained by correcting the eigenvalues of the sample covariance, $\mathbf{c}_s$, it is the **spectrally-corrected allocation**. The corresponding expected portfolio return is:

$$R_s = \mathbf{c}_\mathbf{s}' \,\boldsymbol{\mu},$$

which is the **spectrally-corrected return**. We state the formula in the following proposition:

PROPOSITION 5.1 *Under Assumption (I), we have:*

$$R_s = \begin{cases} \sigma_0\sqrt{\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}} & \text{if} \quad \dfrac{\sigma_0 \mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}{\sqrt{\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}} < 1\,, \\ \dfrac{\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}}{\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}} + b_s\left(\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu} - \dfrac{(\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu})^2}{\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}}\right) & \text{if} \quad \dfrac{\sigma_0 \mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}{\sqrt{\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}} > 1\,, \end{cases} \tag{5.2}$$

in which $b_s = b(\boldsymbol{\mu}, \widetilde{\mathbf{S}}_n)$. *In addition, the* **spectrally-corrected risk** *(that is, the risk of the spectrally-corrected allocation) is:*

$$\begin{aligned} Risk_\mathbf{c}^s &= \hat{\mathbf{c}}_s'\Sigma\hat{\mathbf{c}}_s \\ &= \begin{cases} \dfrac{\sigma_0^2\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}{\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}} & \text{if} \quad \dfrac{\sigma_0 \mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}{\sqrt{\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}} < 1\,, \\ [\mathbb{A}' + b_s(\mathbb{B}' - \mathbb{C}')]\,\Sigma\,[\mathbb{A} + b_s(\mathbb{B} - \mathbb{C})] & \text{if} \quad \dfrac{\sigma_0 \mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}{\sqrt{\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}} > 1\,, \end{cases} \end{aligned} \tag{5.3}$$

where $\mathbb{A} = \frac{\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}}{\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}}$, $\mathbb{B} = \widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}$, *and* $\mathbb{C} = \frac{\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}{\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}}\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}$.

Next, we examine the asymptotic behavior of R_s and $Risk_\mathbf{c}^s$ in the following subsections.

5.1. *The limiting behavior of the spectrally-corrected expected return*

According to R_s in (5.2), the limiting behavior depends on the quadratic forms, namely $\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}$, $\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}$, and $\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}$. In order to obtain a better comparison, we examine the limiting behavior of the quadratic forms for their corresponding parameters, namely $\mathbf{1}'\Sigma^{-1}\mathbf{1}$, $\mathbf{1}'\Sigma^{-1}\boldsymbol{\mu}$, and $\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}$. As both $||\mathbf{1}||$ and $||\boldsymbol{\mu}||$ tend to infinity as $p \to \infty$, it is necessary to standardize the two vectors as $\mathbf{1}/\sqrt{p}$ and $\boldsymbol{\mu}/||\boldsymbol{\mu}||$, respectively. We derive the following theorem to state the asymptotic properties of the standardized terms $\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}$, $\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}$, and $\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}$, and the quadratic forms of their corresponding parameters:

Lemma 5.1 *Under Assumptions (I) and (II), for the three pairs of sequences* $\left(\frac{\mathbf{1}}{\sqrt{p}}, \frac{\mathbf{1}}{\sqrt{p}}\right)$, $\left(\frac{\mathbf{1}}{\sqrt{p}}, \frac{\boldsymbol{\mu}}{||\boldsymbol{\mu}||}\right)$ *and* $\left(\frac{\boldsymbol{\mu}}{||\boldsymbol{\mu}||}, \frac{\boldsymbol{\mu}}{||\boldsymbol{\mu}||}\right)$ *in* Ω*, we have:*

$$\frac{\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}}{p} \to \varsigma_{\mathbf{1},\mathbf{1}}, \quad \frac{\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}{\sqrt{p}||\boldsymbol{\mu}||} \to \varsigma_{\mathbf{1},\mu}\,, \quad \textit{and} \quad \frac{\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}{||\boldsymbol{\mu}||^2} \to \varsigma_{\mu,\mu}, \tag{5.4}$$

in which $\gamma = 1/(1-y)$ $(0 < y < 1)$.

With the aid of Lemma 5.1, we can derive the asymptotic properties of the limiting behavior of the spectrally-corrected return, R_s, for the optimal portfolio, and its corresponding theoretical optimal return, R, as shown in the following theorem:

Theorem 5.1 *Under the conditions of Theorem 4.2, and given the definitions in* (2.13) *and* (5.4)*, if* $\sigma_0 = \xi_0/\sqrt{p}$ *and* $||\boldsymbol{\mu}||/\sqrt{p} = \xi_\mu + o(1)$*, we have:*

a. the theoretical optimal return, R, exists and satisfies:

$$R \longrightarrow \begin{cases} \xi_0\xi_\mu\sqrt{\varsigma^0_{\mu,\mu}} & \textit{if} \quad \xi_0\varsigma^0_{1,\mu}/\varsigma^0_{\mu,\mu} < 1, \\ \xi_\mu\frac{\varsigma^0_{1,\mu}}{\varsigma^0_{1,\mathbf{1}}} + \xi_\mu\sqrt{\frac{\varsigma^0_{\mathbf{1},\mathbf{1}}\xi_0 - 1}{\varsigma^0_{\mu,\mu}\varsigma^0_{1,1} - (\varsigma_{1,\mu})^2}}\left(\varsigma^0_{\mu,\mu} - \frac{(\varsigma^0_{1,\mu})^2}{\varsigma^0_{1,1}}\right) & \textit{if} \quad \xi_0\varsigma^0_{1,\mu}/\varsigma^0_{\mu,\mu} > 1; \end{cases}$$

b. the limit of the spectrally-corrected return, R_s, for the optimal portfolio exists and follows:

$$R_s \longrightarrow \begin{cases} \xi_0\xi_\mu\sqrt{\varsigma_{\mu,\mu}} & \textit{if} \quad \xi_0\varsigma_{1,\mu}/\varsigma_{\mu,\mu} < 1, \\ \xi_\mu\frac{\varsigma_{1,\mu}}{\varsigma_{1,1}} + \xi_\mu\sqrt{\frac{\varsigma_{1,1}\xi_0 - 1}{\varsigma_{\mu,\mu}\varsigma_{1,1} - (\varsigma_{1,\mu})^2}}\left(\varsigma_{\mu,\mu} - \frac{(\varsigma_{1,\mu})^2}{\varsigma_{1,1}}\right) & \textit{if} \quad \xi_0\varsigma_{\mathbf{1},\mu}/\varsigma_{\mu,\mu} > 1. \end{cases}$$

By using Theorem 2.2, we can show that both Lemma 5.1 and Theorem 5.1 hold.

In simulations, we compute $\varsigma(\mathbf{a}, \mathbf{b})$ in Table II and show that Conjecture 4.3 holds for a general sequence pair of $\mathbf{a}_p$ and $\mathbf{b}_p$. Under the assertions in Conjecture 4.3, $(\varsigma_{\mathbf{1},\mathbf{1}}, \varsigma_{\mathbf{1},\mu}, \varsigma_{\mu,\mu})$ is closer to $(\varsigma^0_{\mathbf{1},\mathbf{1}}, \varsigma^0_{\mathbf{1},\mu}, \varsigma^0_{\mu,\mu})$ than is $\gamma(\varsigma^0_{\mathbf{1},\mathbf{1}}, \varsigma^0_{\mathbf{1},\mu}, \varsigma^0_{\mu,\mu})$ under the Euler distance which, in turn, implies that R_s will be closer to R than to R_p. The result is confirmed by the results in Table IV, namely that R_s is close to R, on average, with a smaller standard deviation. We discuss the issue further in the simulation section. We conclude that Lemma 5.1, Theorem 5.1 and the simulation results in Table IV support the conjecture that R_s is *proportionally consistent* with the theoretical optimal expected return R, under appropriate regularity conditions.

5.2. *The limiting behavior of the spectrally-corrected risk*

According to the expression of $\hat{\sigma}^2$ in (2.3), the limiting behavior depends on three quadratic forms, namely $\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}$, $\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}$, and $\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}$. As both $\|\mathbf{1}\|$ and $\|\boldsymbol{\mu}\|$ tend to infinity as $p \to \infty$, it is necessary to standardize these two vectors as $\mathbf{1}/\sqrt{p}$ and $\boldsymbol{\mu}/\|\boldsymbol{\mu}\|$, respectively. We constrain the vector sequences $\mathbf{1}/\sqrt{p}$ and $\boldsymbol{\mu}/\|\boldsymbol{\mu}\|$ satisfying Ω, and develop the limiting properties for the standardized terms of $\varphi_{\mathbf{1}}^{\mathbf{1}}(\widetilde{\mathbf{S}}) = \frac{1}{p}\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}$, $\varphi_{\mathbf{1}}^{\boldsymbol{\mu}}(\widetilde{\mathbf{S}}) = \frac{1}{\sqrt{p}\|\boldsymbol{\mu}\|}\mathbf{1}'\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}$ and $\varphi_{\boldsymbol{\mu}}^{\boldsymbol{\mu}}(\widetilde{\mathbf{S}}) = \frac{1}{\|\boldsymbol{\mu}\|^2}\boldsymbol{\mu}'\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}$, as shown in the following lemma:

LEMMA 5.2 *Under Assumptions (I) to (III), for three pairs of sequences* $\left(\frac{\mathbf{1}}{\sqrt{p}}, \frac{\mathbf{1}}{\sqrt{p}}\right)$, $\left(\frac{\mathbf{1}}{\sqrt{p}}, \frac{\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|}\right)$ *and* $\left(\frac{\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|}, \frac{\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|}\right)$ *in* Ω*, we have:*

$$\frac{\mathbf{1}^T\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\mathbf{1}}{p} \to \varrho_{\mathbf{1},\mathbf{1}}, \quad \frac{\mathbf{1}^T\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}{\sqrt{p}\|\boldsymbol{\mu}\|} \to \varrho_{\mathbf{1},\mu}, \quad \textit{and} \quad \frac{\boldsymbol{\mu}^T\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\boldsymbol{\mu}}{\|\boldsymbol{\mu}\|^2} \to \varrho_{\mu,\mu}, \tag{5.5}$$

in which $\gamma = 1/(1-y)$ $(0 < y < 1)$.

With Lemma 5.2, we develop the asymptotic property for the risk of the spectrally-corrected portfolio, as shown in the following theorem:

THEOREM 5.2 *Under the conditions stated in Theorem 4.2, and the definitions in* (2.13), (5.4) *and* (5.5)*, if* $\sigma_0 = \xi_0/\sqrt{p}$, $\|\boldsymbol{\mu}\|/\sqrt{p} = \xi_\mu + o(1)$*, we have:*

a. *when* $\xi_0\varsigma_{\mathbf{1},\mu}/\varsigma_{\mu,\mu} < 1$, $p \cdot Risk_{\mathbf{c}}^s \to \xi_0\varrho_{\mu,\mu}/\varsigma_{\mu,\mu}$ *almost surely; and*

b. *when* $\xi_0\varsigma_{\mathbf{1},\mu}/\varsigma_{\mu,\mu} > 1$, $p \cdot Risk_{\mathbf{c}}^s$ *converges to:*

$$\begin{aligned} &\frac{\varrho_{\mathbf{1},\mathbf{1}}}{\varsigma_{\mathbf{1},1}} + 2\sqrt{\frac{\varsigma_{\mathbf{1},\mathbf{1}}\xi_0 - 1}{\varsigma_{\mu,\mu}\varsigma_{\mathbf{1},\mathbf{1}} - (\varsigma_{\mathbf{1},\mu})^2}}\left(\frac{\varrho_{1,\mu}}{\varsigma_{1,1}} - \frac{\varsigma_{1,\mu}\varrho_{1,1}}{\varsigma_{1,1}^2}\right) \\ &\quad + \frac{\varsigma_{\mathbf{1},\mathbf{1}}\xi_0 - 1}{\varsigma_{\mu,\mu}\varsigma_{\mathbf{1},\mathbf{1}} - (\varsigma_{\mathbf{1},\mu})^2}\left(\varrho_{\mu,\mu} - 2\frac{\varsigma_{\mathbf{1},\mu}\varrho_{\mathbf{1},\mu}}{\varsigma_{\mathbf{1},1}} + \left(\frac{\varsigma_{\mathbf{1},\mu}}{\varsigma_{\mathbf{1},1}}\right)^2\varrho_{\mathbf{1},\mathbf{1}}\right) \end{aligned}$$

almost surely.

In the simulations, we compute $\varrho(\mathbf{a}, \mathbf{b})$ and verify Conjecture 4.5 in Table III. According to Conjecture 4.5, $(\varrho_{\mathbf{1},\mathbf{1}}, \varrho_{\mathbf{1},\mu}, \varrho_{\mu,\mu})$ is closer to $(\varsigma_{\mathbf{1},\mathbf{1}}^0, \varsigma_{\mathbf{1},\mu}^0, \varsigma_{\mu,\mu}^0)$ than is $\gamma^3(\varsigma_{\mathbf{1},\mathbf{1}}^0, \varsigma_{\mathbf{1},\mu}^0, \varsigma_{\mu,\mu}^0)$. Combined with the conjecture that $(\varsigma_{\mathbf{1},\mathbf{1}}, \varsigma_{\mathbf{1},\mu}, \varsigma_{\mu,\mu})$ is closer to $(\varsigma_{\mathbf{1},\mathbf{1}}^0, \varsigma_{\mathbf{1},\mu}^0, \varsigma_{\mu,\mu}^0)$ than is $\gamma(\varsigma_{\mathbf{1},\mathbf{1}}^0, \varsigma_{\mathbf{1},\mu}^0, \varsigma_{\mu,\mu}^0)$, $p \cdot Risk_{\mathbf{c}}^s$ will be smaller than $p \cdot Risk_{\mathbf{c}}^p$, as verified in Table IV.

REMARK 5.1 *Comparing* $p \cdot Risk_{\mathbf{c}}^{p}$ *in* (2.18) *with* $p \cdot Risk_{\mathbf{c}}^{s}$ *in Theorem* 5.2*,* $p \cdot Risk_{\mathbf{c}}^{p} \to \infty$ *as* $y \to 1$*, while* $p \cdot Risk_{\mathbf{c}}^{s}$ *is stable for large* $y \in (0,1)$*. Thus,* $Risk_{\mathbf{c}}^{s}$ *performs better than does* $Risk_{\mathbf{c}}^{p}$*.*

In the section of Simulation Study, we compute $\varrho(\mathbf{a},\mathbf{b})$ and verify Conjecture 4.5 in Table III. According to Conjecture 4.5, it is reasonable to conjecture that $(\varrho_{\mathbf{1},\mathbf{1}}, \varrho_{\mathbf{1},\mu}, \varrho_{\mu,\mu})$ is closer to $(\varsigma^0_{\mathbf{1},\mathbf{1}}, \varsigma^0_{\mathbf{1},\mu}, \varsigma^0_{\mu,\mu})$ than is $\gamma^3(\varsigma^0_{\mathbf{1},\mathbf{1}}, \varsigma^0_{\mathbf{1},\mu}, \varsigma^0_{\mu,\mu})$. Together with the conjecture that $(\varsigma_{\mathbf{1},\mathbf{1}}, \varsigma_{\mathbf{1},\mu}, \varsigma_{\mu,\mu})$ is closer to $(\varsigma^0_{\mathbf{1},\mathbf{1}}, \varsigma^0_{\mathbf{1},\mu}, \varsigma^0_{\mu,\mu})$ than is $\gamma(\varsigma^0_{\mathbf{1},\mathbf{1}}, \varsigma^0_{\mathbf{1},\mu}, \varsigma^0_{\mu,\mu})$, $p \cdot Risk_{\mathbf{c}}^{s}$ will be smaller than $p \cdot Risk_{\mathbf{c}}^{p}$, which is verified in Table IV. As $p \cdot Risk_{\mathbf{c}}^{b}$ is $O(\gamma^2)$, the same as $p \cdot Risk_{\mathbf{c}}^{p}$ in (3.2), $p \cdot Risk_{\mathbf{c}}^{b}$ is greater than $p \cdot Risk_{\mathbf{c}}^{s}$ as $\gamma = 1/(1-y) \to \infty$. From Remark 5.1 and the simulation results in Table IV, $Risk_{\mathbf{c}}^{s}$ is the smallest among $Risk_{\mathbf{c}}^{w}$ $(w = s, p, b)$.

6. SIMULATION STUDY

According to Proposition 2.1, the main factors to decide the performance of the optimal portfolio estimation are the quadratic forms $\mathbf{a}'\hat{\Sigma}^{-1}\mathbf{b}$ and $\mathbf{a}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mathbf{b}$. In Section 2, we deduced their limiting behavior when $\hat{\Sigma} = \mathbf{S}_n$ in Theorem 2.1. In Section 4, we deduced their limiting behavior when $\hat{\Sigma} = \widetilde{\mathbf{S}}_n$ in Theorems 4.2 and 4.4. We also conjectured the relationships between $\mathbf{a}'\mathbf{S}_n^{-1}\mathbf{b}$, and between $\mathbf{a}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}$, $\mathbf{a}'\mathbf{S}_n^{-1}\Sigma\mathbf{S}_n^{-1}\mathbf{b}$ and $\mathbf{a}'\widetilde{\mathbf{S}}_n^{-1}\Sigma\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}$, in Conjectures 4.3 and 4.5.

In the next subsection, we conduct simulations to support the assertions in Theorem 4.2, in general, and examine whether the assertions in Conjecture 4.3 and 4.5 hold. Thereafter, we will conduct simulations to check whether the assertions made in Theorem 4.4 and Conjecture 4.5 hold.

6.1. *Simulations for* $\mathbf{a}'\hat{\Sigma}^{-1}\mathbf{b}$ *and* $\mathbf{a}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mathbf{b}$

Step 1: Set $\vec{\tau} = (\tau_1, ..., \tau_L)$ and $\vec{w} = (p_1/p, p_2/p, ..., p_L/p)$, with $p = p_1 + p_2 + ... + p_L$, and obtain: $\Sigma_p = diag\left(\tau_1 I_{p_1}, \tau_2 I_{p_2}, ..., \tau_L I_{p_L}\right) := \Sigma_p(\vec{\tau}, \vec{w}, p)$.

Step 2: Select the vector pair $(\mathbf{a}_p, \mathbf{b}_p)$ such that:

$$\sum_{i=p_1+...+p_{k-1}+1}^{p_1+...+p_k} a_i b_i = d_k \quad (p_0 = 0), \tag{6.1}$$

for any given d_k $(k = 1, 2, ..., L)$, in which $\mathbf{a}_p = (a_1, ..., a_p)$ and $\mathbf{b}_p = (b_1, ..., b_p)$. Here, d_k is the inner product of $\mathbf{a}_p$ and $\mathbf{b}_p$ on the subspace extended by the columns of $\mathbf{V}_{k,p_k}$ which is given in condition b of Theorem 2.1. For any vector pair, $(\tilde{\mathbf{a}}_p, \tilde{\mathbf{b}}_p)$, we can construct $\mathbf{a}_p = \tilde{\mathbf{a}}_p$ and:

$$\mathbf{b}_p = \left(\frac{d_1}{\tilde{d}_1}\tilde{b}_1, ..., \frac{d_1}{\tilde{d}_1}\tilde{b}_{p_1}, ..., \frac{d_k}{\tilde{d}_k}\tilde{b}_{\tilde{p}_{k-1}+1}, ..., \frac{d_k}{\tilde{d}_k}\tilde{b}_{\tilde{p}_k}, ..., \frac{d_L}{\tilde{d}_L}\tilde{b}_{\tilde{p}_{L-1}+1}, ..., \frac{d_L}{\tilde{d}_L}\tilde{b}_{\tilde{p}_L}\right), \tag{6.2}$$

in which $\tilde{\mathbf{a}}_p = (\tilde{a}_1, ..., \tilde{a}_p)$, $\tilde{\mathbf{b}}_p = (\tilde{b}_1, ..., \tilde{b}_p)$, and $\sum_{i=p_1+...+p_{k-1}+1}^{p_1+...+p_k} \tilde{a}_i\tilde{b}_i = \tilde{d}_k$.

Step 3: Vary the ratio $y = p/n$ from 0.1 to 0.9. For each value of y, generate the sample $\mathbf{x}_1, ..., \mathbf{x}_n$, and evaluate the values $\varsigma(\mathbf{a}, \mathbf{b})$ and $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) = \mathbf{a}_p'\mathbf{A}^{-1}\mathbf{b}_p$ ($A = \Sigma, \widetilde{\mathbf{S}}_n$ and $\mathbf{S}_n$, respectively). Thus, according to step 2, $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) = \tau_1 d_1 + ... + \tau_L d_L$ is fixed for $\vec{\tau}$.

Step 4: Repeat steps 1 to 3 a total of $N = 10,000$ times, and obtain the mean and standard deviation of the simulated values for each y.

We first use the steps in Simulation 6.1 to conduct simulations to check the assertions made in Conjecture 4.3, and whether the assertions in Theorems 2.1, and 4.2 hold. In order to check the above, in the simulations we compute $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)(= \mathbf{a}_p'\Sigma^{-1}\mathbf{b}_p)$, $\varsigma(\mathbf{a}, \mathbf{b})$, $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)(= \mathbf{a}_p'\mathbf{S}_n^{-1}\mathbf{b}_p)$, $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)(= \mathbf{a}_p'\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}_p)$, and $\gamma\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$, and report the results in Table II. From the table, we obtain the following observations:

Observation 6.1

a. *Confirm Conjecture 4.3 that* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) < \varsigma(\mathbf{a}, \mathbf{b}) < \gamma\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*;*

b. $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ *is close to* $\varsigma(\mathbf{a}, \mathbf{b})$*, and* $\gamma\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ *is far from* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*;*

c. $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$ *and* $\gamma\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ *are the terms in the first limiting equation (2.10) of Theorem 2.1. We observe that the mean of* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$ *is close to* $\gamma\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*, with standard deviation (sd) less than 0.82 for* $y \leq 0.5$*. When* $y > 0.5$*, the mean of* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$ *is still close to* $\gamma\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*, with higher sd, but it is still less than* 0.8 *times* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*. Thus, the results confirm the assertion, in the first limiting equation (2.10) of Theorem 2.1, that* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n) \to \gamma\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*. Overall,* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n) \to \gamma\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*, with a much higher sd than that of* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$*.*

d. $\varsigma(\mathbf{a}, \mathbf{b})$ *and* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$ *are the terms in the limiting equation (4.4) in Theorem 4.2. The value of* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$ *is very close to* $\varsigma(\mathbf{a}, \mathbf{b})$ *in mean, with the sd bounded by* 0.41*. Thus, the results confirm the assertion in Theorem 4.2. In addition, compared with* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$*,* $\phi_{\mathbf{a}_p}^{\mathbf{b}}(\widetilde{\mathbf{S}}_n)$ *has a smaller sd, and is obvious for large y. Overall, we find that* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n) \to \varsigma(\mathbf{a}, \mathbf{b})$*, with small sd.*

These observations confirm that $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$ is a better estimate of $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ than is $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$. For example, when $y = 0.9$ in Panel B of Table II, $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) = 1.7$, $\varsigma = 2.0066$, $\gamma\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) = 17$, $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n) = 1.9514$, with sd = 0.2913, while $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n) = 19.060$, with sd = 11.968.

In order to examine whether the assertions in both Theorem 4.4 and Conjecture 4.5 hold, we can use the steps in Simulation 6.1 to compute $\varrho(\mathbf{a}, \mathbf{b})$ and $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{A}) = \mathbf{a}_p'\mathbf{A}^{-1}\Sigma\mathbf{A}^{-1}\mathbf{b}_p$ ($\mathbf{A} = \Sigma, \widetilde{\mathbf{S}}_n$

and $\mathbf{S}_n$, respectively) in Step 3. The simulation results are reported in Table III. From the table, we obtain the following observations:

Observation 6.2

a. *Confirm Conjecture 4.5 that* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) < \varrho(\mathbf{a},\mathbf{b}) < \gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*;*
b. $\varrho(\mathbf{a},\mathbf{b})$ *is close to* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ *and* $\gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ *is further from* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*;*
c. $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$ *and* $\gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ *are the terms in the second liming equation* (2.10) *of Theorem 2.1. We observe that the mean of* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$ *is close to* $\gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*, with sd less than* 7.2 *for* $y \le 0.5$*. When* $y > 0.5$*, the mean of* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$ *is still close to* $\gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*, but the sd increases with* y*, and reaches more than 3 times* $\gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*. Thus, the results confirm the assertion in the second limiting equation* (2.10) *of Theorem 2.1 that* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n) \to \gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*.*
d. $\varrho(\mathbf{a},\mathbf{b})$ *and* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$ *are the terms in the limiting equation* (4.7) *in Theorem 4.4. The value of* $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$ *is very close to* $\varsigma(\mathbf{a},\mathbf{b})$ *in mean, and with sd bounded by* 2.2*. Thus, the results confirm the assertion in Theorem 4.4. In addition, compared with* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$*,* $\varphi_{\mathbf{a}_p}^{\mathbf{b}}(\widetilde{\mathbf{S}}_n)$ *has a smaller sd, which is obvious for large* y*.*
e. *From c and d,* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n) \to \gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$*, with a much higher sd than that of* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$ *while* $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n) \to \varrho(\mathbf{a},\mathbf{b})$ *with a small sd.*

These observations confirm that $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n)$ is a better estimate of $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$ than is $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$. For example, when $y = 0.9$ in Panel C of Table III, $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma) = 2.2666$, $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\widetilde{\mathbf{S}}_n) = 4.7502$, with $sd = 1.1209$, while $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n) = 3617.4$, with $sd = 7589.3$.

Now we are ready to conduct simulations in the next subsection to compare both return and risk performances of the proposed spectrally-corrected estimates with those of the plug-in and bootstrap-corrected estimates. In order to do so, we compare the performance of $\mathbf{c}_s$ with $\mathbf{c}_p$ and $\mathbf{c}_b$ in equations (2.2) and (2.3) in terms of expected return and risk.

6.2. *Simulations for the optimal portfolio estimates*

Given a p-dimension nonzero vector, $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_p)'$, and a positive definite matrix, $\Sigma = (\sigma_{ij})$, which is assumed to be diagonal for simplicity, we state the simulation procedure as follows:

Step 1: Generate n vectors of returns, $\mathbf{r} = (r_1, \cdots, r_p)$, for the p-branch of assets from a population with mean, $\boldsymbol{\mu}$, and covariance matrix, Σ.

Step 2: Use equations (2.2) and (2.3) to compute the optimal allocation, $\mathbf{c}$, and the expected return, R, for the plug-in, bootstrap-corrected, and the proposed spectrally-corrected estimates, as follows:

(i) use equation (2.11) to compute $\mathbf{c}_p$, the first equation in (3.1) to compute $\mathbf{c}_b$, and equation (5.1) to compute $\mathbf{c}_s$; then

(ii) substitute $\mathbf{c}_w$ into the formula, $R_w = \mathbf{c}'_w \boldsymbol{\mu}$, to obtain the corresponding expected return, $R_w = \mathbf{c}'_w \boldsymbol{\mu}$, for $w = p, s, b$.

Step 3: Compute $R_w - R$, $\|\mathbf{c}_w - \mathbf{c}\|$ and $\mathbf{c}'_w \Sigma \mathbf{c}_w$ ($w = p, s, b$).

Step 4: Repeat Steps 1 to 3 a total of N times, and calculate the means and standard deviations for R_w, $R_w - R$, $\|\mathbf{c}_w - \mathbf{c}\|$ and $\mathbf{c}'_w \Sigma \mathbf{c}_w$ ($w = p, s, b$).

Select a random vector as the population, $\boldsymbol{\mu}$, and consider three different, Σ, where each Σ contains three or four different eigenvalues. For each set of $\boldsymbol{\mu}$ and Σ, conduct simulations according to the above steps, and compute the means and standard deviations of R_w, $R_w - R$, and $\mathbf{c}'_w \Sigma \mathbf{c}_w$ ($w = p, s, b$), in which p is fixed and $y = p/n$ increases from 0.1 to 0.9. In order to make comparisons easier, we compute the percentage of the means of $R_s - R$ over R. In Table IV, we present the simulated results for the three different populations in Panels A, B and C, respectively.

We first compare the expected returns of the optimal portfolio estimates. From all the panels, we have the following observations: (1) the mean of the spectrally-corrected portfolio return, R_s, is the closest estimate to the expected return, R, of the theoretical MV optimal portfolio, followed by that of the bootstrap-corrected portfolio return, R_b, then the mean of the plug-in portfolio return, R_p, with $|R_s - R|$ as the smallest, followed by $|R_b - R|$, and $|R_p - R|$ is the largest for any $y = 0.1$ to 0.9; (2) the sd of $|R_s - R|$ is the smallest, followed by $|R_p - R|$, while the sd of $|R_b - R|$ is the largest for any $y = 0.1$ to 0.9; (3) both the spectrally-corrected portfolio return, R_s, and the bootstrap-corrected portfolio return, R_b, underestimate the expected return, R, of the theoretical MV optimal portfolio, while the plug-in portfolio return, R_p, overestimates the expected return, R, for any $y = 0.1$ to 0.9; (4) the underestimation of the spectrally-corrected portfolio return, R_s, is very small (from 0.01% to 1.58%) for any y; (5) the underestimation of the bootstrap-corrected portfolio return, R_b, could be small for small y, but large for large y (from 0.27% to 115.9%); (6) the overestimation of the plug-in portfolio return, R_p, is very large (from 5.25% to 159.4%). Now we compare the risk for the different portfolio estimates. In Table IV, we set the risk level at $\sigma_0^2 = 1$. From the table, we notice that: (7) all the $Risk_w$ are larger than σ_0^2 for any $w = p, b, s$, and for any $y = 0.1$ to 0.9. Thus, one should select the portfolio estimate in which the risk is not too far from σ_0^2.

From Table IV, we have the following observations: (8) the spectrally-corrected risk, $Risk_s$, is the smallest, followed by the plug-in risk, $Risk_p$, while the bootstrap-corrected risk, $Risk_b$, is the largest for any $y = 0.1$ to 0.9; (9) the sd of the spectrally-corrected risk, $Risk_s$, is the smallest for any $y = 0.1$ to 0.9; (10) comparing the sd of the plug-in risk, $Risk_p$, and of the bootstrap-corrected risk, $Risk_b$, the former is smaller for small y (< 0.3) and large y (> 0.7), while the latter is smaller for $y = 0.3$ and 0.4 to 0.7.

Now we use the results show in Table IV to illustrate the above observations, especially to show that the spectrally-corrected estimates are the best of the three estimates. In each panel, $p = 100$ is given, and n varies such that $y = p/n$ increases from 0.1 to 0.9. As the conclusions drawn from the other panels are the same as that drawn from Panel A, we illustrate the above observations by analyzing the results from only Panel A of Table IV, as follows:

(1) The spectrally-corrected estimates perform the best in terms of the expected return:

   (a) When $y = 0.1$, R_s is only 0.14% (with sd=0.0132) below R, R_p is 5.25% (with sd=0.0242) higher than R, and R_b is 0.31%(with sd=0.0344) higher than R, on average. On the other hand, when $y = 0.9$, R_s is still only 1.5% (with sd=0.0641) below R, R_p is 159.41% (with sd=1.2518) higher than R, and R_b is 81.62% (with sd=1.8346) below R on average.

   (b) The ratio y has the smallest influence on the expected return of the spectrally-corrected portfolio when compared with the plug-in and bootstrap-corrected estimates. When y increases from 0.1 to 0.9, the range of $|(R_{\hat{\mathbf{c}}} - R)/R|$ for $\hat{\mathbf{c}} = \mathbf{c}_s$ is the smallest, from 0.14% to 1.58%, with sd from 0.0132 to 0.0641, the range for $\mathbf{c}_p$ is from 5.25% to 159.41%, with sd from 0.0242 to 1.2518, while that for $\mathbf{c}_b$ is from 0.31% to 81.62%, with sd from 0.0344 to 1.8346.

(2) The spectrally-corrected estimation performs the best in term of risk:

   (a) When $y = 0.1$, $Risk_s$ is 1.0771 (with sd=0.0312), $Risk_p = 1.2323$ (with sd=0.0609), and $Risk_b = 1.2452$ (with sd=0.0806). On the other hand, when $y = 0.9$, $Risk_s$ is still small at 2.2 (with sd=0.5822), $Risk_p$ increases to 86.581 (with sd=86.581), and $Risk_b$ goes beyond 150 (with sd 170.23).

(b) The ratio y has the smallest influence on the risk of the spectrally-corrected portfolio, when compared with the plug-in and bootstrap-corrected estimates because $Risk_{\mathbf{c}_s}$ is the smallest, from 1.0771 to 2.1382, with sd from 0.0312 to 0.5822. On the other hand, both $Risk_{\mathbf{c}_b}$ and $Risk_{\mathbf{c}_p}$ are very large, from 1.2452 to 151.27, with sd from 0.0806 to 151.27, for $\mathbf{c}_b$, and from 1.2323 to 86.581, with sd from 0.0609 to 78.657, for $\mathbf{c}_p$.

(c) For a given method, such as plug-in estimate, the ratio y is smaller, the performance of $\mathbf{c}_s$ is better which also holds for the other two methods. For example, the percentage of the absolute value of the err ratio, $(R_{\hat{\mathbf{c}}} - R_{\mathbf{c}})/R_c$ increases from 0.14% (sd=0.0132) to 1.58% (sd=0.0641), and of $Risk_{\hat{c}}$ from 1.0771 (sd=0.0312) to 2.1382 (0.5822), as $\hat{\mathbf{c}} = \mathbf{c}_s$.

From the above discussion, we conclude that, as R_p has an unacceptably high level of overestimation, and high risk, R_p is not as stable as R_s or R_b. On the other hand, R_b corrects the overestimation of R_p, but: (a) the sd of $\left|R_b - R\right|$ is the largest; (b) the bootstrap-corrected risk, $Risk_b$, is the largest for $y = 0.1$ to 0.9; (c) the sd of $Risk_b$ is the largest for small y, as well as for large y. Thus, we conclude that R_b is not a good choice for the optimal portfolio return. In addition, we conclude that the spectrally-corrected portfolio return, R_s, is the best estimate for the optimal portfolio return when compared with R_p and R_b because: (a) although R_s underestimates the expected return, R, of the theoretical MV optimal portfolio, the underestimation is the smallest for any $y = 0.1$ to 0.9; (b) the mean of R_s is the closest estimate to R, with the sd of $\left|R_s - R\right|$ the smallest for any $y = 0.1$ to 0.9 when compared with both R_p and R_b; (c) the spectrally-corrected risk, $Risk_s$, is the smallest, and its sd is also the smallest for any $y = 0.1$ to 0.9.

7. EMPIRICAL ILLUSTRATION

In this section, we compare the performance of different optimal portfolio estimates for real data from S&P 500. We choose the largest 500 stocks with the highest capitalization from the S&P 500 index from January 1, 2004 to December 31, 2013, and compute their weekly logarithmic returns.

As we will compare the performance of $\hat{\mathbf{c}}_s$ with the portfolios $\hat{\mathbf{c}}_p$, $\hat{\mathbf{c}}_b$ and $\hat{\mathbf{c}}_0$ for different numbers, p, of stocks, we set $p = 50$, 100, 200, and 300 (that is, $y = 0.1$, 0.2, 0.4 and 0.6), respectively, and choose p stocks randomly from 500 stocks in the S&P 500 index. We select the sample mean, $\hat{\boldsymbol{\mu}}$, as the estimate of the $\boldsymbol{\mu}$ vector, and then obtain $\hat{\mathbf{c}}_w$ ($w = p, s, b$). Thereafter, we estimate the expected returns of each $\mathbf{c}_w$ as $\hat{R}_w = \hat{\mathbf{c}}_w' \hat{\boldsymbol{\mu}}$. It is reasonable for the influence of the sample means in these three portfolios to be the same.

As some studies, for example, Frankfurter et al. (1971), find the portfolio, $\hat{\mathbf{c}}_p$, to be less effective than an equally weighted portfolio, we include the estimates of the equally-weighted (EW) portfolio in the empirical illustrations for purposes of comparison. We denote the "equally weighted portfolio estimates" as $\hat{\mathbf{c}}_0 = \mathbf{1}/p$, repeat the procedure N times, and plot the results in Figures 3 to 6. In these figures, we denote $\hat{R}_w$ ($w = p, s, b, 0$) as the SC, PI, BC and EW returns, respectively. The line axes are the repeating time, and the pairs of means and standard deviations are reported for each type of expected return estimates.

The existing literature on the portfolio optimization theory shows that: (1) the plug-in estimates over estimate the theoretical expected return of the optimal portfolio; and (2) the plug-in estimates are likely not as effective as the equally-weighted estimates; as shown in Section 6, (3) bootstrap-corrected estimates under estimate the theoretical expected return of the optimal portfolio; as discussed in Section 2, (4) spectrally-corrected estimates provide consistent estimates for the theoretical expected return of the optimal portfolio. The results shown in the figures support the above findings that the plug-in optimal returns are the largest, while the bootstrap-corrected optimal returns are the smallest, with the equally-weighted and spectrally-corrected optimal returns lying in between.

In addition, we observe that the difference between the plug-in, bootstrap-corrected, equally-weighted, and spectrally-corrected optimal returns are small for $y = 0.1$ and 0.2, and increase for $y = 0.4$ and 0.6. The plug-in return is always larger than the other three estimates, and increases faster than the spectrally-corrected return as y increases. In order to compare variability, as expected, the sd of the equally weighted return is the smallest. On the other hand, when $y = 0.2$, the sd of the plug-in return is smaller than that of the spectrally-corrected return. However, when y increases, the sd of the spectrally-corrected estimate is smaller. In addition, from the figures, the bootstrap-corrected return is always less than zero, while the equally-weighted return is close to zero. All of these observations are consistent with the estimation theory of portfolio optimization and the simulations, as discussed in Sections 2 and 6.

8. CONCLUSION

The purpose of the paper was to solve the "Markowitz optimization enigma" by developing new covariance estimates to capture the essence of portfolio selection. By using large dimensional data analysis, we proved that the expected return of the plug-in allocation is always larger than that of the optimal portfolio in most situations when the number of assets is large. We note that Bai et al. (2009) proved a similar result under a much tighter condition, while in this paper we develop more general results under weaker conditions. For example, we proved that in certain situations, the expected return of the plug-in allocation is $\sqrt{\gamma} = \sqrt{1/(1-y)}$ times greater than that of the optimal portfolio while, in other situations, it is still greater than the optimal portfolio.

In the Markowitz MV portfolio optimization problem, the key issue is how to estimate the population covariance matrix accurately. In this paper, we introduced the spectrally-corrected covariance matrix to correct the sample covariance matrix, and derived important theoretical results. We constructed the spectrally-corrected covariance, $\widetilde{\mathbf{S}}_n$, as the estimate of the population covariance matrix, and provided the limiting behavior of $\mathbf{a}'\widetilde{\mathbf{S}}_n\mathbf{b}$ for different bounded vectors $\mathbf{a}$ and $\mathbf{b}$ when p goes to infinity, with n increasing proportionally. Our simulations demonstrated that $\mathbf{a}'\widetilde{\mathbf{S}}_n\mathbf{b}$ estimated $\mathbf{a}'\Sigma\mathbf{b}$ accurately.

According to the theory developed in the paper, we constructed the spectrally-corrected estimates, which performed more accurately than both the plug-in and the bootstrap-corrected estimates, not only for the expected return but also for risk. As our approach is easy to implement in practice, the efficient frontier of estimates can be constructed analytically. Thus, our proposed estimator facilitates the Markowitz MV optimization procedure, making it useful in practice. In addition, the essence of the portfolio analysis problem can be adequately captured by our proposed approach, which enhances the practical use of the Markowitz mean-variance optimization procedure.

We note that the optimal expected return estimate proposed in the paper not only represents the optimal expected return for the best combination of stocks, but also for the best combination of risk-free assets, bonds, stocks, and other assets. We note that normality is typically assumed in the MV optimization problem (see, for example, Leung et al. (2012), and the references cited therein for further information). However, in the proposed theory, we relax the normality assumption to allow for the existence of fourth moments, so that the proposed spectrally-corrected estimates could be obtained for the high-dimensional Markowitz MV portfolio optimization when the expected returns of the assets are derived under the existence of fourth moments.

Although we have developed several important theoretical results in the paper, there are further results for which we might conduct simulations. Further research could include developing such relationships theoretically. The theory developed in the paper could be applied to many related theories. For example, Korkie and Turtle (2002) established a theory for the optimal return of self-financing portfolios, for which the estimation approach developed in the paper might be extended.

The El Karoui (2008) algorithm of estimating the population eigenvalues of large dimensional covariance matrices, and the nonlinear shrinkage estimation of large dimensional covariance matrices and their inverses, developed in Ledoit and Wolf (2012), could be extended for some weaker conditions. Extensions could include incorporating their covariance estimates to develop new estimates for the high dimensional Markowitz MV portfolio optimization. Menchero et al. (2011) introduced a method called the eigen-adjusted covariance matrices, without using random matrix theory, and presented some simulation results showing its optimality versus that of alternative approaches. The theory developed in the paper improves their approach by incorporating random matrix theory into the adjustment of eigenvalues of the covariance matrices. Thus, our approach could obtain efficient estimates of the optimal return and its corresponding allocation that circumvent all four defects, namely the overprediction, underprediction and allocation estimation problem, as well as the problem of big risk in the Markowitz portfolio optimization.

Jacobs et al. (2005) argue that the model in (1.1), with c_i interpreted as a short position, is not a realistic model. They suggest that a realistic model of short constraints can be formulated as having $2n$ nonnegative "investments", with the first n being long positions and the second n being short positions. Thus formulated, it is a special case of what Markowitz (1959) (Chapter 8 and Appendix A) defines as the "general MV portfolio problem," namely, to find MV efficient portfolios subject to zero or more linear equality and/or (weak) inequality constraints. This could be considered an extension of the problem given in (1.1). Random matrix theory may not be able to solve this problem, but one could apply the least absolute shrinkage and selection operator (LASSO) (see Tibshirani (1996)) to solve the problem. This would be a good direction for purposes of extending the results in the paper.

9. APPENDIX

9.1. *Preliminaries*

Before the proof of Theorem 2.1, we introduce some notation and basic facts which will be used in the remaining parts.

Under Assumption I, let $\mathbf{r}_i = \frac{1}{\sqrt{n}}\mathbf{x}_i$, and $\widetilde{\mathbf{S}}_n = \sum_{i=1}^n \mathbf{r}_i\mathbf{r}_i'$. Denote $\widetilde{\mathbf{S}}_{n,i} = \widetilde{\mathbf{S}}_n - \mathbf{r}_i\mathbf{r}_i'$, $\widetilde{\mathbf{S}}_{n,ij} = \widetilde{\mathbf{S}}_n - \mathbf{r}_i\mathbf{r}_i' - \mathbf{r}_j\mathbf{r}_j'$ and $\delta_i = \mathbf{r}_i'\dot{\mathbf{S}}_{n,i}^{-1}\mathbf{r}_i - n^{-1}\operatorname{tr}\dot{\mathbf{S}}_{n,i}^{-1}$. Define:

$$\beta_j = \frac{1}{1+\mathbf{r}_j^T\widetilde{\mathbf{S}}_{n,j}^{-1}\mathbf{r}_j} \quad \text{and} \quad \beta_{ij} = \frac{1}{1+\mathbf{r}_j^T\widetilde{\mathbf{S}}_{n,ij}^{-1}\mathbf{r}_j},$$
$$\breve{\beta}_j = \frac{1}{1+n^{-1}\operatorname{tr}\widetilde{\mathbf{S}}_{n,j}^{-1}} \quad \text{and} \quad \breve{\beta}_{ij} = \frac{1}{1+n^{-1}\operatorname{tr}\widetilde{\mathbf{S}}_{n,ij}^{-1}},$$
$$b_n = \frac{1}{1+n^{-1}E\operatorname{tr}\widetilde{\mathbf{S}}_{n,1}^{-1}} \quad \text{and} \quad \breve{b}_n = \frac{1}{1+n^{-1}E\operatorname{tr}\widetilde{\mathbf{S}}_{n,12}^{-1}}.$$

Further, for any $p\times p$ symmetric matrix $\mathbf{A}$ and $\mathbf{v}\in\mathbb{R}^p$, the following two identities hold:

$$\mathbf{v}(\mathbf{A}+\mathbf{v}\mathbf{v}')^{-1} = \frac{\mathbf{v}'\mathbf{A}^{-1}}{1+\mathbf{v}'\mathbf{A}\mathbf{v}} \quad \text{and} \quad (\mathbf{A}+\mathbf{v}\mathbf{v}')^{-1} - \mathbf{A}^{-1} = -\frac{\mathbf{A}^{-1}\mathbf{v}\mathbf{v}'\mathbf{A}^{-1}}{1+\mathbf{v}'\mathbf{A}^{-1}\mathbf{v}} \tag{9.1}$$

(see (2.2) and Lemma 2.6 of Siverstein (1995)).

Lemma 9.1 *Theorem 2 in Bai and Yin (1993): Let $\mathbf{X} = [X_{u,v}; u = 1,...,p; v = 1,...,n]$ be a random matrix in which $X_{u,v}$s are i.i.d. random variables with zero mean and unit variance, and $\mathbf{S} = (1/n)\mathbf{X}\mathbf{X}'$. Then, if $E|X|^4 < \infty$, as $p, n\to\infty$, $p/n \to y \in (0,1)$,*

$$\lim \lambda_{min} = (1-\sqrt{y})^2 \quad \textit{and} \quad \lim \lambda_{max} = (1+\sqrt{y})^2,$$

where λ_{min} and λ_{max} are the smallest and largest eigenvalues of $\mathbf{S}$, respectively.

Lemma 9.2 *Lemma 2.1 of Bai and Silverstein (2004): Let $(X_i)_{i=1}^n$ be a complex martingale difference sequence with respect to an increasing σ-field $\{\mathscr{F}_i\}$. Then, for any $k > 1$:*

$$E\left|\sum_{i=1}^{\ell} X_i\right|^k \le KE\left(\sum_{i=1}^n |X_i|^2\right)^{k/2}.$$

Lemma 9.3 *Lemma 2.7 of Bai and Silverstein (1998) Suppose $\mathbf{x} = (x_1,...,x_p)'$, where x_j's are i.i.d. random variables with zero mean and unit variance, and $\mathbf{B}$ is a deterministic $n\times n$ matrix. Then for any $\alpha\ge 2$, we have:*

$$E|\mathbf{x}'\mathbf{B}\mathbf{x} - \operatorname{tr}\mathbf{B}|^\alpha \le K_\alpha\left(\left(E|x_1|^4\operatorname{tr}(\mathbf{B}^2)\right)^{\alpha/2} + E|x_1|^{2\alpha}\operatorname{tr}(\mathbf{B}^\alpha)\right).$$

Lemma 9.4 *Lemma 2.3 in Bai and Silverstein (2004): Let $f_n(\cdot), n = 1,2,\cdots$, be analytic in D, a connected open set of $\mathbb{C}$, satisfying $|f_n(z)| \le M$ for every n and z in D, and $f_n(z)$ converges* *for each z in a subset of D having a limit point in D. Then there exists a function f analytic in D, such that $f_n(z) \to f(z)$ and $f_n'(z) \to f'(z)$ for all $z \in D$. Moreover, on any set bounded by a contour interior to D, the convergence is uniform and $\{f_n'(z)\}$ is uniformly bounded by $2M/\tilde{\varepsilon}$, where $\tilde{\varepsilon}$ is the distance between the contour and the boundary of D.*

LEMMA 9.5 *Theorem 1.1 in Bai and Silverstein (1998): Under Assumption (I), assume F^{Σ_p} converges to a given distribution function H (see (4.3)). Then for any interval $[a, b]$ $(a > 0)$ lying outside the support of H, we have:*

$$P\,(\textit{no eigenvalues of } \mathbf{T}_n \textit{ appears in } [a,b] \textit{ for all } p) = 1,$$

in which $\mathbf{T}_n = \frac{1}{n}\Sigma^{1/2}\mathbf{Z}_p\mathbf{Z}_p'\Sigma^{1/2}$.

LEMMA 9.6 *Theorem 1.2 Bai and Silverstein (1999): Under Assumption (I), assume F^{Σ_p} converges to a given distribution function H. Then if $[a, b]$ $(a > 0)$ lying outside the support of H and not contained in $[0, x_0]$, where x_0 is the greatest lower bound of Θ, we have:*

$$P\left(\lambda_{i_n}^{\mathbf{S}_n} > b \ \textit{ and } \ \lambda_{i_n+1}^{\mathbf{S}_n} < a \ \ \textit{for all large} \ \ n\right) = 1,$$

in which i_n is satisfied such that:

$$\lambda_{i_n}^{\Sigma_p} > -1/m(b) \quad \textit{and} \quad \lambda_{i_n+1}^{\Sigma_p} < -1/m(a),$$

in which m is the unique solution of (2.8).

9.2. *Proof of Theorem 2.1*

Part I: In this part, we prove $\left|\mathbf{a}_p'\mathbf{S}_n^{-1}\mathbf{b}_p - \mathbf{a}_p'\Sigma_p^{-1}\mathbf{b}_p/(1-y)\right| \to 0$ in probability. Without loss of generality, supposing $\boldsymbol{\mu} = 0$, we only need to prove the following two results:

$$\left|\mathbf{a}_p'\mathbf{S}_n^{-1}\mathbf{b}_p - \frac{n-1}{n}\mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p\right| \to 0 \ \text{ and } \ \left|\mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p - \frac{1}{1-y}\mathbf{a}_p'\Sigma_p^{-1}\mathbf{b}_p\right| \to 0 \tag{9.2}$$

in probability, where $\dot{\mathbf{S}}_n = \frac{1}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i'$.

According to (9.1), rewrite:

$$\left|\mathbf{a}_p'\mathbf{S}_n^{-1}\mathbf{b}_p - \frac{n-1}{n}\mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p\right| = \frac{n-1}{n}\left|\frac{\mathbf{a}_p\dot{\mathbf{S}}_n^{-1}\overline{\mathbf{x}}\cdot\overline{\mathbf{x}}'\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p}{1+\overline{\mathbf{x}}'\dot{\mathbf{S}}_n^{-1}\overline{\mathbf{x}}}\right| \le \left|\mathbf{a}_p\dot{\mathbf{S}}_n^{-1}\overline{\mathbf{x}}\right|\cdot\left|\overline{\mathbf{x}}'\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p\right|.$$

Then the first condition in (9.2) is proved only if $E\left(\mathbf{a}_p\widetilde{\mathbf{S}}_n^{-1}\overline{\mathbf{x}}\right)^2 = 0$.

By rewriting $\overline{\mathbf{x}} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{r}_i$ and (9.1), we have:

$$E\left(\mathbf{a}_p\widetilde{\mathbf{S}}_n^{-1}\overline{\mathbf{x}}\right)^2 = \frac{1}{n}\sum_{i=1}^n E\left(\beta_i\mathbf{a}_p\widetilde{\mathbf{S}}_{n,i}^{-1}\mathbf{r}_i\right)^2 + \frac{1}{n}\sum_{i\neq j} E\left(\beta_i\beta_j\mathbf{a}_p'\dot{\mathbf{S}}_{n,i}^{-1}\mathbf{r}_i\mathbf{r}_j'\dot{\mathbf{S}}_{n,j}^{-1}\mathbf{a}_p\right). \tag{9.3}$$

From Lemma 9.1 and Assumption (I), we have:

$$E\left(\beta_i\mathbf{a}_p\dot{\mathbf{S}}_{n,i}^{-1}\mathbf{r}_i\right)^2 \le \frac{1}{n}E\left(\mathbf{a}_p'\dot{\mathbf{S}}_{n,i}^{-1}\Sigma_p\dot{\mathbf{S}}_{n,i}^{-1}\mathbf{a}_p\right) \le O(n^{-1}).$$

Now we only need consider the second part of (9.3). Rewriting $\beta_i = \breve{\beta}_i - \breve{\beta}_i^2\delta_i$, we have:

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{i\neq j} E\left(\beta_i\beta_j \mathbf{a}_p'\dot{\mathbf{S}}_{n,i}^{-1}\mathbf{r}_i\mathbf{r}_j'\dot{\mathbf{S}}_{n,j}^{-1}\mathbf{a}_p\right)\right| \\
\leq\ & \frac{1}{n}\left|\sum_{i\neq j}\breve{\beta}_i\breve{\beta}_j E\left(\mathbf{a}_p'\dot{\mathbf{S}}_{n,i}^{-1}\mathbf{r}_i\mathbf{r}_j'\dot{\mathbf{S}}_{n,j}^{-1}\mathbf{a}_p\right)\right| + \frac{1}{n}\left|\sum_{i\neq j} E\left((\beta_i+\breve{\beta}_i)\breve{\beta}_j^2\delta_j\mathbf{a}_p'\dot{\mathbf{S}}_{n,i}^{-1}\mathbf{r}_i\mathbf{r}_j'\dot{\mathbf{S}}_{n,j}^{-1}\mathbf{a}_p\right)\right| := A_1 + A_2.
\end{aligned}$$

By the Cauchy-Schwarz inequality and (9.1), we have:

$$\begin{aligned}
A_1 &\leq \frac{1}{n}\left|\sum_{i\neq j}\breve{\beta}_i\breve{\beta}_j E\left(\mathbf{a}_p'\left(\dot{\mathbf{S}}_{n,i}^{-1} - \dot{\mathbf{S}}_{n,ij}^{-1}\right)\mathbf{r}_i\mathbf{r}_j'\left(\dot{\mathbf{S}}_{n,j}^{-1} - \dot{\mathbf{S}}_{n,ij}^{-1}\right)\mathbf{b}_p\right)\right| \\
&\leq \frac{1}{n}\sum_{i\neq j} E\left|\mathbf{a}_p'\dot{\mathbf{S}}_{n,ij}\mathbf{r}_j(\mathbf{r}_j'\dot{\mathbf{S}}_{n,ij}^{-1}\mathbf{r}_i)^2\mathbf{r}_i'\dot{\mathbf{S}}_{n,ij}^{-1}\mathbf{b}_p\right| \\
&\leq \frac{1}{n}\sum_{i\neq j}\left(E\left(\mathbf{a}_p'\dot{\mathbf{S}}_{n,ij}^{-1}\mathbf{r}_j\right)^4\right)^{1/4}\left(E\left(\mathbf{r}_j'\dot{\mathbf{S}}_{n,ij}^{-1}\mathbf{r}_i\right)^4\right)^{1/2}\left(E\left(\mathbf{b}_p'\dot{\mathbf{S}}_{n,ij}^{-1}\mathbf{r}_i\right)^4\right)^{1/4}.
\end{aligned}$$

By Lemmas 9.1 and 9.3, we can deduce for any $\alpha \geq 4$:

$$E\left(\mathbf{a}_p'\dot{\mathbf{S}}_{n,i}^{-1}\mathbf{r}_i\right)^{\alpha} = O(n^{-\alpha/2}) \text{ and } E\left(\mathbf{r}_j'\dot{\mathbf{S}}_{n,ij}^{-1}\mathbf{r}_i\right)^{\alpha} = O(n^{-\alpha/2}). \tag{9.4}$$

Therefore, $A_1 = O(n^{-1})$.

For A_2, compute:

$$\begin{aligned}
&\frac{1}{n}\left|\sum_{i\neq j} E\left((\beta_i+\breve{\beta}_i)\breve{\beta}_j^2\delta_j\mathbf{a}_p'\dot{\mathbf{S}}_{n,i}^{-1}\mathbf{r}_i\mathbf{r}_j'\dot{\mathbf{S}}_{n,j}^{-1}\mathbf{a}_p\right)\right| \\
\leq\ & \frac{1}{n}\sum_{i\neq j} E\left|\mathbf{a}_p'\left(\dot{\mathbf{S}}_{n,i}^{-1} - \dot{\mathbf{S}}_{n,ij}^{-1}\right)\mathbf{r}_i\mathbf{r}_j'\left(\dot{\mathbf{S}}_{n,j}^{-1} - \dot{\mathbf{S}}_{n,ij}^{-1}\right)\mathbf{a}_p\delta_i\right| \\
& + \frac{2}{n}\sum_{i\neq j} E\left|\mathbf{a}_p'\dot{\mathbf{S}}_{n,ij}^{-1}\mathbf{r}_i\mathbf{r}_j'\left(\dot{\mathbf{S}}_{n,j}^{-1} - \dot{\mathbf{S}}_{n,ij}^{-1}\right)\mathbf{a}_p\delta_i\right| + \frac{1}{n}\sum_{i\neq j} E\left|\mathbf{a}_p'\dot{\mathbf{S}}_{n,ij}^{-1}\mathbf{r}_i\mathbf{r}_j'\dot{\mathbf{S}}_{n,ij}^{-1}\mathbf{a}_p\delta_i\right| \\
:=\ & A_{21} + A_{22} + A_{23}.
\end{aligned}$$

From the Cauchy-Schwarz inequality and Lemma 9.1, we have:

$$A_{21} = O(n^{-1}), \quad A_{22} = O(n^{-1}) \text{ and } A_{23} = O(n^{-1}),$$

that is, $A_2 = O(n^{-1})$. Further according to (9.3), we have (9.2).

Now we focus on the limit of $\mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p$. Rewriting:

$$\mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p = \frac{1}{4}\left[\left(\mathbf{a}_p + \mathbf{b}_p\right)'\dot{\mathbf{S}}_n^{-1}\left(\mathbf{a}_p + \mathbf{b}_p\right) - \left(\mathbf{a}_p - \mathbf{b}_p\right)'\dot{\mathbf{S}}_n^{-1}\left(\mathbf{a}_p - \mathbf{b}_p\right)\right],$$

we consider the limit of $\mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\mathbf{a}_p$. According to Theorem 1 in Bai et al. (2007), we have:

$$\left|\mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\mathbf{a}_p - \frac{1}{1-y}\mathbf{a}_p'\Sigma_p^{-1}\mathbf{a}_p\right| \to 0$$

in probability.

Part II: In this part, we prove $\left|\mathbf{a}_p'\mathbf{S}_n^{-1}\Sigma_p\mathbf{S}_n^{-1}\mathbf{b}_p - \mathbf{a}_p'\Sigma_p^{-1}\mathbf{b}_p/(1-y)^3\right| \to 0$ in probability. From (9.2) and Lemma 9.1, we have:

$$\left|\mathbf{a}_p'\mathbf{S}_n^{-1}\Sigma_p\mathbf{S}_n^{-1}\mathbf{b}_p - \mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\Sigma_p\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p\right| \leq K\left|\mathbf{a}_p'\mathbf{S}_n^{-1}\mathbf{b}_p - \mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p\right|\left\|\mathbf{S}_n^{-1}\right\|\left\|\dot{\mathbf{S}}_n^{-1}\right\|\left\|\Sigma_p\right\| \to 0$$

in probability. Write:

$$\begin{aligned}\mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\Sigma_p\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p &= \lim_{v\to 0}\mathbf{a}_p'\Sigma_p^{-1/2}\left(\Sigma_p^{-1/2}\dot{\mathbf{S}}_n^{-1}\Sigma_p^{-1/2} - v\cdot iI_p\right)^{-2}\Sigma_p^{-1/2}\mathbf{a}_p \\ &= \lim_{v\to 0}\frac{d}{i\cdot dv}\left(\mathbf{a}_p'\Sigma_p^{-1/2}\left(\Sigma_p^{-1/2}\dot{\mathbf{S}}_n^{-1}\Sigma_p^{-1/2} - v\cdot iI_p\right)^{-1}\Sigma_p^{-1/2}\mathbf{a}_p\right) \\ &:= \lim_{v\to 0}\frac{d}{i\cdot dv}f_p(v).\end{aligned}$$

According to Theorem 2 in Bai et al. (2011), we have:

$$f_p(v) - m(vi)\mathbf{a}_p'\Sigma_p^{-1}\mathbf{a}_p \to 0 \;\; a.s.,$$

in which $m(z)$ is the Stieltjes transform of the MP-Law (see Marcenko and Pastur (1967)). From Lemmas 9.1 and 9.4, we have:

$$f_p'(v) - m'(vi)\mathbf{a}_p'\Sigma_p^{-1}\mathbf{a}_p \to 0 \;\; a.s.$$

Since $f_p'(v)$ and $m'(vi)$ are continuous about $v \in [0, \varepsilon]$ for small enough $\varepsilon > 0$ with probability 1, then according to the Lemma in Bai and Yin (1993) and the dominant convergence theorem, we have:

$$\mathbf{a}_p'\dot{\mathbf{S}}_n^{-1}\Sigma_p\dot{\mathbf{S}}_n^{-1}\mathbf{b}_p + \left.\frac{d(im(vi))}{dv}\right|_{v=0}\mathbf{a}_p'\Sigma_p^{-1}\mathbf{a}_p \to 0$$

in probability. Here:

$$\left.\frac{d(m(vi))}{dv}\right|_{v=0} = \frac{i}{2\pi y}\int_a^b x^{-3}\sqrt{(b-x)(x-a)}dx = \frac{i}{(1-y)^3},$$

in which $a = (1-\sqrt{y})^2$ and $b = (1+\sqrt{y})^2$. Part II is now finished. □

9.3. *Proof of Theorem 4.2*

Since Θ is the splitting support of $F^{\boldsymbol{\tau},\mathbf{w}}$, there exists $\mathbf{t} = \{t_0, t_1, ..., t_L\}$ such that $\mathbf{t}\cap\Theta = \emptyset$ and $\cup_{i=1}^L(t_{i-1}, t_i)\cap\Theta = \Theta$. Now rewrite (4.1) as:

$$\mathbf{S}_n = \mathbf{U}_{1,p_1}\Lambda_{1,p_1}\mathbf{U}_{1,p_1} + \cdots + \mathbf{U}_{L,p_L}\Lambda_{L,p_L}\mathbf{U}_{L,p_L},$$

in which Λ_{i,p_i} is the i-th $p_i\times p_i$ diagonal matrix of Λ_p satisfying $\Lambda_p = \mathrm{diag}(\Lambda_{1,p_1}, ..., \Lambda_{L,p_L})$, and $\mathbf{U}_i$ is the corresponding eigenvectors matrix, satisfying $U_p = (U_{1,p_1}, ..., U_{L,p_L})$ $(i = 1, ..., L)$. We can obtain, from Lemma 9.5, for large p:

$$P\left(\text{eigenvalues of } \Lambda_{i,p_i} \text{ belongs to } (t_{i-1}, t_i)\right) = 1 \quad (i = 1, ..., L).$$

Then for large enough p, we have:

$$\mathbf{a}_p'U_{i,p_i}U_{i,p_i}\mathbf{b}_p = F^{\mathbf{S}_n}(t_i) - F^{\mathbf{S}_n}(t_{i-1}) \;\; (i = 1, ..., L),$$

in which $F^{\mathbf{S}_n}(t_0) = 0$ and $F^{\mathbf{S}_n}(t_L) = 1$. Thus:

$$\begin{aligned}\mathbf{a}_p{}'\widetilde{\mathbf{S}}_{\mathbf{n}}^{-1}\mathbf{b}_p &= \frac{1}{\lambda_1}\mathbf{a}_p'\mathbf{U}_{1,p_1}\mathbf{U}_{1,p_1}'\mathbf{b}_p + \frac{1}{\lambda_2}\mathbf{a}_p'\mathbf{U}_{2,p_2}\mathbf{U}_{2,p_2}'\mathbf{b}_p + ... + \frac{1}{\lambda_L}\mathbf{a}_p'\mathbf{U}_{L,p_L}\mathbf{U}_{p_L}'\mathbf{b}_p \\ &= \sum_{j=1}^L\frac{1}{\lambda_j - z}\int_{t_{j-1}}^{t_j}dF^{\mathbf{S}_n}(x) \to \sum_{j=1}^L\frac{1}{\lambda_j - z}\int_{t_{j-1}}^{t_j}dF^{\boldsymbol{\tau},\mathbf{w}}(x) \quad \text{a.s.}\end{aligned}$$

Denote $m(z) = \int_\Theta (t-z)^{-1} dF^{\boldsymbol{\tau},\mathbf{w}}$ as the Stieltjes transform of $F^{\boldsymbol{\tau},\mathbf{w}}$, where $m(z)$ is the unique solution of (2.8) for $z \in \mathbb{C}^+$, with $H(x) = \sum_{i=1}^L w_i I_{\tau_i \le x}$. Then we have:

$$\mathbf{a}_p{}'\widetilde{\mathbf{S}}_{\mathbf{n}}^{-1}\mathbf{b}_p \quad\to\quad \sum_{j=1}^{L}\frac{1}{\lambda_j}\left(-\frac{1}{2\pi i}\oint_{\mathscr{C}_j} m(z)dz\right) \quad \text{a.s.},$$

in which $\mathscr{C}_j$ is the min complex open set that includes the real set from t_{j-1} to t_j. In addition, supposing that $u = -\left(-\frac{1-y}{z} + ym(z)\right)$, we have:

$$\mathbf{a}_p{}'\widetilde{\mathbf{S}}_{\mathbf{n}}^{-1}\mathbf{b}_p \to \sum_{jk}\frac{d_k}{\lambda_j}\frac{1}{2\pi i}\oint_{\Gamma_j}\frac{1-y\int\frac{t^2dH(t)}{(u-t)^2}}{\left(1+y\int\frac{tdH(t)}{u-t}\right)(u-\lambda_k)}du,$$

where $H(x) = \sum_{i=1}^L w_i I_{\tau_i \le x}$ and Γ_j is a contour of the image of $\mathscr{C}_j$ by u. We note that, for each z with $\Im(z) \neq 0$, there is a unique solution to (2.8) whose imaginary part has the same sign as z. Therefore, the contour Γ_j is well defined. According to the Residue Theorem, we have:

$$\frac{1}{2\pi i}\oint_{\Gamma_j}\frac{1-y\int\frac{t^2dH(t)}{(u-t)^2}}{\left(1+y\int\frac{tdH(t)}{u-t}\right)(u-\lambda_k)}du = \begin{cases} -\frac{\lambda_j}{\lambda_j-\lambda_k}+\frac{u_j}{u_j-\lambda_k}, & k \neq j, \\ \frac{u_j}{u_j-\lambda_j}+\frac{1}{yc_j}\left(1+y\sum_{t\neq j}\frac{c_t\lambda_t}{\lambda_j-\lambda_t}\right), & k = j. \end{cases}$$

Finally, we have:

$$\mathbf{a}_p{}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{b}_p \to \sum_{j=1}^{L}\frac{d_j}{\lambda_j} + \sum_{k=1}^{L}\left(\sum_{j\neq k}\frac{d_k\lambda_k(\lambda_j-u_j)}{\lambda_j(u_j-\lambda_k)(\lambda_j-\lambda_k)} - \frac{d_k}{\lambda_k}\sum_{j\neq k}\frac{c_j\lambda_j(\lambda_j-u_k)}{c_k(\lambda_k-\lambda_j)(u_k-\lambda_j)}\right). \tag{9.5}$$

Let $\|\tilde{\mathbf{a}}_p\| = 1$ and $\tilde{\mathbf{a}}_p'\mathbf{U}_{p_k}\mathbf{U}_{p_k}'\tilde{\mathbf{a}}_p = 1$, we deduce:

$$1 \quad = \quad \sum_{j\neq k}\frac{\lambda_k(\lambda_j-u_j)}{(\lambda_j-\lambda_k)(u_j-\lambda_k)} + 1 - \frac{1}{c_k}\sum_{t\neq k}\frac{c_t\lambda_t(\lambda_k-u_k)}{(\lambda_k-\lambda_t)(u_k-\lambda_t)},$$

by setting $\Delta_p = \mathbf{I}$, and so:

$$\sum_{j\neq k}\frac{c_j\lambda_j(\lambda_k-u_k)}{c_k(\lambda_k-\lambda_j)(u_k-\lambda_j)} = \sum_{j\neq k}\frac{\lambda_k(\lambda_j-u_j)}{(\lambda_j-\lambda_k)(u_j-\lambda_k)}, \qquad \text{for all } k. \tag{9.6}$$

From (9.5) and (9.6), we have:

$$\mathbf{a}_p{}'\widetilde{\mathbf{S}}_{\mathbf{n}}^{-1}\mathbf{b}_p = \sum_{k=1}^{L} d_k \sum_{j=1}^{L}\frac{(u_j-\lambda_j)}{\lambda_j(u_j-\lambda_k)}\ .$$

This completes the proof of Theorem 4.2. □

9.4. *Proof of Theorem 4.4*

We first consider the case of $\mathbf{a}_p = \mathbf{b}_p = \mathbf{x}_p$. Then rewrite:

$$\begin{aligned}\mathbb{I}_p &= \mathbf{x}_p{}'\widetilde{\mathbf{S}}_n^{-1}\Sigma_p\widetilde{\mathbf{S}}_n^{-1}\mathbf{x}_p \\ &= \sum_{k=1}^{L}\tau_k\mathbf{x}_p{}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{U}_{p_k}\mathbf{U}_{p_k}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{x}_p \\ &= \sum_{k=1}^{L}\tau_k\left\|\mathbf{U}_{p_k}\mathbf{U}_{p_k}'\widetilde{\mathbf{S}}_n^{-1}\mathbf{x}_p\right\|^2\end{aligned}$$

$$= \sum_{k=1}^{L} \tau_k \left(\sup_{\mathbf{y}_{p,k} \in E_k; \|\mathbf{y}_{p,k}\|=1} \mathbf{x}_p' \widetilde{\mathbf{S}}_n^{-1} \mathbf{y}_{p,k} \right)^2,$$

where E_k is the kth eigenspace of Σ_p associated with τ_k. Then we have:

$$\liminf \mathbb{I}_p \geq \sum_{k=1}^{L} \tau_k \liminf \left(\mathbf{x}_p' \widetilde{\mathbf{S}}_n^{-1} \mathbf{y}_{p,k} \right)^2,$$

for any sequence of vectors $\{\mathbf{y}_{p,k}\}$. Select a special sequence of vectors $\{\mathbf{y}_{p,k}\}$ such that:

$$\lim \mathbf{x}_p' \mathbf{U}_{p_k} \mathbf{U}_{p_k}' \mathbf{y}_{p,k} = c_k.$$

Then, according to Theorem 4.2, we have:

$$\lim \left(\mathbf{x}_p' \widetilde{\mathbf{S}}_n^{-1} \mathbf{y}_{p,k} \right)^2 = c_k^2 \left(\sum_{j=1}^{L} \frac{u_j - \lambda_j}{\lambda_j (u_j - \lambda_j)} \right)^2.$$

Since $\mathbf{y}_{p,k} \in E_k$ and $\|\mathbf{y}_{p,k}\| = 1$ for all p and k, we have $c_k \in \left[-\sqrt{d_k},\ \sqrt{d_k} \right]$. Then:

$$\liminf \mathbb{I}_p \geq \sum_{k=1}^{L} d_k \lambda_k \left(\sum_{j=1}^{L} \frac{u_j - \lambda_j}{\lambda_j (u_j - \lambda_j)} \right)^2.$$

As the subset of unit vectors in E_k is compact, for each $\varepsilon > 0$, there exists a unit vector $\tilde{\mathbf{y}}_{p,k}$ such that:

$$\sup_{\mathbf{y}_{p,k} \in E_k; \|\mathbf{y}_{p,k}\|=1} \left(\mathbf{x}_p' \mathbf{B}_p^{-1} \mathbf{y}_{p,k} \right)^2 \leq \left(\mathbf{x}_p' \mathbf{B}_p^{-1} \tilde{\mathbf{y}}_{p,k} \right)^2 + \varepsilon.$$

Let $\tilde{c}_k = \lim \mathbf{x}_p' \mathbf{U}_{p_k} \mathbf{U}_{p_k}' \tilde{\mathbf{y}}_{p,k}$. Then:

$$\begin{aligned}
& \limsup \sup_{\mathbf{y}_{p,k} \in E_k; \|\mathbf{y}_{p,k}\|=1} \left(\mathbf{x}_p' \mathbf{B}_p^{-1} \mathbf{y}_{p,k} \right)^2 \\
\leq\ & \limsup \left(\mathbf{x}_p' \mathbf{B}_p^{-1} \tilde{\mathbf{y}}_{p,k} \right)^2 + \varepsilon \\
=\ & \tilde{c}_k^2 \sum_{j=1}^{L} \frac{(u_j - \lambda_j)}{\lambda_j (u_j - \lambda_k)} + \varepsilon \\
\leq\ & d_k \sum_{j=1}^{L} \frac{(u_j - \lambda_j)}{\lambda_j (u_j - \lambda_k)} + \varepsilon,
\end{aligned}$$

that is:

$$\limsup \mathbb{I}_p = \sum_{k=1}^{L} d_k \lambda_k \sum_{j=1}^{L} \frac{u_j - \lambda_j}{\lambda_j (u_j - \lambda_j)}.$$

The general case is obtained by applying this result to the "squares":

$$(\mathbf{a}_p + \mathbf{b}_p)' \mathbf{B}_p^{-1} \Sigma \mathbf{B}_p^{-1} (\mathbf{a}_p + \mathbf{b}_p), \quad (\mathbf{a}_p - \mathbf{b}_p)' \mathbf{B}_p^{-1} \Sigma \mathbf{B}_p^{-1} (\mathbf{a}_p - \mathbf{b}_p), \tag{9.7}$$

and using the parallelogram law:

$$\mathbf{a}_p' \mathbf{B}_p^{-1} \Sigma \mathbf{B}_p^{-1} \mathbf{b}_p = \frac{1}{4} \left((\mathbf{a}_p + \mathbf{b}_p)' \mathbf{B}_p^{-1} \Sigma \mathbf{B}_p^{-1} (\mathbf{a}_p + \mathbf{b}_p) - (\mathbf{a}_p - \mathbf{b}_p)' \mathbf{B}_p^{-1} \Sigma \mathbf{B}_p^{-1} (\mathbf{a}_p - \mathbf{b}_p) \right). \tag{9.8}$$

The proof of Theorem 4.4 is complete. □

REFERENCES

Bai, Z.D., Chen, J. and Yao, J.F. (2010). *On estimation of the population spectral distribution from a high-dimensional sample covariance matrix*. Australian & New Zealand Journal of Statistics, **52** 423-437.

Bai, Z.D., Liu, H.X. and Wong, W.K. (2009). *Enhancement of the Applicability of Markowitz's Portfolio Optimization by Utilizing Random Matrix Theory*. Mathematical Finance, **19(4)** 639-667.

Bai, Z.D., Liu, H.X. and Wong, W.K. (2011). *Asymptotic properties of eigenmatrices of a large sample covariance matrix*. Annals of Applied Probability, 1994-2015.

Bai, Z.D., Miao, B.Q. and Pan, G.M. (2007). *On the asymptotics of eigenvectors of large sample covariance matrix*. Annals of Probability, **35(4)** 1532-1572.

Bai, Z.D. and Silverstein, J.W. (2004), *CLT for linear spectral statistics of large-dimensional sample covariance matrices.* Annals of Probability, **32** 553–605.

Bai, Z.D. and Silverstein, J.W. (1998) *No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices.* Annals of Probability, 316-345.

Bai, Z.D. and Silverstein, J.W.(1999) *Exact separation of eigenvalues of large dimensional sample covariance matrices.* Annals of Probability, 1536-1555.

Bai, Z.D. and Yin, Y.Q. (1993) *Limit of the smallest eigenvalue of a large dimensional sample covariance matrix.* Annals of Probability,1275–1294.

Bickel, P. J. and Levina, E. (2008). *Regularized estimation of large covariance matrices.* Annals of Statistics, **36** 199-227.

Cass, David and Stiglitz, Joseph E. (1970). *The structure of investor preferences and asset returns, and separability in portfolio allocation: A contribution to the pure theory of mutual funds.*

Cai, T. Tony and Zhou, Harrison H. (2012). *Minimax estimation of large covariance matrices under l1 norm.* Statistica Sinica, **22(4)**, 1319–1349.

El Karoui, N. (2008). *Spectrum estimation for large dimensional covariance matrices using random matrix theory.* Annals of Statistics, **36(6)** 2757-2790.

Fan, J., Fan, Y. and Lv, J. (2008). *High dimensional covariance matrix estimation using a factor model.* Journal of Econometrics, *147* 186-197.

Frankfurter, G.M., Phillips, H.E. and Seagle, J.P. (1971). *Portfolio selection: The effects of uncertain means, variances and covariances*. Journal of Financial and Quantitative Analysis, **6** 1251-1262.

Green, J., Hand, J.R.M. and Zhang, X.F. (2013). *The supraview of return predictive signals*. Review of Accounting Studies, **18** 692-730.

Jacobs, B.J., Levy, K.N. and Markowitz, H.M. (2005). *Portfolio optimization with factors, scenarios, and realistic short positions*. Operations Research, **53**, 586-599.

Jorion, P. (1985). *International portfolio diversification with estimation risk*. Journal of Business, **58(3)** 259-278.

Konno, H. and Yamazaki, H. (1991). *Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock arket*. Management Science, **37(5)** 519-531.

Korkie and Turtle, H.J. (2002). *A mean-variance analysis of self-financing portfolios*. Management Science, **48(3)** 427-443.

Kroll, Y. , Levy, H. and Markowitz, H.M. (1984). *Mean-variance versus direct utility maximization*. Journal of Finance, **39(1)** 47-61.

Laloux, L., Cizeau, P., Bouchaud, J.P. and Potters, M. (1999). *Noise dressing of financial correlation matrices*. Physical Review Letters, **83** 1467-1470.

Ledoit, O. and Wolf, M. (2004). *A well-conditioned estimator for large-dimensional covariance matrices*. Journal of Multivariate Analysis, **88(2)** 365-411.

Ledoit, O. and Wolf, M. (2012). *Nonlinear shrinkage estimation of large-dimensional covariance matrices*. Annals of Statistics, **40(2)** 1024-1060.

Leung, P.L., Ng, H.Y. and Wong, W.K. (2012). *An improved estimation to make Markowitz's portfolio optimization theory users friendly and estimation accurate with application on the US stock market investment*. European Journal of Operational Research, **222** 85-95.

Li, W.M., Chen, J.Q., Qin, Y.L., Yao, J.F. and Bai, Z.D. (2013). *Estimation of the population spectral distribution from a large dimensional sample covariance matrix*. Journal of Statistical Planning and Inference, **143(11)** 1887-1897.

Li, W.M. and Yao, J.F. (2014). *A local moment estimator of the spectrum of a large dimensional covariance matrix*. Statistica Sinica, **24** 919-936.

Khare, K. and Rajaratnam, B. (2011). *Wishart distributions for decomposable covariance graph models.* Annals of Statistics, **39** 514-555.

Marcenko, V.A. and Pastur, L.A. (1967). *Distribution for some sets of random matrices*. Mathematics of the USSR-Sbornik, 1(4), 457-483.

Markowitz, H.M. (1952). *Portfolio Selection*. Journal of Finance, **7** 77-91.

Markowitz, H.M. (1959). *Portfolio Selection*. New York: Wiley.

Markowitz, H.M. and Perold, A.F. (1981). *Portfolio analysis with factors and scenarios*, Journal of Finance, **36(4)**, 871-877.

Markowitz, H.M. (1991). *Foundations of Portfolio Theory*. Journal of Finance, **46(2)** 469-77.

Mestre, X. (2008). *Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates.* Information theory, IEEE Transactions on, **54(11)** 5113-5129.

Menchero, J., Wang, J. and Orr, D.J. (2011). *Eigen-adjusted covariance matrices*. MSCI Barra Research Paper, No. 2011-14.

Merton, R.C.(1972) *An analytic derivation of the efficient portfolio frontier*. Journal of Financial and Quantitative Analysis, **7(04)** 1851-1872.

Michaud, R (1989). *The Markowitz optimization enigma: Is optimized optimal?*, Financial Analysts Journal, **45** 31-42.

Rajaratnam, B., Massam, H. and Carvalho, C.M. (2008). *Flexible covariance estimation in graphical Gaussian models.* Annals of Statistics, **36** 2818-2849.

Rao, N.R., Mingo, J.A., Speicher, R. and Edelman, A. (2008). *Statistical eigen-inference from large Wishart matrices.* Annals of Statistics, **36**, 2850-2885.

Ravikumar, P.,Wawinwright, M., Raskutti, G. and Yu, B. (2008). *High-dimensional covariance estimation by minimizing 1-penalized log-determinant divergence* Technical Report 797, Department of Statistics, University of California, Berkeley.

Rohde, A. and Tsybakov, A.B. (2011). *Estimation of high-dimensional low-rank matrices.* Annals of Statistics, **39(2)** 887-930.

Simaan, Y. (1997). *Estimation risk in portfolio selection: The mean variance model versus the mean absolute deviation model*. Management Science, **43(10)** 1437-1446.

Siverstein, J.W. (1995). *Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices*. Journal of Multivariate Analysis, **5** 331-339.

Silverstein, J.W. and Bai, Z.D. (1995). *On the empirical distribution of eigenvalues of a class of large dimensional random matrices*. Journal of Multivariate Analysis, **54(2)** 175-192.

Siverstein, J.W. (1990). *Weak convergence of random functions defined by the eigenvectors of sample covariance matrices*. Annals of Probability, **8(18)** 1174-1193.

Siverstein, J.W. (1989). *On the eigenvectors of large-dimensional sample covariance matrices*. Journal of Multivariate Analysis, **30** 1-16.

Siverstein, J.W. (1984). *Some limit theorems on the eigenvectors of large-dimensional sample covariance matrices*. Journal of Multivariate Analysis, **15** 295-324.

Tibshirani, R. (1996). *Regression shrinkage and selection via the Lasso*. Journal of the Royal Statistical Society, Series B, **58** 267-288.

Wong, F., Carter, C.K. and Kohn, R. (2003). *Efficient estimation of covariance selection models*. Biometrika, **90** 809-830.

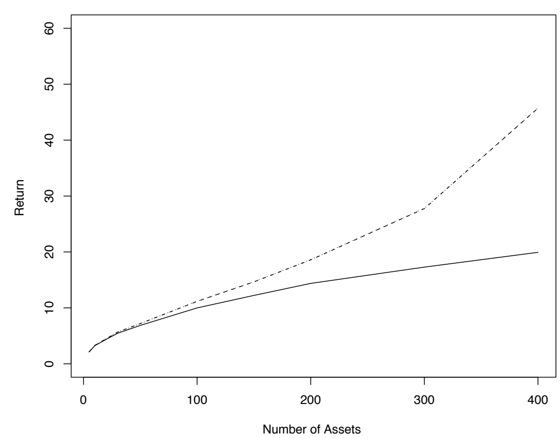
Yao, Jianfeng, Kammoun, Abla and Najim, Jamal.(2012). *Eigenvalue estimation of parameterized covariance matrices of large dimensional data.* IEEE Transactions on,Signal Processing. **60(11)** 5893-5905.

TABLE I: Risk of plug-in allocation estimates and bootstrap-corrected allocation estimates for different values of p and p/n.

p	p/n	$risk_{\mathbf{c}}^{p}$	$risk_{\mathbf{c}}^{b}$	p	p/n	$risk_{\mathbf{c}}^{p}$	$risk_{\mathbf{c}}^{b}$
100	0.5	3.1847	3.9066	252	0.5	3.9408	4.2223
200	0.5	3.7771	4.3980	252	0.6	6.2286	6.2474
300	0.5	3.7881	3.8970	252	0.7	12.8308	13.5662
400	0.5	3.9907	4.4726	252	0.8	17.4854	18.7490
500	0.5	3.2959	3.6370	252	0.9	100.1979	103.5917

Note: The table compares the risk between $\hat{\mathbf{c}}_p$ and $\hat{\mathbf{c}}_b$ for the same p/n ratio with different different number of assets, p, and for same p with different p/n ratio, where n is the size of the sample.

Figure 1: The theoretical optimal return R and the corresponding plug-in return R_p.



Note: The solid and dashed lines denote the values of the theoretical optimal return, R, and the corresponding plug-in return, R_p, respectively, as defined in Theorem 2.2.

TABLE II: Comparison of the performances of $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}\left(\widetilde{\mathbf{S}}_n\right)$, $\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$, and $\varsigma(\mathbf{a},\mathbf{b})$ for $p = 100$

y	$\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$	$\varsigma(\mathbf{a},\mathbf{b})$	$\phi_{\mathbf{a}_p}^{\mathbf{b}_p}\left(\widetilde{\mathbf{S}}_n\right)$	$\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$	$\gamma\phi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$
A: $\vec{\tau} = (25, 10, 5, 1)$, $\vec{w} = \frac{1}{4}(1, 1, 1, 1)$.					
0.1	1.86	1.8857	1.8832(0.0938)	2.0667(0.1308)	2.066
0.2	1.86	1.9153	1.9175(0.1330)	2.3315(0.2095)	2.325
0.3	1.86	1.9497	1.9482(0.1644)	2.6678(0.3085)	2.657
0.4	1.86	1.9896	1.9840(0.2065)	3.1142(0.4673)	3.1
0.5	1.86	2.0370	2.0253(0.2459)	3.7495(0.7119)	3.72
0.6	1.86	2.0953	2.0822(0.2783)	4.7594(1.0897)	4.65
0.7	1.86	2.1661	2.1402(0.3138)	6.4346(1.8411)	6.2
0.8	1.86	2.2479	2.2027(0.3458)	9.6998(3.7428)	9.3
0.9	1.86	2.3540	2.2479(0.4005)	20.638(14.465)	18.6
B: $\vec{\tau} = (10, 5, 1)$, $\vec{w} = \frac{1}{10}(4, 3, 3)$.					
0.1	1.7	1.7161	1.7159(0.0783)	1.8914(0.1124)	1.888
0.2	1.7	1.7348	1.7348(0.1149)	2.1294(0.1921)	2.125
0.3	1.7	1.7567	1.7574(0.1527)	2.4432(0.3064)	2.428
0.4	1.7	1.7823	1.7829(0.1719)	2.8605(0.4222)	2.833
0.5	1.7	1.8126	1.8105(0.1938)	3.4308(0.5982)	3.4
0.6	1.7	1.8498	1.8452(0.2431)	4.3315(1.0416)	4.25
0.7	1.7	1.8943	1.8846(0.2519)	5.9039(1.6676)	5.666
0.8	1.7	1.9444	1.9236(0.2736)	8.9074(3.4104)	8.5
0.9	1.7	2.0066	1.9514(0.2913)	19.060(11.968)	17
C: $\vec{\tau} = (5, 3, 1)$, $\vec{w} = \frac{1}{10}(4, 3, 3)$.					
0.1	2.2666	2.3016	2.3017(0.1102)	2.5216(0.1528)	2.5185
0.2	2.2666	2.3421	2.3396(0.1563)	2.8384(0.2550)	2.8333
0.3	2.2666	2.3892	2.3862(0.2061)	3.2562(0.4079)	3.2380
0.4	2.2666	2.4435	2.4343(0.2265)	3.8107(0.5633)	3.7777
0.5	2.2666	2.5066	2.4757(0.2483)	4.5773(0.8110)	4.5333
0.6	2.2666	2.5809	2.5069(0.2810)	5.7787(1.3933)	5.6666
0.7	2.2666	2.6643	2.5382(0.2793)	7.8695(2.2318)	7.5555
0.8	2.2666	2.7502	2.5699(0.2882)	11.881(4.5272)	11.333
0.9	2.2666	2.8458	2.5890(0.2989)	25.446(16.054)	22.666

Note: Here $\gamma = 1/(1-y)$ and $\phi_{\mathbf{a}_p}^{\mathbf{b}}(\mathbf{A}) = \mathbf{a}_p'\mathbf{A}^{-1}\mathbf{b}_p$. Refer to section 6.1 for the description of the terms used in the table.

TABLE III: Comparison of $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}\left(\widetilde{\mathbf{S}}_n\right)$, $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$, and $\varrho(\mathbf{a},\mathbf{b})$ for $p = 100$

y	$\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$	$\varrho(\mathbf{a},\mathbf{b})$	$\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}\left(\widetilde{\mathbf{S}}_n\right)$	$\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{S}_n)$	$\gamma^3\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\Sigma)$
A: $\vec{\tau} = (25, 10, 3, 1)$, $\vec{w} = \frac{1}{4}(1, 1, 1, 1)$.					
0.1	2.1266	2.1914	2.3659 (0.2314)	2.9287(0.3562)	2.9171
0.2	2.1266	2.2740	2.6595 (0.3718)	4.1816 (0.7598)	4.1535
0.3	2.1266	2.3809	3.0391 (0.5710)	6.3114 (1.6365)	6.2000
0.4	2.1266	2.5198	3.5281 (0.7717)	10.139 (3.2253)	9.8454
0.5	2.1266	2.7045	4.1181 (1.0169)	17.554 (6.6398)	17.012
0.6	2.1266	2.9593	4.7613 (1.3859)	35.643 (19.184)	33.228
0.7	2.1266	3.3045	5.4097 (1.5618)	90.808 (59.328)	78.763
0.8	2.1266	3.7423	6.1136 (1.8169)	313.58 (280.67)	265.82
0.9	2.1266	4.3561	6.7951 (2.1544)	3422.9 (7450.3)	2126.6
B: $\vec{\tau} = (10, 5, 1)$, $\vec{w} = \frac{1}{10}(4, 3, 3)$.					
0.1	1.9666	2.0169	2.1625 (0.2026)	2.7086 (0.3294)	2.6977
0.2	1.9666	2.0828	2.4020 (0.3240)	3.8685 (0.7095)	3.8410
0.3	1.9666	2.1696	2.7037 (0.4896)	5.8330 (1.5159)	5.7335
0.4	1.9666	2.2835	3.0818 (0.6354)	9.3717 (2.9528)	9.1046
0.5	1.9666	2.4349	3.4763 (0.7940)	16.243 (6.1401)	15.732
0.6	1.9666	2.6405	3.8436 (0.9811)	32.984 (17.572)	30.728
0.7	1.9666	2.9098	4.1985 (1.0618)	83.963 (54.907)	72.837
0.8	1.9666	3.2343	4.5451 (1.1707)	289.59 (263.82)	245.82
0.9	1.9666	3.6602	4.8461 (1.2957)	3134.7 (6476.9)	1966.6
C: $\vec{\tau} = (5, 3, 1)$, $\vec{w} = \frac{1}{10}(4, 3, 3)$.					
0.1	2.2666	2.3459	2.5079 (0.2419)	3.1210 (0.3839)	3.1091
0.2	2.2666	2.4587	2.7755 (0.3769)	4.4565 (0.8244)	4.4270
0.3	2.2666	2.6135	3.1020 (0.5570)	6.7186 (1.7533)	6.6081
0.4	2.2666	2.8173	3.4696 (0.6975)	10.786 (3.4074)	10.494
0.5	2.2666	3.0817	3.8066 (0.8334)	18.729 (7.1874)	18.133
0.6	2.2666	3.4268	4.0860 (0.9681)	38.021 (20.461)	35.416
0.7	2.2666	3.8566	4.3398 (1.0042)	96.768 (63.820)	83.948
0.8	2.2666	4.3472	4.5702 (1.0590)	333.82 (307.84)	283.33
0.9	2.2666	4.9539	4.7502 (1.1209)	3617.4 (7589.3)	2266.6

Note: Here $\gamma = 1/(1-y)$ and $\varphi_{\mathbf{a}_p}^{\mathbf{b}_p}(\mathbf{A}) = \mathbf{a}_p'\mathbf{A}^{-1}\Sigma\mathbf{A}^{-1}\mathbf{b}_p$. Refer to section 6.1 for the description of the terms used in the table.

TABLE IV: Comparison of spectrally-corrected estimates with the plug-in and Bootstrap-corrected estimates

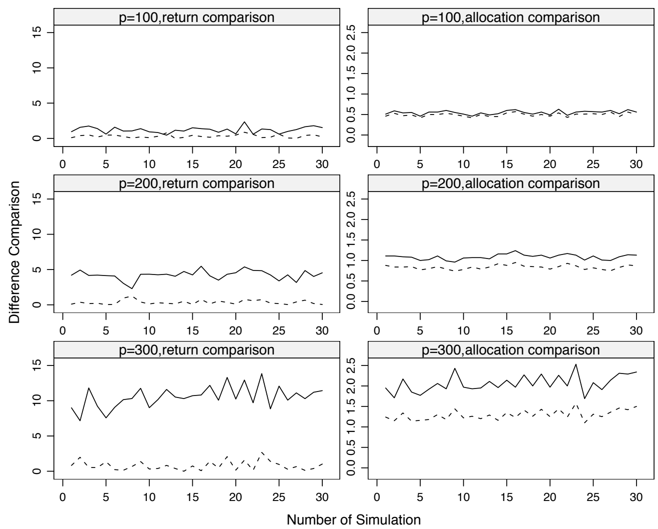
Panel A: $R(\boldsymbol{\mu}, \Sigma) = 3.8190$ and $\sigma^2 = 1$ $\vec{\tau} = (25, 10, 5, 1)$, $\vec{w} = (0.25, 0.25, 0.25, 0.25)$				
y	$\hat{\mathbf{c}}$	$R_{\hat{\mathbf{c}}}$	$100(R_{\hat{\mathbf{c}}} - R)/R\%$	$\hat{\mathbf{c}}'\Sigma\hat{\mathbf{c}}$
0.1	$\mathbf{c}_s$	3.8138 (0.0503)	-0.14(0.0132)	1.0771 (0.0312)
	$\mathbf{c}_p$	4.0197 (0.0924)	5.25(0.0242)	1.2323 (0.0609)
	$\mathbf{c}_b$	3.8071 (0.1312)	-0.31(0.0344)	1.2452 (0.0806)
0.2	$\mathbf{c}_s$	3.8069 (0.0742)	-0.32(0.0194)	1.1675 (0.0536)
	$\mathbf{c}_p$	4.2539 (0.1482)	11.39(0.0388)	1.5553 (0.1219)
	$\mathbf{c}_b$	3.7960 (0.2074)	-11.60(0.0543)	1.5848 (0.1516)
0.3	$\mathbf{c}_s$	3.7973 (0.0948)	-0.57(0.0248)	1.2729 (0.0797)
	$\mathbf{c}_p$	4.5373(0.2235)	11.80(0.0585)	2.0276 (0.2342)
	$\mathbf{c}_b$	3.7727 (0.3165)	-1.21(0.0829)	2.0751 (0.2609)
0.4	$\mathbf{c}_s$	3.7857(0.1128)	-0.87(0.0295)	1.3939 (0.1121)
	$\mathbf{c}_p$	4.8701 (0.3401)	27.52(0.0891)	2.7319 (0.4441)
	$\mathbf{c}_b$	3.7381 (0.5096)	-2.12(0.1134)	2.8165 (0.4297)
0.5	$\mathbf{c}_s$	3.7800 (0.1343)	-1.02(0.0352)	1.5416 (0.1637)
	$\mathbf{c}_p$	5.2814 (0.5721)	38.29(0.1498)	3.8820 (0.9076)
	$\mathbf{c}_b$	3.6502 (0.9054)	-4.42(0.2371)	4.0793 (0.7797)
0.6	$\mathbf{c}_s$	3.7679 (0.1640)	-1.34(0.0429)	1.7010 (0.2492)
	$\mathbf{c}_p$	5.8286 (0.8879)	52.62(0.2325)	6.0203 (1.8452)
	$\mathbf{c}_b$	3.5030 (1.3923)	-8.27(0.3646)	6.5127 (1.6391)
0.7	$\mathbf{c}_s$	3.7626 (0.1891)	-1.48(0.0495)	1.8649 (0.3548)
	$\mathbf{c}_p$	6.5938 (1.4396)	72.66(0.3770)	10.6988 (4.3778)
	$\mathbf{c}_b$	3.2346 (2.1844)	-15.30(0.5720)	12.1496 (4.3399)
0.8	$\mathbf{c}_s$	3.7605 (0.2130)	-1.53(0.0558)	2.0102 (0.4625)
	$\mathbf{c}_p$	7.6161 (2.4100)	99.42(0.6311)	22.22 (12.515)
	$\mathbf{c}_b$	2.5653 (3.5775)	-32.83(0.9368)	28.768 (15.926)
0.9	$\mathbf{c}_s$	3.7585 (0.2449)	-1.58(0.0641)	2.1382 (0.5822)
	$\mathbf{c}_p$	9.9073 (4.7808)	159.41(1.2518)	86.581 (78.657)
	$\mathbf{c}_b$	0.7019 (7.0065)	-81.62(1.8346)	151.27 (170.23)

Panel B: $R(\boldsymbol{\mu}, \Sigma) = 4.0247$ and $\sigma^2 = 1$ $\vec{\tau} = (10, 5, 1)$, $\vec{w} = (0.4, 0.3, 0.3)$				
y	$\hat{\mathbf{c}}$	$R_{\hat{\mathbf{c}}}$	$100(R_{\hat{\mathbf{c}}} - R)/R\%$	$\hat{\mathbf{c}}'\Sigma\hat{\mathbf{c}}$
	$\mathbf{c}_s$	4.0196 (0.0541)	-0.12(0.0134)	1.0708 (0.0312)
0.1	$\mathbf{c}_p$	4.2379 (0.0981)	5.29(0.0244)	1.2326 (0.0611)
	$\mathbf{c}_b$	4.0140 (0.1391)	-0.27(0.0346)	1.2439 (0.0808)
	$\mathbf{c}_s$	4.0122 (0.0789)	-0.31(0.0196)	1.1524 (0.0520)
0.2	$\mathbf{c}_p$	4.4835 (0.1619)	11.40(0.0402)	1.5532 (0.1270)
	$\mathbf{c}_b$	3.9983 (0.2322)	-0.65(0.0577)	1.5798 (0.1531)
	$\mathbf{c}_s$	4.0034 (0.1008)	-0.53(0.0250)	1.2444 (0.0759)
0.3	$\mathbf{c}_p$	4.7775 (0.2618)	18.70(0.0650)	2.0194 (0.2572)
	$\mathbf{c}_b$	3.9629 (0.3950)	-1.53(0.0981)	2.0667 (0.2655)
	$\mathbf{c}_s$	3.9933 (0.1196)	-0.78(0.0297)	1.3462 (0.1075)
0.4	$\mathbf{c}_p$	5.1088 (0.4302)	26.94(0.1069)	2.6997 (0.5118)
	$\mathbf{c}_b$	3.8888 (0.6871)	-3.38(0.1707)	2.8007 (0.4346)
	$\mathbf{c}_s$	3.9909 (0.1410)	-0.84(0.0350)	1.4652 (0.1629)
0.5	$\mathbf{c}_p$	5.5241 (0.7044)	37.25(0.1750)	3.8153 (1.0081)
	$\mathbf{c}_b$	3.7612 (1.1514)	-6.55(0.2861)	4.0675 (0.7844)
	$\mathbf{c}_s$	3.9828 (0.1678)	-1.04(0.0417)	1.5793 (0.2406)
0.6	$\mathbf{c}_p$	6.0615 (1.0906)	50.61(0.2710)	5.8713 (2.0261)
	$\mathbf{c}_b$	3.5352 (1.7415)	-12.16(0.4327)	6.5447 (1.7161)
	$\mathbf{c}_s$	3.9844 (0.1908)	-1(0.0474)	1.6811 (0.3263)
0.7	$\mathbf{c}_p$	6.8264 (1.7075)	69.61(0.4243)	10.393 (4.6787)
	$\mathbf{c}_b$	3.1870 (2.6091)	-20.81(0.6483)	12.336 (4.5793)
	$\mathbf{c}_s$	3.9842 (0.2094)	-1(0.0520)	1.7668 (0.3981)
0.8	$\mathbf{c}_p$	7.8668 (2.7378)	95.46(0.6802)	21.589 (12.998)
	$\mathbf{c}_b$	2.4225 (4.0791)	-39.81(1.0135)	29.425 (16.636)
	$\mathbf{c}_s$	3.9903 (0.2342)	-0.85(0.0582)	1.8290 (0.4788)
0.9	$\mathbf{c}_p$	10.147 (5.2831)	152.13(1.3127)	83.53 (79.77)
	$\mathbf{c}_b$	0.2299 (7.7471)	-94.29(1.9249)	156.9 (177.2)

Panel C: $R(\boldsymbol{\mu}, \Sigma) = 4.3376$ and $\sigma^2 = 1$ $\vec{\tau} = (5, 3, 1)$, $\vec{w} = (0.4, 0.3, 0.3)$				
y	$\hat{\mathbf{c}}$	$R_{\hat{\mathbf{c}}}$	$100(R_{\hat{\mathbf{c}}} - R)/R\%$	$\hat{\mathbf{c}}'\Sigma\hat{\mathbf{c}}$
	$\mathbf{c}_s$	4.3266 (0.0679)	-0.25(0.0157)	1.0673 (0.0352)
0.1	$\mathbf{c}_p$	4.5684 (0.1088)	5.32(0.0251)	1.2319 (0.0634)
	$\mathbf{c}_b$	4.3260 (0.1572)	-0.27(0.0363)	1.2412 (0.0824)
	$\mathbf{c}_s$	4.3122 (0.0974)	-0.59(0.0225)	1.1367 (0.0585)
0.2	$\mathbf{c}_p$	4.8172 (0.2181)	11.06(0.0503)	1.5382 (0.1637)
	$\mathbf{c}_b$	4.2767 (0.3580)	-1.40(0.0825)	1.5646 (0.1711)
	$\mathbf{c}_s$	4.3022 (0.1265)	-0.82(0.0292)	1.2044 (0.0930)
0.3	$\mathbf{c}_p$	5.0988 (0.4000)	17.55(0.0922)	1.9699 (0.3554)
	$\mathbf{c}_b$	4.1712 (0.6833)	-3.84(0.1575)	2.0361 (0.2906)
	$\mathbf{c}_s$	4.3006 (0.1552)	-0.85(0.0358)	1.2601 (0.1349)
0.4	$\mathbf{c}_p$	5.4127 (0.6435)	24.79(0.1484)	2.5992 (0.6690)
	$\mathbf{c}_b$	4.0139 (1.1002)	-7.46(0.2536)	2.7718 (0.4480)
	$\mathbf{c}_s$	4.3104 (0.1756)	-0.63(0.0405)	1.3051 (0.1783)
0.5	$\mathbf{c}_p$	5.8043 (0.9996)	33.81(0.2305)	3.6282 (1.2445)
	$\mathbf{c}_b$	3.7780 (1.6819)	-12.50(0.3878)	4.0675 (0.8115)
	$\mathbf{c}_s$	4.3161 (0.1981)	-0.50(0.0457)	1.3257 (0.2105)
0.6	$\mathbf{c}_p$	6.3027 (1.4782)	45.30(0.3408)	5.5166 (2.3755)
	$\mathbf{c}_b$	3.4027 (2.4206)	-21.55(0.5581)	6.6470 (1.8593)
	$\mathbf{c}_s$	4.3282 (0.2110)	-0.22(0.0486)	1.3450 (0.2420)
0.7	$\mathbf{c}_p$	7.0149 (2.2115)	61.72(0.5098)	9.6467 (5.2082)
	$\mathbf{c}_b$	2.8573 (3.4346)	-34.13(0.7918)	12.79 (5.0708)
	$\mathbf{c}_s$	4.3301 (0.2216)	-0.17(0.0511)	1.3621 (0.2642)
0.8	$\mathbf{c}_p$	8.0686 (3.3101)	86.01(0.7631)	20.1585 (13.6538)
	$\mathbf{c}_b$	1.9350 (4.9736)	-55.39(1.1466)	30.8030 (18.0832)
	$\mathbf{c}_s$	4.3371 (0.2342)	-0.01(0.0540)	1.3754 (0.2839)
0.9	$\mathbf{c}_p$	10.35 (6.0308)	138.79(1.3904)	77.81 (79.22)
	$\mathbf{c}_b$	-0.6901 (8.9579)	-115.91(2.0652)	166.1 (188.6)

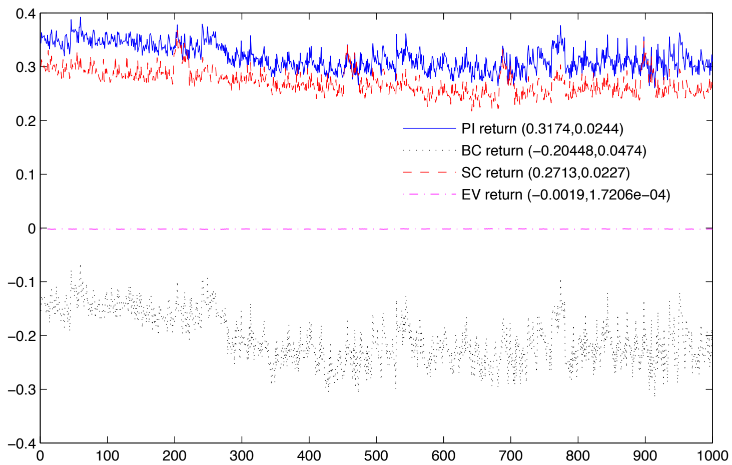
Note: $p = 100$ and $N = 10000$. Here $\mathbf{c}_s$ and $\mathbf{c}_p$ represent as $\mathbf{c}(\boldsymbol{\mu}, \widetilde{\mathbf{S}}_n)$ and $\mathbf{c}(\boldsymbol{\mu}, \mathbf{S}_n)$, respectively.

Figure 2: Comparison between the Empirical and Corrected Portfolio Allocation and Returns



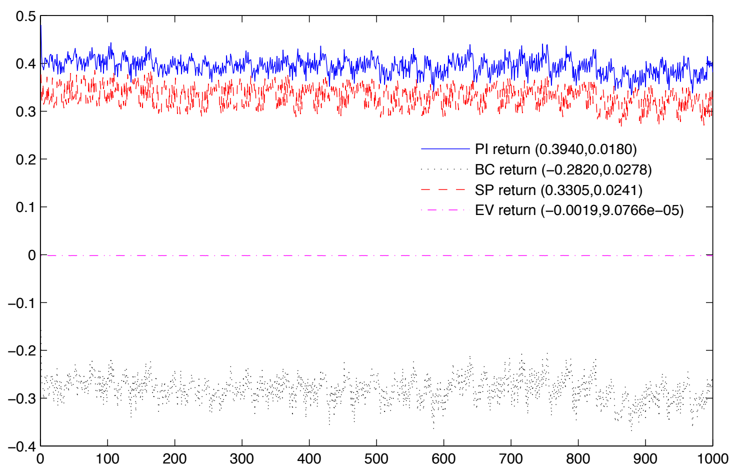
Note: Here, $d_R^b = \hat{R}_b - R$, $d_R^p = \hat{R}_p - R$, $d_{\mathbf{c}}^b = \|\hat{\mathbf{c}}_b - \mathbf{c}\|$, and $d_{\mathbf{c}}^p = \|\hat{\mathbf{c}}_p - \mathbf{c}\|$. Solid line is the absolute values of d_p^c and d_p^R, respectively; Dashed line is the absolute values of $d_{\mathbf{c}}^b$ and $d_{\mathbf{c}}^p$, respectively. The top, middle and bottom two sub-figures are the plots for $p = 100, 200, 300$ and $n = 500$, respectively. The plots on the left are the plots for d_R^p and d_R^b, while the plots on the right are the plots for $d_{\mathbf{c}}^p$ and $d_{\mathbf{c}}^b$, respectively. Here, the population is given according to a multivariate normal distribution with $\mu = (\mu_1, ..., \mu_p)^T$ and $\Sigma = \mathbb{I}$.

Figure 3: Comparison of Different Returns for 50 stocks in the S&P 500 as $y = 0.1$



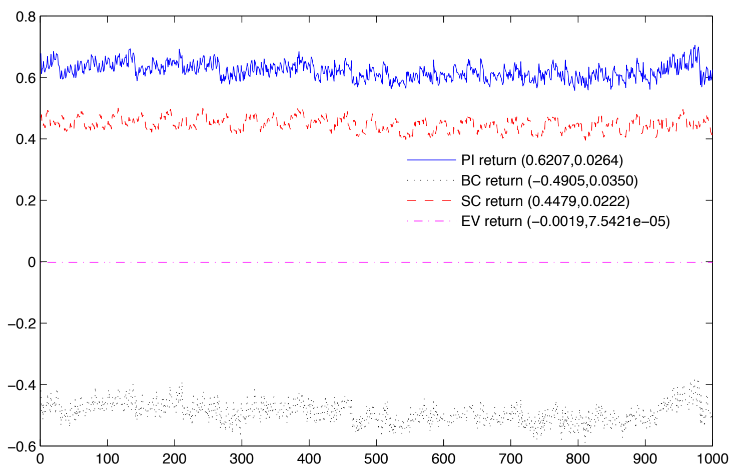
Note: We denote the Plug-in, Bootstrap-corrected, equally weighted, and Spectrally-corrected returns as PI, BC, EW, and SC, respectively.

Figure 4: Comparison of Different Returns for 100 stocks in the S&P 500 as $y = 0.2$



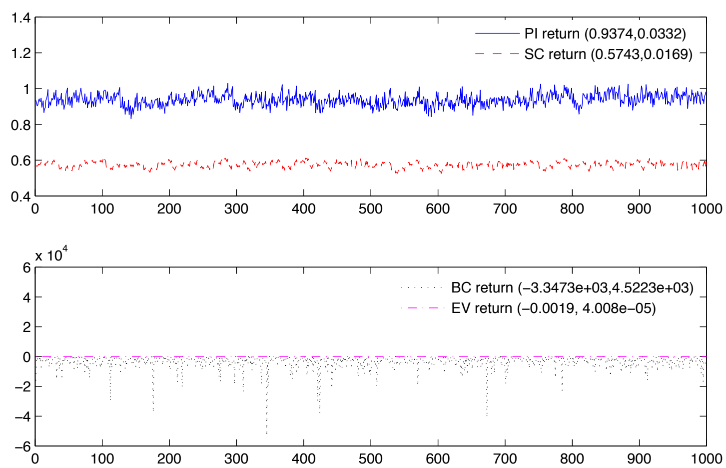
Note: We denote the Plug-in, Bootstrap-corrected, equally weighted, and Spectrally-corrected returns as PI, BC, EW, and SC, respectively.

Figure 5: Comparison of Different Returns for 200 stocks in the S&P 500 as $y = 0.4$



Note: We denote the Plug-in, Bootstrap-corrected, equally weighted, and Spectrally-corrected returns as PI, BC, EW, and SC, respectively.

Figure 6: Comparison of Different Returns for 300 stocks in the S&P 500 as $y = 0.6$



Note: We denote the Plug-in, Bootstrap-corrected, equally weighted, and Spectrally-corrected returns as PI, BC, EW, and SC, respectively.