Scandinavian Journal of Statistics Theory and Applications

Scandinavian Journal of Statistics, Vol. 41: 845–865, 2014
doi: 10.1111/sjos.12066

Statistical Inference for High-Dimensional Global Minimum Variance Portfolios

KONSTANTIN GLOMBEK

Seminar für Wirtschafts- und Sozialstatistik, Universität zu Köln

ABSTRACT. Many studies demonstrate that inference for the parameters arising in portfolio optimization often fails. The recent literature shows that this phenomenon is mainly due to a high-dimensional asset universe. Typically, such a universe refers to the asymptotics that the sample size $n + 1$ and the sample dimension d both go to infinity while $d/n \to c \in (0, 1)$. In this paper, we analyze the estimators for the excess returns' mean and variance, the weights and the Sharpe ratio of the global minimum variance portfolio under these asymptotics concerning consistency and asymptotic distribution. Problems for stating hypotheses in high dimension are also discussed. The applicability of the results is demonstrated by an empirical study.

Key words: general asymptotics, high-dimensional inference, minimum variance portfolio, portfolio optimization

1. Introduction

The task of optimal asset allocation is a classical problem in finance. Since the seminal work of Markowitz (1952), a vast amount of literature about optimizing the mean-variance profile of the excess return of a portfolio has been published (see, e.g., Tobin (1958), Sharpe (1964), Lintner (1965) and the books by Cochrane (2005) and Elton *et al.* (2007)). Although this mean-variance optimization would work well if the theoretical mean and covariance matrix of the excess returns of the assets within the portfolio were known, this procedure often fails in practice. The reason for this failure is that all theoretical quantities are usually unknown and therefore have to be estimated. Several studies show that this leads to estimates which are often far away from their theoretical counterparts (see the introduction in Bai *et al.* (2009)). The former literature focuses on estimation errors caused by estimating the mean by the sample mean (Merton (1980), Chopra & Ziemba (1993)). In contrast, the current literature recognizes that estimating the covariance matrix by the sample covariance matrix is highly problematic if the sample size $n + 1$ is of the same order as d, the number of assets in the portfolio. Therefore, one can either try to improve this estimation by using better estimators (e.g., Jagannathan & Ma (2003), Ledoit & Wolf (2003), Frahm & Memmel (2010)) or to incorporate knowledge about the behaviour of the sample covariance matrix for large n, d.

The latter possibility relies on the spectral properties of the sample covariance matrix under this 'large n and large d'–setting which has been studied by many authors (e.g., Girko (1990, 1995), Serdobolskii (2000), Marčenko & Pastur (1967), Yin (1986), Bai & Yin (1993), Silverstein (1995), Bai *et al.* (2007)). In particular, the eigenvalues of the sample covariance matrix differ substantially from the eigenvalues of the true covariance matrix for n and d both being large. Thus, modern portfolio optimization methods involving the sample covariance matrix should take this difficulty into account. This is performed in Bai *et al.* (2009) and El Karoui (2010) on the basis of the asymptotics that $n, d \to \infty$ and $d/n \to c \in (0, 1)$, the so-called (n, d)– or general asymptotics. Both papers analyze the shortcomings of portfolio optimization in the context of these asymptotics and provide bootstrap methods in order to correct for estimation errors due to high-dimensionality. Whereas the first paper deals with optimizing the mean return of a portfolio under a bounded portfolio return variance, the second paper focuses on quadratic programming and investigates the minimization of the portfolio return variance under certain linear constraints. Both papers concentrate on (n, d)-consistency of selected estimators and do not discuss (n, d)-asymptotic hypothesis testing.

The present paper complements the current literature in several points. First, the estimators for the variance and mean of the excess return of the global minimum variance portfolio (MP) as well as the corresponding Sharpe ratio are analyzed concerning consistency under the (n, d)-asymptotics. Furthermore, the (n, d)-asymptotic distributions of these estimators are given. Inference for the weights of the MP under the (n, d)-asymptotics is discussed as well. In particular, we will see that the new results extend well-known χ^2 and F-tests. Certain difficulties in formulating high-dimensional hypotheses are also discussed.

This paper is organized as follows: Section 2 gives all definitions and assumptions which are needed to derive the main results in section 3. An empirical study in section 4 shows the applicability of the results, which is followed by a conclusion in section 5. Selected proofs are presented in Appendix A.

2. Preliminaries

2.1. Definitions and assumptions

We assume that an investor invests in d risky assets. In particular, the possibility to invest in a risk-free asset is excluded. A portfolio P is a linear combination of these d assets. The vector of weights of this linear combination will be denoted by $w_{\mathrm{P}} \in \mathbb{R}^d$ and fulfills the budget constraint $w_{\mathrm{P}}^t \mathbb{1} = 1$, where $\mathbb{1} = (1, \ldots, 1)^t \in \mathbb{R}^d$. Note that this implies that $\mathbb{1} w_{\mathrm{P}}^t$ is a projector. We allow for short-selling; that is, the weight vector w_{P} may have negative components. In the following, the equality $\mathrm{P}_1 = \mathrm{P}_2$ for two portfolios $\mathrm{P}_1, \mathrm{P}_2$ stands for the equality of the corresponding weight vectors.

Let $R = (R_1, \ldots, R_d)^t$ be the excess returns of the d assets. We assume that R has a d-dimensional distribution with mean $\mu \in \mathbb{R}^d$ and positive definite covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$. Thus, the excess return of a portfolio P is given by $w_{\mathrm{P}}^t R$. The variance and mean of $w_{\mathrm{P}}^t R$ are

$$\sigma_{\mathrm{P}}^2 := \mathbb{V}ar\left(w_{\mathrm{P}}^t R\right) = w_{\mathrm{P}}^t \Sigma w_{\mathrm{P}},$$
$$\mu_{\mathrm{P}} := \mathbb{E}\left(w_{\mathrm{P}}^t R\right) = w_{\mathrm{P}}^t \mu,$$

and the corresponding Sharpe ratio equals

$$SR_{\mathrm{P}} = \frac{\mu_{\mathrm{P}}}{\sigma_{\mathrm{P}}} = \frac{w_{\mathrm{P}}^t \mu}{\sqrt{w_{\mathrm{P}}^t \Sigma w_{\mathrm{P}}}}.$$

The variance of the portfolio excess return is uniquely minimized by the *global minimum variance portfolio* (MP) which is characterized by the weights

$$w_{\mathrm{MP}} = \frac{\Sigma^{-1} \mathbb{1}}{\mathbb{1}^t \Sigma^{-1} \mathbb{1}}.$$

So, beside for w_{MP}, we are interested in inference for the following quantities:

$$\sigma^2_{\mathrm{MP}} = w^t_{\mathrm{MP}} \Sigma w_{\mathrm{MP}} = \frac{1}{\mathbb{1}^t \Sigma^{-1} \mathbb{1}},$$
$$\mu_{\mathrm{MP}} = w^t_{\mathrm{MP}} \mu = \frac{\mathbb{1}^t \Sigma^{-1} \mu}{\mathbb{1}^t \Sigma^{-1} \mathbb{1}},$$
$$SR_{\mathrm{MP}} = w^t_{\mathrm{MP}} \mu \sqrt{\mathbb{1}^t \Sigma^{-1} \mathbb{1}} = \frac{\mathbb{1}^t \Sigma^{-1} \mu}{\sqrt{\mathbb{1}^t \Sigma^{-1} \mathbb{1}}}.$$

In the following, inference will be based on a sample $R_1, \ldots, R_{n+1}$ which is drawn from R. The mean μ shall be estimated by the sample mean

$$\hat{\mu} = \overline{R} = \frac{1}{n+1} \sum_{i=1}^{n+1} R_i$$

and the covariance matrix by the unbiased sample covariance matrix

$$\begin{aligned} \widehat{\Sigma} := S &:= \frac{1}{n} \sum_{i=1}^{n+1} (R_i - \overline{R})(R_i - \overline{R})^t \\ &= \Sigma^{1/2} \tilde{S} \Sigma^{1/2}, \end{aligned}$$

where $\tilde{S}$ is an analogous estimator for the covariance matrix of $\tilde{R} \sim (0, I)$, I being the identity.

We will make the following assumptions:

Assumptions.

(1) The sample of excess returns is independent and identically normally distributed, that is, $R_i \overset{\text{i.i.d.}}{\sim} N_d(\mu, \Sigma), i = 1, \ldots, n+1$, where the mean $\mu \in \mathbb{R}^d$ and the positive definite covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$ are unknown.

(2) The (n, d)-asymptotics are given by $n, d \to \infty$ and $d/n \to c \in (0, 1)$.

(3) The (n, d)-limit

$$w^t_{\mathrm{MP}} \mu \longrightarrow a_1 \in \mathbb{R} \setminus \{0\}$$

exists.

(4) We have either

(a) $\sigma^2_{\mathrm{MP}} \longrightarrow 0$ or

(b) $\sigma^2_{\mathrm{MP}} \longrightarrow a_2 \in (0, \infty)$

under the (n, d)-asymptotics.

Although often criticized, the assumption of normality is very common in the literature. If the frequency of the data is not too high (e.g., monthly), this assumption may hold. A justification for that can be found in Frahm & Memmel (2010) and the references therein. Note that the normality assumption implies that $n\widehat{\Sigma}$ has a d-dimensional Wishart distribution with scale matrix Σ and n degrees of freedom.

The second assumption guarantees that the sample covariance matrix remains regular in the high-dimensional framework. In the following, the classical setting of $c = 0$ will be treated separately for each discussed estimator as well.

If we look at assumption 3, we see that it can be regarded as an integrability assumption, meaning that the components of the vector μ are summable with respect to the signed measure which is induced by the weight vector w_{MP}.

Limit (4a) refers to the following difficulty: If $\lambda_{\max}(\Sigma) = o(d)$ under the (n, d)-asymptotics, where $\lambda_{\max}(\Sigma)$ denotes the largest eigenvalue of Σ, then

$$\sigma^2_{\mathrm{MP}} = \frac{d}{\mathbb{1}^t \Sigma^{-1} \mathbb{1}} \frac{1}{d} \in \frac{1}{d} [\lambda_{\min}(\Sigma), \lambda_{\max}(\Sigma)] \longrightarrow \{0\} \tag{1}$$

because of the properties of the Rayleigh quotient; see also Schott (2005), theorem 3.16. Economically, this effect may be interpreted as perfect diversification as the variance of the portfolio return can be made arbitrarily small by increasing the number of considered assets. However, this situation is rather theoretical for the following two reasons: First, inference for σ^2_{MP} is impossible as the corresponding test statistic will require a division by the null value of zero. Second, according to principal component analysis, the largest eigenvalues of Σ equal the variances of the main 'risk driving factors' in the market. These variances should increase at least linearly with respect to d as the corresponding principal components reflect most of the information of the linearly growing number of excess returns. An important example for this case is Σ being an equicorrelation matrix; that is,

$$\Sigma = \Sigma_{\rho,\sigma^2} = \sigma^2 \left((1-\rho) I + \rho \mathbb{1}\mathbb{1}^t \right),$$

where $\sigma^2 > 0$ and $\rho \in (-1/(d-1), 1)$. The largest eigenvalue of Σ_{ρ,σ^2} equals $\sigma^2(1 + (d-1)\rho)$ which is of the order $\mathcal{O}(d)$ as $d \to \infty$. In this case, we have that $w_{\mathrm{MP}} = \mathbb{1}/d$ and

$$\sigma^2_{\mathrm{MP}} = \sigma^2 \frac{1 + (d-1)\rho}{d} \longrightarrow \sigma^2 \rho > 0$$

as $d \to \infty$ for $\rho > 0$. Thus, we will refer to the limit (4b) as to the 'natural' case because the positivity of this limit requires $\lambda_{\max}(\Sigma) = \mathcal{O}(d)$ under the (n, d)-asymptotics. But for the sake of completeness, we will include both cases (4a) and (4b) in the following discussion.

Let us define the set of admissible portfolios as

$$\mathcal{P} := \left\{ \mathrm{P} \;\middle|\; \lim_{d\to\infty} \frac{\sigma^2_{\mathrm{P}}}{\sigma^2_{\mathrm{MP}}} \text{exists and is finite, } \lim_{d\to\infty} w^t_{\mathrm{P}} \mu \text{ exists and is non-zero} \right\}.$$

Note that $\mathcal{P}$ is non-empty due to assumptions 3 and 4. This set will be important in order to formulate hypotheses and performing tests in the following.

2.2. Basic facts

Next, we present several facts which are important in the following discussion. First, we provide a strong law of large numbers for quadratic forms involving a Wishart matrix.

Proposition 1. *Let $n\tilde{S}$ have a d-dimensional standard Wishart distribution with n degrees of freedom and $u \in \mathbb{R}^d$ with $\|u\|_2 = 1$. Then, we have under the (n, d)-asymptotics the following almost sure limits:*

$$u^t \tilde{S} u \xrightarrow{\text{a.s.}} 1, \qquad u^t \tilde{S}^{-1} u \xrightarrow{\text{a.s.}} \frac{1}{1-c}.$$

Proof. The result follows from Bai *et al.* (2007), corollary 2. □

The next theorem extends proposition 1 to the case of a stochastic u.

Proposition 2. *Under assumptions 1, 3 and 4, we have the following almost sure (n, d)-limit:*

$$\frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \hat{\mu}}{\mathbb{1}^t \Sigma^{-1} \mu} \xrightarrow{\text{a.s.}} \frac{1}{1-c},$$

provided that $\mathbb{1}^t \Sigma^{-1} \mu \neq 0$.

Proof. See the Appendix. □

A central limit theorem for quadratic forms involving a Wishart matrix is given by the following theorem.

Proposition 3. *Let $n\tilde{S}$ have a d-dimensional standard Wishart distribution with n degrees of freedom and $u, v \in \mathbb{R}^d$ with $\|u\|_2 = \|v\|_2 = 1$. Then, we have under the (n, d)-asymptotics the following weak convergence:*

$$\sqrt{n} \begin{pmatrix} u^t \tilde{S} u - 1 \\ v^t \tilde{S}^{-1} v - \frac{1}{1-d/n} \end{pmatrix} \Longrightarrow N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & -\frac{2}{1-c} \\ -\frac{2}{1-c} & \frac{2}{(1-c)^3} \end{bmatrix} \right) \tag{2}$$

Proof. See the Appendix. □

Note that the asymptotic covariance matrix given in proposition 3 is regular for $c \in (0, 1)$, whereas it becomes singular for $c = 0$ and does not exist for $c = 1$. In the following, we will apply the delta method to proposition 3, which is non-problematic for $c \in (0, 1)$. In the classical case of $c = 0$, the delta method can only be applied to a single margin of the random vector in (2). For $c = 1$, the delta method is not applicable. A further discussion about the usage of the delta method under (n, d)-asymptotics can be found in Birke & Dette (2005).

3. Inference under (n, d)-asymptotics

We begin with inference for the variance of the excess return of the MP.

3.1. Inference for the variance

The usual estimator for the variance of the excess return of the MP is given by

$$\hat{\sigma}^2_{\mathrm{MP}} := \frac{1}{\mathbb{1}^t \widehat{\Sigma}^{-1} \mathbb{1}}.$$

The exact distribution of $\hat{\sigma}^2_{\mathrm{MP}}$ for finite $n > d - 1$ is

$$n \frac{\hat{\sigma}^2_{\mathrm{MP}}}{\sigma^2_{\mathrm{MP}}} \sim \chi^2_{n-d+1}; \tag{3}$$

see, for example, Frahm (2010), theorem 2. Hence, the estimator $\hat{\sigma}^2_{\mathrm{MP}}$ is biased because of

$$\mathbb{E}\left(\hat{\sigma}^2_{\mathrm{MP}}\right) = \left(1 - \frac{d-1}{n}\right) \sigma^2_{\mathrm{MP}}.$$

Further, we immediately see that this estimator is weakly n-consistent for d being fixed.

Now, we want to discuss hypothesis tests for σ^2_{MP} on the basis of (3). Since $\sigma^2_{\mathrm{MP}} < \sigma^2_{\mathrm{P}} = w^t_{\mathrm{P}} \Sigma w_{\mathrm{P}}$ for any portfolio P $\neq$ MP by definition of the MP, the only possible (and economically meaningful) hypothesis is

$$H_0 : \sigma^2_{\mathrm{MP}} = \sigma^2_0 \quad \text{versus} \quad H_1 : \sigma^2_{\mathrm{MP}} < \sigma^2_0 \tag{4}$$

for some $\sigma_0^2 > 0$. The power function for this test can be calculated as

$$\mathrm{pow}\left(\sigma_{\mathrm{MP}}^2; \alpha\right) = F_{\chi^2_{n-d+1}}\left(\frac{c\sigma_0^2}{\sigma_{\mathrm{MP}}^2}\right),$$

where $\sigma_{\mathrm{MP}}^2 \in (0, \sigma_0^2]$, $F_{\chi^2_{n-d+1}}$ is the distribution function of the χ^2 distribution with $n-d+1$ degrees of freedom, $c = F^{-1}_{\chi^2_{n-d+1}}(\alpha)$, and $\alpha \in (0,1)$ is the level of the test. From this result, we immediately obtain unbiasedness of this test.

However, proving (n,d)-consistency of this test based on its power function is not immediate because the χ^2 distribution with $n-d+1$ degrees of freedom degenerates under the (n,d)-asymptotics. Another problem in practice is that the variance of (3) tends to infinity in that case so that the variance test based on this statistic becomes infeasible for large n and d.

Thus, we seek for an asymptotic behaviour of $\hat{\sigma}_{\mathrm{MP}}^2$ under the (n,d)-asymptotics. The next theorem shows that $\hat{\sigma}_{\mathrm{MP}}^2$ is not only a biased but also an inconsistent estimator for σ_{MP}^2 under the (n,d)-asymptotics unless $c = 0$.

Theorem 4. *Let assumption 1 be true. Then, we have under the (n,d)-asymptotics that*

$$\frac{\hat{\sigma}_{\mathrm{MP}}^2}{\sigma_{\mathrm{MP}}^2} \xrightarrow{\text{a.s.}} 1 - c\,.$$

Proof. The estimator for the variance of the MP can be represented as

$$\begin{aligned}
\hat{\sigma}_{\mathrm{MP}}^2 &= \frac{1}{\mathbb{1}^t \widehat{\Sigma}^{-1} \mathbb{1}} = \frac{1}{\mathbb{1}^t \Sigma^{-1/2} \tilde{S}^{-1} \Sigma^{-1/2} \mathbb{1}} \\
&= \frac{1}{\|z\|_2^2}\left(\frac{z^t}{\|z\|_2} \tilde{S}^{-1} \frac{z}{\|z\|_2}\right)^{-1} = \sigma_{\mathrm{MP}}^2 \left(\frac{z^t}{\|z\|_2} \tilde{S}^{-1} \frac{z}{\|z\|_2}\right)^{-1} \\
\Leftrightarrow \frac{\hat{\sigma}_{\mathrm{MP}}^2}{\sigma_{\mathrm{MP}}^2} &= \left(\frac{z^t}{\|z\|_2} \tilde{S}^{-1} \frac{z}{\|z\|_2}\right)^{-1},
\end{aligned} \tag{5}$$

where $z := \Sigma^{-1/2}\mathbb{1}$, and $\tilde{S}$ has been defined in section 2.1. The proof of this theorem follows immediately from this representation and proposition 1. □

The reason to consider the ratio in theorem 4 is the possibility of the theoretical case of assumption 4a. The inconsistency of this estimator is therefore to be understood in the sense of theorem 4. Another consequence of this observation is that it is in general not reasonable to test for a hypothesis of the form (4). Instead, one should reformulate the hypothesis as

$$H_0 : \xi := \frac{\sigma_0^2}{\sigma_{\mathrm{MP}}^2} = 1 \quad \text{versus} \quad H_1 : \xi > 1, \tag{6}$$

where σ_0^2 is associated to some portfolio $\mathrm{P} \in \mathcal{P}$. Note that ξ remains finite as $d \to \infty$ because of the definition of $\mathcal{P}$.

Next, we investigate the distribution of $\hat{\sigma}_{\mathrm{MP}}^2$ under the (n,d)-asymptotics. We know from the classical central limit theorem that

$$\frac{\chi_k^2 - k}{\sqrt{2k}} \Longrightarrow N(0,1)$$

as $k \to \infty$ for a random variable χ_k^2 which is χ^2 distributed with k degrees of freedom. We have that

$$\frac{n\frac{\hat{\sigma}^2_{\mathrm{MP}}}{\sigma^2_{\mathrm{MP}}}-(n-d+1)}{\sqrt{n+1}} = \sqrt{n}\left(\frac{\hat{\sigma}^2_{\mathrm{MP}}}{\sigma^2_{\mathrm{MP}}}-\left(1-\frac{d}{n}\right)\right)\frac{\sqrt{n}}{\sqrt{n+1}}-\frac{1}{\sqrt{n+1}} \tag{7}$$
$$\stackrel{\mathrm{d}}{=} \sqrt{2\left(1-\frac{d}{n+1}\right)}\frac{\chi^2_{n-d+1}-(n-d+1)}{\sqrt{2(n-d+1)}}.$$

It is therefore conceivable that the limiting distribution of (7) under the (n,d)-asymptotics is $N(0,2(1-c))$. However, since the classical central limit theorem only considers n-asymptotics, it is not clear that this asymptotic distribution readily follows. But the following theorem ensures that this conjecture is true.

Theorem 5. *Under assumption 1, we have the following weak convergence under the* (n,d)-*asymptotics:*

$$\sqrt{n}\left(\frac{\hat{\sigma}^2_{\mathrm{MP}}}{\sigma^2_{\mathrm{MP}}}-\left(1-\frac{d}{n}\right)\right) \Longrightarrow N(0,2(1-c)) \tag{8}$$

Proof. We apply the delta method to proposition 3 using $f(x,y)=1/y$ and (5). Here, the argument x refers to the first component of the random vector in (2), whereas y refers to the second component. Note that $f(1,1/(1-d/n)) = 1-d/n$ and $f'(1,1/(1-d/n)) = [0,-(1-d/n)^2]$. The asymptotic variance of (8) is obtained from calculating the (n,d)-limit of

$$\left[0,-(1-d/n)^2\right]\begin{bmatrix} 2 & -\frac{2}{1-c} \\ -\frac{2}{1-c} & \frac{2}{(1-c)^3}\end{bmatrix}\left[0,-(1-d/n)^2\right]^t.$$

□

In the classical setting where d remains fixed and $n \to \infty$, we have that

$$\sqrt{n}\left(\frac{\hat{\sigma}^2_{\mathrm{MP}}}{\sigma^2_{\mathrm{MP}}}-1\right) \Longrightarrow N(0,2), \tag{9}$$

which can be deduced from (7) by applying the classical central limit theorem. However, this setting may lead to wrong test decisions as we will see in section 4.

The last theorem allows us to investigate the asymptotic power function of the variance test.

Corollary 6. *The test for the hypothesis (6) based on the statistic (8) is* (n,d)-*consistent and asymptotically unbiased.*

Proof. A direct calculation shows that the asymptotic power function of this test is given by

$$\mathrm{pow}(\xi;\alpha) = \Phi\left(c\xi+\sqrt{\frac{n(1-c)}{2}}(\xi-1)\right),$$

where Φ is the standard normal distribution function, $c=\Phi^{-1}(\alpha)$, $\alpha \in (0,1)$ the level of the test, and ξ is defined in (6). We immediately see that the test is (n,d)-consistent and that the infimum of $\mathrm{pow}(\xi;\alpha)$ on $[1,\infty)$ is taken at $\xi=1$. □

3.2. Inference for the mean

The mean of the excess return of the MP is usually estimated by

$$\hat{\mu}_{\mathrm{MP}} = \frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \hat{\mu}}{\mathbb{1}^t \widehat{\Sigma}^{-1} \mathbb{1}} .$$

Frahm (2010) derives this estimator as a least squares estimator for μ_{MP} in the context of linear regression theory and provides its exact distribution for finite $n > d - 1$ as

$$\sqrt{n+1}\sqrt{1 - \frac{d}{n+1}} \frac{\hat{\mu}_{\mathrm{MP}} - \mu_{\mathrm{MP}}}{\sqrt{\hat{\sigma}^2_{\mathrm{MP}} \left(\frac{n}{n+1} + \hat{\mu}^t \widehat{\Sigma}^{-1} \hat{\mu} \right) - \hat{\mu}^2_{\mathrm{MP}}}} \sim t_{n-d+1}, \tag{10}$$

where t_{n-d+1} denotes the t-distribution with $n-d+1$ degrees of freedom. Thus, this estimator is in particular unbiased if $n > d$. The distribution of $\hat{\mu}_{\mathrm{MP}}$ is also derived by Bodnar & Okhrin (2011) who consider the null hypothesis $\mu_{\mathrm{MP}} = 0$. The next theorem shows that $\hat{\mu}_{\mathrm{MP}}$ is even a strongly (n, d)-consistent estimator for μ_{MP}.

Theorem 7. *We have under assumptions 1, 3 and 4 and the* (n, d)*-asymptotics that*

$$\hat{\mu}_{\mathrm{MP}} \xrightarrow{\text{a.s.}} a_1 .$$

Proof. We have that

$$\begin{aligned} \hat{\mu}_{\mathrm{MP}} &= \frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \hat{\mu}}{\mathbb{1}^t \widehat{\Sigma}^{-1} \mathbb{1}} = \underbrace{\frac{\mathbb{1}^t \Sigma^{-1} \mu}{\mathbb{1}^t \Sigma^{-1} \mathbb{1}}}_{=\mu_{\mathrm{MP}}} \left(\frac{z^t}{\|z\|_2} \tilde{S}^{-1} \frac{z}{\|z\|_2} \right)^{-1} \frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \hat{\mu}}{\mathbf{1}^t \Sigma^{-1} \mu} \\ &\xrightarrow{\text{a.s.}} a_1 (1 - c) \frac{1}{1-c} = a_1, \end{aligned}$$

where we have used propositions 1 and 2 and set again $z = \Sigma^{-1/2} \mathbb{1}$. □

Note that the (n, d)-limit of μ_{MP}, which is a_1, exists by assumption 3. Furthermore, in contrast to the foregoing variance test, we do not have to state hypotheses for μ_{MP} under the (n, d)-asymptotics in terms of a ratio because of this assumption. The next theorem gives the (n, d)-asymptotic distribution of $\hat{\mu}_{\mathrm{MP}}$.

Theorem 8. *Let assumptions 1 and 3 be true. Then, we have under the* (n, d)*-asymptotics the following weak convergence:*

$$\sqrt{n} \frac{\hat{\mu}_{\mathrm{MP}} - \mu_{\mathrm{MP}}}{\sqrt{\hat{\sigma}^2_{\mathrm{MP}} (1 + \hat{\mu}^t \widehat{\Sigma}^{-1} \hat{\mu}) - \hat{\mu}^2_{\mathrm{MP}}}} \Longrightarrow N\left(0, \frac{1}{1-c}\right) \tag{11}$$

Proof. The assertion follows immediately from (10). □

The classical setting of $c = 0$ can also be taken into account in theorem 8. In this case, it is

$$\sqrt{n} \frac{\hat{\mu}_{\mathrm{MP}} - \mu_{\mathrm{MP}}}{\sqrt{\hat{\sigma}^2_{\mathrm{MP}} (1 + \hat{\mu}^t \widehat{\Sigma}^{-1} \hat{\mu}) - \hat{\mu}^2_{\mathrm{MP}}}} \Longrightarrow N(0, 1) . \tag{12}$$

Since $c \approx d/n > 0$ in every practical situation, this setting may lead to wrong conclusions when performing tests based on (11) as we will see in section 4.

Note that each element of the alternative of every test for μ_{MP} (one-sided or two-sided) based on (11) must correspond to a portfolio $\mathrm{P} \in \mathcal{P}$. Furthermore, the corresponding asymptotic power functions can be easily calculated using (11), which shows (n, d)-consistency and asymptotic unbiasedness of these tests.

3.3. *Inference for the portfolio weights*

The weights of the MP are usually estimated by

$$\hat{w}_{\mathrm{MP}} = \frac{\widehat{\Sigma}^{-1}\mathbb{1}}{\mathbb{1}^t\widehat{\Sigma}^{-1}\mathbb{1}}.$$

The finite sample distribution of $\hat{w}_{\mathrm{MP}}$ is derived in Frahm (2010) as

$$\hat{w}_{\mathrm{MP}} \sim t_d\left(w_{\mathrm{MP}}, \left(\sigma^2_{\mathrm{MP}}\Sigma^{-1} - w_{\mathrm{MP}}w^t_{\mathrm{MP}}\right)/(n-d+2), n-d+2\right),$$

which is the d-dimensional t-distribution with $n-d+2$ degrees of freedom, location vector w_{MP} and dispersion matrix $\left(\sigma^2_{\mathrm{MP}}\Sigma^{-1} - w_{\mathrm{MP}}w^t_{\mathrm{MP}}\right)/(n-d+2)$ (see also Kotz & Nadarajah (2004), chapter 1, for a definition of this distribution). Thus, if $n > d-1$, then $\hat{w}_{\mathrm{MP}}$ is an unbiased estimator for w_{MP}. However, the covariance matrix of $\hat{w}_{\mathrm{MP}}$ (which exists for $n > d$) is singular because

$$\sigma^2_{\mathrm{MP}}\Sigma^{-1} - w_{\mathrm{MP}}w^t_{\mathrm{MP}} = \sigma^2_{\mathrm{MP}}\Sigma^{-1}\left(I - \mathbb{1}w^t_{\mathrm{MP}}\right),$$

and $I - \mathbb{1}w^t_{\mathrm{MP}}$ is a projector satisfying $(I - \mathbb{1}w^t_{\mathrm{MP}})\mathbb{1} = 0$. Note that this problem is induced by the linear constraint $\hat{w}^t_{\mathrm{MP}}\mathbb{1} = 1$. It can be overcome if one omits one component of $\hat{w}_{\mathrm{MP}}$. The omitted component can then be recovered from $\hat{w}^t_{\mathrm{MP}}\mathbb{1} = 1$ so that there is no loss of information. Without loss of generality, we omit the first component of $\hat{w}_{\mathrm{MP}} = (\hat{w}^1_{\mathrm{MP}}, \ldots, \hat{w}^d_{\mathrm{MP}})^t$ and define $\hat{w}^{\mathrm{ex}}_{\mathrm{MP}} := \left(\hat{w}^2_{\mathrm{MP}}, \ldots, \hat{w}^d_{\mathrm{MP}}\right)^t$. The finite sample distribution of $\hat{w}^{\mathrm{ex}}_{\mathrm{MP}}$ can be calculated as

$$\hat{w}^{\mathrm{ex}}_{\mathrm{MP}} \sim t_{d-1}\left(w^{\mathrm{ex}}_{\mathrm{MP}}, \sigma^2_{\mathrm{MP}}\Omega^{-1}/(n-d+2), n-d+2\right),$$

where $\Omega = \Delta\Sigma\Delta^t$, $\Delta := [\mathbb{1}\ \ -I] \in \mathbb{R}^{(d-1)\times d}$, and $w^{\mathrm{ex}}_{\mathrm{MP}}$ consists of the components of w_{MP} except for the first one. This result was firstly proven in Okhrin & Schmid (2006). Note that Ω is always regular as the rank of Δ is $d-1$ and Σ is positive definite. A well-known test for the portfolio weights of the MP is the F-test. It is a two-sided test which considers the hypothesis

$$H_0 : w_{\mathrm{MP}} = w_0 \ \text{ versus } \ H_1 : w_{\mathrm{MP}} \neq w_0 \tag{13}$$

and is based on the statistic

$$\frac{n-d+1}{d-1}\,\frac{(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})^t\widehat{\Sigma}(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})}{\hat{\sigma}^2_{\mathrm{MP}}} \sim F^{d-1}_{n-d+1}; \tag{14}$$

see, for example, Frahm (2010), where F^{d-1}_{n-d+1} is the F-distribution with $d-1$ and $n-d+1$ degrees of freedom.

This test has two shortcomings under the (n, d)-asymptotics. First, the hypothesis (13) must be reformulated. The reason is the following: Suppose that w_{MP} does not only fulfill $w^t_{\mathrm{MP}}\mathbb{1} = 1$ but also is an element of the sphere which is generated by the 1-norm; that is, $w_{\mathrm{MP}} \in \{w_{\mathrm{P}} \mid |w_{\mathrm{P}}|^t\mathbb{1} = 1\}$, where $|w_{\mathrm{P}}|$ is the vector of the componentwise absolute values of w_{P}. In particular, this situation occurs if short selling is not allowed. Since this sphere converges to the set $\{0\} \subset \mathbb{R}^\infty$ as $d \to \infty$, we obtain that w_{MP} converges to a sequence of zeros in the space $\mathbb{R}^\infty$ (here, we embed $\mathbb{R}^d \times \{0\} \times \{0\} \times \ldots \hookrightarrow \mathbb{R}^\infty$). Thus, we must characterize w_{MP} in another way.

The second problem is that the null-distribution of the statistic (14) degenerates under the (n,d)-asymptotics because we have for $n > d+3$ that

$$\begin{aligned}\mathbb{V}ar(Y) &= \frac{2(n-d+1)^2(n-2)}{(d-1)(n-d-1)^2(n-d-3)}\\ &= \frac{2n^3(1-(d-1)/n)^2(1-2/n)}{n^4(d-1)/n(1-(d+1)/n)^2(1-(d+3)/n)} = \mathcal{O}(n^{-1}),\end{aligned}$$

where $Y \sim F^{d-1}_{n-d+1}$. Thus, we want to derive an alternative standardization of the F-statistic which can cope with (n,d)-asymptotics. First, we give the almost sure limit of the F-statistic.

Theorem 9. *Under assumption 1, the following almost sure convergence holds under the* (n,d)-*asymptotics:*

$$\frac{(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})^t \widehat{\Sigma} (\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})}{\hat{\sigma}^2_{\mathrm{MP}}} \xrightarrow{\text{a.s.}} \frac{c}{1-c}$$

Proof. We have that

$$\hat{w}^t_{\mathrm{MP}} \widehat{\Sigma} w_{\mathrm{P}} = \hat{\sigma}^2_{\mathrm{MP}}$$

for any portfolio weight vector $w_{\mathrm{P}} \in \mathbb{R}^d$ with $w_{\mathrm{P}}^t \mathbb{1} = 1$. It follows that

$$\begin{aligned}\frac{(\hat{w}_{\mathrm{MP}} - w_{\mathrm{P}})^t \widehat{\Sigma} (\hat{w}_{\mathrm{MP}} - w_{\mathrm{P}})}{\hat{\sigma}^2_{\mathrm{MP}}} &= \frac{\hat{w}^t_{\mathrm{MP}} \widehat{\Sigma} \hat{w}_{\mathrm{MP}} - 2\hat{w}^t_{\mathrm{MP}} \widehat{\Sigma} w_{\mathrm{P}} + w^t_{\mathrm{P}} \widehat{\Sigma} w_{\mathrm{P}}}{\hat{\sigma}^2_{\mathrm{MP}}}\\ &= \frac{w^t_{\mathrm{P}} \widehat{\Sigma} w_{\mathrm{P}}}{\hat{\sigma}^2_{\mathrm{MP}}} - 1 = \frac{\sigma^2_{\mathrm{P}}}{\sigma^2_{\mathrm{MP}}} \left(\frac{\tilde{z}^t}{\|\tilde{z}\|_2} \tilde{S} \frac{\tilde{z}}{\|\tilde{z}\|_2} \right) \left(\frac{z^t}{\|z\|_2} \tilde{S}^{-1} \frac{z}{\|z\|_2} \right) - 1,\end{aligned} \tag{15}$$

where $z = \Sigma^{-1/2}\mathbb{1}$, $\tilde{z} := \Sigma^{1/2} w_{\mathrm{P}}$, and $\tilde{S}$ has been defined in section 2.1. If P = MP, we have that

$$\begin{aligned}\frac{(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})^t \widehat{\Sigma} (\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})}{\hat{\sigma}^2_{\mathrm{MP}}} &= \left(\frac{\tilde{z}^t}{\|\tilde{z}\|_2} \tilde{S} \frac{\tilde{z}}{\|\tilde{z}\|_2} \right) \left(\frac{z^t}{\|z\|_2} \tilde{S}^{-1} \frac{z}{\|z\|_2} \right) - 1\\ &\xrightarrow{\text{a.s.}} 1 \frac{1}{1-c} - 1 = \frac{c}{1-c}\end{aligned}$$

under the (n,d)-asymptotics because of proposition 1, which proves the assertion. □

This result as such is interesting: Since

$$\frac{(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})^t \widehat{\Sigma} (\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})}{\hat{\sigma}^2_{\mathrm{MP}}} = \frac{(\hat{w}^{\mathrm{ex}}_{\mathrm{MP}} - w^{\mathrm{ex}}_{\mathrm{MP}})^t \widehat{\Omega} (\hat{w}^{\mathrm{ex}}_{\mathrm{MP}} - w^{\mathrm{ex}}_{\mathrm{MP}})}{\hat{\sigma}^2_{\mathrm{MP}}},$$

where $\widehat{\Omega} = \Delta \widehat{\Sigma} \Delta^t$ (Frahm & Memmel (2010)), the F-statistic is an estimator for the squared Mahalanobis distance of $\hat{w}^{\mathrm{ex}}_{\mathrm{MP}}$ to its mean $w^{\mathrm{ex}}_{\mathrm{MP}}$ (see also Mahalanobis (1936)). We see that this distance does not converge to zero under the (n,d)-asymptotics unless $c = 0$. From (15), we see further that

$$\frac{(\hat{w}_{\mathrm{MP}} - w_0)^t \widehat{\Sigma} (\hat{w}_{\mathrm{MP}} - w_0)}{\hat{\sigma}^2_{\mathrm{MP}}} \stackrel{\mathrm{d}}{=} \frac{\sigma^2_0}{\sigma^2_{\mathrm{MP}}} \left(u^t \tilde{S} u \right) \left(v^t \tilde{S}^{-1} v \right) - 1,$$

where σ_0^2 is the excess return variance with respect to the weight vector w_0 (which itself is associated to some $\mathrm{P} \in \mathcal{P}$), $\tilde{S}$ is defined in section 2, and $u, v \in \mathbb{R}^d$ are arbitrary unit vectors. Thus, the distribution of the F-statistic under the null and the alternative only depends on the ratio $\sigma_0^2/\sigma_{\mathrm{MP}}^2$. This is why the F-test can be seen as another variance test, that is, a test for the hypothesis

$$H_0 : \frac{\sigma_0^2}{\sigma_{\mathrm{MP}}^2} = 1 \text{ versus } H_1 : \frac{\sigma_0^2}{\sigma_{\mathrm{MP}}^2} > 1 . \tag{16}$$

Due to the uniqueness of the MP, this hypothesis gives an equivalent and reasonable way to reformulate hypothesis (13) in the high-dimensional setting.

Now, we turn to the asymptotic distribution of the F-statistic under the (n, d)-asymptotics. From the central limit theorem (see also formula 26.6.13 in Abramowitz & Stegun (1972)), we have the weak convergence

$$\frac{F_l^k - 1}{\sqrt{2(k+l)/kl}} \Longrightarrow N(0,1)$$

for large k, l, where F_l^k denotes an F_l^k-distributed random variable. We conclude that

$$\begin{aligned}
&\sqrt{n}\left(\frac{(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})^t\,\widehat{\Sigma}(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})}{\hat{\sigma}_{\mathrm{MP}}^2} - \frac{d-1}{n-d+1}\right)\\
&= \sqrt{\frac{2n}{(d-1)(n-d+1)}}\sqrt{n}\frac{d-1}{n-d+1}\frac{\left(\frac{n-d+1}{d-1}\frac{(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})^t\,\widehat{\Sigma}(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})}{\hat{\sigma}_{\mathrm{MP}}^2} - 1\right)}{\sqrt{\frac{2n}{(d-1)(n-d+1)}}}\\
&\stackrel{\mathrm{d}}{=} \sqrt{\frac{2(d-1)/n}{(1-(d-1)/n)^3}}\frac{F_{n-d+1}^{d-1} - 1}{\sqrt{\frac{2n}{(d-1)(n-d-1)}}}
\end{aligned}$$

for large n, d. Thus, we suspect that the F-statistic limits a normal distribution under the (n, d)-asymptotics. The next theorem shows that this is indeed the case.

Theorem 10. *Under assumption 1, we have the following weak convergence under the (n, d)-asymptotics:*

$$\sqrt{n}\left(\frac{(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})^t\,\widehat{\Sigma}(\hat{w}_{\mathrm{MP}} - w_{\mathrm{MP}})}{\hat{\sigma}_{\mathrm{MP}}^2} - \frac{d/n}{1-d/n}\right) \Longrightarrow N\left(0, \frac{2c}{(1-c)^3}\right) \tag{17}$$

Proof. We apply the delta method to proposition 3 again. Here, we use (15) and set $f(x, y) = xy - 1$ with x referring to the first and y to the second component of the random vector in (2). We obtain $f(1, 1/(1-d/n)) = d/n/(1-d/n)$ and $f'(1, 1/(1-d/n)) = \left[1/(1-d/n), 1\right]$. The asymptotic variance of (17) is the (n, d)-limit of

$$\left[1/(1-d/n), 1\right]\begin{bmatrix} 2 & -\frac{2}{1-c} \\ -\frac{2}{1-c} & \frac{2}{(1-c)^3}\end{bmatrix}\left[1/(1-d/n), 1\right]^t .$$

□

The proof of theorem 10 requires $c > 0$ (see also the comments after proposition 3) so that the classical setting $c = 0$ must be excluded here.

Although the asymptotic power function of the F-test cannot be derived explicitly from theorem 10, we can show the following theorem.

Theorem 11. *The portfolio weight test based on the hypothesis (16) and the statistic (17) is (n,d)-consistent.*

Proof. Since it is hard to calculate the power function of the F-test explicitly, we adopt a method by Ledoit & Wolf (2002) to prove consistency. The idea is to investigate the probability limit of

$$\frac{(\hat{w}_{\mathrm{MP}} - w_0)^t \widehat{\Sigma}(\hat{w}_{\mathrm{MP}} - w_0)}{\hat{\sigma}^2_{\mathrm{MP}}}$$

for some weight vector $w_0 \in \mathbb{R}^d$ which is associated to some $\mathrm{P} \in \mathcal{P}$. Note that this limit always exists under the (n,d)-asymptotics because of the definition of $\mathcal{P}$, (15) and proposition 1. Under the null, that is, $\sigma_0^2/\sigma^2_{\mathrm{MP}} = 1$, the limit is $c/(1-c)$ due to theorem 9. Now, the test is (n,d)-consistent if and only if this limit can be obtained for $\sigma_0^2/\sigma^2_{\mathrm{MP}} = 1$, which is obviously the case because of the uniqueness of the MP. Thus, this proof is completed. □

3.4. Inference for the Sharpe ratio

Now, we turn to high-dimensional inference for the Sharpe ratio of the MP which has been defined in section 2.1. This Sharpe ratio is usually estimated by

$$\widehat{SR}_{\mathrm{MP}} := \frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \hat{\mu}}{\sqrt{\mathbb{1}^t \widehat{\Sigma}^{-1} \mathbb{1}}} .$$

The next theorem shows that this estimator is inconsistent under the (n,d)-asymptotics unless $c = 0$.

Theorem 12. *Under assumptions 1, 3 and 4, we have under the (n,d)-asymptotics that*

$$\frac{\widehat{SR}_{\mathrm{MP}}}{SR_{\mathrm{MP}}} \xrightarrow{\text{a.s.}} \frac{1}{\sqrt{1-c}} .$$

Proof. The assertion follows immediately from theorems 4 and 7. □

We see that, whereas $\hat{\sigma}^2_{\mathrm{MP}}$ underestimates σ^2_{MP} under the (n,d)-asymptotics, SR_{MP} is overestimated by $\widehat{SR}_{\mathrm{MP}}$.

As for the variance test, we have to formulate the aforementioned theorem in terms of a ratio due to the possibility of the limit (4a). Therefore, hypotheses for SR_{MP} have to be formulated in terms of SR_{MP}/SR_0, where SR_0 is a null Sharpe ratio associated to some portfolio $\mathrm{P} \in \mathcal{P}$.

Next, we provide the (n,d)-asymptotic distribution of $\widehat{SR}_{\mathrm{MP}}$.

Theorem 13. *If assumptions 1 and 3 hold, we have under the (n,d)-asymptotics the following weak convergence:*

$$\sqrt{n}\left(\frac{\widehat{SR}_{\mathrm{MP}}}{SR_{\mathrm{MP}}} - \frac{1}{\sqrt{1-d/n}}\right) \Longrightarrow N\left(0, \frac{1}{2(1-c)^2}\frac{2SR^2_{\mathrm{TP}} - SR^2_{\mathrm{MP}} + 2}{SR^2_{\mathrm{MP}}}\right) \tag{18}$$

Proof. See the Appendix. □

Here, $SR_{\mathrm{TP}} = \sqrt{\mu^t \Sigma^{-1} \mu}$ denotes the Sharpe ratio of the *tangency portfolio* (TP).

Let us assume that the (n, d)-limit of the mean excess return of the TP, $\mu_{\text{TP}} = \mu^t \Sigma^{-1} \mu / \mathbb{1}^t \Sigma^{-1} \mu$, is finite and unequal to zero. Then, the (n, d)-limit of the asymptotic variance in (18) is positive and finite because of $SR_{\text{TP}}^2 \geq SR_{\text{MP}}^2$, assumption 3 and the relationship

$$\sigma_{\text{TP}}^2 = \frac{\mu_{\text{TP}}}{\mu_{\text{MP}}} \sigma_{\text{MP}}^2,$$

where σ_{TP}^2 denotes the variance of the excess return of the TP. We do not formulate the variance in (18) in terms of the limits $a_i, i \in \{1, 2\}$, for reasons of clarity and readability.

In the theoretical case, that is, if assumption 4a holds, adding 2 to the numerator of the variance in (18) is not necessary because $2/SR_{\text{MP}}^2 \to 0$ as $d \to \infty$ due to section 2.1. But because the convergence of $SR_{\text{MP}} \to \infty$ may be very slow, this term should be added in order to obtain a better approximation of the distribution of $\widehat{SR}_{\text{MP}}/SR_{\text{MP}}$ for finite n, d.

Theorem 13 coincides with well-known results on the n-asymptotic distribution of the Sharpe ratio in the univariate case; that is, one has $d = 1$ and $c = 0$. Denote the Sharpe ratio of some univariate asset by SR and let $\widehat{SR}$ be the corresponding estimator. It is well known that, under assumption 1,

$$\sqrt{n}(\widehat{SR} - SR) \Longrightarrow N\left(0, 1 + \frac{1}{2} SR^2\right) \tag{19}$$

as $n \to \infty$ (see, e.g., Schmid & Schmidt (2010) and the references therein). Since $SR_{\text{MP}} = SR_{\text{TP}}$ for $d = 1$, theorem 13 is an extension of this result.

The asymptotic variance of (18) causes a problem for inference in practice: In order to perform tests for SR_{MP}, an estimate for this variance is needed which is independent of $\widehat{SR}_{\text{MP}}$. In the case of n-asymptotics and $d = 1$, this problem can be solved by using a transform for $\widehat{SR}_{\text{MP}}$ so that an application of the delta method drops the unknown variance out (see again Schmid & Schmidt (2010)). However, such a transform cannot exist in the current situation because the variance of (18) contains two different Sharpe ratios which cannot be eliminated both by using the delta method. Thus, future research will be dedicated to estimation methods for the unknown variance, for example, bootstrapping methods which have also been proposed in Bai *et al.* (2009) and El Karoui (2010).

4. Empirical study

In this section, we apply the results from the last section to data. Two main objectives are pursued: First, it is shown that tests based on (8), (11) and (17) lead to the same test decisions as the existing exact tests based on (3), (10) and (14). Second, as it has already been pointed out in section 2, we illustrate that hypotheses may depend on the number d of assets in the portfolio. In the following, we will assume that assumption 4b, that is, the natural case, holds.

This study is comparable with that in Frahm (2010) and is based on monthly excess return data of all firms which have been part of the S&P 500 index from January 1980 to December 2011. The return data are taken from the the Center for Research in Security Prices database, and the risk free interest rate is set to the monthly secondary market 3-month US treasury bill rate. In total, $d = 233$ firms have been listed in the S&P 500 index during the observation period of $n+1 = 32 \cdot 12 = 384$ months. Thus, the setting is high-dimensional with $c = d/n = 233/383$.

Now, we test for the following hypotheses at confidence level $\alpha = 0.05$:

$$\begin{aligned}
&H_0^1 : \sigma_{\text{MP}}^2 = (0.2)^2/12 = 0.33\% && \text{versus} && H_1^1 : \sigma_{\text{MP}}^2 < 0.33\%, \\
&H_0^2 : \mu_{\text{MP}} \leq 0.02/12 = 0.17\% && \text{versus} && H_1^2 : \mu_{\text{MP}} > 0.17\%, \\
&H_0^3 : w_{\text{MP}} = \mathbb{1}/d && \text{versus} && H_1^3 : w_{\text{MP}} \neq \mathbb{1}/d\ .
\end{aligned}$$

These hypotheses are also considered in Frahm (2010) with respect to a 10-dimensional portfolio, that is, a low-dimensional setting. We obtain from the exact tests that

$$\begin{aligned}
n\frac{\hat{\sigma}^2_{\mathrm{MP}}}{0.0033} &= 27.9445 < 123.5968 = F^{-1}_{\chi^2_{151}}(\alpha),\\
\sqrt{n-d+1}\frac{\hat{\mu}_{\mathrm{MP}} - 0.0017}{\sqrt{\hat{\sigma}^2_{\mathrm{MP}}\left(\frac{n}{n+1} + \hat{\mu}^t\widehat{\Sigma}^{-1}\hat{\mu}\right) - \hat{\mu}^2_{\mathrm{MP}}}} &= 3.1752 > 1.6550 = F^{-1}_{t,151}(1-\alpha),\\
\frac{n-d+1}{d-1}\frac{(\hat{w}_{\mathrm{MP}} - \mathbb{1}/d)^t\widehat{\Sigma}(\hat{w}_{\mathrm{MP}} - \mathbb{1}/d)}{\hat{\sigma}^2_{\mathrm{MP}}} &= 5.0130 > 1.2809 = F^{-1}_{F,232,151}(1-\alpha)\,.
\end{aligned}$$

Thus, all null hypotheses can be rejected. The rejections of H_0^1 and H_0^3 are in line with the results of Frahm (2010), whereas the rejection of H_0^2 is not. In Frahm (2010), the observation period is from January 1980 to November 2003. Thus, the financial crisis in 2008 is not taken into account, which may be the reason for these two different findings.

The aforementioned results can be replicated with the new test statistics. It is

$$\begin{aligned}
\sqrt{n}\left(\frac{\hat{\sigma}^2_{\mathrm{MP}}}{0.0033} - (1-c)\right) &= -6.2367\\
< \Phi^{-1}(\alpha)\sqrt{2(1-c)} &=: q_1(\alpha) = -1.4558,\\
\sqrt{n}\frac{\hat{\mu}_{\mathrm{MP}} - 0.0017}{\sqrt{\hat{\sigma}^2_{\mathrm{MP}}\left(1 + \hat{\mu}^t\widehat{\Sigma}^{-1}\hat{\mu}\right) - \hat{\mu}^2_{\mathrm{MP}}}} &= 5.0528\\
> \Phi^{-1}(1-\alpha)\sqrt{\frac{1}{1-c}} &=: q_2(1-\alpha) = 2.6283,\\
\sqrt{n}\left(\frac{(\hat{w}_{\mathrm{MP}} - \mathbb{1}/d)^t\widehat{\Sigma}(\hat{w}_{\mathrm{MP}} - \mathbb{1}/d)}{\hat{\sigma}^2_{\mathrm{MP}}} - \frac{c}{1-c}\right) &= 120.3343\\
> \Phi^{-1}(1-\alpha)\sqrt{\frac{2c}{(1-c)^3}} &=: q_3(1-\alpha) = 7.4026,
\end{aligned}$$

where Φ is the standard normal distribution function. Again, all hypotheses can be rejected (here, the hypotheses are transformed or identified, respectively, according to the discussion in section 3). Further, the p-values of all considered tests are close to zero, which shows again that the exact and (n, d)-asymptotic tests give the same results.

Next, we consider exact and asymptotic one-sided confidence intervals for σ^2_{MP} and μ_{MP} at level $1-\alpha = 0.95$, that is, $\alpha = 0.05$. It is $\hat{\sigma}^2_{\mathrm{MP}} = 0.0243\%$ and $\hat{\mu}_{\mathrm{MP}} = 0.6773\%$ so that they are given by

$$\begin{aligned}
\left[0, \frac{n\hat{\sigma}^2_{\mathrm{MP}}}{F^{-1}_{\chi^2_{151}}(\alpha)}\right] &= [0, 0.0754\%],\\
\left[0, \frac{\sqrt{n}\hat{\sigma}^2_{\mathrm{MP}}}{q_1(\alpha) + \sqrt{n}(1-c)}\right] &= [0, 0.0767\%]
\end{aligned}$$

concerning σ^2_{MP} and by

$$\left[\hat{\mu}_{\mathrm{MP}} - F^{-1}_{t,151}(1-\alpha)\frac{\sqrt{\hat{\sigma}^2_{\mathrm{MP}}\left(\frac{n}{n+1} + \hat{\mu}^t\widehat{\Sigma}^{-1}\hat{\mu}\right) - \hat{\mu}^2_{\mathrm{MP}}}}{\sqrt{n-d+1}}, \infty\right) = [0.4111\%, \infty),$$

$$\left[\hat{\mu}_{\mathrm{MP}} - q_2(1-\alpha)\frac{\sqrt{\hat{\sigma}^2_{\mathrm{MP}}\left(1 + \hat{\mu}^t\widehat{\Sigma}^{-1}\hat{\mu}\right) - \hat{\mu}^2_{\mathrm{MP}}}}{\sqrt{n}}, \infty\right) = [0.4117\%, \infty)$$

concerning μ_{MP}. Thus, the relative approximation errors of the right end and the left end, respectively, of the asymptotic confidence intervals are

$$\frac{|0.0754\% - 0.0767\%|}{0.0754\%} = 1.7172\%$$

and

$$\frac{|0.4111\% - 0.4117\%|}{0.4111\%} = 0.1307\%,$$

respectively. The first error is less than 2% and therefore small enough to indicate that the central limit theorem in proposition 3 holds. As expected, the second relative approximation error is also small as the t-distribution reaches the normal distribution for large degrees of freedom. So, the asymptotic tests for σ^2_{MP} and μ_{MP} give proper alternatives for the exact tests in the high-dimensional setting.

In order to see that the n-asymptotic approach is not suitable here, we compare the (n, d)-asymptotic confidence intervals for σ^2_{MP} and μ_{MP} with the corresponding n-asymptotic ones. They are based on (9) and (12) and are given by

$$\left[0, \frac{\sqrt{n}\hat{\sigma}^2_{\mathrm{MP}}}{\Phi^{-1}(\alpha)\sqrt{2} + \sqrt{n}}\right] = [0, 0.0276\%]$$

for σ^2_{MP} and by

$$\left[\hat{\mu}_{\mathrm{MP}} - \Phi^{-1}(1-\alpha)\frac{\sqrt{\hat{\sigma}^2_{\mathrm{MP}}\left(1 + \hat{\mu}^t\widehat{\Sigma}^{-1}\hat{\mu}\right) - \hat{\mu}^2_{\mathrm{MP}}}}{\sqrt{n}}, \infty\right) = [0.5111\%, \infty)$$

for μ_{MP}. The length of the n-asymptotic confidence interval for σ^2_{MP} is about three times smaller than the (n, d)-asymptotic one. We see further that the n-asymptotic confidence interval for μ_{MP} is also smaller than the (n, d)-asymptotic one in the sense that $[0.5111\%, \infty) \subset [0.4117\%, \infty)$. Hence, one should be very careful with the n-asymptotic approach as it leads under high-dimensionality to wrong rejections of tests for σ^2_{MP} and μ_{MP}.

Next, we analyze the exact and asymptotic confidence ellipsoids at level $1-\alpha$ of w_{MP} which are given by

$$\left\{ w \in \mathbb{R}^d \;\middle|\; \frac{n-d+1}{d-1} \frac{(\hat{w}_{\mathrm{MP}} - w)^t \widehat{\Sigma} (\hat{w}_{\mathrm{MP}} - w)}{\hat{\sigma}^2_{\mathrm{MP}}} \le F^{-1}_{F,233,151}(1-\alpha) \right\}$$

$$= \left\{ w \in \mathbb{R}^d \;\middle|\; (\hat{w}_{\mathrm{MP}} - w)^t \left[\underbrace{\frac{d-1}{n-d+1} F^{-1}_{F,233,151}(1-\alpha) \hat{\sigma}^2_{\mathrm{MP}} \widehat{\Sigma}^{-1}}_{=:A_1} \right]^{-1} (\hat{w}_{\mathrm{MP}} - w) \le 1 \right\}$$

and

$$\left\{ w \in \mathbb{R}^d \;\middle|\; \sqrt{n} \left(\frac{(\hat{w}_{\mathrm{MP}} - w)^t \widehat{\Sigma} (\hat{w}_{\mathrm{MP}} - w)}{\hat{\sigma}^2_{\mathrm{MP}}} - \frac{c}{1-c} \right) \le q_3(1-\alpha) \right\}$$

$$= \left\{ w \in \mathbb{R}^d \;\middle|\; (\hat{w}_{\mathrm{MP}} - w)^t \left[\underbrace{\left(\frac{c}{1-c} + \frac{q_3(1-\alpha)}{\sqrt{n}} \right) \hat{\sigma}^2_{\mathrm{MP}} \widehat{\Sigma}^{-1}}_{=:A_2} \right]^{-1} (\hat{w}_{\mathrm{MP}} - w) \le 1 \right\}.$$

The volumes of these ellipsoids are

$$\frac{\pi^{d/2}}{\Gamma(d/2+1)} \sqrt{\det(A_1)} = 1.0424 \cdot 10^{-218}$$

and

$$\frac{\pi^{d/2}}{\Gamma(d/2+1)} \sqrt{\det(A_2)} = 1.1882 \cdot 10^{-219},$$

where Γ denotes the gamma function. Here, we obtain a very large relative approximation error of 88.6011%, although the absolute approximation error is only $9.2355 \cdot 10^{-219}$. Since the asymptotic ellipsoid has a smaller volume than the exact ellipsoid, the asymptotic test appears to be more conservative than the exact one. However, the volumes of the exact and the asymptotic confidence ellipsoids for w_{MP} attain a level which makes their elements almost indistinguishable from a sequence of zeros. This finding is in line with the considerations in section 3.3. Thus, for large d, one should be very careful with both portfolio weight tests as a sequence of zeros does not uniquely define a certain portfolio weight vector. Problems of this kind for the F-test are well known and have been extensively discussed in the literature (see, e.g., Goeman *et al.* (2006)). A possible theoretical solution to this problem is given in section 3.3, where it has been pointed out that one should rather regard the test for w_{MP} as a further variance test.

5. Conclusion

This paper analyzes inference for the MP under (n,d)-asymptotics and the assumption of normality. Although the estimator for the mean excess return of the MP remains consistent under this high-dimensional setting, the estimators for the excess return variance and Sharpe ratio of the MP suffer from certain inconsistencies. Since the exact distribution of some test statistics degenerates under the (n,d)-asymptotics, the corresponding (n,d)-asymptotic distributions are derived which turn out to be natural extensions of the existing exact ones. The correct statement of hypotheses for large-dimensional portfolios is another concern of this paper. Although hypotheses for the mean excess return can be stated as usual, hypotheses for the corresponding variances and Sharpe ratios should be formulated in terms of ratios. Since hypotheses for the portfolio weights of the MP remain problematic, these hypotheses are identified by scalar ones in terms of the corresponding excess return variance. At last, an empirical study shows the usefulness of the new asymptotic approach as it reproduces the exact tests' inferences. Relaxing the normality assumption will be one of the main concerns of future research.

Acknowledgements

The author would like to thank Paavo Salminen and two anonymous referees for their helpful comments and suggestions which helped to improve this paper. The valuable advice of Karl Mosler is gratefully acknowledged as well. Many thanks also belong to Stefan Erdorf and Oliver Grothe for fruitful discussions.

References

Abramowitz, M. & Stegun, I. (eds). (1972). *Handbook of mathematical functions with formulas, graphs and mathematical tables*, (10th ed.), National Bureau of Standards, Washington.

Bai, Z., Liu, H. & Wong, W.-K. (2009). Enhancement of the applicability of Markowitz's portfolio optimization by utilizing random matrix theory. *Math. Finance* **19**, 639–667.

Bai, Z., Liu, H. & Wong, W.-K. (2011). Asymptotic properties of eigenmatrices of large sample covariance matrix. *Ann. Appl. Probab.* **21**, 1994–2015.

Bai, Z., Miao, B. Q. & Pan, G. M. (2007). On asymptotics of eigenvectors of large sample covariance matrix. *Ann. Probab.* **35**, 1532–1572.

Bai, Z. & Silverstein, J. W. (2010). *Spectral analysis of large dimensional random matrices*, (2nd ed.), Springer, New York.

Bai, Z. & Yin, Y. Q. (1993). Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. *Ann. Probab.* **21**, 1275–1294.

Birke, M. & Dette, H. (2005). A note on testing the covariance matrix for large dimension. *Stat. Probabil. Lett.* **74**, 281–289.

Bodnar, T. & Okhrin, Y. (2011). On the product of inverse Wishart and normal distributions with applications to discriminant analysis and portfolio theory. *Scand. J. Stat.* **38**, 311–331.

Chopra, V. K. & Ziemba, W. T. (1993). The effects of errors in means, variances and covariances on optimal portfolio choice. *J. Port. Management* **19**, 6–11.

Cochrane, J. H. (2005). *Asset pricing*, Princeton University Press, New Jersey. revised edition.

Das Gupta, S. (1968). Some aspects of discrimination function coefficients. *Sankhyā A* **30**, 387–400.

El Karoui, N. (2010). High-dimensionality effects in the Markowitz problem and other quadratic programs with linear constraints: risk underestimation. *Ann. Statist.* **38**, 3487–3566.

Elton, E. J., Gruber, M. J., Brown, S. J. & Goetzmann, W. N. (2007). *Modern portfolio theory and investment analysis*, (7th ed.), Wiley & Sons.

Frahm, G. (2010). Linear statistical inference for global and local minimum variance portfolios. *Statist. Papers* **51**, 789–812.

Frahm, G. & Memmel, C. (2010). Dominating estimators for minimum-variance portfolios. *J. Econometrics* **159**, 289–302.

Girko, V. L. (1990). *Theory of random determinants*, (1st ed.), Kluwer Academic Publishers, Boston.

Girko, V. L. (1995). *Statistical analysis of observations of increasing dimension*, (1st ed.), Kluwer Academic Publishers, Boston.

Goeman, J. J., van de Geer, S. A. & van Houwelingen, H. C. (2006). Testing against a high dimensional alternative. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **68**, 477–493.

Jagannathan, R. & Ma, T. (2003). Risk reduction in large portfolios: why imposing the wrong constraints helps. *J. Finance* **58**, 1651–1684.

Kollo, T. & von Rosen, D. (2005). *Advanced multivariate statistics with matrices*, (1st ed.), Springer, Dordrecht, The Netherlands.

Kotz, S. & Nadarajah, S. (2004). *Multivariate t distributions and their applications*, (1st ed.), Cambridge University Press, Cambridge.

Ledoit, O. & Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. Statist.* **30**, 1081–1102.
Ledoit, O. & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empirical Finance* **10**, 603–621.
Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Rev. Economics and Statistics* **47**, 13–37.
Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proc. Nat. Instit. Sci. India* **2**, 49–55.
Markowitz, H. (1952). Portfolio selection. *J. Finance* **7**, 77–91.
Marčenko, V. A. & Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Math. USSR-Sb.* **1**, 457–483.
Merton, R. C. (1980). On estimating the expected return on the market: an exploratory investigation. *J. Financial Economics* **8**, 323–361.
Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*, (1st ed.), Wiley & Sons, New York.
Okhrin, Y. & Schmid, W. (2006). Distributional properties of portfolio weights. *J. Econometrics* **134**, 235–256.
Schmid, F. & Schmidt, R. (2010). Statistical inference for Sharpe's ratio. In *Interest rate models, asset allocation and quantitative techniques for central banks and sovereign wealth funds* (eds Berkelaar, A., Coche, J. & Nyholm, K.), Palgrave Macmillan, Basingstoke; 337–357.
Schott, J. (2005). *Matrix analysis for statistics*, (2nd ed.), Wiley & Sons, New York.
Serdobolskii, V. (2000). *Multivariate statistical analysis - a high-dimensional approach*, (1st ed.), Kluwer Academic Publishers, Boston.
Sharpe, W. F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *J. Finance* **19**, 425–442.
Silverstein, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *J. Multivariate Anal.* **55**, 331–339.
Siskind, V. (1972). Second moments of inverse Wishart-matrix elements. *Biometrika* **59**, 690–691.
Tobin, J. (1958). Liquidity preference as behavior towards risk. *Rev. Economic Studies* **25**, 65–86.
Yin, Y. Q. (1986). Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.* **20**, 50–68.

Received May 2012, in final form September 2013

Konstantin Glombek, Seminar für Wirtschafts- und Sozialstatistik, Universität zu Köln, Albertus-Magnus-Platz, 50923 Köln, Germany.
E-mail: glombek@wiso.uni-koeln.de

Appendix A: Proofs

This appendix provides all skipped proofs. We begin with the proof of proposition 2.

A1. *Proof of proposition 2*

Proof. From the proof of part (c) of lemma 3.1 in Bai *et al.* (2009), we have that

$$\frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \hat{\mu}}{\mathbb{1}^t \Sigma^{-1} \mu} = \frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \mu}{\mathbb{1}^t \Sigma^{-1} \mu} + \frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \hat{\mu}^*}{\mathbb{1}^t \Sigma^{-1} \mu},$$

where $\hat{\mu}^*$ is the sample mean of $R_i^* \overset{\text{i.i.d.}}{\sim} N_d(0, \Sigma)$, $1 \leq i \leq n+1$. The mentioned lemma gives

$$\frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \mu}{\mathbb{1}^t \Sigma^{-1} \mu} \xrightarrow{\text{a.s.}} \frac{1}{1-c} .$$

It remains to investigate the second summand. From the Cauchy–Schwarz inequality, we obtain:

$$\left|\frac{\mathbb{1}^t\widehat{\Sigma}^{-1}\hat{\mu}^*}{\mathbb{1}^t\Sigma^{-1}\mu}\right|^2 \le \frac{\mathbb{1}^t\widehat{\Sigma}^{-1}\mathbb{1}}{\mathbb{1}^t\Sigma^{-1}\mu}\,\frac{\hat{\mu}^{*t}\widehat{\Sigma}^{-1}\hat{\mu}^*}{\mathbb{1}^t\Sigma^{-1}\mu} = \underbrace{\frac{z^t}{\|z\|_2}\tilde{S}^{-1}\frac{z}{\|z\|_2}}_{\xrightarrow{\text{a.s.}} 1/(1-c)}\underbrace{\frac{\mathbb{1}^t\Sigma^{-1}\mathbb{1}}{\mathbb{1}^t\Sigma^{-1}\mu}}_{\xrightarrow{\text{a.s.}} 1/a_1}\frac{\hat{\mu}^{*t}\widehat{\Sigma}^{-1}\hat{\mu}^*}{\mathbb{1}^t\Sigma^{-1}\mu},$$

where $z := \Sigma^{-1/2}\mathbb{1}$, and $\tilde{S}$ has been defined in section 2.1. From the proof of (A.9) in Bai *et al.* (2009), we have that $\hat{\mu}^{*t}\widehat{\Sigma}^{-1}\hat{\mu}^* \xrightarrow{\text{a.s.}} 0$. Thus, we should exclude the case of $\mathbb{1}^t\Sigma^{-1}\mu \to 0$ as $d \to \infty$. Assume that this case is true. Then, we conclude from assumption 3 that $\mathbb{1}^t\Sigma^{-1}\mathbb{1} \to 0$ as $d \to \infty$. This implies that $\sigma^2_{\text{MP}} = 1/\mathbb{1}^t\Sigma^{-1}\mathbb{1} \to \infty$, which contradicts assumption 4. Thus, this proof is completed. □

Next, we prove proposition 3.

A2. *Proof of proposition 3*

First, we need to know about the first two ordinary and inverse moments of the standard Marčenko–Pastur law.

Lemma 14. *Let $F_c(x)$ be the standard Marčenko–Pastur law; that is,*

$$\mathrm{d}F_c(x) = \frac{1}{2\pi xc}\sqrt{(a_+ - x)(x - a_-)}\,\mathbb{1}_{[a_-,a_+]}(x)\,\mathrm{d}x,$$

where $a_\pm = (1 \pm \sqrt{c})^2$ and $c \in (0,1)$. We have the following identities:

$$\int x\,\mathrm{d}F_c(x) = 1, \qquad \int x^2\,\mathrm{d}F_c(x) = 1 + c,$$
$$\int \frac{1}{x}\,\mathrm{d}F_c(x) = \frac{1}{1-c}, \qquad \int \frac{1}{x^2}\,\mathrm{d}F_c(x) = \frac{1}{(1-c)^3}\,.$$

Proof. The first and the second identity are given in Bai & Silverstein (2010), lemma 3.1, and the third one in Bai *et al.* (2009). The last identity is obtained from

$$\begin{aligned}
\int \frac{1}{x^2}\,\mathrm{d}F_c(x) &= \frac{1}{2\pi c}\int_{a_-}^{a_+}\frac{\sqrt{(a_+ - x)(x - a_-)}}{x^3}\,\mathrm{d}x \\
&\overset{(*)}{=} \frac{1}{2\pi c}\int_0^{\pi}\frac{\sqrt{4c\sin^2(\theta)}}{(1 + c - 2\sqrt{c}\cos(\theta))^3}2\sqrt{c}\sin(\theta)\,\mathrm{d}\theta \\
&= \frac{2}{\pi}\int_0^{\pi}\frac{\sin^2(\theta)}{(1 + c - 2\sqrt{c}\cos(\theta))^3}\,\mathrm{d}\theta \\
&= \frac{1}{\pi}\int_0^{2\pi}\frac{\sin^2(\theta)}{(1 + c - 2\sqrt{c}\cos(\theta))^3}\,\mathrm{d}\theta \\
&= \frac{1}{\pi}\int_0^{2\pi}\frac{1}{(1 + c - \sqrt{c}(\mathrm{e}^{i\theta} + \mathrm{e}^{-i\theta}))^3}\left(\frac{\mathrm{e}^{i\theta} - \mathrm{e}^{-i\theta}}{2i}\right)^2\mathrm{d}\theta \\
&= -\frac{1}{4\pi i}\oint_{|z|=1}\frac{1}{(1 + c - \sqrt{c}(z + 1/z))^3}\left(z - \frac{1}{z}\right)^2\frac{1}{z}\,\mathrm{d}z \\
&= -\frac{1}{4\pi i}\oint_{|z|=1}\frac{(z^2 - 1)^2}{(1 + c - \sqrt{c}(z + 1/z))^3 z^3}\,\mathrm{d}z
\end{aligned}$$

$$= \frac{1}{4\pi i c^{3/2}} \oint_{|z|=1} \frac{(z^2-1)^2}{(z-\sqrt{c})^3 (z - 1/\sqrt{c})^3} \, \mathrm{d}z$$
$$\overset{(**)}{=} \frac{1}{2c^{3/2}} \frac{2c^{3/2}}{(1-c)^3} = \frac{1}{(1-c)^3}.$$

$(*)$: Substitute $x(\theta) = 1 + c - 2\sqrt{c}\cos(\theta)$. Then: $x(0) = a_-, x(\pi) = a_+, x'(\theta) = 2\sqrt{c}\sin(\theta)$. $(**)$: The integral is calculated using the residue theorem. The integrand has two poles of order three in $\sqrt{c}$ and $1/\sqrt{c}$. Since $c < 1$, we have that $1/\sqrt{c}$ does not lie within the unit circle so that the corresponding residue is not of interest. □

Now, we come to the proof of proposition 3.

Proof. From theorem 3 in Bai *et al.* (2007), we have under the (n, d)-asymptotics that

$$\sqrt{n} \begin{pmatrix} u^t \tilde{S} u - 1 \\ v^t \tilde{S}^{-1} v - \frac{1}{1-d/n} \end{pmatrix}$$

converges weakly to a bivariate normal with mean zero and covariance matrix

$$\frac{2}{c} \begin{bmatrix} \int x^2 \, \mathrm{d}F_c(x) - \left(\int x \, \mathrm{d}F_c(x)\right)^2 & \int \mathrm{d}F_c(x) - \int x \, \mathrm{d}F_c(x) \int \frac{1}{x} \, \mathrm{d}F_c(x) \\ \int \mathrm{d}F_c(x) - \int x \, \mathrm{d}F_c(x) \int \frac{1}{x} \, \mathrm{d}F_c(x) & \int \frac{1}{x^2} \, \mathrm{d}F_c(x) - \left(\int \frac{1}{x} \, \mathrm{d}F_c(x)\right)^2 \end{bmatrix}.$$

Inserting the identities from lemma 14 gives the result. □

A3. *Proof of theorem 13*

In order to prove theorem 13, we need the following well-known results about Wishart matrices.

Lemma 15. *The following assertions hold under assumption 1:*

(1.) $\hat{\mu}$ *and* $\widehat{\Sigma}$ *are stochastically independent.*
(2.) $\mathbb{E}\left(\widehat{\Sigma}^{-1}\right) = \frac{n}{n-d-1} \Sigma^{-1}$, *provided that* $n - d - 1 > 0$.
(3.) $\mathbb{C}ov\left(\widehat{\Sigma}^{-1}\hat{\mu}\right) = \frac{n^2\left(\mu^t \Sigma^{-1} \mu \Sigma^{-1} + \frac{n-1}{n+1} \Sigma^{-1} + \frac{n-d+1}{n-d-1} \Sigma^{-1} \mu \mu^t \Sigma^{-1}\right)}{(n-d)(n-d-1)(n-d-3)}$, *provided that* $n - d - 3 > 0$.
(4.) $\mathbb{E}\left(\widehat{\Sigma}^{-1} a b^t \widehat{\Sigma}^{-1}\right) = \frac{n^2 \Sigma^{-1} a b^t \Sigma^{-1}}{(n-d)(n-d-3)} + \frac{n^2 \Sigma^{-1} a^t \Sigma^{-1} b}{(n-d)(n-d-1)(n-d-3)}$ *for* $a, b \in \mathbb{R}^d$, *provided that* $n - d - 3 > 0$.

Proof. The first and second assertions are well known and can be found, for example, in Muirhead (1982), chapter 3. The third one is due to Das Gupta (1968), corollary 2.1.2, and the fourth one is a variant of Siskind (1972). The last two assertions can also be concluded from more general results given by Kollo & von Rosen (2005), p. 257. □

Now, we provide the proof of theorem 13 which is based on the following theorem.

Theorem 16. *Under assumptions 1 and 3, we have the following weak convergence under the* (n, d)*-asymptotics:*

$$\sqrt{n} \begin{pmatrix} \frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \hat{\mu}}{\mathbb{1}^t \Sigma^{-1} \mu} - \frac{1}{1-d/n} \\ \frac{\mathbb{1}^t \widehat{\Sigma}^{-1} \mathbb{1}}{\mathbb{1}^t \Sigma^{-1} \mathbb{1}} - \frac{1}{1-d/n} \end{pmatrix} \Longrightarrow N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \frac{1}{(1-c)^3} \begin{bmatrix} 1 + \frac{SR_{\mathrm{TP}}^2 + 1}{SR_{\mathrm{MP}}^2} & 2 \\ 2 & 2 \end{bmatrix}\right),$$

provided that $\mathbb{1}^t \Sigma^{-1} \mu \neq 0$.

Proof. Since $\hat{\mu}$ is independent from $\widehat{\Sigma}$, the asymptotic normality with mean zero of

$$\sqrt{n}\begin{pmatrix} \dfrac{\mathbb{1}^t\widehat{\Sigma}^{-1}\hat{\mu}}{\mathbb{1}^t\Sigma^{-1}\mu} - \dfrac{1}{1-d/n} \\ \dfrac{\mathbb{1}^t\widehat{\Sigma}^{-1}\mathbb{1}}{\mathbb{1}^t\Sigma^{-1}\mathbb{1}} - \dfrac{1}{1-d/n} \end{pmatrix} \tag{20}$$

follows from Bai *et al.* (2011), corollary 1. There, a central limit theorem under (n,d)-asymptotics is stated for a transformed version of a random vector of the type (20) given $\hat{\mu} = \mu$. Due to proposition 3, it remains to calculate the non-transformed asymptotic variance of the first component of (20) and the corresponding asymptotic covariance. Due to lemma 15, 3., it is

$$\begin{aligned}
\mathbb{V}ar\left(\sqrt{n}\frac{\mathbb{1}^t\widehat{\Sigma}^{-1}\hat{\mu}}{\mathbb{1}^t\Sigma^{-1}\mu}\right) &= \frac{n}{(\mathbb{1}^t\Sigma^{-1}\mu)^2}\mathbb{1}^t\mathbb{C}ov\left(\widehat{\Sigma}^{-1}\hat{\mu}\right)\mathbb{1} \\
&= \frac{n}{(\mathbb{1}^t\Sigma^{-1}\mu)^2}\mathbb{1}^t\frac{n^2\left(\mu^t\Sigma^{-1}\mu\Sigma^{-1} + \frac{n-1}{n+1}\Sigma^{-1} + \frac{n-d+1}{n-d-1}\Sigma^{-1}\mu\mu^t\Sigma^{-1}\right)}{(n-d)(n-d-1)(n-d-3)}\mathbb{1} \\
&= \frac{n^3\left(\mu^t\Sigma^{-1}\mu\mathbb{1}^t\Sigma^{-1}\mathbb{1} + \frac{n-1}{n+1}\mathbb{1}^t\Sigma^{-1}\mathbb{1} + \frac{n-d+1}{n-d-1}\left(\mathbb{1}^t\Sigma^{-1}\mu\right)^2\right)}{(n-d)(n-d-1)(n-d-3)\left(\mathbb{1}^t\Sigma^{-1}\mu\right)^2} \\
&\longrightarrow \frac{1}{(1-c)^3}\left(1 + \frac{SR_{\mathrm{TP}}^2 + 1}{SR_{\mathrm{MP}}^2}\right).
\end{aligned}$$

Further, we have that

$$\begin{aligned}
&\mathbb{C}ov\left(\sqrt{n}\frac{\mathbb{1}^t\widehat{\Sigma}^{-1}\hat{\mu}}{\mathbb{1}^t\Sigma^{-1}\mu}, \sqrt{n}\frac{\mathbb{1}^t\widehat{\Sigma}^{-1}\mathbb{1}}{\mathbb{1}^t\Sigma^{-1}\mathbb{1}}\right) \\
&= \frac{n}{\mathbb{1}^t\Sigma^{-1}\mu\mathbb{1}^t\Sigma^{-1}\mathbb{1}}\mathbb{1}^t\left[\mathbb{E}\left(\widehat{\Sigma}^{-1}\hat{\mu}\mathbb{1}^t\widehat{\Sigma}^{-1}\right) - \mathbb{E}\left(\widehat{\Sigma}^{-1}\hat{\mu}\right)\mathbb{E}\left(\mathbb{1}^t\widehat{\Sigma}^{-1}\right)\right]\mathbb{1} \\
&= \frac{n}{\mathbb{1}^t\Sigma^{-1}\mu\mathbb{1}^t\Sigma^{-1}\mathbb{1}}\mathbb{1}^t\left[\mathbb{E}_{\widehat{\Sigma}}\left(\mathbb{E}\left(\widehat{\Sigma}^{-1}\hat{\mu}\mathbb{1}^t\widehat{\Sigma}^{-1}\,\middle|\,\widehat{\Sigma}^{-1}\right)\right) - \mathbb{E}\left(\widehat{\Sigma}^{-1}\right)\mathbb{E}(\hat{\mu})\,\mathbb{1}^t\mathbb{E}\left(\widehat{\Sigma}^{-1}\right)\right]\mathbb{1} \\
&= \frac{n}{\mathbb{1}^t\Sigma^{-1}\mu\mathbb{1}^t\Sigma^{-1}\mathbb{1}}\mathbb{1}^t\left[\mathbb{E}\left(\widehat{\Sigma}^{-1}\mathbb{E}(\hat{\mu})\mathbb{1}^t\widehat{\Sigma}^{-1}\right) - \frac{n^2\Sigma^{-1}\mu\mathbb{1}^t\Sigma^{-1}}{(n-d-1)^2}\right]\mathbb{1} \\
&= \frac{n}{\mathbb{1}^t\Sigma^{-1}\mu\mathbb{1}^t\Sigma^{-1}\mathbb{1}}\mathbb{1}^t\left[\frac{n^2\Sigma^{-1}\mu\mathbb{1}^t\Sigma^{-1}}{(n-d)(n-d-3)} + \frac{n^2\Sigma^{-1}\mu^t\Sigma^{-1}\mathbb{1}}{(n-d)(n-d-1)(n-d-3)}\right. \\
&\qquad\qquad \left. - \frac{n^2\Sigma^{-1}\mu\mathbb{1}^t\Sigma^{-1}}{(n-d-1)^2}\right]\mathbb{1} \\
&= \frac{n^3 2(n-d)}{(n-d)(n-d-1)^2(n-d-3)} \longrightarrow \frac{2}{(1-c)^3},
\end{aligned}$$

where we have used lemma 15, 1., 2. and 4. □

Note that the covariance matrix given in theorem 16 is always non-singular under assumption 4b. It is also non-singular under assumption 4a unless MP = TP. However, this can only happen in the theoretical case that the slope of the capital market line approaches infinity.

The proof of theorem 13 is an application of the delta method to theorem 16 using the function $f(x,y) = x/\sqrt{y}$, where x refers to the first and y to the second component of the random vector in that theorem.