WORKING PAPER 11/2023 (ECONOMICS AND STATISTICS)

A test on the location of tangency portfolio for small sample size and singular covariance matrix

Svitlana Drin, Stepan Mazur and Stanislas Muhinyuza

ISSN 1403-0586

Örebro University School of Business
SE-701 82 Örebro, Sweden

A test on the location of tangency portfolio for small sample size and singular covariance matrix

Svitlana Drin[1,2], Stepan Mazur[1,3], and Stanislas Muhinyuza[3]

[1]School of Business, Örebro University, 70182 Örebro, Sweden
[2]Department of Mathematics, National University of Kyiv-Mohyla Academy, 04070 Kyiv, Ukraine
[3]School of Business and Economics, Linnaeus University, 35195 Växjö, Sweden

October 18, 2023

Abstract

In this paper, we propose the test for the location of the tangency portfolio on the set of feasible portfolios when both the population and the sample covariance matrices of asset returns are singular. We derive the exact distribution of the test statistic under both the null and alternative hypotheses. Furthermore, we establish the high-dimensional asymptotic distribution of that test statistic when both the portfolio dimension and the sample size increase to infinity. We complement our theoretical findings by comparing the high-dimensional asymptotic test with an exact finite sample test in the numerical study. A good performance of the obtained results is documented.

Keywords: Tangency portfolio, Hypothesis testing, Singular Wishart distribution, Singular covariance matrix, Moore-Penrose inverse, High-dimensional asymptotics.

1 Introduction

Modern portfolio theory, which was introduced by Harry Max Markowitz in the 1950s, has been a fundamental concept in investment theory. It aims to find an investment portfolio that considers the investor's beliefs regarding risk and return. The optimal portfolio, known as the tangency portfolio (TP), determines how an investor should allocate their wealth between risk-free asset and risky assets. TP plays a crucial role in the financial literature as it is the only portfolio that maximizes the Sharpe ratio and is often used as the market portfolio in the capital asset pricing model. Therefore, a perfect understanding of its properties under all real conditions is of great importance for any financial actor. When constructing a portfolio, investors typically rely on data estimation rather than specifying all parameters explicitly, introducing estimation uncertainty into the allocation process. Quantifying this uncertainty is crucial for investors, as their expectations may not align with the actual performance of the portfolio. Moreover, communicating these uncertainties to stakeholders and adhering to regulatory frameworks is easier with analytic results.

Modern portfolio theory, introduced by Markowitz (1952), marked an early milestone in the formalization of the asset allocation decision-making process. Over the following decades, researchers have continued to advance this theory, enhancing methods for portfolio assessment and management. The extensive body of literature on modern portfolio theory has extensively investigated the ramifications of estimation uncertainty in a general context. Notable studies include the works of Okhrin and Schmid (2006), Bai et al. (2009), Ledoit and Wolf (2017), and Li et al. (2022), among many others. Research on the topic of the TP can be traced back to the late 1970s, with contributions from Winkler (1973), Klein and Bawa (1976), and Jorion (1986) conducting Bayesian analyses of the TP. Jobson and Korkie (1980) provided approximations for the mean and variance of estimated TP weights, while Britten-Jones (1999) derived a statistical test for these weights. Okhrin and Schmid (2006) furthered this line of inquiry by deriving the asymptotic distribution for portfolio weights. Subsequently, Kan and Zhou (2007) characterized the moments of TP weights assuming normally distributed returns. Bodnar and Okhrin (2011) developed statistical tests for the composite hypothesis of TP weights, while Kotsiuba and Mazur (2016) approximated the density using Taylor expansion. Palczewski and Palczewski (2014) investigated sampling distributions from the perspective of the mean squared error loss function. Bauder et al. (2018) used a Bayesian approach to examine the characteristics of TP weights.[1] Muhinyuza et al. (2020) and Muhinyuza (2020) derived a statistical test for the TP in small and large dimensions to determine the efficiency of a portfolio. Karlsson et al. (2021) delivered the high-dimensional asymptotic distribution of the estimated TP weights and developed an asymptotic test for linear combinations of the elements of the TP. Analytical expressions for the higher-order moments of the estimated TP weights are obtained by Javed et al. (2021), while Bodnar et al. (2022a) investigated the distribution of weights for a wide range of portfolios, including the normalized TP weights, in both small and large dimensions. More recently, Javed et al. (2023) studied the distributional properties of the estimated TP weights in small and large dimensions assuming that the asset returns follow a matrix-variate closed skew-normal distribution.[2]

The above-mentioned papers focus on the case when the number of assets, n, is greater than the portfolio size, k, and the population covariance matrix is positive definite. In this setting, the sample covariance matrix is non-singular. However, one can face cases when the population and/or sample covariance matrices are singular. The case of a singular population covariance matrix may arise due to the multicollinearity and correlations of asset returns. Another source of singularity can arise in situations where the sample size is smaller than the portfolio size, i.e., $n < k + 1$. This case leads to a singular sample covariance matrix. These sources of singularity have recently garnered considerable attention in academic literature, prompting the development of diverse methodologies.

[1]Let us note that the posterior distribution of the TP weights is proportional to the product of a (singular) Wishart matrix and (singular) Gaussian vector. The statistical properties of these products in various scenarios have also been investigated by Bodnar et al. (2013, 2014, 2019a) and Yonenaga and Suzukawa (2023).

[2]Bodnar et al. (2019b) delivered the asymptotic distributions for functionals of the sample covariance matrix and mean vector in the high-dimensional regime under the assumption that the matrix of observations has a matrix-variate location mixture of normal distributions. These functionals have numerous important applications, especially for the TP weights.

Pappas et al. (2010), Gulliksson and Mazur (2020) and Gulliksson et al. (2023) proposed the mathematical solutions to the mean-variance portfolio problem with a singular population covariance matrix. Bodnar et al. (2016, 2017, 2022b) provided statistical analysis of the mean-variance portfolio weights as well as portfolio compositions under both singular population and sample covariance matrices. For the TP weights, Bodnar et al. (2019c) delivered statistical inference in small and large dimensions by considering scenarios when both the population and sample covariance matrices are singular. Lastly, Alfelt and Mazur (2022) investigated the mean and variance of the TP weights in the case of positive definite population covariance matrix and singular sample covariance matrix.

The present paper assumes that the asset returns are independently and identically distributed and follow multivariate Gaussian distribution with an underlying singular covariance matrix. Furthermore, it also assumes that the number of observations is less than the number of assets, i.e. $n < k + 1$. Under these settings, we contribute to the existing literature in the following way. First, we deliver the extension of the test on the existence of the TP on the set of feasible portfolios and provide its distribution under both null and alternative hypotheses. Second, we give a simple and accurate approximation of the obtained results in the high-dimensional setting.

The rest of the paper is organized as follows. In Section 2, we establish the test statistic and its exact distribution under both null and alternative hypotheses. Section 3 focuses on the asymptotic distribution of the test statistic in the high-dimensional asymptotic regime. Section 4 provides the results of the numerical study. Finally, Section 5 gives concluding remarks.

2 Exact test

Let $\mathbf{x}_t = (x_{1t}, \dots, x_{kt})'$ be a k-dimensional vector of returns of the risky assets at time point $t = 1, \dots, n$. Throughout the paper, it is assumed that $\mathbf{x}_1, \dots, \mathbf{x}_n$ are independently and identically normally distributed with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Additionally, it is assumed that $\boldsymbol{\Sigma}$ is singular with $rank(\boldsymbol{\Sigma}) = r_n < n < k+1$. Furthermore, let $\mathbf{w} = (w_1, \dots, w_k)'$ be a k-dimensional vector of portfolio weights, where w_i is the portion of the wealth allocated to the i-th asset. The expected return and variance of the portfolio are denoted by $R = \mathbf{w}'\boldsymbol{\mu}$ and $V = \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$, respectively.

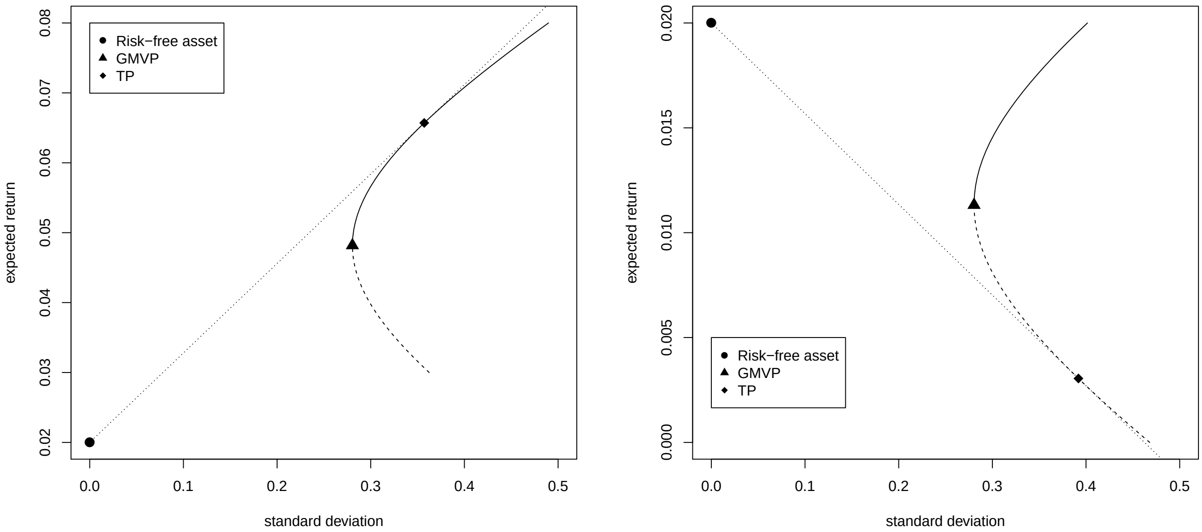
The optimal portfolios as proposed by Markowitz's theory lie on the upper part of the parabola in the mean-variance space. This parabola is known as the efficient frontier (EF) and, if $\boldsymbol{\Sigma}$ is positive definite, is given by

$$(R - R_{GMV})^2 = s(V - V_{GMV}) \tag{1}$$

where

$$R_{GMV} = \frac{\mathbf{1}_k'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}_k'\boldsymbol{\Sigma}^{-1}\mathbf{1}_k} \qquad \text{and} \qquad V_{GMV} = \frac{1}{\mathbf{1}_k'\boldsymbol{\Sigma}^{-1}\mathbf{1}_k} \tag{2}$$

are the expected return and variance of the portfolio with the smallest variance among the efficient portfolios which is called the global minimum variance portfolio (GMVP). Here, the symbol $\mathbf{1}_k$ stands for the k-dimensional vector of ones. The parameter



(a) $R_{GMV} \geq r_f$ (b) $R_{GMV} < r_f$

Figure 1: Location of the tangency portfolio on the set of feasible portfolios in the two cases: Figure 1(a) $R_{GMV} \geq r_f$ and Figure 1(b) $R_{GMV} < r_f$.

$$s = \boldsymbol{\mu}'\mathbf{R}\boldsymbol{\mu} \qquad \text{with} \qquad \mathbf{R} = \boldsymbol{\Sigma}^{-1} - \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}_k\mathbf{1}_k'\boldsymbol{\Sigma}^{-1}}{\mathbf{1}_k'\boldsymbol{\Sigma}^{-1}\mathbf{1}_k} \tag{3}$$

stands for the slope coefficient of the parabola.

On the other hand, if $\boldsymbol{\Sigma}$ is singular, the EF is constructed by replacing the inverse with the Moore-Penrose inverse. Then the EF parameters become

$$R_{GMV} = \frac{\mathbf{1}_k'\boldsymbol{\Sigma}^{+}\boldsymbol{\mu}}{\mathbf{1}_k'\boldsymbol{\Sigma}^{+}\mathbf{1}_k}, \qquad V_{GMV} = \frac{1}{\mathbf{1}_k'\boldsymbol{\Sigma}^{+}\mathbf{1}_k}, \qquad s = \boldsymbol{\mu}'\mathbf{R}\boldsymbol{\mu} \tag{4}$$

with $\mathbf{R} = \boldsymbol{\Sigma}^{+} - \frac{\boldsymbol{\Sigma}^{+}\mathbf{1}_k\mathbf{1}_k'\boldsymbol{\Sigma}^{+}}{\mathbf{1}_k'\boldsymbol{\Sigma}^{+}\mathbf{1}_k}$. Let us note that a number of papers have applied the Moore-Penrose inverse in the portfolio theory, see for example, Pappas et al. (2010), Bodnar et al. (2016, 2019c, 2022b). We notice that the relations in (4) can only be used under the condition that $\mathbf{1}_k'\boldsymbol{\Sigma}^{+}\mathbf{1}_k \neq 0$, which is assumed throughout the paper.

If there is a possibility to invest in a risk-free asset, one may choose to put a portion of his/her investment into a risk-free asset, henceforth, the efficient frontier becomes a straight line in the mean-variance space passing through the return of the risk-free asset and tangent to the parabola in (1). This tangent point is also known as the tangency portfolio (TP), see for example, Ingersoll (1987). The optimality/efficiency of the TP depends crucially on the relation between the return of the GMVP, R_{GMV}, and the return of the risk-free asset, r_f, as can be seen in Figure 1. The mean-variance efficiency of TP is then observed when the GMVP return is greater than the return of the risk-free asset, i.e. $R_{GMV} > r_f$. This can be formulated as a statistical test with the hypotheses expressed as

$$H_0 : R_{GMV} \leq r_f \text{ against } H_1 : R_{GMV} > r_f. \tag{5}$$

The rejection of the null hypothesis suggests that the TP lies on the upper part of the efficient frontier as shown in Figure 1(a). On the other hand, if the null hypothesis in (5) cannot be rejected as in Figure 1(b), then the investor cannot be certain of the optimality of the TP, and allocation into the risk-free asset could be considered as a suitable alternative.

Muhinyuza et al. (2020) and Muhinyuza (2020) constructed the test statistic for testing the hypotheses in (5) and derived its distribution for both finite and high-dimensional settings assuming that $n > k+1$ and positive definiteness of the population covariance matrix, $\mathbf{\Sigma}$. We extend those results for testing (5) in case of $n < k+1$ and singular $\mathbf{\Sigma}$ with $rank(\mathbf{\Sigma}) = r_n < n$ by considering the following test statistic

$$T = \sqrt{\frac{n-r_n}{n-1}} \frac{\widehat{R}_{GMV} - r_f}{\sqrt{1+\frac{n}{n-1}\widehat{s}}\sqrt{\frac{\widehat{V}_{GMV}}{n}}}, \tag{6}$$

where

$$\widehat{R}_{GMV} = \frac{\mathbf{1}_k'\mathbf{S}^+\overline{\mathbf{x}}}{\mathbf{1}_k'\mathbf{S}^+\mathbf{1}_k}, \qquad \widehat{V}_{GMV} = \frac{1}{\mathbf{1}_k'\mathbf{S}^+\mathbf{1}_k}, \qquad \widehat{s} = \overline{\mathbf{x}}'\widehat{\mathbf{R}}\overline{\mathbf{x}} \tag{7}$$

with $\widehat{\mathbf{R}} = \mathbf{S}^+ - \frac{\mathbf{S}^+\mathbf{1}_k\mathbf{1}_k'\mathbf{S}^+}{\mathbf{1}_k'\mathbf{S}^+\mathbf{1}_k}$ are the sample estimators of R_{GMV}, V_{GMV} and s, while

$$\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \qquad \text{and} \qquad \mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})'$$

are the sample estimators of $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$, respectively.

The following theorem provides distribution of T under both the null and alternative hypotheses. Note that f sub-indexed by a distribution stands for the density of that distribution.

Theorem 1. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d random vectors with $\mathbf{x}_1 \sim \mathcal{N}_k(\boldsymbol{\mu}, \mathbf{\Sigma}), k > n-1$ and $rank(\mathbf{\Sigma}) = r_n < n$. Then the density of T is given by*

$$f_T(x) = \frac{n(n-r_n+1)}{(r_n-1)(n-1)} \int_0^\infty f_{t_{n-r_n,\delta(y)}}(x) f_{\mathcal{F}_{r_n-1,n-r_n+1,ns}}\left(\frac{n(n-r_n+1)}{(r_n-1)(n-1)}y\right) dy \tag{8}$$

where $\delta(y) = \sqrt{\frac{n}{1+n/(n-1)y}}S_{GMV}$ with $S_{GMV} = \frac{R_{GMV}-r_f}{\sqrt{V_{GMV}}}$ which is the Sharpe ratio of the GMVP.

Proof. The density function of the test statistic T in (6) is obtained by utilizing the distributional properties of essential quantities $\widehat{R}_{GMV}$, $\widehat{V}_{GMV}$ and $\widehat{s}$ as presented in Bodnar et al. (2022b). In particular, we make use of the following properties

(P1) $\widehat{R}_{GMV}|\widehat{s} \sim \mathcal{N}\left(R_{GMV}, (1+\frac{n}{n-1}\widehat{s})\frac{V_{GMV}}{n}\right)$;

(P2) $\frac{n(n-r_n+1)}{(n-1)(r_n-1)}\widehat{s} \sim \mathcal{F}_{r_n-1,n-r_n+1,ns}$;

(P3) $(n-1)\widehat{V}_{GMV}/V_{GMV} \sim \chi^2_{n-r_n}$;

(P4) $\widehat{V}_{GMV}$ is independent of $(\widehat{R}_{GMV}, \widehat{s})$.

Now adding and subtracting R_{GMV} on the numerator and dividing both the numerator and denominator by $\sqrt{V_{GMV}}$ of the test statistic in (6), and rearranging it, we get

$$T = \left(\frac{\widehat{R}_{GMV} - R_{GMV}}{\sqrt{1 + \frac{n}{n-1}\widehat{s}}\sqrt{\frac{V_{GMV}}{n}}} + \frac{R_{GMV} - r_f}{\sqrt{1 + \frac{n}{n-1}\widehat{s}}\sqrt{\frac{V_{GMV}}{n}}} \right) \frac{1}{\sqrt{\frac{n-1}{n-r_n}\frac{\widehat{V}_{GMV}}{V_{GMV}}}}.$$

Applying properties (P1), (P3) and (P4) and using the definition of non-central t-distribution, we obtain that

$$T|\widehat{s} = y \sim t_{n-r_n,\delta(y)} \qquad \text{with} \qquad \delta(y) = \frac{R_{GMV} - r_f}{\sqrt{1 + \frac{n}{n-1}y}\sqrt{\frac{V_{GMV}}{n}}}.$$

Applying property (P2) and computing the unconditional distribution of T, we arrive at the statement of Theorem 1. □

In Theorem 1, we can observe that the density function of the test statistic T is expressed as a one-dimensional integral of the product of two well-known univariate density functions. This formula can be easily computed in many mathematical software such as, for example, R and Mathematica. From the proof of Theorem 1, it can be also seen that the test statistic T may be represented as a mixture of a non-central t-distribution with $n - r_n$ degrees of freedom and a non-centrality parameter $\delta(y)$. Now having the density function of T, we can derive the critical value for the test (5) at significance level α. This result is provided in the next theorem.

Theorem 2. *Under the conditions of the Theorem 1, it holds that*

$$\sup_{V_{GMV}>0, s\geq 0, R_{GMV}\leq r_f} G_{T,\alpha,t_{n-r_n;1-\alpha}}\left(S_{GMV}, s\right) \leq \mathbb{P}_{H_0: R_{GMV}=r_f}\left(T > t_{n-r_n;1-\alpha}\right) = \alpha,$$

where

$$G_{T,\alpha,c}\left(S_{GMV}, s\right) = \mathbb{P}\left(T > c\right) = \int_c^\infty f_T(x)dx$$

and the symbol $t_{n-r_n;1-\alpha}$ stands for the $(1-\alpha)$ quantile of the t-distribution with $n - r_n$ degrees of freedom.

Proof. Using Theorem 1 and for a given constant c, we have that

$$\begin{aligned}
& G_{T,\alpha,c}\left(S_{GMV}, s\right) = \mathbb{P}\left(T > c\right) = \int_c^\infty f_T(x)dx \\
= & \frac{n(n-r_n+1)}{(r_n-1)(n-1)} \int_c^\infty \int_0^\infty f_{t_{n-r_n,\delta(y)}}(x) f_{\mathcal{F}_{r_n-1,n-r_n+1,ns}}\left(\frac{n(n-r_n+1)}{(r_n-1)(n-1)}y\right) dydx \\
= & \frac{n(n-r_n+1)}{(r_n-1)(n-1)} \int_0^\infty \left(\int_c^\infty f_{t_{n-r_n,\delta(y)}}(x)dx\right) f_{\mathcal{F}_{r_n-1,n-r_n+1,ns}}\left(\frac{n(n-r_n+1)}{(r_n-1)(n-1)}y\right) dy \\
= & \frac{n(n-r_n+1)}{(r_n-1)(n-1)} \int_0^\infty \left(1 - F_{t_{n-r_n,\delta(y)}}(c)\right) f_{\mathcal{F}_{r_n-1,n-r_n+1,ns}}\left(\frac{n(n-r_n+1)}{(r_n-1)(n-1)}y\right) dy,
\end{aligned}$$

where $F_{t_{n-r_n,\delta(y)}}(\cdot)$ stands for the cumulative distribution function of the non-central t-distribution with $n - r_n$ degrees of freedom and a non-centrality parameter $\delta(y)$. Since $1 - F_{t_{n-r_n,\delta(y)}}(c) \leq 1 - F_{t_{n-r_n,0}}(c)$ for all $y \geq 0$ and $R_{GMV} < r_f$, we obtain that

$$
\begin{aligned}
& G_{T,\alpha,c}\left(S_{GMV}, s\right) \\
\leq \; & \frac{n(n-r_n+1)}{(r_n-1)(n-1)} \int_0^{\infty} \left(1 - F_{t_{n-r_n,0}}(c)\right) f_{\mathcal{F}_{r_n-1,n-r_n+1,ns}}\left(\frac{n(n-r_n+1)}{(r_n-1)(n-1)} y\right) dy \\
= \; & \left(1 - F_{t_{n-r_n,0}}(c)\right) \underbrace{\frac{n(n-r_n+1)}{(r_n-1)(n-1)} \int_0^{\infty} f_{\mathcal{F}_{r_n-1,n-r_n+1,ns}}\left(\frac{n(n-r_n+1)}{(r_n-1)(n-1)} y\right) dy}_{1} \\
= \; & 1 - F_{t_{n-r_n},0}(c) = \alpha
\end{aligned}
$$

with $c = t_{n-r_n;1-\alpha}$. The proof of the theorem is completed. □

Theorem 2 delivers us the message that the test of (5) rejects H_0 in favour of H_1 as $T \geq t_{n-r_n;1-\alpha}$. We can also see that the power of the test based on the test statistic T is given by

$$
\begin{aligned}
& G_{T,\alpha,t_{n-r_n;1-\alpha}}(S_{GMV}, s) = \mathbb{P}\left(T > t_{n-r_n;1-\alpha}\right) \\
= \; & \frac{n(n-r_n+1)}{(r_n-1)(n-1)} \int_0^{\infty} \left(1 - F_{t_{n-r_n,\delta(y)}}(t_{n-r_n;1-\alpha})\right) f_{\mathcal{F}_{r_n-1,n-r_n+1,ns}}\left(\frac{n(n-r_n+1)}{(r_n-1)(n-1)} y\right) dy.
\end{aligned}
$$

It is noted that the power function depends on $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ through the quantities S_{GMV} and s. This fact simplifies considerably the study of the power of the test. In Figure 2, we present the power of the test as a function of S_{GMV} with fixed $s \in \{1, 5, 10\}$. We also set $n \in \{50, 250\}$, $r_n = 0.5n$ and $\alpha = 5\%$. We can observe that the power of the test increases as s decreases and that the suggested test rejects the null hypothesis for small values of S_{GMV}.

Since a statistical test and interval estimation are related, we can construct a $(1 - \alpha)$ one-sided confidence interval for the risk-free rate r_f. Namely, if r_f belongs to this interval, a conclusion about the investment into the TP can be made. For the upper one-sided test, this interval is expressed as

$$
I_{1-\alpha} = \left[\widehat{R}_{GMV} - t_{n-r_n;1-\alpha}\sqrt{\frac{n-1}{n-r_n}}\sqrt{1 + \frac{n}{n-1}\widehat{s}}\sqrt{\frac{\widehat{V}_{GMV}}{n}}, +\infty\right)
$$

while for the lower one-sided test, we have that

$$
\breve{I}_{1-\alpha} = \left(-\infty, \widehat{R}_{GMV} - t_{n-r_n;\alpha}\sqrt{\frac{n-1}{n-r_n}}\sqrt{1 + \frac{n}{n-1}\widehat{s}}\sqrt{\frac{\widehat{V}_{GMV}}{n}}\right].
$$

Therefore, it leads us to the conclusion that for all $r_f \notin I_{1-\alpha}$ the TP lies on the EF, while for $r_f \notin \breve{I}_{1-\alpha}$ the TP lies on the lower part of the set of feasible portfolios.

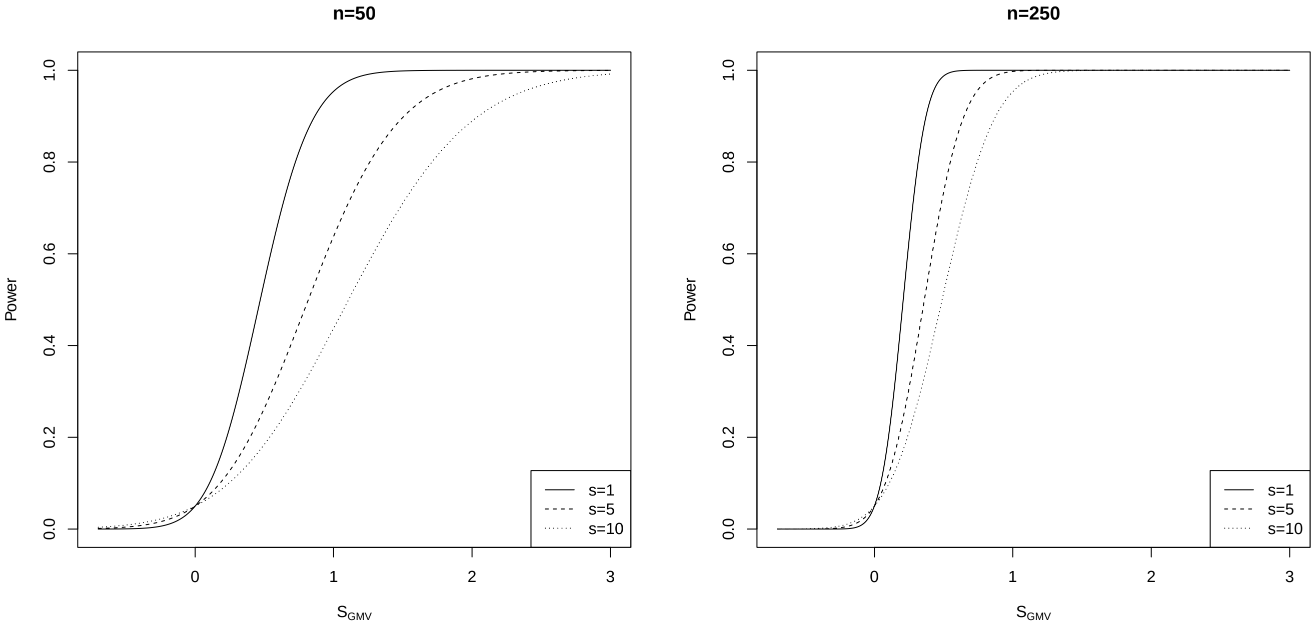


Figure 2: Power of the test statistic T as a function of S_{GMV} for $s \in \{1,5,10\}, n \in \{50,250\}, r_n = 0.5n$ and $\alpha = 5\%$.

3 High-dimensional asymptotics

In this section, we derive the high-dimensional asymptotic distribution of test statistic given in (6) under both the null and alternative hypothesis. We treat the rank r_n of the population covariance matrix $\mathbf{\Sigma}$ as the actual dimension of the data-generating process. Furthermore, we assume that $r_n/n \to c \in (0,1)$ as $n \to \infty$. Let us note that we don't assume a relationship between the portfolio dimension k and the sample size n except for $k > n$. It means that k can grow to infinity much faster than n, then one can consider, for example, exponential growth which is of great importance in economics.

In the following theorem, we derive the high-dimensional asymptotic distribution of the test statistic T given in (6).

Theorem 3. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d random vectors with $\mathbf{x}_1 \sim \mathcal{N}_k(\boldsymbol{\mu}, \mathbf{\Sigma}), k > n-1$ and $rank(\mathbf{\Sigma}) = r_n < n$. Let also $c_n := r_n/n \to c \in (0,1)$ as $n \to \infty$. Then*

(a) the asymptotic distribution of T is given by

$$\sigma_T^{-1}\left(T - \sqrt{n}\frac{S_{GMV}}{\sqrt{1 + \frac{r_n-1}{n-r_n+1}\left(1 + \frac{n}{r_n-1}s\right)}}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

where

$$\sigma_T^2 = 1 + \frac{S_{GMV}^2}{2(1+s)}\left(1 + \frac{s^2+2s+c}{(1+s)^2}\right).$$

(b) under the null hypothesis it holds that $T \sim \mathcal{N}(0,1)$.

Proof. From the proof of Theorem 1, we have that

$$T|\widehat{s}=y \sim t_{n-r_n,\delta(y)}$$

with $\delta(y)=\sqrt{\frac{n}{1+n/(n-1)y}}S_{GMV}$. Additionally, it holds that $u=\frac{n(n-r_n+1)}{(n-1)(r_n-1)}\widehat{s}\sim \mathcal{F}_{r_n-1,n-r_n+1,ns}$. Consequently, the stochastic representation of T is given by

$$T \stackrel{d}{=} \sqrt{\frac{n-r_n}{\xi}}\left(z_0+\frac{\sqrt{n}S_{GMV}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}u}}\right)$$

where $z_0\sim\mathcal{N}(0,1)$, $\xi\sim\chi^2_{n-r_n}$ and $u\sim\mathcal{F}_{r_n-1,n-r_n+1,ns}$; moreover, z_0,ξ and u are mutually independently distributed. Now it holds that

$$\begin{aligned}
&T-\sqrt{n}\frac{S_{GMV}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)}}\\
=&\sqrt{\frac{n-r_n}{\xi}}\left(z_0+\frac{\sqrt{n}S_{GMV}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}u}}\right)-\sqrt{n}\frac{S_{GMV}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)}}\\
=&\sqrt{\frac{n-r_n}{\xi}}z_0+\frac{S_{GMV}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}u}}\left[\sqrt{n}\left(\sqrt{\frac{n-r_n}{\xi}}-1\right)+\sqrt{n}\left(1-\frac{\sqrt{1+\frac{r_n-1}{n-r_n+1}u}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)}}\right)\right],
\end{aligned}$$

where the last expression is obtained by adding and subtracting $\sqrt{n}\frac{S_{GMV}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}u}}$, factoring out $\frac{S_{GMV}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}u}}$, and rearranging. Let us note that

$$\begin{aligned}
&1-\frac{\sqrt{1+\frac{r_n-1}{n-r_n+1}u}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)}}\\
=&\frac{1}{\sqrt{1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)}}\frac{\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s-u\right)}{\sqrt{1+\frac{r_n-1}{n-r_n+1}u}+\sqrt{1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)}}
\end{aligned}$$

From the proof of Theorem 5 in Bodnar et al. (2019c) and the proof of Theorem 4 in Bodnar et al. (2022b), we have that

$$\begin{aligned}
\frac{\xi}{n-r_n}-1 &\stackrel{a.s.}{\to} 0,\\
\sqrt{n}\left(\frac{\xi}{n-r_n}-1\right) &\stackrel{\mathcal{D}}{\to} \mathcal{N}\left(0,\frac{2}{1-c}\right),
\end{aligned}$$

and

$$\begin{aligned}
u-1-\frac{n}{r_n-1}s &\stackrel{a.s.}{\to} 0,\\
\sqrt{n}\left(u-1-\frac{n}{r_n-1}s\right) &\stackrel{\mathcal{D}}{\to} \mathcal{N}\left(0,\sigma_u^2\right)
\end{aligned}$$

with $\sigma_u^2 = \frac{2}{c}\left(1+2\frac{s}{c}\right)+\frac{2}{1-c}\left(1+\frac{s}{c}\right)^2$, for $r_n/n \to c \in (0,1)$ as $n \to \infty$. It is also well-known that

$$\sqrt{n}\left(\frac{z_0}{\sqrt{n}}\right) \overset{\mathcal{D}}{\to} \mathcal{N}(0,1).$$

Finally, putting all the above together and applying Slutsky's lemma (see, e.g., Theorem 2.8 in Van der Vaart (2000)), we arrive at the first part of the theorem. By setting $S_{GMV}=0$, we get the second part of the theorem under the null hypothesis. The theorem is proved.

□

Having the high-dimensional asymptotic distribution of test statistic in Theorem 3, we can easily compute the power function of that test. Therefore, it is given by

$$G_{T,\alpha,z_{1-\alpha}}(S_{GMV},s) = 1-\Phi\left(\frac{z_{1-\alpha}-\sqrt{n}\frac{S_{GMV}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)}}}{\sigma_T}\right),$$

where $z_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the standard normal distribution and $\Phi(\cdot)$ stands for the distribution function of the standard normal distribution.

4 Simulation study

In this section, we compare the power functions of the exact test and the high-dimensional asymptotic test which are delivered in Theorem 1 and 3, respectively. Let us recall that both expressions depend on the slope parameter of the efficient frontier, s, and the Sharpe ratio of the GMVP, S_{GMV}. In what follows, we set s to be equal to 1, i.e. $s=1$. The significance level is taken to be $\alpha = 5\%$. We consider several values for the sample size such as $n \in \{50, 120, 250, 500\}$ which approximately corresponds to the length of one year, two years, five years, and ten years of weekly financial data.

In Figures 3-6, we present the results of the simulation study for $c \in \{0.1, 0.4, 0.7, 0.9\}$. The dashed black line represents the power function of the exact test, while the power function of the high-dimensional test is indicated by a solid black line. The power of the asymptotic test is almost indistinguishable from the exact one. It is remarkable that the high-dimensional asymptotic test is properly sized for all values of n and the differences between the two tests are observable only for the case of $n=50$ and $c=0.9$. We also note that for $c=0.9$, the exact power coincides with the asymptotic power as the sample size increases.

5 Conclusions

The role of the TP has become indescribable for both researcher and practitioner actors in finance. Hence, having complete comprehension of the TP properties under possible conditions is vital for any financial strategist. In this paper, we deal with the test on the mean-variance efficiency of the TP when both the population and sample covariance

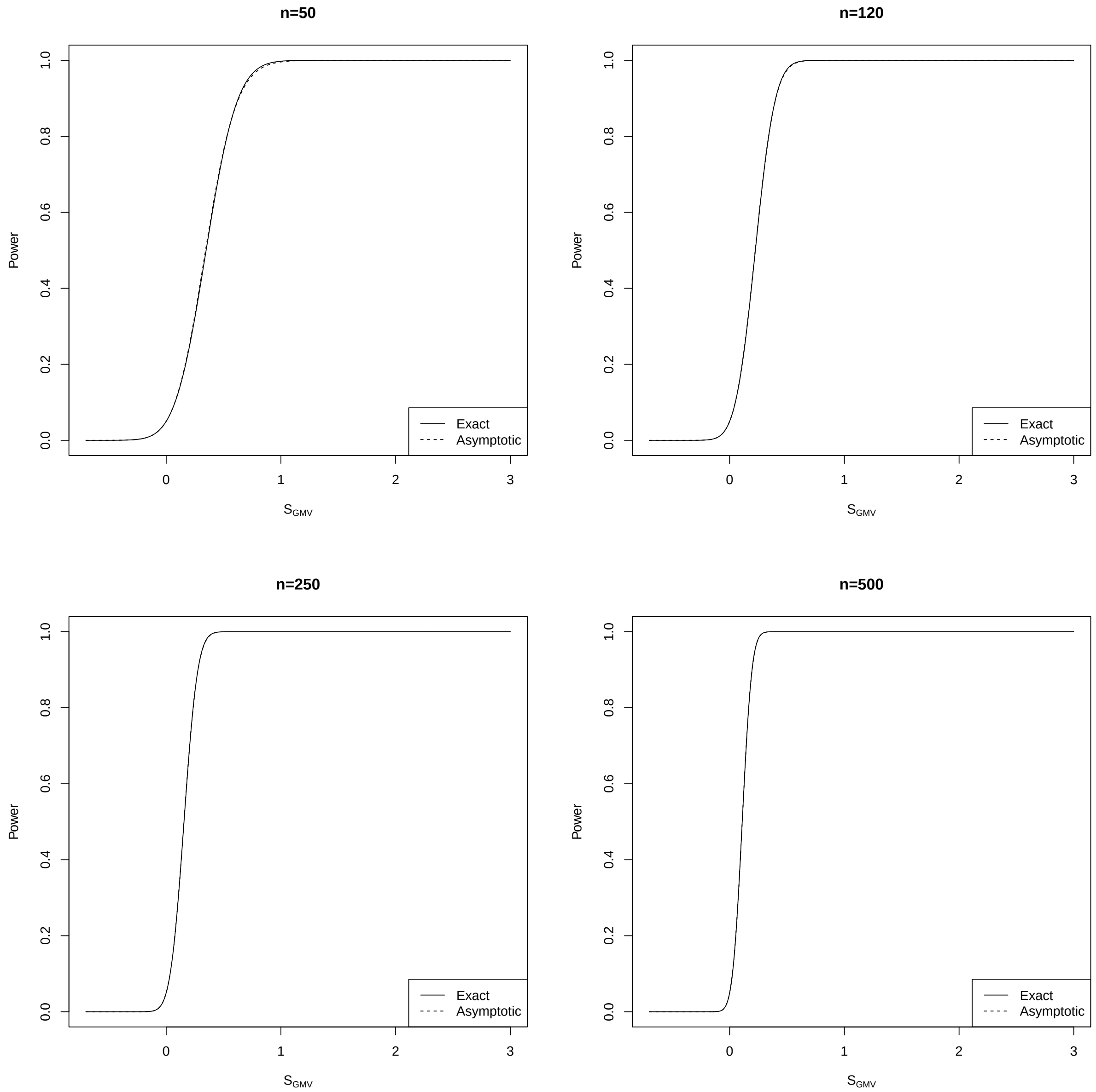


Figure 3: Powers of the exact test and the high-dimensional asymptotic test as a function of S_{GMV} based on statistic T for $c = 0.1$ with $s = 1$ and $\alpha = 5\%$.

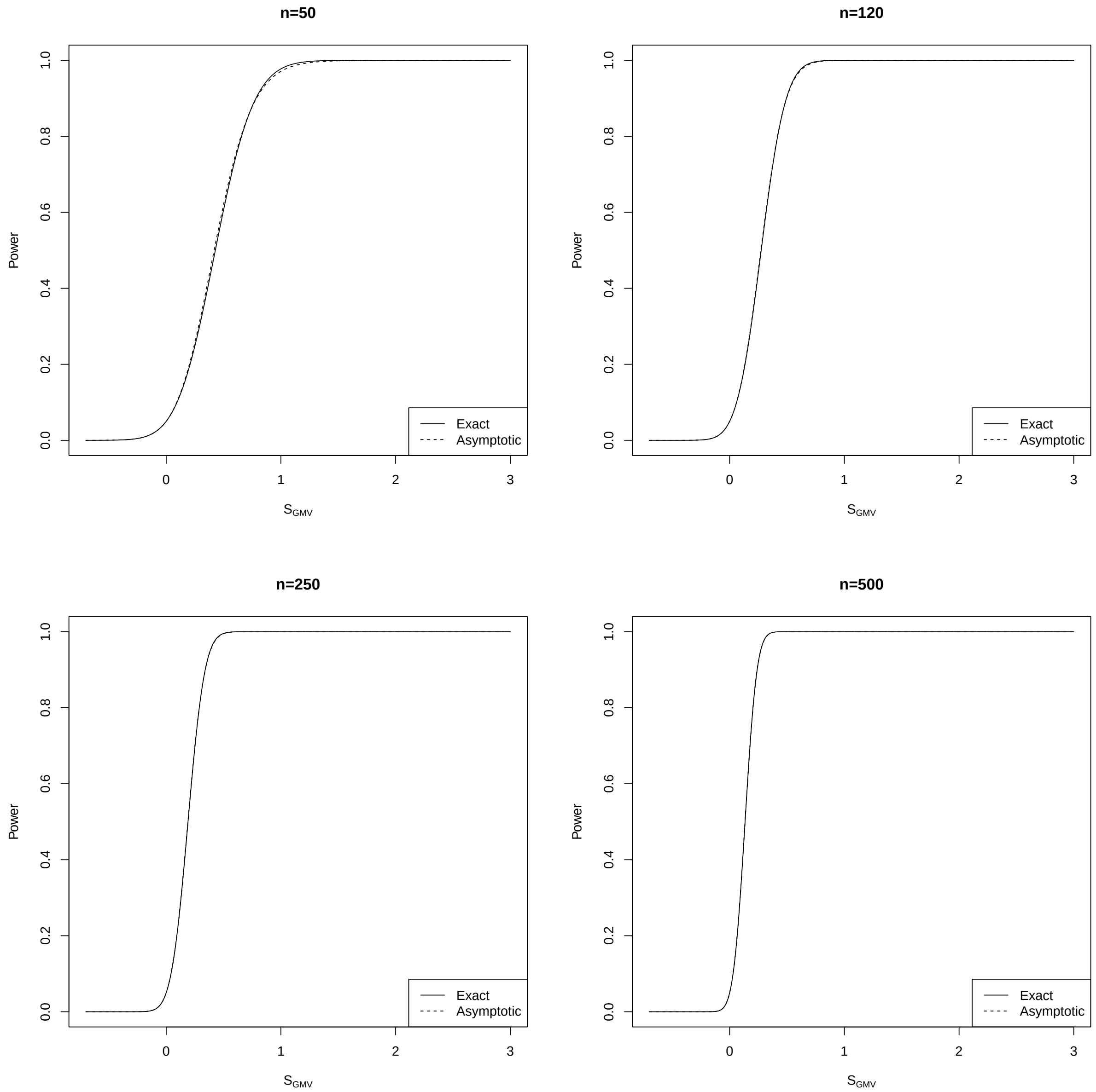


Figure 4: Powers of the exact test and the high-dimensional asymptotic test as a function of S_{GMV} based on statistic T for $c = 0.4$ with $s = 1$ and $\alpha = 5\%$.

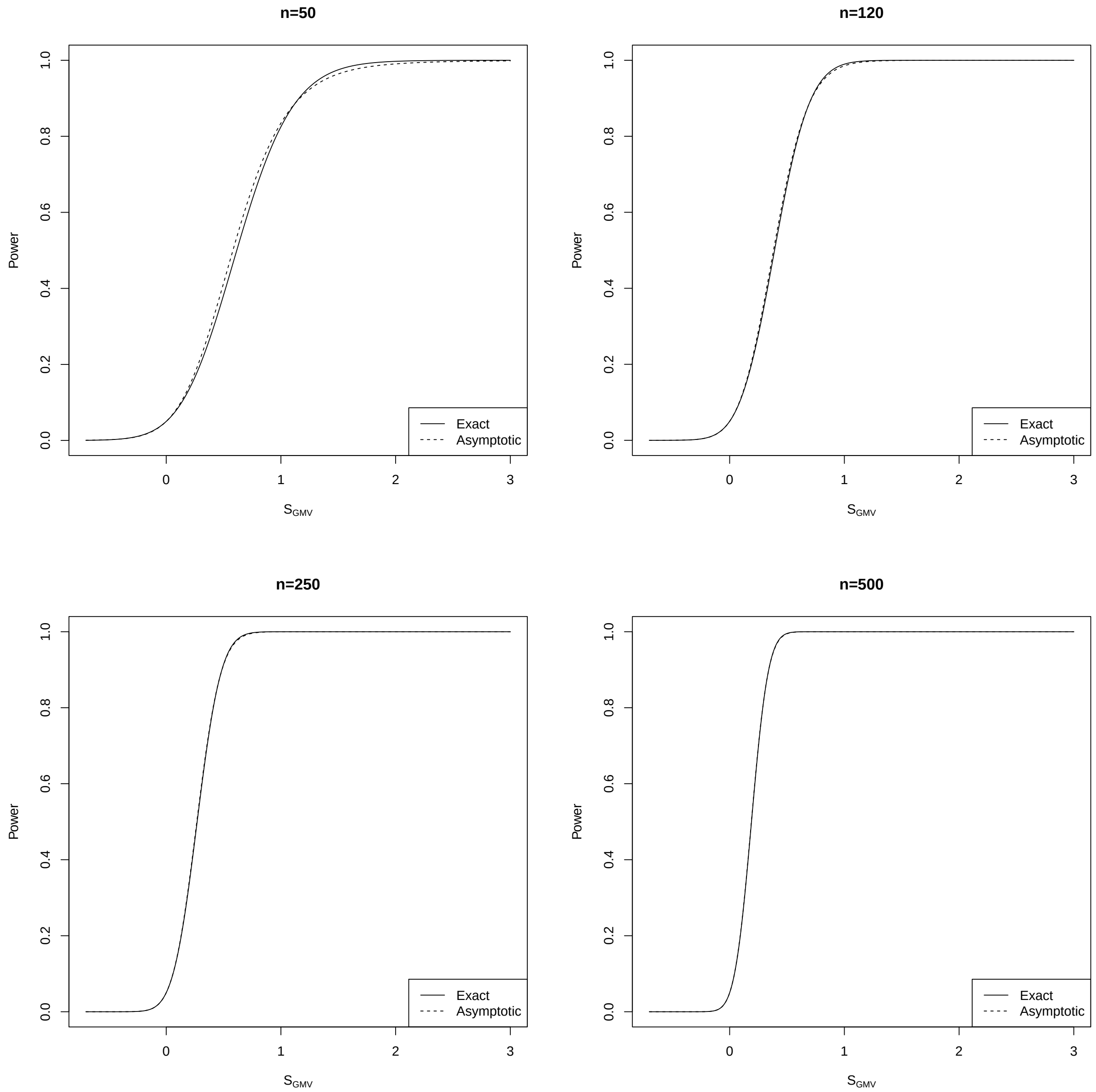


Figure 5: Powers of the exact test and the high-dimensional asymptotic test as a function of S_{GMV} based on statistic T for $c = 0.7$ with $s = 1$ and $\alpha = 5\%$.

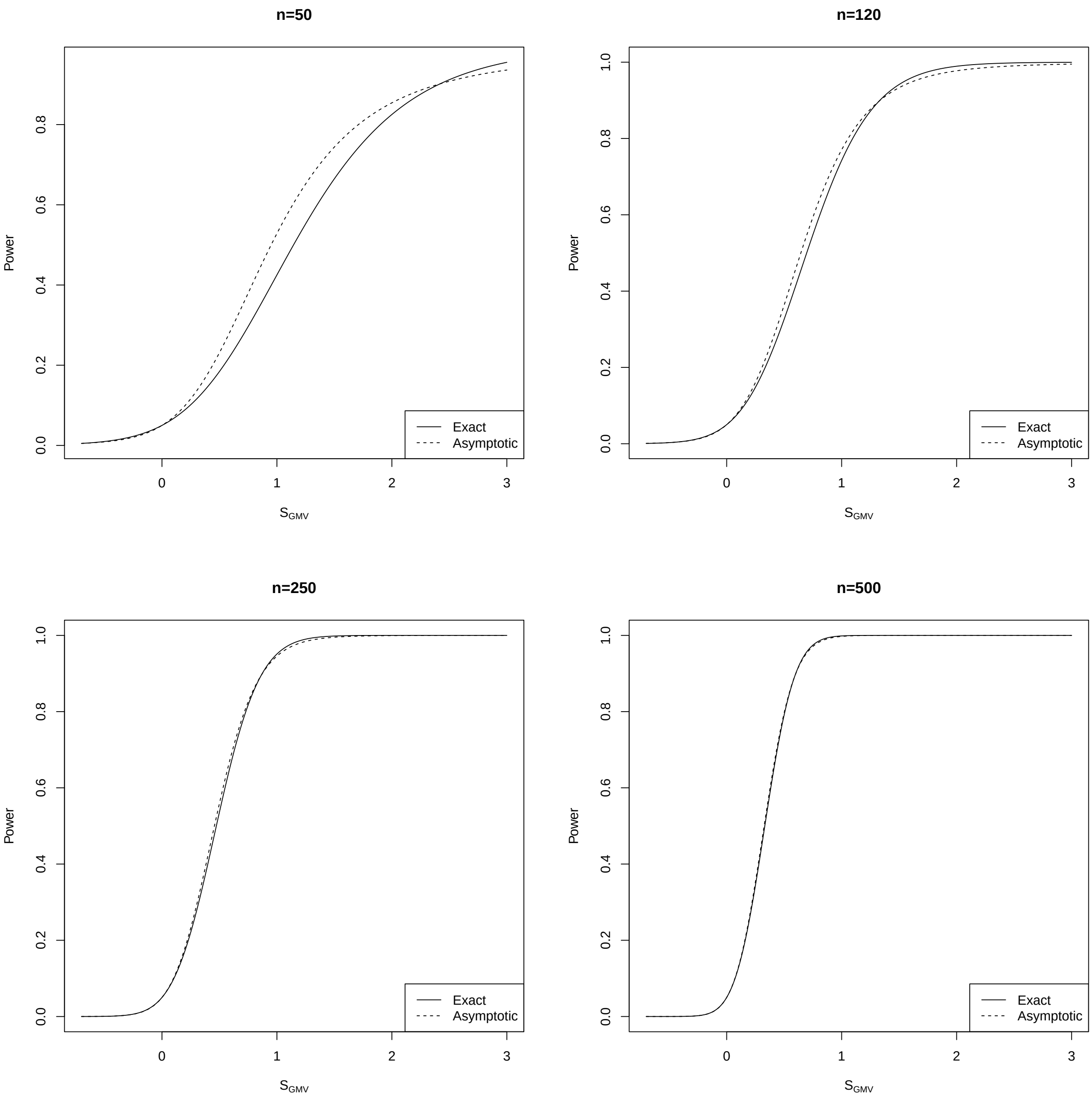


Figure 6: Powers of the exact test and of the high-dimensional asymptotic test as a function of S_{GMV} based on statistic T for $c = 0.9$ with $s = 1$ and $\alpha = 5\%$.

matrices are singular. Under these conditions, we deliver the finite sample test statistic and its distribution under both the null and alternative hypotheses. We also derive the high-dimensional asymptotic distribution of the considered test statistic under the null hypothesis as well as for the alternative hypothesis. Through the simulation studies, we observe a good performance of the derived theoretical results, that is, the high-dimensional asymptotic test is properly sized for all values of n and the differences between two tests are observable only for the case of $n = 50$ and $c = 0.9$.

Acknowledgement

Svitlana Drin acknowledges financial support from the Knowledge Foundation Grant (Dnr: 20220115). Stepan Mazur acknowledges financial support from the internal research grants at Örebro University.

References

Alfelt, G. and Mazur, S. (2022). On the mean and variance of the estimated tangency portfolio weights for small samples. *Modern Stochastics: Theory and Applications*, 9(4):453–482.

Bai, Z., Liu, H., and Wong, W.-K. (2009). Enhancement of the applicability of markowitz's portfolio optimization by utilizing random matrix theory. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 19(4):639–667.

Bauder, D., Bodnar, T., Mazur, S., and Okhrin, Y. (2018). Bayesian inference for the tangent portfolio. *International Journal of Theoretical and Applied Finance*, 21(08):1850054.

Bodnar, T., Dette, H., Parolya, N., and Thorsén, E. (2022a). Sampling distributions of optimal portfolio weights and characteristics in small and large dimensions. *Random Matrices: Theory and Applications*, 11(01):2250008.

Bodnar, T., Mazur, S., Muhinyuza, S., and Parolya, N. (2019a). On the product of a singular wishart matrix and a singular gaussian vector in high dimension. *Theory of Probability and Mathematical Statistics*, 99:39–52.

Bodnar, T., Mazur, S., and Nguyen, H. (2022b). Estimation of optimal portfolio compositions for small sample and singular covariance matrix. Working paper **15**, School of Business, Orebro University, Sweden.

Bodnar, T., Mazur, S., and Okhrin, Y. (2013). On the exact and approximate distributions of the product of a wishart matrix with a normal vector. *Journal of Multivariate Analysis*, 122:70–81.

Bodnar, T., Mazur, S., and Okhrin, Y. (2014). Distribution of the product of singular wishart matrix and normal vector. *Theory of Probability and Mathematical Statistics*, 91:1–15.

Bodnar, T., Mazur, S., and Parolya, N. (2019b). Central limit theorems for functionals of large sample covariance matrix and mean vector in matrix-variate location mixture of normal distributions. *Scandinavian Journal of Statistics*, 46(2):636–660.

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

Bodnar, T., Mazur, S., and Podgórski, K. (2017). A test for the global minimum variance portfolio for small sample and singular covariance. *AStA Advances in Statistical Analysis*, 101(3):253–265.

Bodnar, T., Mazur, S., Podgórski, K., and Tyrcha, J. (2019c). Tangency portfolio weights for singular covariance matrix in small and large dimensions: estimation and test theory. *Journal of Statistical Planning and Inference*, 201:40–57.

Bodnar, T. and Okhrin, Y. (2011). On the product of inverse wishart and normal distributions with applications to discriminant analysis and portfolio theory. *Scandinavian Journal of Statistics*, 38(2):311–331.

Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *The Journal of Finance*, 54(2):655–671.

Gulliksson, M. and Mazur, S. (2020). An iterative approach to ill-conditioned optimal portfolio selection. *Computational Economics*, 56(4):773–794.

Gulliksson, M., Oleynik, A., and Mazur, S. (2023). Portfolio selection with a rank-deficient covariance matrix. *Computational Economics*, 1.

Ingersoll, J. E. (1987). *Theory of financial decision making*, volume 3. Rowman & Littlefield.

Javed, F., Mazur, S., and Ngailo, E. (2021). Higher order moments of the estimated tangency portfolio weights. *Journal of Applied Statistics*, 48(3):517–535.

Javed, F., Mazur, S., and Thorsén, E. (2023). Tangency portfolio weights under a skew-normal model in small and large dimensions. *Journal of the Operational Research Society*, 0(0):1–12.

Jobson, J. D. and Korkie, B. (1980). Estimation for markowitz efficient portfolios. *Journal of the American Statistical Association*, 75(371):544–554.

Jorion, P. (1986). Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative analysis*, 21(3):279–292.

Kan, R. and Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 42(3):621–656.

Karlsson, S., Mazur, S., and Muhinyuza, S. (2021). Statistical inference for the tangency portfolio in high dimension. *Statistics*, 55(3):532–560.

Klein, R. W. and Bawa, V. S. (1976). The effect of estimation risk on optimal portfolio choice. *Journal of financial economics*, 3(3):215–231.

Kotsiuba, I. and Mazur, S. (2016). On the asymptotic and approximate distributions of the product of an inverse wishart matrix and a gaussian vector. *Theory of Probability and Mathematical Statistics*, 93:103–112.

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

Li, H., Bai, Z., Wong, W.-K., and McAleer, M. (2022). Spectrally-corrected estimation for high-dimensional markowitz mean-variance optimization. *Econometrics and Statistics*, 24:133–150.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Muhinyuza, S. (2020). A test on mean-variance efficiency of the tangency portfolio in high-dimensional setting. *Theory of Probability and Mathematical Statistics*, 103:103–119.

Muhinyuza, S., Bodnar, T., and Lindholm, M. (2020). A test on the location of the tangency portfolio on the set of feasible portfolios. *Applied Mathematics and Computation*, 386:125519.

Okhrin, Y. and Schmid, W. (2006). Distributional properties of portfolio weights. *Journal of Econometrics*, 134(1):235–256.

Palczewski, A. and Palczewski, J. (2014). Theoretical and empirical estimates of mean–variance portfolio sensitivity. *European Journal of Operational Research*, 234(2):402–410.

Pappas, D., Kiriakopoulos, K., and Kaimakamis, G. (2010). Optimal portfolio selection with singular covariance matrix. In *International Mathematical Forum*, volume 5, pages 2305–2318.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, volume 3. Cambridge university press.

Winkler, R. L. (1973). Bayesian models for forecasting future security prices. *The Journal of Financial and Quantitative Analysis*, 8(3):387–405.

Yonenaga, K. and Suzukawa, A. (2023). Distribution of the product of a wishart matrix and a normal vector. *Theory of Probability and Mathematical Statistics*, 108:209–224.