Stockholms universitet

# Statistical Aspects of Sustainability in Optimal Portfolio Theory

Vivi Wong

Masteruppsats i matematisk statistik
Master Thesis in Mathematical Statistics

Masteruppsats 2020:4
Matematisk statistik
Juni 2020

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Stockholm University

Mathematical Statistics
Stockholm University
Master Thesis **2020:4**
http://www.math.su.se

# Statistical Aspects of Sustainability in Optimal Portfolio Theory

Vivi Wong*

June 2020

**Abstract**

The world strives to satisfy sustainability conditions, such as human rights and environmental protection, among others, nowadays. Hence, investors would find it interesting to know whether investing in ethical portfolios will worsen the investment opportunities, compared to investing in portfolios constructed by both sustainable and unsustainable assets. One can choose to analyze this topic based on classical optimal portfolio analysis, such as studying whether the difference between the minimum-variance frontier of sustainable assets and the minimum-variance frontier of those assets and some additional unsustainable assets is statistically significant by applying mean-variance spanning tests. In this thesis, we used monthly and weekly returns, respectively, of 21 stocks in the OMX Stockholm 30 (OMXS30) index over the period 2008-2019, and we performed four screenings to obtain different numbers of stocks that are considered more sustainable. Mean-variance spanning tests were then applied to study whether the differences between the minimum-variance frontier of all the 21 stocks and each of the minimum-variance frontiers of the assets obtained after screening are statistically significant or not. The results of the spanning tests showed statistically nonsignificant differences between the minimum-variance frontiers. Hence, our study suggests that an investor would not obtain better investment opportunities when she, in addition to the considered more sustainable assets in the OMXS30 index, adds the stocks that are considered more unsustainable into the investment portfolio.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: viviwong.96@gmail.com. Supervisor: Taras Bodnar.

# Preface

This is a master's degree thesis worth 30 credits (ECTS) at Stockholm University.

I am extremely thankful to my supervisor Taras Bodnar at Stockholm University, for his patient guidance, advice, and for introducing this project to me.

# Abbreviations

| | |
|---|---|
| GMM | Generalized methods of moments |
| MVP | Minimum-variance portfolio |
| TP | Tangency portfolio |
| OMXS30 | OMX Stockholm 30 |

# Contents

# List of Tables

## List of Figures

# 1 Introduction

Over the past few years, the demand for sustainability has been increasing in various fields in the world, and one of them is the financial field. Many investors are therefore using socially responsible criteria to make investment decisions. One usually refers to three main areas when social responsibility is mentioned, and these are the following: environmental (E), social (S), and governance (G). Hence, ESG is used for short when talking about socially responsible investments [15, 33]. The environmental factor focuses on environmental protection, while the social part focuses on, e.g., human rights, and the governance factor concentrates on corporate governance [25]. In order to know if an investment might be considered sustainable, one can analyze a company's risk of encountering ESG issues, which can be measured by different kinds of ESG scores. One can, for example, receive ESG scores from Sustainalytics and MSCI, among others, which are leading ESG rating providers [15, 30].

Except for making a difference to the society and environment when one invests according to ESG criteria, investors might also wonder whether one will gain or lose money from it. Hence, it might, therefore, be interesting to study if the returns are statistically significantly different for socially responsible funds and conventional funds, or if the efficient frontier of assets with low ESG scores is statistically significantly different from the efficient frontier of assets with both low and high ESG scores [15].

Previous researches on this subject have mostly compared socially responsible indexes with conventional indexes. Bauer, Koedijk, and Otter [3] studied ethical and conventional funds from the German, UK and US markets over the period 1990-2001, and they did not find support that the returns of the ethical and conventional funds are statistically significantly different after controlling for some common factors. Another paper by Ortas, Burritt, and Moneva [26] studied the Dow Jones Sustainability Asia Pacific Index (DJSI-AP) and the Dow Jones Global Total Stock Market Index (DJ-G), and the result showed that the risk-adjusted returns of the DJSI-AP are not statistically different from the risk-adjusted returns of the DJ-G. A study, based on a sample of socially responsible mutual funds and conventional funds during the period 2000-2001, by Nofsinger and Varma [24] showed that socially responsible mutual funds outperform conventional funds in crisis periods, but the results show the opposite during non-crisis periods. Henke [14] studied several funds in the US and Eurozone, respectively, and saw that socially responsible funds outperform conventional funds during the period 2001-2014. Furthermore, Herzel, Nicolosi, and Stărică [15] used the mean-variance spanning test based on the Wald test to study whether the minimum-variance frontier of socially responsible assets and the minimum-variance frontier of those assets and some additional conventional assets are statistically significantly different. The screening procedure was performed in their study to obtain socially responsible assets. They saw that the efficient frontiers of the screened assets and that of the whole investment universe are statistically different when short selling is not allowed and when the screening procedure was based on ESG scores that focus on the environmental factor.

In this thesis, the aim is to understand the asymptotic and exact mean-variance spanning tests based on the likelihood ratio test introduced by Huberman and Kandel [16], and the spanning tests based on the Wald test and Lagrange multiplier test, respectively, purposed by Kan and Zhou [17], and apply them to real data. The data we use consists of the returns from 21 stocks that are included in the OMXS30 index over the period 2008-2019. We will study whether the differences between the efficient frontiers of considered more sustainable assets and the efficient frontier of all 21 assets are statistically significant.

# 2 Optimal portfolio theory and sustainability

In this section, we mainly follow Capinski and Zastawniak [8] to introduce portfolio theory and several types of optimal portfolios used in the development of theoretical findings of the thesis.

## 2.1 Optimal portfolio theory

An investor can choose to invest in one or more securities, either in risky assets (e.g., stocks) or risk-free assets such as bank deposits or bonds, among others, or a combination of both. A risky asset might have a higher expected return, but there are chances of loss when one invests in such assets. However, one might want to take some risk at a reasonable level to have a chance of getting higher expected returns by investing in a risky asset. Furthermore, one has probably heard of the saying "do not put all eggs in one basket" before, and hence investors usually choose to invest in more than one asset. Nevertheless, investing in different assets does not imply that there would be no losses [32]. The proportion of investing in a specific asset is also known as the weight of that asset in portfolio analysis. The pair of each such weights for every asset we invest in is called a portfolio, and each possible portfolio based on the same assets might have different expected returns and risks. Investors would, therefore, like to opt for a portfolio with the highest expected return and lowest risk. Risks in portfolio theory are commonly measured with standard deviation because it can measure volatility in returns. Hence, many investors might try to find the optimal portfolio by mean-variance analysis.

### 2.1.1 Minimum-variance portfolio

The minimum-variance portfolio (MVP), is the portfolio with the lowest variance out of all feasible portfolios. Let us first set up some notations, and then show how the MVP can be obtained. We consider a portfolio consisting of $N$ risky assets, which can be expressed with the weights collected into the following vector

$$w = (w_1, w_2, ..., w_N)^\top,$$

where $w_i$ denotes the weight of the $i$th risky asset, and the sum of the $N$ weights should be equal to 1. We can write this condition in matrix form as

$$w^\top 1_N = 1, \tag{1}$$

where $1_N$ is a vector of size $N$ containing only ones. Let us also denote the expected returns of the $N$ risky assets in vector form as

$$\mu = (\mu_1, \mu_2, ..., \mu_N)^\top,$$

and let the following matrix denote the covariance matrix of the returns (of the $N$ assets)

$$V = \begin{pmatrix} V_{11} & V_{12} & \cdots & V_{1N} \\ V_{21} & V_{22} & \cdots & V_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ V_{N1} & V_{N2} & \cdots & V_{NN} \end{pmatrix},$$

so, $V_{ij}$ denotes the covariance between the returns of the $i$th and $j$th assets for $i = 1, 2, ..., N$ and $j = 1, 2, ..., N$. The matrix $V$ is symmetric ($V = V^\top$) and positive semi-definite. Assume that the determinant of the covariance matrix $V$ is not equal to zero, so that its inverse matrix, $V^{-1}$, exists. Moreover, under the assumption of non-singularity, we get that $V$ is positive definite.

In order to find the MVP, we have to compute the minimum of the variance of a portfolio, $w^\top V w$, subject to the constraint $w^\top 1_N = 1$ (where all weights sum up to 1). This minimization problem can be expressed as

$$\min_w w^\top V w,$$

$$\text{s.t } w^\top 1_N = 1.$$

To solve this, one can use the Lagrange multipliers method. Let $\lambda$ be a Lagrange multiplier, then we can set up the following Lagrange function

$$F(w, \lambda) = w^\top V w - \lambda(w^\top 1_N - 1).$$

Taking the gradient of $F(w, \lambda)$ with respect to $w$, and setting that expression equal to a zero-vector of size $N$ gives the following system of linear equations

$$2Vw - \lambda 1_N = 0_N,$$

and solving for $w$ gives

$$w = \frac{\lambda}{2} V^{-1} 1_N. \tag{2}$$

By plugging in Equation (2) into Equation (1), we obtain that

$$\frac{\lambda}{2}(V^{-1}1_N)^\intercal 1_N = \frac{\lambda}{2} 1_N^\intercal \left(V^{-1}\right)^\intercal 1_N = \frac{\lambda}{2} 1_N^\intercal V^{-1} 1_N = 1,$$

and solving for the Lagrange multiplier $\lambda$ gives

$$\lambda = 2\left(1_N^\intercal V^{-1} 1_N\right)^{-1}.$$

The weights of the MVP are then obtained by plugging in the formula for $\lambda$ into Equation (2), and hence we get that

$$w_{\text{MVP}} = \frac{V^{-1} 1_N}{1_N^\intercal V^{-1} 1_N}.$$

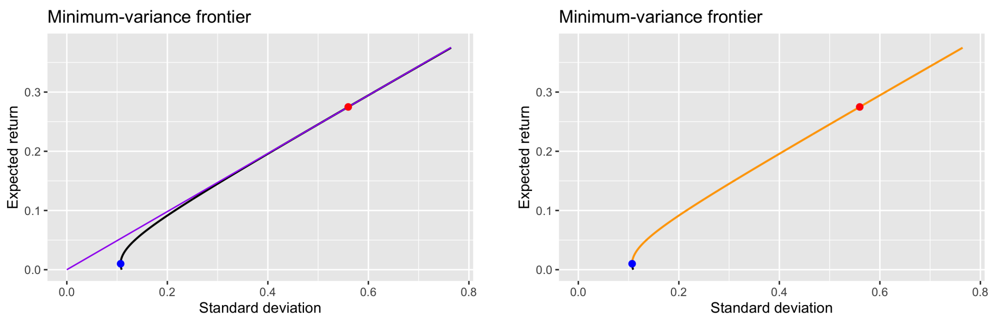
#### 2.1.2 Tangency portfolio

Assume that we have a portfolio that is constructed by $N$ risky assets and one risk-free asset, and let $\mu_P$ and $\sigma_P$ be the expected return and standard deviation, respectively, of a feasible portfolio $P$ constructed by the $N$ risky assets. Let also the vectors $w$ and $\mu$ consist of the weights and expected returns of each risky asset, as in the previous section. Let also $V$ be the covariance matrix of the risky assets as before. Assume furthermore that the expected return of the risk-free asset is denoted by $R_f$, and that its standard deviation is equal to zero. When a risk-free asset exists, investors would be interested in the tangency portfolio (TP), which is the portfolio $P$ on the efficient frontier that is tangent by the steepest possible straight line, referred to as a capital market line, that starts at the point $(0, R_f)$. The slope of this line is called the Sharpe ratio and can be obtained by using the following formula

$$\frac{\mu_P - R_f}{\sigma_P} = \frac{w^\intercal \mu - R_f}{\sqrt{w^\intercal V w}}.$$

An example of a TP and the steepest capital market line is shown in red and purple, respectively, in Figure 1 (a). Furthermore, an efficient frontier consists of all possible portfolios on the minimum-variance frontier that has a higher expected return than that of the MVP (see Figure 1 (b)). The minimum-variance frontier is the set of all feasible portfolios that gives the lowest risk for every attainable expected return.

Figure 1: The blue dots represent the MVP, and the red dots represent the TP in both plots. The straight purple line in the left figure (a) shows a capital market line where the expected return of the risk-free asset is equal to zero (or in the absence of the risky asset, $R_f = 0$). Furthermore, the black curved line represents the minimum-variance frontier in the left plot (a). In the right figure (b), we can see the minimum-variance frontier, and the orange part of the minimum-variance frontier represents the efficient frontier.




(a) (b)

To get the weights of the TP, we need to find the maximum of the Sharpe ratio subject to the constraint $w^\top 1_N = 1$. In other words, we want to solve for this maximization problem

$$\max_w \frac{w^\top \mu - R_f}{\sqrt{w^\top V w}},$$

$$\text{s.t. } w^\top 1_N = 1.$$

One can solve this problem by first setting up the Lagrange function as follows

$$G(w, \lambda) = \frac{w^\top \mu - R_f}{\sqrt{w^\top V w}} - \lambda \left( w^\top 1_N - 1 \right),$$

where $\lambda$ denotes a Lagrange multiplier. Then we find the gradient of $G(w, \lambda)$ with respect to $w$, and set it equal to a zero-vector of size $N$

$$\nabla_w G(w, \lambda) = \frac{\sqrt{w^\top V w}\mu - \frac{(w^\top \mu - R_f) V w}{\sqrt{w^\top V w}}}{w^\top V w} - \lambda 1_N = 0_N,$$

which can be written by

$$\frac{\mu_P - R_f}{\sigma_P^2} V w = \mu - \lambda \sigma_P 1_N, \tag{3}$$

where we used that $w^\top V w = \sigma_P^2$ and $w^\top \mu = \mu_P$. Multiplying Equation (3) by $w^\top$ on the left gives

$$\frac{\mu_P - R_f}{\sigma_P^2} w^\top V w = w^\top \mu - \lambda \sigma_P w^\top 1_N,$$

which is equivalent to

$$\mu_P - R_f = \mu_P - \lambda \sigma_P,$$

and solving for $\lambda$ gives that $\lambda = R_f/\sigma_P$. By plugging in $\lambda = R_f/\sigma_P$ into Equation (3), we obtain

$$\frac{\mu_P - R_f}{\sigma_P^2} V w = \mu - R_f 1_N.$$

Let us multiply both sides with the inverse of the covariance matrix, $V^{-1}$, on the left. Then,

$$\frac{\mu_P - R_f}{\sigma_P^2} w = V^{-1}(\mu - R_f 1_N). \tag{4}$$

To solve for $w$, we can first multiply both sides by $1_N^\top$ on the left, and the reason why we do not divide both sides by $(\mu_P - R_f)/\sigma_P^2$ to obtain the weights is that the expression $(\mu_P - R_f)/\sigma_P^2$ depends on $w$. Hence, we obtain

$$\frac{\mu_P - R_f}{\sigma_P^2} = 1_N^\top V^{-1}(\mu - R_f 1_N),$$

and by plugging in this expression into Equation (4) and then solving for $w$ gives that the weights of the TP are given by

$$w_{\mathrm{TP}} = \frac{V^{-1}(\mu - R_f 1_N)}{1_N^\top V^{-1}(\mu - R_f 1_N)}.$$

### 2.1.3 Two-fund theorem

Let us assume as before that the covariance matrix of $N$ risky assets is denoted as $V$ and that the determinant of this covariance matrix is not equal to zero, i.e., $|V| \neq 0$. Assume also that the two vectors $\mu$ and $1_N$ of size $N$, consisting of the expected returns of the $N$ risky assets and only ones respectively, are linearly independent. Then, one can obtain a vector of weights $w$ for the $N$ assets, which describes a portfolio that lies on the minimum-variance frontier with a specified expected return of the portfolio, say $\mu_P$, by solving the following optimization problem

$$\min_w w^\top V w$$

$$\text{s.t. } w^\top \mu = \mu_P$$

$$\text{s.t. } w^\top 1_N = 1.$$

One can do it by first setting up the Lagrange function with $\lambda$ and $\delta$ as the Lagrange multipliers

$$L(w, \lambda, \delta) = w^\top V w - \lambda(w^\top \mu - \mu_P) - \delta(w^\top 1_N - 1),$$

and then find the gradient of $L(w, \lambda, \delta)$ with respect to vector $w$, $\lambda$, and $\delta$. We then set the gradient equal to a zero-vector which leads to

$$2Vw - \lambda\mu - \delta 1_N = 0_N, \tag{5}$$

$$w^\top \mu - \mu_P = 0, \tag{6}$$

$$w^\top 1_N - 1 = 0. \tag{7}$$

We can rewrite Equation (5) as

$$2w = \lambda V^{-1}\mu + \delta V^{-1} 1_N, \tag{8}$$

and by multiplying this expression by $\mu^\top$ and $1_N^\top$ on the left, respectively, gives

$$\lambda \mu^\top V^{-1}\mu + \delta \mu^\top V^{-1} 1_N = 2\mu^\top w = 2\mu_P, \tag{9}$$

$$\lambda 1_N^\top V^{-1}\mu + \delta 1_N^\top V^{-1} 1_N = 2 \cdot 1_N^\top w = 2, \tag{10}$$

where $\mu^\top w = w^\top \mu = \mu_P$ and $1_N^\top w = w^\top 1_N = 1$ because of Equation (6) and (7), and the property of multiplying two vectors of the same size. We can express the equation system (9) and (10) in matrix form as

$$2\begin{pmatrix} \mu_P \\ 1 \end{pmatrix} = \begin{pmatrix} \mu^\top V^{-1}\mu & \mu^\top V^{-1} 1_N \\ 1_N^\top V^{-1}\mu & 1_N^\top V^{-1} 1_N \end{pmatrix} \begin{pmatrix} \lambda \\ \delta \end{pmatrix},$$

and we can solve for the Lagrange multipliers using the following expression

$$\begin{pmatrix} \lambda \\ \delta \end{pmatrix} = 2\begin{pmatrix} \mu^\top V^{-1}\mu & \mu^\top V^{-1} 1_N \\ 1_N^\top V^{-1}\mu & 1_N^\top V^{-1} 1_N \end{pmatrix}^{-1} \begin{pmatrix} \mu_P \\ 1 \end{pmatrix}.$$

Let us for simplicity write the inverse matrix as

$$\begin{pmatrix} \mu^\top V^{-1}\mu & \mu^\top V^{-1} 1_N \\ 1_N^\top V^{-1}\mu & 1_N^\top V^{-1} 1_N \end{pmatrix}^{-1} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}.$$

Then we obtain that $\lambda = 2(\mu_P M_{11} + M_{12})$, and $\delta = 2(\mu_P M_{21} + M_{22})$. By plugging in these into Equation (8), and dividing both sides by 2, we get that the weights of the portfolio on the minimum-variance frontier with the expected return $\mu_P$ is given by

$$\begin{aligned} w &= (\mu_P M_{11} + M_{12}) V^{-1}\mu + (\mu_P M_{21} + M_{22}) V^{-1} 1_N \\ &= \mu_P \left( M_{11} V^{-1}\mu + M_{21} V^{-1} 1_N \right) + M_{21} V^{-1}\mu + M_{22} V^{-1} 1_N, \end{aligned}$$

which can be written on this form

$$w = \mu_P a + b,$$

where $a = M_{11}V^{-1}\mu + M_{21}V^{-1}1_N$ and $b = M_{21}V^{-1}\mu + M_{22}V^{-1}1_N$.

The two-fund theorem is then defined as the following. Let $w_1$ and $w_2$ be two vectors of size $N$ containing the weights of two different portfolios, constructed by the $N$ risky assets, that lie on the minimum-variance frontier. Let also the expected return of the two portfolios be denoted by $\mu_1$ and $\mu_2$, respectively, where $\mu_1 \neq \mu_2$. One can then obtain the weights of a third portfolio, $w_3$, on the minimum-variance frontier with the expected return $\mu_3$ by using the weights and expected returns of the two known portfolios (i.e., $w_1$, $w_2$, $\mu_1$ and $\mu_2$) with the following formula

$$w_3 = \gamma w_1 + (1 - \gamma) w_2, \tag{11}$$

where $\gamma$ is a constant given by $\gamma = (\mu_3 - \mu_2)/(\mu_1 - \mu_2)$. The proof can be seen in Appendix A.1.1.

## 2.2 Sustainability

Sustainability, sometimes also referred to as socially responsible or ESG (environmental, social, and governance), seen from the business perspective mainly focuses on the following three dimensions [15, 25]:

- **Environmental (E):** The environmental factor focuses on how well companies are trying to protect the environment, i.e., this criterion concentrates on companies' ways of dealing with several kinds of issues that may affect the environment, such as climate change, pollution, and waste, among others.

- **Social (S):** The social criteria takes social aspects in companies such as human rights, labor rights, working conditions, and diversity, among others, into account.

- **Governance (G):** This dimension concentrates on corporate governance, and some aspects in this matter might include companies' taxes, executive compensations (e.g., salaries, and insurance, among others), and other responsibilities of company management, such as being aware of anti-corruption laws and staying away from bribery and corruption.

These three dimensions above will, therefore, contribute to measurements of sustainability. Investors with morality would preferably choose to invest in companies that take the three important factors above into account. It is common to measure a company's risk of facing ESG issues related to a specific dimension, and how well the company can manage them by using some kind of ESG scores that focus on the desired factor such as the environment risk score, social risk score, and governance risk score. Furthermore, one might, for example, want to make an overall judgment of a company's risk of facing and dealing with all three ESG dimensions instead of a particular dimension when one is studying sustainability. This is what total ESG scores and total ESG risk ratings do, to measure sustainability when all three dimensions are taken into account. Hence, we will use the total ESG scores and total ESG risk ratings, respectively, as the measurements for sustainability in the thesis.

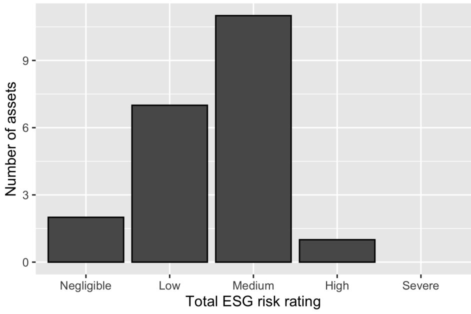
The total ESG scores and total ESG risk ratings we used were fetched from Yahoo Finance, and they are powered by Sustainalytics, a global leader in ESG research and ratings. Sustainalytics measures a company's total ESG risk rating by using a two-dimensional framework to assess what kind of industry-specific material ESG issues a company faces, and how well the company is managing those problems. Moreover, the Sustainalytics' total ESG risk rating is measured with the total ESG score, a score between 0 and 100. The score is then divided into five categories to describe the risk rating, where a score between 0-10 implies negligible risk, 10-20 means low risk, 20-30 medium risk, 30-40 high risk, and 40-100 severe risk. Hence, a higher value implies more unmanaged ESG risks [31]. We ended up with a dataset containing the total ESG scores and the total ESG risk ratings for 21 out of the 30 stocks in the OMXS30 index after some data cleaning. These scores and ratings were updated in January 2020.

The distribution of the total ESG risk ratings for the 21 assets can be seen in Figure 2 (a). We can, for example, see that most of the companies seem to have low to medium risk of facing ESG related issues, and none of the companies seem to be at severe risk. Furthermore, we can see from Figure 2 (b) that there are two assets with a total ESG score of 20, which means that the corresponding companies for these two assets might have a low or medium risk of facing ESG problems. However, from the same plot, we know that there are ten assets that have a total ESG score over 20 and below 30, and we can see from Figure 2 (a) that there are eleven stocks in total that have a total ESG risk rating of medium. Hence, one of the two assets with a total ESG score of 20 has a low risk of facing ESG issues, and the other has a medium risk.

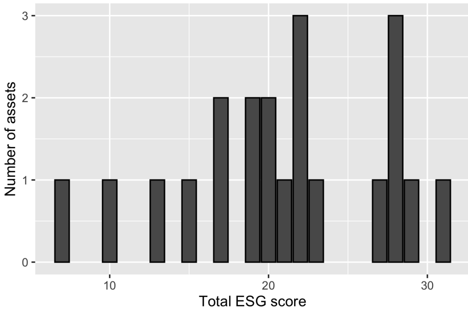
## 2.3 Screening

To obtain assets, where all the corresponding firms satisfy the ESG-criteria fairly well, one can exclude the firms where most of the ESG-criteria are not met. This procedure is called a screening [15].

Figure 2: The left plot (a) shows the number of assets out of 21 (in the OMXS30 index) that falls in each of the possible total ESG risk rating categories, and the right plot (b) visualizes the distribution of the total ESG scores for those assets.




(a) Total ESG risk rating




(b) Total ESG score

# 3 Mean-variance spanning test

The two papers by Huberman and Kandel [16], and Kan and Zhou [17], respectively, are the main references for the following subsections about mean-variance spanning tests. We will also follow the notations and steps closely from Kan and Zhou [17] in these subsections.

## 3.1 Spanning test based on regression

Huberman and Kandel [16] purposed and described a method called the mean-variance spanning test, which tests if the difference between the minimum-variance frontier of some assets and the minimum-variance frontier of those assets and some additional assets is statistically significant [15].

Let us assume that we have $K$ number of risky assets and that we would like to know whether the set of these assets spans the set consisting of these $K$ assets and $N$ other risky assets. Note that the set of the $K$ risky assets is said to span the set of the $K + N$ risky assets if the minimum-variance frontier of the $K$ risky assets and the minimum-variance frontier of the $K+N$ assets are identical. One usually calls the $K$ risky assets as the benchmark assets, and the $N$ risky assets are referred to as the test assets.

If there, for example, exists a risk-free asset, then investors that focus on the mean-variance analysis of portfolios will be interested in the TP of the risky assets. Furthermore, the interest of the investors is then to know if the TP obtained by using the set of $K$ benchmarks assets is significantly equivalent to the TP obtained by using $K+N$ risky assets (the $K$ benchmark assets and additionally $N$ test assets).

If risk-free assets are absent, on the other hand, then investors are rather interested to know whether the two minimum-variance frontiers are identical. If this is the case, let $R_{\mathrm{Tot},t} = (R_t^\top, r_t^\top)^\top$ denote an $(K + N) \times 1$-vector of the total returns of the $K + N$ assets at time $t$, where $R_t$ is a $K \times 1$-vector consisting of the returns of the $K$ benchmark assets at time $t$, and $r_t$ is an $N \times 1$-vector of the returns of the $N$ test assets at time $t$. The expected returns of the $K + N$ assets are given by $\mu = E[R_{\text{Tot},t}] = (\mu_{R_t}^\top, \mu_{r_t}^\top)^\top$, where $\mu_{R_t}$ and $\mu_{r_t}$ denote a $K \times 1$-vector and an $N \times 1$-vector of the expected returns of the benchmark assets and the test assets respectively. Let us furthermore define the covariance matrix of the returns of the $K + N$ assets as the following block matrix

$$V = Var(R_{\text{Tot},t}) = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix},$$

where we assume that the covariance matrix $V$ is nonsingular, and $V_{11}$, $V_{12}$, $V_{21}$, and $V_{22}$ denote a $K \times K$-matrix, a $K \times N$-matrix, an $N \times K$-matrix and an $N \times N$-matrix respectively.

Let $\alpha$ and $\epsilon_t$ denote two $N \times 1$-vectors, where the first-mentioned is a vector of intercepts, and the latter is a vector of disturbances at time $t$. Let furthermore $\beta$ be an $N \times K$-matrix consisting of regression coefficients. Assume that the expectation of vector $\epsilon_t$ is a vector of size $N$ consisting of only zeros, $E[\epsilon_t] = 0_N$. Moreover, assume that the expectation of $\epsilon_t R_t^\top$ is equal to an $N \times K$-matrix consisting of zeros, $E[\epsilon_t R_t^\top] = 0_{N \times K}$. Huberman and Kandel [16] then considered the following linear regression model

$$r_t = \alpha + \beta R_t + \epsilon_t. \tag{12}$$

Here, $r_t$, $R_t$, and $\epsilon_t$ are random vectors, and it is assumed that $\epsilon_t$ is not correlated with $R_t$.

**Proposition 1.** *The vector $\alpha$ and the matrix $\beta$ can be obtained by using the following formulas*

$$\alpha = \mu_{r_t} - \beta \mu_{R_t},$$

*and*

$$\beta = V_{21} V_{11}^{-1}.$$

*Proof.* The first expression can be shown by first taking the expectation of Equation (12)

$$E[r_t] = E[\alpha + \beta R_t + \epsilon_t] = \alpha + \beta E[R_t] + 0_N = \alpha + \beta \mu_{R_t}.$$

Using that $E[r_t] = \mu_{r_t}$, and solving for $\alpha$ in the expression above will give us that $\alpha = \mu_{r_t} - \beta \mu_{R_t}$. To prove that $\beta = V_{21} V_{11}^{-1}$, we can multiply the vector $R_t^\top$ on both sides (on the right) of Equation (12), and then take the expectation of that expression

$$\begin{aligned} E[r_t R_t^\top] &= E[\alpha R_t^\top + \beta R_t R_t^\top + \epsilon_t R_t^\top] \\ &= \alpha E[R_t^\top] + \beta E[R_t R_t^\top] + 0_{N \times K} \\ &= (\mu_{r_t} - \beta \mu_{R_t}) \mu_{R_t}^\top + \beta E[R_t R_t^\top] \\ &= \mu_{r_t} \mu_{R_t}^\top + \beta V_{11}, \end{aligned}$$

where we have used that $E[\epsilon_t R_t^\top] = 0_{N\times K}$ in the second equality, and the expression $\alpha = \mu_{r_t} - \beta\mu_{R_t}$ in the third equality. The last equality holds because $E[R_t R_t^\top] - \mu_{R_t}\mu_{R_t}^\top = V_{11}$. Solving for $\beta$ yields

$$\beta = \left(E[r_t R_t^\top] - \mu_{r_t}\mu_{R_t}^\top\right) V_{11}^{-1} = V_{21}V_{11}^{-1}.$$

Note that $E[r_t R_t^\top] - \mu_{r_t}\mu_{R_t}^\top$ is an $N \times K$-matrix containing covariances between the elements in the vectors $r_t$ and $R_t$. Hence,

$$E[r_t R_t^\top] - \mu_{r_t}\mu_{R_t}^\top = V_{21}.$$

□

The spanning test will verify the null-hypothesis of spanning, which can be expressed as

$$H_0 : \alpha = 0_N \quad \text{and} \quad 1_N - \beta 1_K = 0_N, \tag{13}$$

against the alternative-hypothesis

$$H_1 : \alpha \neq 0_N \quad \text{or} \quad 1_N - \beta 1_K \neq 0_N,$$

where $1_N$ and $1_K$ denote an $N \times 1$-vector and a $K \times 1$-vector consisting of ones respectively.

By testing $\alpha = 0_N$, we test whether the TP has zero weights in the $N$ test assets, and we test whether the MVP has zero weights in the $N$ test assets by testing $1_N - \beta 1_K = 0_N$. To motivate this, let us assume that we have two portfolios located on the minimum-variance frontier of the $K + N$ risky assets and that the expected return of the MVP is not equal to zero, i.e., $\frac{\mu^\top V^{-1} 1_{K+N}}{1_{K+N}^\top V^{-1} 1_{K+N}} \neq 0$, where $1_{K+N}^\top V^{-1} 1_{K+N} \neq 0$ since $V$ is positive definite, hence it implies that $1_{K+N}^\top V^{-1}\mu \neq 0$. The weights of the two portfolios are given by the following expressions (see Section 2.1.1 and Section 2.1.2)

$$w_1 = \frac{V^{-1}\mu}{1_{K+N}^\top V^{-1}\mu},$$

and

$$w_2 = \frac{V^{-1} 1_{K+N}}{1_{K+N}^\top V^{-1} 1_{K+N}}.$$

The first portfolio is the TP when the expected return of the risk-free asset is zero (or in the absence of risk-free assets), and the other portfolio is the MVP. Let us use the formula for the inverse of a partitioned matrix on the nonsingular matrix $V$, and let $\Omega = V_{22} - V_{21}V_{11}^{-1}V_{12}$ for simplicity, then we obtain that

$$\begin{aligned} V^{-1} &= \begin{pmatrix} V_{11}^{-1} + V_{11}^{-1}V_{12}(V_{22} - V_{21}V_{11}^{-1}V_{12})^{-1}V_{21}V_{11}^{-1} & -V_{11}^{-1}V_{12}(V_{22} - V_{21}V_{11}^{-1}V_{12})^{-1} \\ -\left(V_{22} - V_{21}V_{11}^{-1}V_{12}\right)^{-1} V_{21}V_{11}^{-1} & \left(V_{22} - V_{21}V_{11}^{-1}V_{12}\right)^{-1} \end{pmatrix} \\ &= \dots \\ &= \begin{pmatrix} V_{11}^{-1} + \beta^\top\Omega^{-1}\beta & -\beta^\top\Omega^{-1} \\ -\Omega^{-1}\beta & \Omega^{-1} \end{pmatrix}. \end{aligned}$$

The derivation of the inverse covariance matrix can be found in Appendix A.1.2.

**Proposition 2.** *Let $Q = (0_{N\times K}, I_N)$, where $I_N$ is an $N\times N$-identity matrix. Then the expressions of the weights of the $N$ test assets in the two portfolios (TP and MVP) can be written as*

$$Qw_1 = \frac{\Omega^{-1}\alpha}{1_{K+N}^\top V^{-1}\mu}, \quad \text{and} \quad Qw_2 = \frac{\Omega^{-1}(1_N - \beta 1_K)}{1_{K+N}^\top V^{-1}1_{K+N}}, \tag{14}$$

*which depend on $\alpha$ and $1_N - \beta 1_K$ respectively.*

*Proof.* Recall that $\mu = (\mu_{R_t}^\top, \mu_{r_t}^\top)^\top$ and $\alpha = \mu_{r_t} - \beta\mu_{R_t}$. Hence, the weights of the $N$ test assets in the TP can be expressed as

$$\begin{aligned}
Qw_1 &= \frac{QV^{-1}\mu}{1_{K+N}^\top V^{-1}\mu} \\
&= \begin{pmatrix} 0_{N\times K} & I_N \end{pmatrix} \begin{pmatrix} V_{11}^{-1} + \beta^\top\Omega^{-1}\beta & -\beta^\top\Omega^{-1} \\ -\Omega^{-1}\beta & \Omega^{-1} \end{pmatrix} \frac{\mu}{1_{K+N}^\top V^{-1}\mu} \\
&= \frac{(-\Omega^{-1}\beta, \Omega^{-1})\mu}{1_{K+N}^\top V^{-1}\mu} \\
&= (-\Omega^{-1}\beta, \Omega^{-1}) \begin{pmatrix} \mu_{R_t} \\ \mu_{r_t} \end{pmatrix} \frac{1}{1_{K+N}^\top V^{-1}\mu} \\
&= \frac{-\Omega^{-1}\beta\mu_{R_t} + \Omega^{-1}\mu_{r_t}}{1_{K+N}^\top V^{-1}\mu} \\
&= \frac{\Omega^{-1}(\mu_{r_t} - \beta\mu_{R_t})}{1_{K+N}^\top V^{-1}\mu} \\
&= \frac{\Omega^{-1}\alpha}{1_{K+N}^\top V^{-1}\mu},
\end{aligned}$$

which depends on $\alpha$. Similarly, by first noting that the vector $1_{K+N}$ can be rewritten as $1_{K+N} = (1_K^\top, 1_N^\top)^\top$, then we can express the weights of the $N$ test assets in the MVP as the following

$$\begin{aligned}
Qw_2 &= \frac{QV^{-1}1_{K+N}}{1_{K+N}^\top V^{-1}1_{K+N}} \\
&= \frac{(-\Omega^{-1}\beta, \Omega^{-1})1_{K+N}}{1_{K+N}^\top V^{-1}1_{K+N}} \\
&= \frac{(-\Omega^{-1}\beta, \Omega^{-1})(1_K^T, 1_N^\top)^\top}{1_{K+N}^\top V^{-1}1_{K+N}} \\
&= \frac{-\Omega^{-1}\beta 1_K + \Omega^{-1}1_N}{1_{K+N}^\top V^{-1}1_{K+N}} \\
&= \frac{\Omega^{-1}(1_N - \beta 1_K)}{1_{K+N}^\top V^{-1}1_{K+N}},
\end{aligned}$$

which depends on $1_N - \beta 1_K$. $\square$

Hence, we can see from Equation (14) that testing $\alpha = 0_N$ implies testing whether the TP of the $K+N$ assets has zero weights in the $N$ test assets, and testing $1_N - \beta 1_K = 0_N$ implies testing whether the MVP of the $K+N$ risky assets will have zero weights in the $N$ test assets. From the two-fund theorem, we know that one can obtain each possible portfolio that lies on the minimum-variance frontier if we have the weights and expected return of two portfolios on the same minimum-variance frontier (see Equation (11)). Hence, if both the TP and the MVP on the minimum-variance frontier of the $K+N$ assets have zero weights in the $N$ test assets, then the formula for the two-fund theorem in Equation (11) shows that the weights of the $N$ test assets in each portfolio on the same minimum-variance frontier will also be zero. In other words, the minimum-variance frontier of the $K+N$ assets will be the same as the minimum-variance frontier of the $K$ benchmark assets.

### 3.1.1 Spanning test based on multivariate regression

Recall the regression model described in Equation (12), where we have data of returns of the $K+N$ assets over $T$ time points, or periods. Let us assume that $\alpha$ and $\beta$ are constant over time. We can then estimate them by using a multivariate regression model expressed as

$$Y = XB + E,$$

where $Y$ denotes a $T \times N$-matrix of the returns of the $N$ test assets, $r_t$, over a time series where $t = 1, 2, ..., T$. The matrix $X$ is a $T \times (K+1)$-matrix consisting of ones in the first column, and the returns of the $K$ benchmark assets, $R_t$ for $t = 1, 2, ..., T$, in the other columns. In other words, the $t$th row in $X$ is given by $[1, R_t^\top]$, where $R_t$ is a vector of size $K$ consisting of the returns of the $K$ benchmark assets at the time point, or period $t$. Furthermore, $B = [\alpha, \beta]^\top$ is a $(K+1) \times N$-matrix, and $E$ is a $T \times N$-matrix, where the $t$th row is given by a vector of the disturbances at the time point or period $t$, $\epsilon_t^\top$, for $t = 1, 2, ..., T$. To clarify, we can write these matrices as

$$Y = \begin{pmatrix} r_1^\top \\ r_2^\top \\ \vdots \\ r_T^\top \end{pmatrix}, \quad X = \begin{pmatrix} 1 & R_1^\top \\ 1 & R_2^\top \\ \vdots & \vdots \\ 1 & R_T^\top \end{pmatrix}, \quad B = \begin{pmatrix} \alpha^\top \\ \beta^\top \end{pmatrix}, \quad E = \begin{pmatrix} \epsilon_1^\top \\ \epsilon_2^\top \\ \vdots \\ \epsilon_T^\top \end{pmatrix},$$

where $r_t$ for $t = 1, 2, ..., T$, is a vector of size $N$ (consisting of the returns of the test asset) as in Section 3.1, $\alpha$ is a vector consisting of $N$ intercepts, and $\beta^\top$ is a $K \times N$-matrix. Moreover, we should also assume that $T \geq N+K+1$. This assumption implies that there exists an inverse of the matrix $X^\top X$, see Theorem 3.1.4 in Muirhead [22]. Additionally, we assume that the disturbances,

$\epsilon_t$ for $t = 1, 2, ..., T$, conditioned on the returns of the $K$ benchmark assets $R_t$ for $t = 1, 2, ..., T$, are independent and identically distributed (i.i.d.), and follow a multivariate normal distribution with a mean-vector of size $N$ containing solely zeros and a covariance matrix denoted as $\Sigma$, which can be written as $N(0_N, \Sigma)$. Hence, $E$ will be $N(0_{T\times N}, I_T \otimes \Sigma)$-distributed under the normality assumption of the disturbances conditioned on the returns of the $K$ risky assets [1], where $\otimes$ denotes the Kronecker product, and $0_{T\times N}$ is a $T \times N$ matrix of zeros [22]. Furthermore, the unconstrained maximum likelihood estimators of the matrices $B$ and $\Sigma$ are obtained by using the following expressions (see Appendix A.1.3 for the derivation of the formulas)

$$\hat{B} \equiv [\hat{\alpha}, \hat{\beta}]^\top = (X^\top X)^{-1} X^\top Y,$$

$$\hat{\Sigma} = \frac{1}{T}(Y - X\hat{B})^\top (Y - X\hat{B}).$$

Under the assumption that the disturbances $\epsilon_t$ are i.i.d. $N(0_N, \Sigma)$-distributed, the maximum likelihood estimator of $B$ conditionally on the returns of the benchmark assets will be normally distributed with the mean-matrix $B$ and covariance matrix $(X^\top X)^{-1} \otimes \Sigma$ according to Muirhead [22]. This can be denoted as

$$\hat{B}|X \sim N[B, (X^\top X)^{-1} \otimes \Sigma],$$

and the vectorization of $\hat{B}^\top$ conditionally on $X$ will then be normally distributed with the vectorization of $B^\top$ as the mean-vector and $(X^\top X)^{-1} \otimes \Sigma$ as the covariance matrix,

$$\text{vec}(\hat{B}^\top)|X \sim N\left[\text{vec}(B^\top), (X^\top X)^{-1} \otimes \Sigma\right].$$

Furthermore, $T\hat{\Sigma}$ will follow an $N$-dimensional central Wishart distribution with $T - K - 1$ degrees of freedom and covariance matrix $\Sigma$ under the normality assumption, which can be denoted as

$$T\hat{\Sigma} \sim W_N\left(T - K - 1, \Sigma\right).$$

Let us define $\Theta = [\alpha, 1_N - \beta 1_K]^\top$, then we can rewrite the null-hypothesis in Equation (13) as

$$H_0 : \Theta = 0_{2\times N}, \tag{15}$$

where $0_{2\times N}$ is a $2 \times N$-matrix with only zeros. The matrix $\Theta$ can also be written in the following form $\Theta = AB + C$, where

$$A = \begin{pmatrix} 1 & 0_K^\top \\ 0 & -1_K^\top \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 0_N^\top \\ 1_N^\top \end{pmatrix}.$$

[1]The unconditional distribution of $E$ will also have this distribution, $N(0_{T\times N}, I_T \otimes \Sigma)$, since the conditional distribution of $E$ does not depend on $X$.

To obtain the maximum likelihood estimator of $\Theta$, one can use the formula $\hat{\Theta} \equiv [\hat{\alpha}, 1_N - \hat{\beta}1_K]^\top = A\hat{B} + C$, where the vectorization of $\hat{\Theta}^\top$ is normally distributed as follows

$$\text{vec}(\hat{\Theta}^\top)|X \sim N\left[\text{vec}(\Theta^\top), (\hat{G}/T) \otimes \Sigma\right]$$

with

$$\hat{G} = TA(X^\top X)^{-1}A^\top = \begin{pmatrix} 1 + \hat{\mu}_{R_t}^\top \hat{V}_{11}^{-1}\hat{\mu}_{R_t} & \hat{\mu}_{R_t}^\top \hat{V}_{11}^{-1} 1_K \\ \hat{\mu}_{R_t}^\top \hat{V}_{11}^{-1} 1_K & 1_K^\top \hat{V}_{11}^{-1} 1_K \end{pmatrix},$$

where $\hat{\mu}_{R_t} = (1/T)\sum_{t=1}^T R_t$ and $\hat{V}_{11} = (1/T)\sum_{t=1}^T (R_t - \hat{\mu}_{R_t})(R_t - \hat{\mu}_{R_t})^\top$.

Remember that conditional on $X$, the matrix $E$ is assumed to follow a $N(0_{T\times N}, I_T \otimes \Sigma)$-distribution, so $Y$ is assumed to be $N(XB, I_T \otimes \Sigma)$-distributed. From Muirhead [22], one can then find that the likelihood function of $Y$ conditionally on $X$ is given by

$$L(B,\Sigma) = (2\pi)^{-NT/2}|\Sigma|^{-T/2}\exp\left\{\text{Trace}\left[-\frac{1}{2}(Y - XB)\Sigma^{-1}(Y - XB)^\top\right]\right\}.$$

Let $\tilde{B}$ and $\tilde{\Sigma}$ denote the constrained maximum likelihood estimators of $B$ and the covariance matrix $\Sigma$, respectively, i.e., the maximum likelihood estimators of $B$ and $\Sigma$ when $H_0$ is true. Then, the asymptotic likelihood ratio test to verify $H_0$ is given by

$$LR = -2\left\{\log\left[L(\tilde{B},\tilde{\Sigma})\right] - \log\left[L(\hat{B},\hat{\Sigma})\right]\right\} = -T\log\left(\frac{|\hat{\Sigma}|}{|\tilde{\Sigma}|}\right) \overset{a}{\sim} \chi^2_{2N}.$$

In other words, the likelihood ratio test statistic will asymptotically follow a $\chi^2$-distribution with $2N$ degrees of freedom. By continue following Kan and Zhou [17], we let $U = |\hat{\Sigma}|/|\tilde{\Sigma}|$. Moreover, they mentioned that it is possible to compute the likelihood ratio test statistic without calculating the constrained maximum likelihood estimator $\tilde{\Sigma}$. We should first note that

$$\tilde{\Sigma} - \hat{\Sigma} = \hat{\Theta}^\top \hat{G}^{-1}\hat{\Theta}$$

according to Seber [29]. By using this expression and some determinant properties, we can rewrite $1/U$ as the following

$$\frac{1}{U} = \frac{|\tilde{\Sigma}|}{|\hat{\Sigma}|} = |\hat{\Sigma}^{-1}\tilde{\Sigma}| = |\hat{\Sigma}^{-1}(\tilde{\Sigma} - \hat{\Sigma} + \hat{\Sigma})| = |\hat{\Sigma}^{-1}(\hat{\Sigma} + \hat{\Theta}^\top \hat{G}^{-1}\hat{\Theta})|,$$

which is equivalent to

$$\frac{1}{U} = |I_N + \hat{\Sigma}^{-1}\hat{\Theta}^\top \hat{G}^{-1}\hat{\Theta}| = |I_2 + \hat{H}\hat{G}^{-1}|,$$

where the second equality holds according to the equation (B.1.16) of [28], and

$$\hat{H} = \hat{\Theta}\hat{\Sigma}^{-1}\hat{\Theta}^\top = \begin{pmatrix} \hat{\alpha}^\top \hat{\Sigma}^{-1}\hat{\alpha} & \hat{\alpha}^\top \hat{\Sigma}^{-1}(1_N - \hat{\beta}1_K) \\ \hat{\alpha}^\top \hat{\Sigma}^{-1}(1_N - \hat{\beta}1_K) & (1_N - \hat{\beta}1_K)^\top \hat{\Sigma}^{-1}(1_N - \hat{\beta}1_K) \end{pmatrix}.$$

Let us denote $\lambda_1$ and $\lambda_2$ as the two eigenvalues[2] of the matrix $\hat{H}\hat{G}^{-1}$, where we assume that $\lambda_1 \geq \lambda_2 \geq 0$. Then we can obtain $1/U$ by using the following formula

$$\frac{1}{U} = (1+\lambda_1)(1+\lambda_2),$$

and the desired likelihood ratio test can then be expressed as

$$LR = T\log\left(\frac{1}{U}\right) = T[\log(1+\lambda_1) + \log(1+\lambda_2)] = T\sum_{i=1}^{2}\log(1+\lambda_i) \overset{a}{\sim} \chi^2_{2N}.$$

Furthermore, Kan and Zhou [17] provided two asymptotic spanning tests, the Wald test and the Lagrangian multiplier test, which are based on the Lawley-Hotelling trace and the Pillai's trace respectively. From Anderson [2], we know that one can test a null-hypothesis by using the Lawley-Hotelling trace, which can be obtained by calculating $\text{Trace}(\hat{H}\hat{G}^{-1})$. Alternatively, one can use the Pillai's trace by computing $\text{Trace}[\hat{H}\hat{G}^{-1}/(I_2 + \hat{H}\hat{G}^{-1})]$ to test for the null-hypothesis. Both test statistics obtained by multiplying the Lawley-Hotelling trace and Pillai's trace by $T$ are asymptotically $\chi^2$-distributed. Hence, the asymptotic Wald test to verify the hypothesis of spanning presented in the paper by Kan and Zhou [17] can be obtained by the following

$$W = T(\lambda_1 + \lambda_2) \overset{a}{\sim} \chi^2_{2N}.$$

To perform the asymptotic Lagrange multiplier test, we can use the following expression

$$LM = T\sum_{i=1}^{2}\frac{\lambda_i}{1+\lambda_i} \overset{a}{\sim} \chi^2_{2N}.$$

All three tests (the Likelihood ratio test, the Wald test, and the Lagrange multiplier test) are asymptotic $\chi^2_{2N}$-distributed. However, the three tests might give different results when one uses them to perform a hypothesis test. Berndt and Savin [4] argued that the following inequality holds when the samples are finite $W \geq LR \geq LM$. Hence, Kan and Zhou [17] stated that it might be important to study and apply all three tests, and not only focusing on the likelihood ratio test.

### 3.1.2 Spanning test for small samples

Asymptotic tests can lead to misleading results if a sample is not large enough, and might, therefore, be unsuitable to use when one has small samples. When there is no risk-free asset, Huberman and Kandel [16] proposed

[2] The eigenvalues, $\lambda_1$ and $\lambda_2$, can be seen as measures of the maximum and minimum differences, respectively, between the potential minimum-variance frontiers in terms of squared sample Sharpe ratios. More details can be seen in Lemma 1 from the paper by Kan and Zhou [17].

the exact distribution of a statistic (which is a monotone transformation of the likelihood ratio test) when one wants to test $H_0$ against $H_1$. However, there is a small typo in the paper by Huberman and Kandel [16], which Kan and Zhou [17] noticed. The exact distribution of the test statistic under $H_0$ when $N \geq 2$ can be written as

$$\left(\frac{1}{\sqrt{U}} - 1\right)\left(\frac{T-K-N}{N}\right) \sim F_{2N,2(T-K-N)},$$

which has an $F$-distribution with $2N$ and $2(T-K-N)$ degrees of freedoms. In the case with only one test asset, $N = 1$, the exact distribution of the test statistic under $H_0$ is given by

$$\left(\frac{1}{U} - 1\right)\left(\frac{T-K-1}{2}\right) \sim F_{2,T-K-1}.$$

Furthermore, from Kan and Zhou [17] we can see that the exact distribution of the Wald test under $H_0$ when one has two or more test assets is

$$\begin{aligned} P(\lambda_1 + \lambda_2 \leq w) &= I_{\frac{w}{2+w}}(N-1, T-K-N) \\ &\quad - \frac{B\left(\frac{1}{2}, \frac{T-K}{2}\right)}{B\left(\frac{N}{2}, \frac{T-K-N+1}{2}\right)}(1+w)^{-\left(\frac{T-K-N}{2}\right)} I_{\left(\frac{w}{2+w}\right)^2}\left(\frac{N-1}{2}, \frac{T-K-N}{2}\right), \end{aligned}$$

where $B(a,b)$ is a beta function with $a$ and $b$ as the parameters, and $I_x(a,b)$ is the regularized incomplete beta function, which is the incomplete beta function $B(x;a,b)$ divided by the beta function $B(a,b)$.

Kan and Zhou [17] also showed that the exact distribution of the Lagrange multiplier test under $H_0$ when one has two or more test assets and for $0 \leq v \leq 2$ is the following

$$\begin{aligned} P\left(\frac{\lambda_1}{1+\lambda_1} + \frac{\lambda_2}{1+\lambda_2} \leq v\right) &= I_{\frac{v}{2}}(N-1, T-K-N+1) \\ &\quad - \frac{\int_{\max(0,v-1)}^{v^2/4} u^{\frac{N-3}{2}}(1-v+u)^{\frac{T-K-N}{2}} du}{2B(N-1, T-K-N+1)}. \end{aligned}$$

### 3.1.3 Spanning test when the normality assumption is violated and in the presence of heteroscedasticity

Kan and Zhou [17] mentioned that the asymptotic mean-variance spanning tests described previously will not be asymptotically $\chi^2_{2N}$-distributed under $H_0$ if the residuals are conditionally heteroscedastic. Furthermore, the distributions of the exact tests will not follow the distributions shown in the previous section in this case. They continued by letting the readers know how one can deal with this issue. The suggested solution is to use the approach in the paper by Ferson, Foerster, and Keim [11], which is based on the generalized methods of moments (GMM) presented in a paper by Hansen [12].

Following the paper by Kan and Zhou [17], we should let $x_t = (1, R_t^\top)^\top$, then the disturbances can be expressed as $\epsilon_t = r_t - B^\top x_t$ for $t = 1, 2, ..., T$. Let furthermore, $g_t = x_t \otimes \epsilon_t$. We should then use the following moment conditions in order to obtain the GMM estimator of matrix $B$,

$$E[g_t] = E[x_t \otimes \epsilon_t] = 0_{(K+1)N}.$$

Then, we would have to make assumptions of the returns of the $K+N$ risky assets, $R_{\text{Tot},t} = (R_t^\top, r_t^\top)^\top$, being stationary, and where the fourth moments are finite. The GMM estimate of $B$ is then obtained by minimizing the following formula

$$\bar{g}_T(B)^\top S_T^{-1} \bar{g}_T(B),$$

where $S_T$ is a consistent estimate of $E[g_t g_t^\top]$ in the absence of serial correlation for $g_t$, and $\bar{g}_T(B)$ is the sample moments, which is given by

$$\bar{g}_T(B) = \frac{1}{T}\sum_{t=1}^{T} x_t \otimes \epsilon_t = \frac{1}{T}\sum_{t=1}^{T} x_t \otimes (r_t - B^\top x_t).$$

Furthermore, Kan and Zhou [17] mentioned that the unconstrained GMM estimate of $B$ is identical to the maximum likelihood estimate $\hat{B}$ in Section 3.1.1 in this case. Hence, $\hat{\Theta}$ based on the GMM estimate of $B$ would also be the same as in Section 3.1.1. The spanning test based on GMM can be seen as a variant of the Wald test. The test statistic can be obtained by the formula as follows, which is asymptotically $\chi^2$-distributed with $2N$ degrees of freedom,

$$W_a = T\text{vec}(\hat{\Theta}^\top)^\top \left[(A_T \otimes I_N) S_T (A_T^\top \otimes I_N)\right]^{-1} \text{vec}(\hat{\Theta}^\top) \overset{a}{\sim} \chi^2_{2N}.$$

The matrix $A_T$ in the expression above is given by

$$A_T = \begin{pmatrix} 1 + \hat{a}_1 & -\hat{\mu}_{R_t}^\top \hat{V}_{11}^{-1} \\ \hat{b}_1 & -1_K^\top \hat{V}_{11}^{-1} \end{pmatrix},$$

where $\hat{a}_1 = \hat{\mu}_{R_t}^\top \hat{V}_{11}^{-1} \hat{\mu}_{R_t}$, $\hat{b}_1 = \hat{\mu}_{R_t}^\top \hat{V}_{11}^{-1} 1_K$, $\hat{\mu}_{R_t} = (1/T)\sum_{t=1}^T R_t$, and $\hat{V}_{11} = (1/T)\sum_{t=1}^T (R_t - \hat{\mu}_{R_t})(R_t - \hat{\mu}_{R_t})^\top$. More details about GMM can be found in Hansen [12].

## 3.2 Robustness analysis: Design of simulation study

In this section, we describe some simulation procedures used in the thesis to study the robustness of non-normality of the asymptotic and exact mean-variance spanning tests. Let $E^*$ and $\epsilon_t^*$ denote a simulated sample of $E$ and

$\epsilon_t$ respectively. One can study the robustness of the tests by first simulating each row in $E^*$, i.e., $\epsilon_t^*$ for $t = 1, 2, ..., T$, from a $N(0_N, \Sigma)$-distribution. Then, one may use the simulated matrix $E^*$ and a matrix $B$, such that $H_0$ (or $H_1$) is satisfied, to obtain new monthly and weekly returns of the test assets, respectively. One can then perform the asymptotic and exact tests as usual with the new simulated return of the test assets. This procedure will be repeated until we have performed 10,000 simulations in total, and we will then compute the actual probabilities of rejecting the null-hypothesis in (13) when $H_0$ (or $H_1$) is true. Likewise, one can perform this simulation procedure but where each row in $E^*$ is drawn from a multivariate t-distribution instead to obtain the probabilities of rejection. One can then study whether the probabilities obtained using multivariate normally distributed disturbances and multivariate t-distributed disturbances, respectively, differ a lot or are quite similar. By doing so, one can study the deviation of the performance of the spanning tests when the residuals are i.i.d. but not normally distributed. Furthermore, one might want to analyze the performance of the tests when the disturbances are independent but possibly not identically distributed. To do this, one can compute the probabilities of rejecting the null-hypothesis in (13) when $H_0$ (or $H_1$) is true by performing a similar simulation procedure as before, but where one applies the residual bootstrap resampling method to obtain $E^*$ instead. Lastly, one would probably also want to study the performance of the spanning tests when the residuals might show heteroscedastic or autocorrelated patterns. One option to do this is to use a non-overlapping block bootstrap in order to resample the matrix $E^*$. These are some methods that will be used in our thesis, and a more detailed description of the simulation procedures when $E^*$ is simulated from a multivariate normal distribution, multivariate t-distribution, residual bootstrap, and non-overlapping block bootstrap can be seen in Section 3.2.1, 3.2.2, 3.2.3 and 3.2.4 respectively.

Let us first make it clear how we obtain some parameter matrices, which will be consistent for almost all the simulation procedures. The first parameter matrix we need in order to perform most of our simulations is the covariance matrix $\Sigma$. We will draw the matrix $\Sigma$ from a Wishart distribution with $N$ degrees of freedom and covariance matrix $I_N$. The $B$-matrix we will use in all the simulation procedures is based on the maximum likelihood estimated matrix $\hat{B}$ obtained using our original dataset with monthly returns. In other words, we set the $\hat{B}$-matrix as the $B$-matrix. Then, we force the matrix $B$ to satisfy $H_0$ (or a possible $H_1$ case):

- $B$ **under** $H_0$: Set the first row of the matrix $B$ to zeros (so $\alpha^\top = 0_N^\top$), and then divide each element in the matrix with its corresponding column sum (hence, $1_K^\top \beta^\top = 1_N^\top$).
- $B$ **under** $H_1$: The matrix $B$ under some possible cases of $H_1$ will build on the matrix $B$ under $H_0$. In other words, we will start by

using the matrix $B$ where we forced it to satisfy $H_0$, and then make some changes in one or two rows, such that the matrix $B$ will satisfy one of the following possibilities of $H_1$:

- **Case 1**: The $B$ matrix under this $H_1$ is very similar to $H_0$, but all the elements in the first row of $B$ are set to be 0.00001 instead of zero, and we also add the value 0.00001 to each element in the second row.
- **Case 2**: Each element in the first row of $B$ is set to be 0.001 instead of zero, and the value 0.001 is added to each element in the second row.
- **Case 3**: The elements in the first row of $B$ are all set to the value 0.1. This value is also added to each element in the second row.
- **Case 4**: We let all the elements in the first row of $B$ to be zero ($\alpha^\top = 0_N^\top$). Furthermore, we add the value 0.00001 to each element in the second row, so each column in $B$ are summed up to 1.00001.
- **Case 5**: The first row of $B$ consists of solely zeros ($\alpha^\top = 0_N^\top$), while we add the value 0.001 to each element in the second row. Hence, we have that each element in $1_K^\top \beta^\top$ will be equal to 1.001.
- **Case 6**: The first row of $B$ consists of only zeros as the previous case ($\alpha^\top = 0_N^\top$), but we add the value 0.1 to each element in the second row instead of 0.001. So, $1_K^\top \beta^\top$ is equal to a vector where each element is equal to 1.1.
- **Case 7**: All the elements in the first row of $B$ are still zeros ($\alpha^\top = 0_N^\top$), but each element in the second row will be added by the value 1, which implies that each element in $1_K^\top \beta^\top$ is equal to 2.
- **Case 8**: All the elements in the first row of $B$ are set to 0.00001, while the other values in the matrix remain the same as under $H_0$. In other words, we have that $1_K^\top \beta^\top = 1_N^\top$.
- **Case 9**: The elements in the first row of $B$ are all equal to 0.001, and the remaining values in the matrix are identical to those under $H_0$.
- **Case 10**: The elements in the first row of $B$ are all equal to 0.1, while the other values in the matrix are the same as those under $H_0$.

Note that we have that both the TPs and the MVPs are different ($\alpha \neq 0_N$ and $1_N - \beta 1_K \neq 0_N$) under cases 1 to 3 and that the TPs are the same while the MVPs differ ($\alpha = 0_N$ and $1_N - \beta 1_K \neq 0_N$) under cases 4 to 7. Furthermore, we have that the TPs differ while the MVPs are identical ($\alpha \neq 0_N$ and $1_N - \beta 1_K = 0_N$) under cases 8 to 10.

### 3.2.1 Simulate the residuals from a multivariate normal distribution

Recall that the covariance matrix $\Sigma$ is simulated from a Wishart distribution and that the matrix $B$ satisfies $H_0$ (or a possible case of $H_1$, depending on the interest). The simulation procedure to obtain the actual probability of rejecting the null-hypothesis in (13) at the 5% significance level when $H_0$ (or $H_1$) is true for each spanning test is the following:

1. Simulate a matrix $E^*$, by simulating each row, $\epsilon_t^*$ for $t = 1, 2, ..., T$, from the $N(0_N, \Sigma)$-distribution.

2. Obtain new returns of the test assets, which we denote as $Y^*$, by using the formula $Y^* = XB + E^*$.

3. Use $X$ and $Y^*$ to compute the maximum likelihood estimators of $B$ and $\Sigma$, respectively, i.e., calculate $\hat{B}^*$ and $\hat{\Sigma}^*$ by using the formulas for $\hat{B}$ and $\hat{\Sigma}$, respectively, in Section 3.1.1, where we set $Y^*$ as $Y$. Then use $\hat{B}^*$ and $\hat{\Sigma}^*$ to obtain the eigenvalues $\lambda_1^*$ and $\lambda_2^*$. These eigenvalues are obtained in the same way as the eigenvalues $\lambda_1$ and $\lambda_2$ in Section 3.1.1, but where $\hat{B}^*$ and $\hat{\Sigma}^*$ are used instead of $\hat{B}$ and $\hat{\Sigma}$.

4. Perform the asymptotic and exact spanning tests using the eigenvalues obtained in the previous step (see Section 3.1.1).

5. Repeat steps 1 to 4 until we have made 10,000 simulations.

6. Count the number of times we reject $H_0$ at the 5% significance level for each spanning test, and then divide each of them by 10,000 to obtain the actual probabilities of rejecting the null-hypothesis in (13) when $H_0$ (or $H_1$) is true for the tests.

### 3.2.2 Simulate the residuals from a multivariate t-distribution

We will also use the covariance matrix $\Sigma$ simulated from a Wishart distribution and the matrix $B$ satisfying $H_0$ (or a possible case of $H_1$) in this simulation procedure. The steps of the procedure to get the actual probability of rejecting the null-hypothesis in (13) at the 5% significance level when $H_0$ (or $H_1$) is true for each test are described below:

1. Simulate a matrix $E^*$, by simulating each $\epsilon_t^*$ for $t = 1, 2, ..., T$, from a multivariate t-distribution with five degrees of freedom [3] and covariance matrix $\Sigma$.

[3] Kan and Zhou. [17] used a multivariate t-distribution with five degrees of freedom when they studied the power of the spanning tests. Hence, we simulate the residuals from a similar distribution

2. Obtain new returns of the test assets, $Y^*$, by using the formula $Y^* = XB + E^*$.

3. Compute the maximum likelihood estimators of $B$ and $\Sigma$, i.e., calculate $\hat{B}^*$ and $\hat{\Sigma}^*$, and use them to obtain the eigenvalues $\lambda_1^*$ and $\lambda_2^*$ (see Section 3.1.1, but where we use $Y^*$, $\hat{B}^*$, and $\hat{\Sigma}^*$ instead of $Y$, $\hat{B}$, and $\hat{\Sigma}$ respectively).

4. Perform the asymptotic and exact spanning tests using the eigenvalues obtained in step 3 (see Section 3.1.1).

5. The steps 1 to 4 will be repeated until we have made 10,000 simulations.

6. Count the number of times one can reject $H_0$ at the 5% significance level for each spanning test. Compute the actual probability of rejecting the null-hypothesis in (13) when $H_0$ (or $H_1$) is true for each test by dividing each count by 10,000.

#### 3.2.3 Residual bootstrap to simulate the residuals

We will use the matrix $B$, which we have forced to satisfy $H_0$ (or a possible case of $H_1$) in this bootstrap procedure. The procedure to compute the actual probabilities of rejecting the null-hypothesis in (13) at the 5% significance level when $H_0$ (or $H_1$) is true for the spanning tests are described below:

1. Use $X$ and $Y$ to fit a multivariate regression model and use the residual matrix of this model $\hat{E}$ to approximate the error matrix $E$.

2. Simulate a $T \times N$-matrix $E^*$, by drawing each row in $E^*$ from the matrix $\hat{E}$ with replacement. In other words, we draw $T$ rows from $\hat{E}$ with replacement.

3. The next step is to obtain new returns of the test assets, $Y^*$, by using the formula $Y^* = XB + E^*$.

4. In this step, we compute the maximum likelihood estimators of $B$ and $\Sigma$, and then the eigenvalues $\lambda_1^*$ and $\lambda_2^*$ by using the formulas in Section 3.1.1, but with $Y^*$ set as $Y$.

5. Perform the asymptotic and exact spanning tests using the eigenvalues obtained in the previous step (see Section 3.1.1).

6. Repeat steps 2 to 5 until we have made 10,000 simulations.

7. We can calculate the actual probabilities of rejecting the null-hypothesis in (13) when $H_0$ (or $H_1$) is true for each spanning test by first counting the number of times we can reject $H_0$ for each test at the 5% significance level. Then, we can obtain the desired probabilities by dividing each count by 10,000.

### 3.2.4 Non-overlapping block bootstrap to simulate the residuals

In this procedure, we will again use the matrix $B$, which was forced to satisfy $H_0$ (or a possible case of $H_1$). To obtain the actual probability of rejecting the null-hypothesis in (13) at the 5% significance level when $H_0$ (or $H_1$) is true for each test, where we simulate $E^*$ by using a non-overlapping block bootstrap method, we can follow the steps presented below:

1. Use $X$ and $Y$ to fit a multivariate regression model and use the residual matrix of this model $\hat{E}$ to approximate the error matrix $E$.

2. Divide the residual matrix $\hat{E}$ into twelve equally large blocks if one uses monthly returns or eleven equally large blocks if one uses weekly returns. Hence, the twelve blocks for monthly returns can be expressed as

$$\hat{E}_{\text{block1}} = \begin{pmatrix} \hat{\epsilon}_1^\top \\ \vdots \\ \hat{\epsilon}_{12}^\top \end{pmatrix}, \quad \hat{E}_{\text{block2}} = \begin{pmatrix} \hat{\epsilon}_{13}^\top \\ \vdots \\ \hat{\epsilon}_{24}^\top \end{pmatrix}, \quad \cdots, \quad \hat{E}_{\text{block12}} = \begin{pmatrix} \hat{\epsilon}_{133}^\top \\ \vdots \\ \hat{\epsilon}_{144}^\top \end{pmatrix},$$

   and the eleven blocks used for weekly returns as

$$\hat{E}_{\text{block1}} = \begin{pmatrix} \hat{\epsilon}_1^\top \\ \vdots \\ \hat{\epsilon}_{57}^\top \end{pmatrix}, \quad \hat{E}_{\text{block2}} = \begin{pmatrix} \hat{\epsilon}_{58}^\top \\ \vdots \\ \hat{\epsilon}_{114}^\top \end{pmatrix}, \quad \cdots, \quad \hat{E}_{\text{block11}} = \begin{pmatrix} \hat{\epsilon}_{571}^\top \\ \vdots \\ \hat{\epsilon}_{627}^\top \end{pmatrix}.$$

   Applying the non-overlapping block bootstrap procedure on $\hat{E}$ could probably preserve some patterns of heteroscedasticity or autocorrelation in $\hat{E}$ in the resampling procedure (if there are any). We choose to divide the blocks into twelve and eleven equally large blocks, respectively, because each block will then consists of twelve and 57 residuals, which are not very small. Too small blocks might not be able to preserve patterns of possible heteroscedasticity or autocorrelation (see chapter 2 in [19] for more details about non-overlapping block bootstrap).

3. Simulate a matrix $E^*$, by drawing twelve or eleven blocks from the blocks described in the previous step (depending on if one uses monthly or weekly returns). We will draw these blocks with replacement.

4. Obtain new returns of the test assets, $Y^*$, by using the formula $Y^* = XB + E^*$.

5. Compute the maximum likelihood estimators of $B$ and $\Sigma$ by using the expressions in Section 3.1.1, where we use $Y^*$ as $Y$ instead. By using these maximum likelihood estimators, we can obtain the eigenvalues $\lambda_1^*$ and $\lambda_2^*$ (see Section 3.1.1).

6. In this step, we perform the asymptotic and exact spanning tests using the eigenvalues, $\lambda_1^*$ and $\lambda_2^*$, obtained in step 5 (see Section 3.1.1).

7. Repeat steps 3 to 6 until we have run these steps 10,000 times in total.

8. Count the number of times one rejects $H_0$ at the 5% significance level for each mean-variance spanning test. Divide each value by 10,000 to get the actual probabilities of rejecting the null-hypothesis in (13) at the 5% significance level when $H_0$ (or $H_1$) is true for the tests.

# 4 Data description

The analysis of our study was performed with the statistical software R (version 1.1.463), where we used the following packages: tidyverse, tidyquant, rvest, httr, xml2, MVN, mvShapiroTest, and gmm.

We have used two datasets containing monthly and weekly returns, respectively, of 21 stocks in the OMXS30 index between the years 2008 to 2019. The OMXS30 index is a market-weighted price index, and it consists of the 30 most actively traded stocks on the OMX Nordic Exchange Stockholm, and these are all considered large-cap stocks on the Stockholm stock market [10, 23]. To obtain the two datasets, we loaded a data consisting of closing prices for each of the 30 stocks between December 2007 to December 2019 into R by using the tq_get() command. The closing prices obtained by using this command are fetched from Yahoo Finance [21]. Furthermore, we also noted that the prices obtained were daily prices. However, we observed that all the stocks in the OMXS30 index are not from different companies. In fact, two of these 30 stocks are from the same company. Hence the OMXS30 index consists of 29 distinct companies. We saw that Atlas Copco AB has a Class A share and a Class B share, and the difference between them is a shareholder's number of voting rights [1]. Furthermore, we collected a dataset containing the total ESG scores and the total ESG risk ratings for the companies in the OMXS30 index, and these scores and ratings were found at Yahoo Finance. We saw that the Atlas Copco AB Class B stock did not have a reported total ESG score or a total ESG risk rating in our ESG dataset, and was therefore excluded from our study. The Atlas Copco AB Class A stock, on the other hand, had a reported total ESG score and total ESG risk rating at Yahoo Finance and was hence included in the study. Furthermore, we noted that six more stocks have missing values in the total ESG score and the total ESG risk rating, and these were therefore also excluded from the study. In other words, all the assets without a reported total

ESG score and total ESG risk rating in Yahoo Finance were not included in our study. Furthermore, we saw that one of the remaining 23 stocks did not have reported daily closing prices for the whole study period (the first reported closing price happened after the year 2010), and another stock had more than 200 missing values of the daily closing prices. Hence, these two assets were also excluded from the study. Moreover, three out of the remaining 21 stocks contained one or two missing values of their daily closing prices, and we solved this issue by manually adding the correctly closing prices into the data, where the prices were found at Avanza. We then divided the dataset containing closing prices into two different datasets, one were we computed monthly closing prices for each of the 21 stocks, and one with weekly closing prices. From these two datasets, we calculated the monthly and weekly returns, respectively, for each stock between the years 2008 to 2019.

To analyze whether the efficient frontier of sustainable stocks and the efficient frontier of all the 21 stocks are statistically significantly different, we applied the screening procedure and mean-variance spanning tests. We screened the assets based on the 90th, 75th, and 50th percentile of the total ESG score respectively. All the stocks with a higher total ESG score than the 90th percentile were screened, and they were considered as the test assets in the first scenario, while the remaining stocks were set as the benchmark assets. We refer this scenario to as 10% screening throughout the thesis. Furthermore, we screened all the stocks with a total ESG score that is higher than the 75th and 50th percentile, respectively, and these stocks were set as our test assets, while the more sustainable assets (those assets that were not screened) were set as the benchmark assets. We call these two scenarios as 25% and 50% screening, respectively, in the thesis. Additionally, we considered another scenario, which we refer to as "All" in our tables. In this scenario, we performed screening on all the stocks with a total ESG risk rating of medium to severe. These stocks were then set as the test assets, while the stocks with a total ESG risk rating of negligible to low were considered as the benchmark assets. The asymptotic and exact mean-variance spanning tests (see Section 3.1.1 and 3.1.2) were then used for each screening scenario to see if the difference between the efficient frontier of the benchmark assets and the efficient frontier of the benchmark assets and the test assets is statistically significant. Using notations from Section 3, we denote the number of test assets as $N$, the number of benchmark assets as $K$, and the length of the time series as $T$. The spanning tests verify $H_0$ (the null-hypothesis of spanning) by testing if the weights of the test assets are all equal to zero in the MVP and in the TP. Moreover, $H_0$ is tested against $H_1$ (the alternative-hypothesis of no spanning).

Furthermore, we checked for the robustness for non-normality thus the spanning tests assume normality in the residuals. Hence, we used a similar setup as in Kan and Zhou [17], where we drew disturbances from a multivariate t-distribution with five degrees of freedom to simulate new returns of the test assets. We simulated 10,000 samples in total by using the simulation procedure presented in Section 3.2.2 to obtain the actual probabilities of rejecting the null-hypothesis in (13) when $H_0$ and $H_1$ is true, respectively, for each spanning test. Likewise, we used the simulation procedure in Section 3.2.1 to obtain the probabilities of rejecting the null-hypothesis in (13) when $H_0$ and $H_1$ is true, respectively, when the normality assumption holds. We then compared these probabilities under the normality assumption with the probabilities obtained using multivariate t-distributed disturbances. We also performed two resampling methods, the residual bootstrap and the non-overlapping block bootstrap (see Section 3.2.3 and 3.2.4), on the residuals to obtain the probabilities of rejecting the null-hypothesis in (13) when $H_0$ and $H_1$ is true, respectively, when the residuals are probably not i.i.d. Then, we studied the differences between the obtained probabilities based on these two procedures and the probabilities obtained under the normality assumption. Furthermore, multivariate normality tests were performed using the quantile-quantile plot (QQ-plot) based on both the chi-squared quantiles and the squared Mahalanobis distance, the Mardia's test, and the generalized Shapiro-Wilk's test, where the latter two are purposed by Mardia [20], and Villaseñor and González-Estrada [34] respectively.

# 5 Results of the empirical illustration and simulation study

The results of the asymptotic mean-variance spanning tests where we used monthly returns and weekly returns are summarized in Table 1 and Table 3 respectively. Furthermore, the results of the exact tests, which were obtained based on monthly returns and weekly returns, can be seen in Table 2 and Table 4 respectively. All the spanning tests that we performed did not show statistically significant difference between two efficient frontiers when we applied 10%, 25%, and 50% screening respectively (at least at the 5% significance level). The results of the six spanning tests when monthly and weekly returns were used, respectively, and when we screened for all the unsustainable stocks (stocks with a total ESG risk rating between medium to severe), showed that we also could not reject $H_0$, that the nine sustainable assets span the 21 assets, at all conventional significance levels. This means that the difference between the minimum-variance frontier of all the 21 assets and each of the minimum-variance frontiers of a smaller number of assets that are considered more sustainable are statistically insignificant. Hence, the differences between the efficient frontiers are not statistically significant as well. An investor that invests in the more sustainable assets might, therefore, not improve the investment opportunity if she also invests in the more unsustainable assets.

Table 1: The tests statistics of the asymptotic spanning tests and their corresponding $p$-values based on the monthly returns of the assets. "LR - test", "W - test", and "LM -test" denote the likelihood ratio test, the Wald test, and the Lagrange multiplier test respectively.

| | | | | **Asymptotic tests** | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | **LR - test** | | **W - test** | | **LM - test** | |
| **Screening** | **N** | **K** | **T** | **LR** | **$p$-value** | **W** | **$p$-value** | **LM** | **$p$-value** |
| 10% | 2 | 19 | 144 | 1.183 | 0.881 | 1.187 | 0.880 | 1.178 | 0.882 |
| 25% | 5 | 16 | 144 | 8.008 | 0.628 | 8.150 | 0.614 | 7.869 | 0.642 |
| 50% | 10 | 11 | 144 | 27.723 | 0.116 | 29.117 | 0.085 | 26.419 | 0.152 |
| All | 12 | 9 | 144 | 30.711 | 0.162 | 32.425 | 0.117 | 29.116 | 0.216 |

Table 2: The test statistics of the three exact spanning tests and their corresponding $p$-values obtained by using the monthly returns of the assets, where "LR - test", "W - test", and "LM -test" denote the likelihood ratio test, the Wald test, and the Lagrange multiplier test respectively.

| | | | | **Exact tests** | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | **LR - test** | | **W - test** | | **LM - test** | |
| **Screening** | **N** | **K** | **T** | **LR** | **$p$-value** | **W** | **$p$-value** | **LM** | **$p$-value** |
| 10% | 2 | 19 | 144 | 0.253 | 0.908 | 0.008 | 0.908 | 0.008 | 0.908 |
| 25% | 5 | 16 | 144 | 0.694 | 0.730 | 0.057 | 0.731 | 0.055 | 0.729 |
| 50% | 10 | 11 | 144 | 1.243 | 0.220 | 0.202 | 0.228 | 0.183 | 0.212 |
| All | 12 | 9 | 144 | 1.153 | 0.287 | 0.225 | 0.296 | 0.202 | 0.277 |

Table 3: The test statistics and the corresponding $p$-values of the three asymptotic spanning tests based on the data with weekly returns, where "LR - test", "W - test", and "LM -test" represent the likelihood ratio test, the Wald test, and the Lagrange multiplier test respectively.

| | | | | **Asymptotic tests** | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | **LR - test** | | **W - test** | | **LM - test** | |
| **Screening** | **N** | **K** | **T** | **LR** | **p-value** | **W** | **p-value** | **LM** | **p-value** |
| 10% | 2 | 19 | 627 | 3.872 | 0.424 | 3.883 | 0.422 | 3.862 | 0.425 |
| 25% | 5 | 16 | 627 | 6.596 | 0.763 | 6.625 | 0.760 | 6.566 | 0.766 |
| 50% | 10 | 11 | 627 | 28.533 | 0.097 | 28.939 | 0.089 | 28.135 | 0.106 |
| All | 12 | 9 | 627 | 31.903 | 0.129 | 32.403 | 0.117 | 31.413 | 0.142 |

Table 4: The test statistics and the corresponding $p$-values of the three exact spanning tests based on the data with weekly returns, where "LR - test", "W - test", and "LM -test" represent the likelihood ratio test, the Wald test, and the Lagrange multiplier test respectively.

| | | | | **Exact tests** | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | **LR - test** | | **W - test** | | **LM - test** | |
| **Screening** | **N** | **K** | **T** | **LR** | **p-value** | **W** | **p-value** | **LM** | **p-value** |
| 10% | 2 | 19 | 627 | 0.937 | 0.441 | 0.006 | 0.441 | 0.006 | 0.440 |
| 25% | 5 | 16 | 627 | 0.639 | 0.781 | 0.011 | 0.780 | 0.010 | 0.782 |
| 50% | 10 | 11 | 627 | 1.395 | 0.115 | 0.046 | 0.114 | 0.045 | 0.115 |
| All | 12 | 9 | 627 | 1.301 | 0.151 | 0.052 | 0.150 | 0.050 | 0.151 |

We saw that $H_0$ can be rejected at the 5% significance level for all six spanning tests when monthly and weekly returns were used respectively. Hence, it seems like the asymptotic spanning tests and the exact tests might give us quite similar results when we use the 5% significance level, even though we have a fairly small data sample of size 144. However, we see from Tables 1 and 3 that the asymptotic W test shows that the differences between the efficient frontiers based on 50% screening using monthly and weekly returns, respectively, are statistically significant at the 10% significance level. Furthermore, the asymptotic LR test using weekly returns does also reject the hypothesis of spanning at the 10% significance level. The $p$-values of the other spanning tests, on the other hand, are all higher than all conventional significance levels and suggest that we can not reject $H_0$. Hence, the asymptotic tests and the exact tests might give slightly different results and one should therefore probably opt to use the exact tests instead of the asymptotic tests, but it might be interesting to compare the different spanning tests further. We performed the six spanning tests on assets of different dimensions and different sample sizes. To do this, we started by excluding the two assets with the highest total ESG scores, which will be our test assets. We then chose the first two assets from the remaining 19 assets as our benchmark assets. Then, we performed the spanning tests with the data consisting of these two benchmark assets and the two test assets over three different periods. We continued this procedure by adding more assets to our test assets and benchmark assets. These analyses were performed on the data with monthly returns and weekly returns respectively. The results of the spanning tests where we used monthly returns can be found in Table 5 and the results based on weekly returns are summarized in Table 6.

We can see from Tables 5 and 6 that all the exact spanning tests using monthly and weekly returns, respectively, do not reject $H_0$ at the 5% significance level, because all the $p$-values from the exact tests are much greater than 0.05, which is the significance level used in this thesis. Furthermore, we see from the tables that some $p$-values from the asymptotic W and LR tests are smaller than 0.05, which means that the difference between the efficient frontiers is statistically significant at the 5% significance level. Generally, the three exact tests give similar $p$-values, while the $p$-values obtained from the asymptotic tests can differ from each other and the exact tests. We see that the $p$-values obtained by the exact tests are usually higher than the $p$-values obtained from the asymptotic tests. For example, from Table 5, we can see that the $p$-values of the three exact tests are around 0.2 when $N = 10$, $K = 2$, and $T = 50$ using monthly returns, which is more than twice as high as the $p$-value of the asymptotic LR test, and approximately ten times as high as the $p$-value of the asymptotic W test. We can also see from Table 6 that the $p$-values of the three exact tests are around 0.16 when $N = 2$, $K = 15$, and $T = 50$ based on weekly returns. This is almost four times as high as the $p$-value of the asymptotic LR test, more than five times as high as the $p$-value of the asymptotic W test, and more than twice as high as the $p$-value of the asymptotic LM test. Hence, one might obtain misleading results when one applies the asymptotic spanning tests, especially when $T$ is not large. Moreover, we can also see that the asymptotic LM test out of the three asymptotic tests seems to give the closest $p$-value to the exact tests from the tables (Tables 5 and 6). One should, therefore, probably opt to apply the asymptotic LM test if one insists on using an asymptotic test to verify $H_0$.

Furthermore, we should note that the residuals were assumed to be normally distributed in the regression model we used to compute the spanning tests. The normality assumption was tested by using the generalized Shapiro-Wilk's test, Mardia's test, and graphically illustrated by chi-squared QQ-plots. The results from the first two tests are summarized in Table 7, and the QQ-plots are shown in Figures 3 and 4. We saw that the hypothesis of normality could be rejected in all four cases (10%, 25%, 50% and All screening) using monthly returns as well as in the four cases using weekly returns from the results of the generalized Shapiro-Wilk's test, Mardia's test and the QQ-plots. The QQ-plots in Figures 3 and 4 show that the squared Mahalanobis distance of the residuals do not fully follow the straight line, which means that the residuals are not normally distributed. Therefore, the normality assumption is likely to be violated. We could also see that the residuals of the monthly returns might be closer to being normally distributed compared to those from the weekly returns, which makes it interesting to make further analysis on both monthly and weekly returns respectively.

Table 5: The $p$-values of the asymptotic and exact spanning tests using monthly returns for different values of $T$, $K$, and $N$.

| | | | Asymptotic tests | | | Exact tests | | |
|---|---|---|---|---|---|---|---|---|
| **N** | **K** | **T** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| 2 | 2 | 50 | 0.113 | 0.091 | 0.138 | 0.139 | 0.136 | 0.143 |
| | | 100 | 0.059 | 0.050 | 0.069 | 0.067 | 0.066 | 0.069 |
| | | 144 | 0.164 | 0.155 | 0.173 | 0.174 | 0.172 | 0.175 |
| | 5 | 50 | 0.184 | 0.161 | 0.209 | 0.249 | 0.247 | 0.251 |
| | | 100 | 0.145 | 0.132 | 0.158 | 0.172 | 0.170 | 0.174 |
| | | 144 | 0.322 | 0.314 | 0.330 | 0.347 | 0.346 | 0.348 |
| | 10 | 50 | 0.744 | 0.738 | 0.750 | 0.826 | 0.826 | 0.826 |
| | | 100 | 0.617 | 0.611 | 0.623 | 0.672 | 0.671 | 0.673 |
| | | 144 | 0.558 | 0.555 | 0.562 | 0.599 | 0.600 | 0.598 |
| | 15 | 50 | 0.730 | 0.723 | 0.737 | 0.851 | 0.850 | 0.852 |
| | | 100 | 0.460 | 0.451 | 0.468 | 0.554 | 0.553 | 0.554 |
| | | 144 | 0.798 | 0.797 | 0.798 | 0.832 | 0.832 | 0.831 |
| | 19 | 50 | 0.575 | 0.561 | 0.588 | 0.789 | 0.788 | 0.790 |
| | | 100 | 0.932 | 0.932 | 0.932 | 0.954 | 0.955 | 0.954 |
| | | 144 | 0.881 | 0.880 | 0.882 | 0.908 | 0.908 | 0.908 |
| 5 | 2 | 50 | 0.109 | 0.058 | 0.177 | 0.168 | 0.157 | 0.181 |
| | | 100 | 0.063 | 0.039 | 0.094 | 0.082 | 0.073 | 0.092 |
| | | 144 | 0.230 | 0.209 | 0.252 | 0.257 | 0.257 | 0.257 |
| | 5 | 50 | 0.123 | 0.074 | 0.184 | 0.235 | 0.232 | 0.238 |
| | | 100 | 0.079 | 0.053 | 0.110 | 0.116 | 0.108 | 0.125 |
| | | 144 | 0.400 | 0.382 | 0.418 | 0.450 | 0.454 | 0.447 |
| | 10 | 50 | 0.593 | 0.548 | 0.635 | 0.800 | 0.801 | 0.798 |
| | | 100 | 0.207 | 0.178 | 0.238 | 0.315 | 0.316 | 0.314 |
| | | 144 | 0.322 | 0.304 | 0.340 | 0.403 | 0.407 | 0.399 |
| | 15 | 50 | 0.736 | 0.712 | 0.759 | 0.927 | 0.929 | 0.925 |
| | | 100 | 0.435 | 0.397 | 0.473 | 0.604 | 0.597 | 0.613 |
| | | 144 | 0.651 | 0.638 | 0.664 | 0.744 | 0.745 | 0.744 |
| 10 | 2 | 50 | 0.083 | 0.026 | 0.191 | 0.212 | 0.234 | 0.189 |
| | | 100 | 0.362 | 0.275 | 0.450 | 0.459 | 0.449 | 0.470 |
| | | 144 | 0.283 | 0.239 | 0.329 | 0.346 | 0.354 | 0.337 |
| | 5 | 50 | 0.141 | 0.055 | 0.273 | 0.394 | 0.412 | 0.373 |
| | | 100 | 0.567 | 0.494 | 0.634 | 0.693 | 0.689 | 0.697 |
| | | 144 | 0.316 | 0.271 | 0.363 | 0.409 | 0.416 | 0.401 |
| | 10 | 50 | 0.306 | 0.151 | 0.483 | 0.745 | 0.730 | 0.759 |
| | | 100 | 0.194 | 0.128 | 0.271 | 0.381 | 0.379 | 0.384 |
| | | 144 | 0.123 | 0.090 | 0.162 | 0.222 | 0.227 | 0.216 |

Table 6: The $p$-values of the asymptotic and exact spanning tests using weekly returns for different values of $T$, $K$, and $N$.

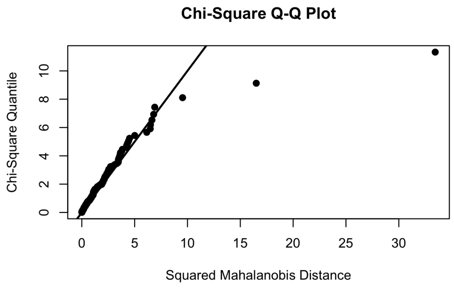
| | | | Asymptotic tests | | | Exact tests | | |
|---|---|---|---|---|---|---|---|---|
| **N** | **K** | **T** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| 2 | 2 | 50 | 0.168 | 0.147 | 0.190 | 0.199 | 0.200 | 0.199 |
| | | 100 | 0.567 | 0.563 | 0.571 | 0.585 | 0.587 | 0.583 |
| | | 627 | 0.625 | 0.622 | 0.627 | 0.636 | 0.637 | 0.635 |
| | 5 | 50 | 0.131 | 0.117 | 0.145 | 0.187 | 0.195 | 0.178 |
| | | 100 | 0.729 | 0.725 | 0.732 | 0.753 | 0.753 | 0.754 |
| | | 627 | 0.644 | 0.641 | 0.647 | 0.664 | 0.664 | 0.664 |
| | 10 | 50 | 0.134 | 0.114 | 0.155 | 0.247 | 0.249 | 0.245 |
| | | 100 | 0.638 | 0.633 | 0.643 | 0.691 | 0.690 | 0.691 |
| | | 627 | 0.708 | 0.706 | 0.711 | 0.740 | 0.740 | 0.740 |
| | 15 | 50 | 0.043 | 0.031 | 0.059 | 0.160 | 0.160 | 0.159 |
| | | 100 | 0.506 | 0.498 | 0.514 | 0.597 | 0.596 | 0.598 |
| | | 627 | 0.629 | 0.625 | 0.633 | 0.682 | 0.682 | 0.683 |
| | 19 | 50 | 0.050 | 0.038 | 0.064 | 0.231 | 0.237 | 0.225 |
| | | 100 | 0.329 | 0.317 | 0.341 | 0.452 | 0.450 | 0.454 |
| | | 627 | 0.556 | 0.550 | 0.561 | 0.630 | 0.629 | 0.630 |
| 5 | 2 | 50 | 0.416 | 0.346 | 0.485 | 0.509 | 0.500 | 0.519 |
| | | 100 | 0.940 | 0.938 | 0.942 | 0.950 | 0.950 | 0.949 |
| | | 627 | 0.934 | 0.932 | 0.936 | 0.941 | 0.942 | 0.941 |
| | 5 | 50 | 0.400 | 0.344 | 0.456 | 0.553 | 0.557 | 0.548 |
| | | 100 | 0.903 | 0.899 | 0.906 | 0.926 | 0.926 | 0.925 |
| | | 627 | 0.862 | 0.857 | 0.866 | 0.883 | 0.884 | 0.883 |
| | 10 | 50 | 0.203 | 0.147 | 0.265 | 0.450 | 0.453 | 0.447 |
| | | 100 | 0.937 | 0.934 | 0.940 | 0.961 | 0.961 | 0.961 |
| | | 627 | 0.898 | 0.895 | 0.901 | 0.925 | 0.925 | 0.924 |
| | 15 | 50 | 0.115 | 0.078 | 0.159 | 0.449 | 0.468 | 0.427 |
| | | 100 | 0.789 | 0.772 | 0.805 | 0.880 | 0.877 | 0.883 |
| | | 627 | 0.884 | 0.879 | 0.889 | 0.924 | 0.923 | 0.924 |
| 10 | 2 | 50 | 0.222 | 0.115 | 0.357 | 0.414 | 0.439 | 0.386 |
| | | 100 | 0.451 | 0.366 | 0.533 | 0.549 | 0.539 | 0.558 |
| | | 627 | 0.809 | 0.784 | 0.832 | 0.848 | 0.847 | 0.849 |
| | 5 | 50 | 0.288 | 0.167 | 0.427 | 0.581 | 0.601 | 0.558 |
| | | 100 | 0.463 | 0.375 | 0.547 | 0.599 | 0.587 | 0.612 |
| | | 627 | 0.795 | 0.768 | 0.820 | 0.851 | 0.850 | 0.852 |
| | 10 | 50 | 0.272 | 0.150 | 0.415 | 0.716 | 0.729 | 0.699 |
| | | 100 | 0.399 | 0.306 | 0.493 | 0.607 | 0.591 | 0.624 |
| | | 627 | 0.461 | 0.404 | 0.517 | 0.601 | 0.596 | 0.607 |

Table 7: Results of the generalized Shapiro-Wilk's test and Mardia's test, which tests for normality. $T = 144$ indicates that the results are based on the monthly returns, and $T = 627$ are based on the weekly returns.

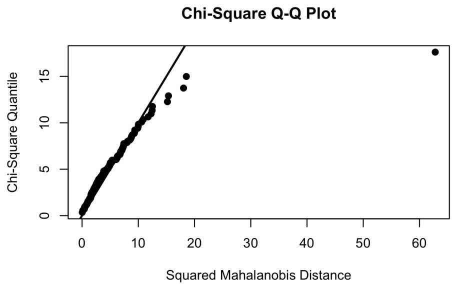
| | | **Geneneralized Shapiro-Wilk's test** | | **Mardia's test** | | | |
|---|---|---|---|---|---|---|---|
| **Screening** | **T** | **Test statistic** | **p-value** | **Test statistic for skewness** | **p-value** | **Test statistic for kurtosis** | **p-value** |
| 10% | 144 | 0.96 | $< 1 \cdot 10^{-6}$ | 41.82 | $< 2 \cdot 10^{-8}$ | 11.04 | 0.000 |
| | 627 | 0.95 | $< 3 \cdot 10^{-16}$ | 39.47 | $< 6 \cdot 10^{-8}$ | 23.36 | 0.000 |
| 25% | 144 | 0.96 | $< 7 \cdot 10^{-12}$ | 327.26 | $< 4 \cdot 10^{-49}$ | 18.78 | 0.000 |
| | 627 | 0.95 | $< 3 \cdot 10^{-16}$ | 151.13 | $< 3 \cdot 10^{-16}$ | 41.64 | 0.000 |
| 50% | 144 | 0.97 | $< 3 \cdot 10^{-14}$ | 668.98 | $< 8 \cdot 10^{-47}$ | 13.95 | 0.000 |
| | 627 | 0.94 | $< 3 \cdot 10^{-16}$ | 2798.81 | 0.000 | 112.21 | 0.000 |
| All | 144 | 0.98 | $< 6 \cdot 10^{-12}$ | 836.34 | $< 4 \cdot 10^{-39}$ | 13.50 | 0.000 |
| | 627 | 0.95 | $< 3 \cdot 10^{-16}$ | 3559.78 | 0.000 | 118.75 | 0.000 |

When the normality assumption of the residuals does not hold, one might want to study the robustness to non-normality for the spanning tests. This was done by simulating new monthly and weekly returns of the test assets, respectively, by drawing disturbances from a multivariate t-distribution with five degrees of freedom (see Section 3.2.2 for more details). We then computed the actual probability of rejecting the null-hypothesis in (13) under $H_0$ and $H_1$, respectively, at the 5% significance level for each spanning test. Likewise, we calculated the actual probability of rejecting the null-hypothesis in (13) under $H_0$ and $H_1$, respectively, at the 5% significance level for each test when the disturbances were drawn from a multivariate normal distribution (see Section 3.2.1). We then compared the probabilities of rejection obtained using multivariate t-distributed disturbances with those obtained using multivariate normally distributed disturbances. The obtained probabilities of rejecting the null-hypothesis in (13) when $H_0$ is true, based on multivariate normally distributed residuals and multivariate t-distributed disturbances, respectively, can be found in Table 8, and those when $H_1$ is true (with different possibilities as $H_1$) are shown in Table 9.

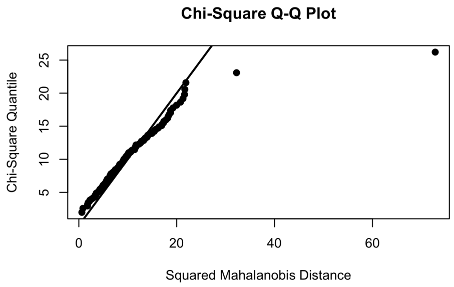
Figure 3: Chi-squared QQ-plots of the residuals from the regression models where monthly returns were used.



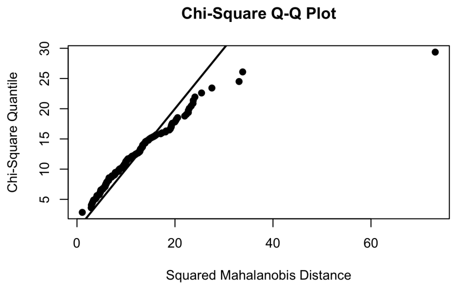

(a) 10% screening




(b) 25% screening

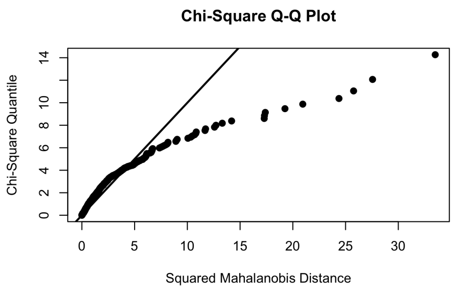



(c) 50% screening.

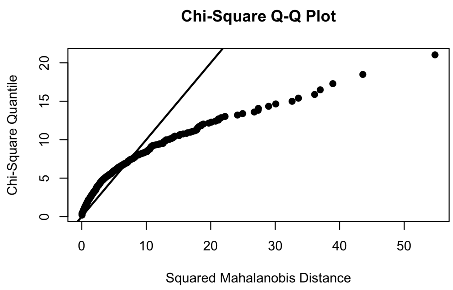



(d) All screening

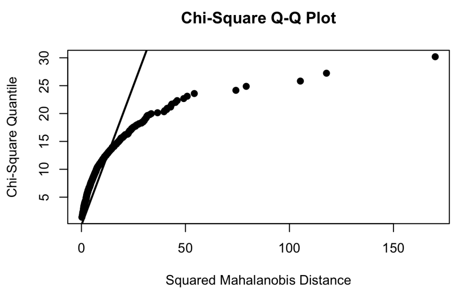
Figure 4: Chi-squared QQ-plots of the residuals from the regression models where we used weekly returns.



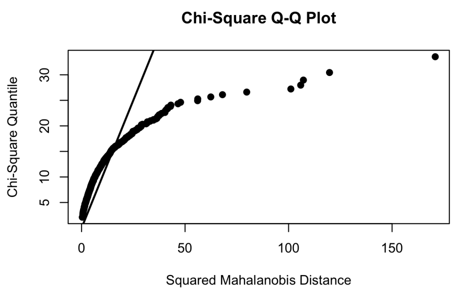

(a) 10% screening




(b) 25% screening




(c) 50% screening.




(d) All screening

When $H_0$ was set to be true, we can see from Table 8 that the actual probability of rejecting $H_0$ for each test based on multivariate t-distributed disturbances are similar to the corresponding probability obtained by using multivariate normally distributed residuals. We also noted that the probability of rejecting $H_0$ for each of the three exact tests under normality gives a value of around 0.05, which is the expected probability, thus we used the 5% significance level. Furthermore, the probabilities of rejecting $H_0$ under $H_0$ using the three exact tests when the disturbances were multivariate t-distributed gave us similar results. Hence, the probability of getting a type I error using the exact tests seems to be very similar to the expected value even if the residuals are i.i.d., but not normally distributed. However, the asymptotic spanning tests seem to give us a higher probability of rejecting $H_0$ than the expected, which can clearly be seen in the probabilities obtained based on monthly returns, where $T$ is fairly small. Nevertheless, this can even be seen for some cases when we used weekly returns, where $T = 627$. Sometimes we obtain a probability of rejection that is two or three times higher than the expected based on the asymptotic tests, e.g., when we performed 50% screening based on the monthly returns and i.i.d. normally distributed residuals, we can see that the probability of rejecting the null-hypothesis in (13) under $H_0$ is 0.116 for the asymptotic LR test, and 0.165 for the asymptotic W test. However, we also see that the actual probabilities of rejecting the null-hypothesis in (13) under $H_0$ are generally closer to the expected value for the three asymptotic tests when we had a greater value of $T$, by using weekly returns ($T = 627$) instead of monthly returns ($T = 144$).

Table 8: The probabilities of rejecting the null-hypothesis in (13) when $H_0$ is true for the spanning tests, and under the assumption of normally distributed and t-distributed disturbances, respectively, using 10,000 simulations. $T = 144$ indicates that the results are based on monthly returns and $T = 627$ are based on weekly returns.

| | | | | **Multivariate normal distribution** | | | | | | **Multivariate t-distribution** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | **Asymptotic test** | | | **Exact test** | | | **Asymptotic test** | | | **Exact test** | | |
| **Screening** | **N** | **K** | **T** | **LR** | **W** | **LM** | **LR** | **W** | **LM** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| 10% | 2 | 19 | 144 | 0.085 | 0.092 | 0.077 | 0.047 | 0.047 | 0.047 | 0.087 | 0.095 | 0.077 | 0.051 | 0.051 | 0.051 |
| | | | 627 | 0.055 | 0.057 | 0.053 | 0.048 | 0.048 | 0.049 | 0.060 | 0.061 | 0.058 | 0.053 | 0.053 | 0.053 |
| 25% | 5 | 16 | 144 | 0.104 | 0.125 | 0.083 | 0.050 | 0.051 | 0.050 | 0.106 | 0.126 | 0.086 | 0.051 | 0.051 | 0.051 |
| | | | 627 | 0.058 | 0.061 | 0.055 | 0.049 | 0.049 | 0.049 | 0.066 | 0.069 | 0.063 | 0.055 | 0.056 | 0.055 |
| 50% | 10 | 11 | 144 | 0.116 | 0.165 | 0.075 | 0.050 | 0.050 | 0.051 | 0.119 | 0.168 | 0.076 | 0.050 | 0.050 | 0.050 |
| | | | 627 | 0.064 | 0.070 | 0.056 | 0.051 | 0.051 | 0.051 | 0.064 | 0.070 | 0.058 | 0.054 | 0.054 | 0.054 |
| All | 12 | 9 | 144 | 0.119 | 0.175 | 0.068 | 0.050 | 0.051 | 0.051 | 0.116 | 0.179 | 0.069 | 0.051 | 0.051 | 0.052 |
| | | | 627 | 0.064 | 0.073 | 0.055 | 0.050 | 0.050 | 0.050 | 0.065 | 0.074 | 0.058 | 0.054 | 0.054 | 0.054 |

Table 9 shows the probabilities of rejecting $H_0$ when $H_1$ was set to be true, and when the disturbances were multivariate normally distributed and multivariate t-distributed, respectively, for four different possibilities as $H_1$. We studied ten different possibilities as $H_1$, which are described in Section 3.2. However, we only showed the results for four of them (cases 3, 6, 7, and 10) in Table 9. The results of the other cases (cases 1, 2, 4, 5, 8, and 9) can be found in Tables 13 and 14 in Appendix A.3. It might be worth mentioning that those cases that are the most similar to $H_0$ (cases 1, 4, and 8) gave us actual probabilities of rejecting $H_0$ that were very similar to those found in Table 8 (when we performed the simulations under $H_0$). Under cases 1, 4, and 8, the actual probability of rejecting $H_0$ for each test when the disturbances are multivariate t-distributed are almost identical to those obtained using multivariate normally distributed disturbances. Furthermore, cases 2, 5, and 9 gave us similar probabilities of rejection to those obtained under cases 1, 4, and 8 respectively (with an exception for cases 2 and 9 when 50% screening was performed based on monthly and weekly returns, respectively, and an explanation to this can be due to the chosen parameters $B$ and $\Sigma$, and the sample size used). The probability of rejection for each spanning test under cases 2, 5, and 9, respectively, when the disturbances are multivariate t-distributed are generally very similar to the corresponding probabilities when the disturbances are multivariate normally distributed (note that the probabilities obtained based on 50% screening are higher when the disturbances are normally distributed compared to when they are multivariate t-distributed).

From Table 9, we see that the probabilities of rejecting $H_0$ under case 6 as $H_1$ when the residuals follow a multivariate normally distribution are very similar to the corresponding probabilities obtained using multivariate t-distributed disturbances. However, we noted that the probability of rejecting $H_0$ under cases 3, 7, and 10 as $H_1$, respectively, for each test might differ fairly much when the disturbances are multivariate t-distributed compared to when they were simulated from a multivariate normal distribution. Overall, under cases 3, 7, and 10, the probabilities of rejection obtained based on multivariate t-distributed disturbances were generally lower than those obtained based on multivariate normally distributed disturbances (note that the probabilities of rejecting $H_0$ based on 50% screening under normality and when the residuals are multivariate t-distributed, respectively, gave us the same values). It also seems like cases 3 and 10 give similar probabilities of rejection, and the probabilities of rejecting $H_0$ obtained under case 7 are closer to the corresponding probabilities under the cases 3 and 10 compared to those obtained under case 6. We can also see that the $\alpha$-vector seems to be the main crucial factor that affects the probabilities of rejecting $H_0$ for the spanning tests. When all the elements of $\alpha$ are equal to 0.1 (cases 3 and 10), we can see that the probability of rejecting $H_0$ for each spanning test does not differ much for different vectors of $1_K^\top \beta^\top$. Moreover, when the $\alpha$-vector consists of solely zeros (case 6 and 7), the probability of rejecting $H_0$ for each test becomes high when $1_K^\top \beta^\top$ is a vector consisting of larger values than 0.1 (the elements in $1_K^\top \beta^\top$ in case 6 are all set to 0.1, while these are all set to 1 in case 7). In other words, the weights of the two tangency portfolios might mainly determine the power of the spanning tests.

Table 9: The probabilities of rejecting the null-hypothesis in (13) when $H_1$ is true for the tests, under the assumption of having normally distributed and t-distributed disturbances, respectively, which were obtained based on 10,000 simulations. $T = 144$ represents that the results were obtained by using monthly returns, and $T = 627$ by using weekly returns.

| | | | | | **Multivariate normal distribution** | | | | | | **Multivariate t-distribution** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | **Asymptotic test** | | | **Exact test** | | | **Asymptotic test** | | | **Exact test** | | |
| **Case** | **Screening** | **N** | **K** | **T** | **LR** | **W** | **LM** | **LR** | **W** | **LM** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| 3 | 10% | 2 | 19 | 144 | 0.423 | 0.441 | 0.406 | 0.327 | 0.328 | 0.328 | 0.302 | 0.318 | 0.286 | 0.218 | 0.217 | 0.218 |
| | | | | 627 | 0.981 | 0.981 | 0.981 | 0.979 | 0.979 | 0.979 | 0.857 | 0.861 | 0.855 | 0.845 | 0.845 | 0.846 |
| | 25% | 5 | 16 | 144 | 0.720 | 0.755 | 0.682 | 0.592 | 0.597 | 0.587 | 0.520 | 0.560 | 0.478 | 0.382 | 0.385 | 0.380 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.996 | 0.997 | 0.996 | 0.996 | 0.996 | 0.996 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 0.430 | 0.538 | 0.323 | 0.278 | 0.283 | 0.273 | 0.316 | 0.412 | 0.220 | 0.183 | 0.186 | 0.181 |
| | | | | 627 | 0.984 | 0.987 | 0.981 | 0.980 | 0.981 | 0.980 | 0.853 | 0.867 | 0.836 | 0.830 | 0.833 | 0.827 |
| 6 | 10% | 2 | 19 | 144 | 0.087 | 0.096 | 0.081 | 0.050 | 0.050 | 0.050 | 0.089 | 0.096 | 0.081 | 0.051 | 0.051 | 0.051 |
| | | | | 627 | 0.079 | 0.081 | 0.078 | 0.070 | 0.070 | 0.071 | 0.075 | 0.076 | 0.073 | 0.066 | 0.066 | 0.067 |
| | 25% | 5 | 16 | 144 | 0.111 | 0.133 | 0.091 | 0.055 | 0.056 | 0.056 | 0.111 | 0.132 | 0.089 | 0.055 | 0.054 | 0.055 |
| | | | | 627 | 0.099 | 0.103 | 0.094 | 0.084 | 0.084 | 0.084 | 0.085 | 0.090 | 0.081 | 0.073 | 0.073 | 0.073 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 0.124 | 0.180 | 0.071 | 0.052 | 0.052 | 0.052 | 0.120 | 0.181 | 0.070 | 0.053 | 0.051 | 0.052 |
| | | | | 627 | 0.081 | 0.090 | 0.072 | 0.066 | 0.066 | 0.066 | 0.075 | 0.085 | 0.066 | 0.063 | 0.063 | 0.062 |
| 7 | 10% | 2 | 19 | 144 | 0.549 | 0.567 | 0.530 | 0.443 | 0.444 | 0.442 | 0.383 | 0.400 | 0.364 | 0.287 | 0.287 | 0.286 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.990 | 0.990 | 0.989 | 0.988 | 0.988 | 0.988 |
| | 25% | 5 | 16 | 144 | 0.874 | 0.893 | 0.848 | 0.791 | 0.794 | 0.788 | 0.674 | 0.708 | 0.632 | 0.537 | 0.541 | 0.533 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 0.556 | 0.657 | 0.429 | 0.386 | 0.390 | 0.379 | 0.400 | 0.500 | 0.291 | 0.247 | 0.250 | 0.241 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.995 | 0.996 | 0.994 | 0.994 | 0.994 | 0.993 |
| 10 | 10% | 2 | 19 | 144 | 0.414 | 0.430 | 0.397 | 0.319 | 0.319 | 0.320 | 0.296 | 0.309 | 0.278 | 0.211 | 0.212 | 0.212 |
| | | | | 627 | 0.978 | 0.979 | 0.977 | 0.974 | 0.974 | 0.974 | 0.843 | 0.846 | 0.841 | 0.831 | 0.832 | 0.832 |
| | 25% | 5 | 16 | 144 | 0.706 | 0.741 | 0.663 | 0.577 | 0.584 | 0.571 | 0.504 | 0.546 | 0.463 | 0.371 | 0.373 | 0.368 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.995 | 0.995 | 0.994 | 0.993 | 0.993 | 0.993 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 0.419 | 0.525 | 0.313 | 0.268 | 0.270 | 0.264 | 0.310 | 0.403 | 0.215 | 0.179 | 0.181 | 0.178 |
| | | | | 627 | 0.979 | 0.983 | 0.976 | 0.974 | 0.975 | 0.974 | 0.833 | 0.849 | 0.817 | 0.811 | 0.814 | 0.809 |

The analysis of the robustness to the assumption of normality made so far has been based on the assumption that the residuals are identically distributed, which might not be satisfied in our data. Hence, we study the actual probabilities of rejecting $H_0$ at the 5% significance level under $H_0$ and $H_1$, respectively, where we simulate new returns (monthly and weekly respectively) of the test assets by applying the residual bootstrap procedure on the residuals (see Section 3.2.3 for more details). In Table 10, the actual probabilities of rejecting the null-hypothesis in (13) under $H_0$ based on the bootstrapped residuals are presented, and these values are quite similar to the actual probabilities we obtained when the residuals were normally distributed, which can be found in Table 8. We can furthermore see from Table 10 that the actual probabilities of rejecting the null-hypothesis in (13) when $H_0$ is true are the lowest for the exact tests. The probabilities of rejection for the exact tests are close to the expected value 0.05. Moreover, the probabilities of rejection for the asymptotic tests exceed the expected value of 0.05 fairly much when we used the monthly returns, while it is closer to 0.05 when we used the weekly returns. The reason might be that we have a larger sample size when we use the dataset consisting of weekly returns compared to when we use monthly returns.

Nonetheless, we obtained the actual probabilities of rejecting the null-hypothesis in (13) under $H_1$ based on the residual bootstrap procedure, where we studied the same ten cases as before as our $H_1$. The probabilities of rejecting $H_0$ using the bootstrap procedure under the cases 3, 6, 7, and 10 are shown in Table 11, while those under the cases 1, 2, 4, 5, 8, and 9 can be found in Tables 15 and 16, respectively, in Appendix A.3. We saw that the probabilities of rejecting $H_0$ under cases 1, 4, 5, and 8, respectively, based on the bootstrap procedure were fairly similar to the corresponding probabilities obtained when the residuals were i.i.d. multivariate normally distributed. Furthermore, the probabilities obtained under cases 2 and 9, respectively, based on the residual bootstrap simulation was generally higher than those obtained when the disturbances were normally distributed. We could also see that the actual probabilities of rejecting $H_0$ under cases 3, 7, and 10, respectively, gave us probabilities equal to one, which are higher or equal to the corresponding probabilities when the residuals follow a multivariate normal distribution. Furthermore, case 6 gave us values close to one, which are also much higher or equal to the corresponding probabilities when the normality assumption holds. Hence, we observed that the probabilities of rejecting $H_0$ at the 5% significance level when $H_1$ is true might differ from what we found when the residuals follows a multivariate normal distribution (see the probabilities of rejection under normality and $H_1$ in Table 9). However, the probabilities of rejecting $H_0$ for the spanning tests where we used bootstrapped residuals seem to be almost as high or higher than when the residuals are multivariate normally distributed. We can also see from Tables 11, 15 and 16 that the probabilities of rejecting $H_0$ might be generally lower under an $H_1$ where the weights of the test assets in the TP are zero ($\alpha = 0_N$) compared to an $H_1$ where the weights of the test assets in the TP are not equal to zero ($\alpha \neq 0_N$).

Table 10: The probabilities of rejecting the null-hypothesis in (13) when $H_0$ is true for the spanning tests, and these were obtained by using 10,000 residual bootstraps.

| | | | **Monthly returns** ($T = 144$) | | | | | | **Weekly returns** ($T = 627$) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | **Asymptotic test** | | | **Exact test** | | | **Asymptotic test** | | | **Exact test** | | |
| **Screening** | **N** | **K** | **LR** | **W** | **LM** | **LR** | **W** | **LM** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| 10% | 2 | 19 | 0.084 | 0.092 | 0.078 | 0.049 | 0.050 | 0.050 | 0.063 | 0.064 | 0.061 | 0.056 | 0.056 | 0.056 |
| 25% | 5 | 16 | 0.110 | 0.133 | 0.090 | 0.055 | 0.055 | 0.056 | 0.058 | 0.060 | 0.055 | 0.050 | 0.050 | 0.050 |
| 50% | 10 | 11 | 0.127 | 0.179 | 0.084 | 0.057 | 0.058 | 0.058 | 0.066 | 0.073 | 0.060 | 0.055 | 0.055 | 0.055 |
| All | 12 | 9 | 0.128 | 0.192 | 0.076 | 0.058 | 0.057 | 0.057 | 0.068 | 0.076 | 0.059 | 0.056 | 0.056 | 0.056 |

Another aspect one should note is that the residuals do not necessarily have to be homoscedastic, and they might also be serially correlated. We used the non-overlapping block bootstrap procedure to simulate new residuals (see Section 3.2.4 for more details), and one reason for using this method is to be able to include some possible heteroscedasticity and autocorrelation patterns (if there are any) in the simulations. We implemented this procedure on the monthly and weekly returns, respectively, for the four different scenarios we obtained by the screening procedures. We presented the probabilities of rejection under $H_0$ and under the following three possibilities of $H_1$: cases 3, 6, and 10, in Table 12. The other results are available from the author on request. We can see from Table 12 that the probabilities of rejecting the null-hypothesis in (13) under $H_1$ are extremely high. Hence, the power of the spanning tests seems to be high when possible patterns such as heteroscedasticity exists in the residuals. However, the probabilities of rejecting the null-hypothesis in (13) when $H_0$ is true are all higher than expected, they all exceed 0.05 fairly much, with the exception when 10% screening is applied on weekly returns. This implies that the nonsignificant results (that the difference between the efficient frontiers is not statistically significant at the 5% significance level) we obtained in Tables 1, 2, 3, and 4 might not be misleading if possible heteroscedasticity exists in the residuals.

Table 11: The probabilities of rejecting the null-hypothesis in (13) when $H_1$ is true for the spanning tests, and the probabilities were obtained by using 10,000 residual bootstraps. $T$ = 144 represents that the probabilities were obtained by using monthly returns, and $T$ = 627 by using weekly returns.

| | | | | | **10,000 bootstrap simulations** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | **Asymptotic test** | | | **Exact test** | | |
| **Case** | **Screening** | **N** | **K** | **T** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| 3 | 10% | 2 | 19 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 25% | 5 | 16 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 6 | 10% | 2 | 19 | 144 | 0.919 | 0.924 | 0.911 | 0.873 | 0.873 | 0.873 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 25% | 5 | 16 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 7 | 10% | 2 | 19 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 25% | 5 | 16 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 10 | 10% | 2 | 19 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 25% | 5 | 16 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Table 12: The probabilities of rejecting the null-hypothesis in (13) when $H_0$ and $H_1$ is true, respectively, for the tests and they were obtained using 10,000 non-overlapping block bootstrap simulations on the residuals. $T = 144$ and $T = 627$ represent that the probabilities were obtained by using monthly returns and weekly returns respectively. The number of equally sized blocks we used was twelve for monthly returns, and eleven for weekly returns.

| | | | | | **Block bootstrap simulations** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | **Asymptotic test** | | | **Exact test** | | |
| **Case** | **Screening** | **N** | **K** | **T** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| $H_0$ | 10% | 2 | 19 | 144 | 0.124 | 0.132 | 0.115 | 0.079 | 0.079 | 0.080 |
| | | | | 627 | 0.044 | 0.046 | 0.043 | 0.039 | 0.039 | 0.040 |
| | 25% | 5 | 16 | 144 | 0.192 | 0.222 | 0.169 | 0.125 | 0.126 | 0.123 |
| | | | | 627 | 0.107 | 0.110 | 0.103 | 0.094 | 0.094 | 0.094 |
| | 50% | 10 | 11 | 144 | 0.237 | 0.294 | 0.179 | 0.145 | 0.147 | 0.142 |
| | | | | 627 | 0.121 | 0.128 | 0.112 | 0.105 | 0.105 | 0.105 |
| | All | 12 | 9 | 144 | 0.250 | 0.319 | 0.183 | 0.156 | 0.158 | 0.154 |
| | | | | 627 | 0.155 | 0.167 | 0.143 | 0.138 | 0.138 | 0.137 |
| $H_1$ (case 3) | 10% | 2 | 19 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 25% | 5 | 16 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| $H_1$ (case 6) | 10% | 2 | 19 | 144 | 0.975 | 0.977 | 0.972 | 0.954 | 0.954 | 0.953 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 25% | 5 | 16 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 0.999 | 0.999 | 0.999 | 0.9990 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| $H_1$ (case 10) | 10% | 2 | 19 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 25% | 5 | 16 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 50% | 10 | 11 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | All | 12 | 9 | 144 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | 627 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Furthermore, Kan and Zhou [17] described a spanning test approach based on GMM, which works well when the residuals show conditional heteroscedasticity. More details about this specific test can be found in their paper or the paper by Hansen [12]. We applied this test based on the monthly and weekly returns, respectively, in order to verify our results, and we observed that the GMM Wald test statistics were all very small, which gave us very high $p$-values (over 0.9) for the eight different scenarios (data of monthly and weekly returns with 10%, 25%, 50%, and All screening). Hence, the hypothesis of spanning can not be rejected for these scenarios, and the asymptotic and exact spanning tests seem to perform well for our data even though the assumptions of the residuals such as being i.i.d. normally distributed might be violated.

# 6 Conclusion

In this thesis, we studied whether investing in a portfolio constructed by more sustainable assets might give us as a high utility as when one is investing in a portfolio consisting of both sustainable assets and unsustainable assets. Sustainability was measured with total ESG scores and total ESG risk ratings, respectively, provided by Sustainalytics and can be found at Yahoo Finance. The data we used consist of the monthly and weekly returns of 21 large-cap stocks that are included in the OMXS30 index over the period 2008-2019. To study whether an investment in all the 21 stocks might give better investment opportunities compared to investments in solely the more sustainable stocks, one might analyze the efficient frontiers of the different numbers of assets. If the difference between the efficient frontiers is statistically significant, one might assume that adding the stocks that are less sustainable into the investment portfolio would improve the investment opportunity. Some methods that can possibly test for this are the mean-variance spanning tests proposed by Huberman and Kandel [16] as well as those spanning tests presented in Kan and Zhou [17]. These tests are based on a multivariate regression and were the methods used in this thesis.

The mean-variance spanning tests based on a multivariate regression assume that the residuals are i.i.d. multivariate normally distributed. However, this assumption might, unfortunately, be violated for most data, which can lead to misleading interpretations. Nevertheless, we saw that the exact spanning tests found in Huberman and Kandel [16], and Kan and Zhou [17] give the best performance overall. Even when most of the assumptions are violated, they still perform well under $H_0$ (with an exception for when possible heteroscedasticity or autocorrelation might be present), but also under $H_1$ when the residuals are independently but not necessarily identically distributed. Nevertheless, the power of the tests seems a little weaker when the disturbances are i.i.d. but not normally distributed compared to when they are i.i.d. normally distributed. Additionally, the main component that affects the power of the spanning tests seems to be the weights of the $N$ test assets in the TP. Furthermore, we saw that the power of the tests seems high if possible heteroscedasticity and autocorrelation exist in our data. However, further studies on these aspects are possibly needed to get a better estimate of the performances of the spanning tests when heteroscedasticity and autocorrelation are present, respectively, in the data.

We saw that the minimum-variance frontier of the 21 stocks and the minimum-variance frontier when we applied 10%, 25%, and 50% screening on data with monthly and weekly returns, respectively, are not significantly different for the whole study period (2008-2019). This means that investors would probably not get better investment opportunities if they, in addition to investing in the more sustainable assets in the OMXS30 index, also invests in those stocks with higher total ESG scores. Furthermore, we saw similar results from the mean-variance spanning tests based on monthly and weekly returns, respectively, when we screened for all the stocks with a total ESG risk rating between medium and severe.

# 7 Discussion

We have studied the robustness to non-normality of the asymptotic and exact mean-variance spanning tests based on various simulations, and by computing the probabilities of rejecting $H_0$ under different possibilities as $H_1$. We saw that the weights of the $N$ test assets in the TP seem to affect the tests the most. However, we would probably need to study more cases of $H_1$ (that both the TPs and the MVPs differ in the efficient frontiers, where $\alpha$ and $1_N - \beta 1_K$ differ from $H_0$, and where one of the two vectors differs more than the other) in order to conclude for that. Furthermore, this finding in our study contradicts to what Kan and Zhou [17] observed in their paper. They mentioned that the spanning tests would place heavy weights on $1_N - \beta 1_K$ and smaller weights on $\alpha$. Their motivation to why the spanning tests mainly rely on $1_N - \beta 1_K$ is that $\beta$ does not depend on the expected returns $\mu$, so $1_N - \beta 1_K$ can be estimated more accurately than $\alpha$ (see, Kan and Zhou [17]). Hence, further studies on the power of the tests will probably be needed if one uses the simulation procedures mentioned in our thesis.

Moreover, we have mentioned that the eigenvalues, $\lambda_1$ and $\lambda_2$, can be seen as measures of the maximum and minimum differences, respectively, between the expected minimum-variance frontiers (based on the estimated parameters $\hat{B}$ and $\hat{\Sigma}$), according to Kan and Zhou [17]. Then, the eigenvalues $\theta_1$ and $\theta_2$ of $H\hat{G}^{-1}$, where $H = \Theta\Sigma^{-1}\Theta^\top$, would be measures of the maximum and minimum differences, respectively, between the population minimum-variance frontiers based on the true parameters $B$ and $\Sigma$. One might then want to make sure to choose a $B$-matrix that satisfies an $H_1$, where both the TPs and the MVPs differ, and a $\Sigma$-matrix such that $\theta_1 \geq \theta_2 > 0$ (thus if the maximum and minimum differences between the minimum-variance frontiers are strictly greater than zero, then it is logical to think that both the TPs and the MVPs are different in the minimum-variance frontiers). Furthermore, when either the TPs or the MVPs are identical, on the other hand, one would desire to have $\theta_1 > 0$ and $\theta_2 = 0$. An interesting aspect to study for the robustness to the non-normality for the spanning tests under $H_0$ and $H_1$, respectively, is then to take these conditions of $\theta_1$ and $\theta_2$ into account when one performs the simulation procedures in Section 3.2. This interesting aspect to study the probabilities of rejecting the null-hypothesis in (13) when $H_0$ or $H_1$ is true is left for future research. However, if the interest is to solely study the power of the asymptotic or exact mean-variance spanning tests and where the normality assumption holds, one would probably prefer to use the simulation method described in Section 3.2 in the paper by Kan and Zhou [17].

It is also worth to mention that we divided the residuals obtained using the monthly returns into twelve equally sized blocks when we applied the non-overlapping block bootstrap on the residuals. Likewise, we divided the residuals of the weekly returns into eleven equally sized blocks. These numbers of blocks might not be the most efficient choices, and there may exist better resampling or simulation methods to use when one wants to preserve, e.g., possible heteroscedasticity or autocorrelation patterns in the data. Hence, the actual probabilities of rejecting the null-hypothesis in (13) under $H_0$ and $H_1$, respectively, when heteroscedasticity or autocorrelation is present might be higher or lower than what we observed. However, one might argue that our results, that the differences in the efficient frontiers are statistically insignificant, obtained by using the asymptotic and exact tests might be correct when heteroscedasticity exists, and in the absence of autocorrelation. The reason is that the results obtained by applying the GMM Wald test for spanning showed nonsignificance differences between the efficient frontiers, and this method considers heteroscedasticity, but not necessarily autocorrelation. Nevertheless, one should note that the GMM Wald test does assume that $x_t \otimes \epsilon_t$ for $t = 1, 2, ..., T$, are not serially correlated, which can be checked by applying the multivariate Ljung-Box test.

Lastly, we should also mention another interesting aspect for future studies if one has, for instance, obtained a result from a mean-variance spanning test, which shows that the difference between the minimum-variance frontiers is statistically significant. If such a result is observed, one might be interested in knowing whether the difference is caused by the MVPs or the TPs. One might then perhaps want to apply the step-down test presented in Kan and Zhou [17], which tests for $\alpha = 0_N$ first and then test for

$1_N - \beta 1_K = 0_N$ conditional on $\alpha = 0_N$ [4]. If one can reject $\alpha = 0_N$ from the first test, then it means that the TPs differ. Moreover, if one can reject $1_N - \beta 1_K = 0_N$ conditional on $\alpha = 0_N$ based on the second test, then it implies that the MVPs are different [9]. Alternatively, one can test whether the difference in the minimum-variance frontiers was caused by the MVPs by using the exact test presented in the paper by Bodnar and Schmid [7]. One can also consider applying the exact test found in Bodnar and Okhrin [6] if one wants to study whether the difference in the minimum-variance frontiers was due to the TPs.

# 8 References


[1] Aktieinformation. `https://www.atlascopcogroup.com/se/investor-relations/atlas-copco-share/share-information`. [Online; accessed 10-05-2020].

[2] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons Inc, 3rd edition, 2003.

[3] R. Bauer, K. Koedijk, and R. Otten. International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29(7):1751 – 1767, 2005.

[4] E. R. Berndt and N. E. Savin. Conflict among Criteria for Testing Hypotheses in the Multivariate Linear Regression Model. *Econometrica*, 45(5):1263–1277, 1977.

[5] H. J. Bierens. The inverse of a partitioned matrix. `http://www.math.chalmers.se/~rootzen/highdimensional/blockmatrixinverse.pdf`, 2013. [Online; accessed 19-02-2020].

[6] T. Bodnar and Y. Okhrin. 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, 2011.

[7] T. Bodnar and W. Schmid. A test for the weights of the global minimum variance portfolio in an elliptical model. *Metrika*, 67(2):127, 2008.

[8] M. Capinski and T. Zastawniak. *Mathematics for Finance : An Introduction to Financial Engineering*. Springer, New York, 2nd edition, 2011.

[9] W.-P. Chen, H. Chung, K.-Y. Ho, and T.-L. Hsu. *Portfolio Optimization Models and Mean–Variance Spanning Tests*, pages 165–184. Springer US, Boston, MA, 2010.


[4] Note that one can also use the step-down test to verify the null-hypothesis of spanning. See Kan and Zhou [17] for more details.

[10] Di. https://www.di.se/bors/large-cap/, 2020. [Online; accessed 17-04-2020].

[11] W. E. Ferson, S. R. Foerster, and D. B. Keim. General Tests of Latent Variable Models and Mean-Variance Spanning. *The Journal of Finance*, 48(1):131–156, 1993.

[12] L. P. Hansen. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50(4):1029–1054, 1982.

[13] H. V. Henderson and S. R. Searle. On Deriving the Inverse of a Sum of Matrices. *SIAM Review*, 23(1):53–60, 1981.

[14] H.-M. Henke. The effect of social screening on bond mutual fund performance. *Journal of Banking & Finance*, 67:69 – 84, 2016.

[15] S. Herzel, M. Nicolosi, and C. Stărică. The cost of sustainability in optimal portfolio decisions. *The European Journal of Finance*, 18(3-4):333–349, 2012.

[16] G. Huberman and S. Kandel. Mean-Variance Spanning. *The Journal of Finance*, 42(4):873–888, 1987.

[17] R. Kan and G. Zhou. Tests of Mean-Variance Spanning. *Annals of Economics and Finance*, 13(1):139–187, 2012.

[18] S. Korkmaz, D. Goksuluk, and G. Zararsiz. MVN: An R Package for Assessing Multivariate Normality. *The R Journal*, 6:151–162, 12 2014.

[19] S. N. Lahiri. *Resampling methods for dependent data*. Springer, New York, 2003.

[20] K. V. Mardia. Measures of Multivariate Skewness and Kurtosis with Applications. *Biometrika*, 57(3):519–530, 1970.

[21] D. V. Matt Dancho. Package 'tidyquant'. https://cran.r-project.org/web/packages/tidyquant/tidyquant.pdf, 2020. [Online; accessed 24-04-2020].

[22] R. J. Muirhead. *Aspects of multivariate statistical theory [E-book]*. Wiley, New York, 1982.

[23] Nasdaq. http://www.nasdaqomxnordic.com/utbildning/optionerochterminer/vadaromxstockholm30index, 2020. [Online; accessed 15-03-2020].

[24] J. Nofsinger and A. Varma. Socially responsible funds and market crises. *Journal of Banking & Finance*, 48:180 – 193, 2014.

[25] Nordea. https://www.nordea.com/sv/hallbarhet/hallbar-verksamhet/vad-ar-ESG/. [Online; accessed 17-02-2020].

[26] E. Ortas, R. L. Burritt, and J. M. Moneva. Socially Responsible Investment and cleaner production in the Asia Pacific: does it pay to be good? *Journal of Cleaner Production*, 52:272 – 280, 2013.

[27] K. B. Petersen and P. Syskind. The Matrix Cookbook. https://www.ics.uci.edu/~welling/teaching/KernelsICS273B/MatrixCookBook.pdf, 2005. [Online; accessed 09-03-2020].

[28] C. Pozrikidis. *An Introduction to Grids, Graphs, and Networks [E-book]*. Oxford University Press, 2014.

[29] G. A. F. Seber. *Multivariate observations [E-book]*. Wiley, New York, 1984.

[30] Sustainalytics. https://www.sustainalytics.com/esg-ratings/#1530569101275-ef381f7e-5014. [Online; accessed 15-04-2020].

[31] Sustainalytics. https://www.sustainalytics.com/esg-ratings/. [Online; accessed 15-03-2020].

[32] Understanding Investing. http://www.understandinginvesting.org/course/diversification-the-art-of-not-putting-all-your-eggs-in-one-basket/. [Online; accessed 13-04-2020].

[33] UNIBusiness Editor. https://business.uni.edu/news-views/why-social-responsibility-has-become-bigger-part-company-planning, 2020. [Online; accessed 15-04-2020].

[34] J. Villaseñor and E. González-Estrada. A Generalization of Shapiro–Wilk's Test for Multivariate Normality. *Communications in Statistics—Theory and Methods*, 38:1870–1883, 07 2009.

# A Appendix

## A.1 Proofs, matrices and maximum likelihood estimators

### A.1.1 Proof of the two-fund theorem

Let us first assume that the inverse of the covariance matrix of returns, $V$, exists, and that the vectors $\mu$ and $1_N$ are linearly independent. Assume furthermore that $w_1$ and $w_2$ denote two vectors consisting of the weights from any two portfolios ($P_1$, and $P_2$) on the minimum-variance frontier with the expected returns $\mu_1$ and $\mu_2$, where the two portfolios do not have the same expected returns. It is then possible to obtain the weights of each portfolio $P_3$ on the minimum-variance frontier by using the following formula

$$w_3 = \gamma w_1 + (1-\gamma) w_2,$$

for some $\gamma \in \mathbb{R}$ [8].

*Proof.* Let us denote the expected return of the portfolio $P_3$ as $\mu_3$, and let us then define $\mu_3$ as a function of the expected returns of the portfolios $P_1$ and $P_2$

$$\mu_3 = \gamma \mu_1 + (1-\gamma)\mu_2. \tag{16}$$

Solving for $\gamma$ gives

$$\gamma = \frac{\mu_3 - \mu_2}{\mu_1 - \mu_2}.$$

Note that $\mu_1 - \mu_2 \neq 0$ thus the expected returns are different. Moreover, when a portfolio with the expected return $\mu_P$ lies on the minimum-variance frontier, then its weights can be expressed in the form $w = \mu_P a + b$ (see Section 2.1.3). Hence, we have that $w_1 = \mu_1 a + b$, $w_2 = \mu_2 a + b$, and $w_3 = \mu_3 a + b$, and we need to show that $w_3 = \gamma w_1 + (1-\gamma) w_2$. By plugging in the expressions for $w_1$ and $w_2$ into $\gamma w_1 + (1-\gamma) w_2$ gives

$$\gamma w_1 + (1-\gamma) w_2 = \gamma(\mu_1 a + b) + (1-\gamma)(\mu_2 a + b) = [\gamma \mu_1 + (1-\gamma)\mu_2] a + b = \mu_3 a + b,$$

which is equal to the weights of portfolio $P_3$, where we used Equation (16) in the last equality [8]. □

### A.1.2 Inverse of a partitioned matrix

A symmetric nonsingular block matrix

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix},$$

where $V_{11}$ is a nonsingular $K \times K$ matrix, $V_{12}$ a $K \times N$ matrix, $V_{21}$ an $N \times K$ matrix, and $V_{22}$ a nonsingular $N \times N$ matrix, has the following inverse matrix

$$V^{-1} = \begin{pmatrix} V_{11}^{-1} + \beta^\top \Omega^{-1} \beta & -\beta^\top \Omega^{-1} \\ -\Omega^{-1} \beta & \Omega^{-1} \end{pmatrix},$$

where $\Omega = V_{22} - V_{21} V_{11}^{-1} V_{12}$, and $\beta = V_{21} V_{11}^{-1}$.

*Proof.* From Henderson 1981 [13], we know that the inverse matrix of $V$ can be written in the following form

$$V^{-1} = \begin{pmatrix} V_{11}^{-1} + V_{11}^{-1} V_{12} (V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} V_{21} V_{11}^{-1} & -V_{11}^{-1} V_{12} (V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} \\ -(V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} V_{21} V_{11}^{-1} & (V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} \end{pmatrix},$$

by using the following expression

$$\begin{pmatrix} V_{11}^{-1} & 0_{K\times N} \\ -V_{21}V_{11}^{-1} & I_N \end{pmatrix}\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} = \begin{pmatrix} I_K & V_{11}^{-1}V_{12} \\ 0_{N\times K} & V_{22} - V_{21}V_{11}^{-1}V_{12} \end{pmatrix},$$

where $I_K$ and $I_N$ are a $K\times K$-identity matrix and an $N\times N$-identity matrix respectively. Hence,

$$V^{-1} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}^{-1} = \begin{pmatrix} I_K & V_{11}^{-1}V_{12} \\ 0_{N\times K} & V_{22} - V_{21}V_{11}^{-1}V_{12} \end{pmatrix}^{-1}\begin{pmatrix} V_{11}^{-1} & 0_{K\times N} \\ -V_{21}V_{11}^{-1} & I_N \end{pmatrix}.$$

We can use the inverse of a partitioned matrix to obtain another expression for the inverse matrix of

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} = \begin{pmatrix} I_K & V_{11}^{-1}V_{12} \\ 0_{N\times K} & V_{22} - V_{21}V_{11}^{-1}V_{12} \end{pmatrix}.$$

Let us follow the steps from Herman J. Bierens [5]. We first note that $M_{11} = I_K$ is not singular, which means that $M_{11}$ is invertible. Assume also that the inverse of $M_{22}$ exists. We also know the following property,

$$MM^{-1} = M^{-1}M = I_{K+N},$$

where $I_{K+N}$ is a $(K+N)\times(K+N)$-identity matrix. Let us write the inverse matrix on this form

$$M^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where $A_{11}$ is a $K\times K$ matrix, $A_{12}$ a $K\times N$ matrix, $A_{21}$ an $N\times K$ matrix, and $A_{22}$ an $N\times N$ matrix. Then,

$$\begin{aligned} MM^{-1} &= \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \\ &= \begin{pmatrix} M_{11}A_{11} + M_{12}A_{21} & M_{11}A_{12} + M_{12}A_{22} \\ M_{21}A_{11} + M_{22}A_{21} & M_{21}A_{12} + M_{22}A_{22} \end{pmatrix} \\ &= \begin{pmatrix} I_K & 0_{K\times N} \\ 0_{N\times K} & I_N \end{pmatrix}. \end{aligned}$$

This implies that we get the following four matrix equations

$$M_{11}A_{11} + M_{12}A_{21} = I_K, \tag{17}$$

$$M_{11}A_{12} + M_{12}A_{22} = 0_{K\times N}, \tag{18}$$

$$M_{21}A_{11} + M_{22}A_{21} = 0_{N\times K}, \tag{19}$$

$$M_{21}A_{12} + M_{22}A_{22} = I_N. \tag{20}$$

Thus, we have assumed that $M_{11}$ and $M_{22}$ are nonsingular, then we can solve for $A_{12}$ and $A_{21}$ in Equation (18) and (19), which gives that

$$A_{12} = -M_{11}^{-1} M_{12} A_{22}, \tag{21}$$

and

$$A_{21} = -M_{22}^{-1} M_{21} A_{11}. \tag{22}$$

The next step is to plug in Equation (21) and (22) into Equation (20) and (17) respectively, and then solve for $A_{11}$ and $A_{22}$. So, we obtain the following

$$A_{11} = \left(M_{11} - M_{12} M_{22}^{-1} M_{21}\right)^{-1},$$

and

$$A_{22} = \left(M_{22} - M_{21} M_{11}^{-1} M_{12}\right)^{-1}.$$

Hence, we have that

$$M^{-1} = \begin{pmatrix} \left(M_{11} - M_{12} M_{22}^{-1} M_{21}\right)^{-1} & -M_{11}^{-1} M_{12} \left(M_{22} - M_{21} M_{11}^{-1} M_{12}\right)^{-1} \\ -M_{22}^{-1} M_{21} \left(M_{11} - M_{12} M_{22}^{-1} M_{21}\right)^{-1} & \left(M_{22} - M_{21} M_{11}^{-1} M_{12}\right)^{-1} \end{pmatrix}. \tag{23}$$

Similarly, using $M^{-1} M = I_{K+N}$ instead, we obtain

$$M^{-1} = \begin{pmatrix} \left(M_{11} - M_{12} M_{22}^{-1} M_{21}\right)^{-1} & -\left(M_{11} - M_{12} M_{22}^{-1} M_{21}\right)^{-1} M_{12} M_{22}^{-1} \\ -\left(M_{22} - M_{21} M_{11}^{-1} M_{12}\right)^{-1} M_{21} M_{11}^{-1} & \left(M_{22} - M_{21} M_{11}^{-1} M_{12}\right)^{-1} \end{pmatrix}. \tag{24}$$

Equation (23) and (24) are equivalent because a nonsingular matrix can only have one inverse.

By using Equation (23) or (24), we obtain the following

$$M^{-1} = \begin{pmatrix} I_K & -V_{11}^{-1} V_{12} (V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} \\ 0_{N\times K} & (V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} \end{pmatrix}.$$

Then,

$$\begin{aligned} V^{-1} &= M^{-1} \begin{pmatrix} V_{11}^{-1} & 0_{K\times N} \\ -V_{21} V_{11}^{-1} & I_N \end{pmatrix} \\ &= \begin{pmatrix} I_K & -V_{11}^{-1} V_{12} (V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} \\ 0_{N\times K} & (V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} \end{pmatrix} \begin{pmatrix} V_{11}^{-1} & 0_{K\times N} \\ -V_{21} V_{11}^{-1} & I_N \end{pmatrix} \\ &= \begin{pmatrix} V_{11}^{-1} + V_{11}^{-1} V_{12} (V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} V_{21} V_{11}^{-1} & -V_{11}^{-1} V_{12} (V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} \\ -(V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} V_{21} V_{11}^{-1} & (V_{22} - V_{21} V_{11}^{-1} V_{12})^{-1} \end{pmatrix} \end{aligned}$$

Let $\Omega = V_{22} - V_{21} V_{11}^{-1} V_{12}$ and $\beta = V_{21} V_{11}^{-1}$. Furthermore, we know that $V$ is a symmetric block matrix, so $V_{21} = V_{12}^{\top}$. We can then rewrite $V^{-1}$ as

$$V^{-1} = \begin{pmatrix} V_{11}^{-1} + \beta^{\top} \Omega^{-1} \beta & -\beta^{\top} \Omega^{-1} \\ -\Omega^{-1} \beta & \Omega^{-1} \end{pmatrix}.$$

□

### A.1.3 Maximum likelihood estimators of $B$ and $\Sigma$

Assume that the desired multivariate regression can be expressed as

$$Y = XB + E,$$

where the matrices $Y$, $X$, $B$ and $E$ are described in Section 3.1.1. Let us assume that $T \geq N + K + 1$ and that $X$ has rank $K + 1$. Furthermore, assume that conditioned on $R_t$, the $\epsilon_t$ are i.i.d. normally distributed with mean $0_N$ and covariance matrix $\Sigma$ for $t = 1, ..., T$. So, the rows in matrix $E$ are i.i.d. $N(0_N, \Sigma)$-distributed thus $E = (\epsilon_1, ..., \epsilon_T)^\top$. Then from Muirhead [22], we know that $E$ is multivariate normally distributed with mean $0_{T\times N}$ and covariance matrix $I_T \otimes \Sigma$ (which can be denoted as $N(0_{T\times N}, I_T \otimes \Sigma)$), where $\otimes$ denotes the Kronecker product. So, $Y$ is $N(XB, I_T \otimes \Sigma)$-distributed, and the density function of $Y$ is given by (note that we have change the notation for determinant to $\det(\cdot)$ in this section)

$$\begin{aligned} f(Y; B, \Sigma) = (2\pi)^{-NT/2}[\det(\Sigma)]^{-T/2}[\det(I_T)]^{-N/2} \\ \exp\left\{\text{Trace}\left[-\frac{1}{2}I_T^{-1}(Y - XB)\Sigma^{-1}(Y - XB)^\top\right]\right\}, \end{aligned}$$

which is equivalent to

$$f(Y; B, \Sigma) = (2\pi)^{-NT/2}[\det(\Sigma)]^{-T/2} \exp\left\{\text{Trace}\left[-\frac{1}{2}(Y - XB)\Sigma^{-1}(Y - XB)^\top\right]\right\}.$$

Note that this is also the likelihood function, so the log-likelihood function can be written as

$$\begin{aligned} l(B, \Sigma) &\propto -\frac{T}{2}\log[\det(\Sigma)] + \text{Trace}\left[-\frac{1}{2}(Y - XB)\Sigma^{-1}(Y - XB)^\top\right] \\ &= -\frac{T}{2}\log[\det(\Sigma)] \\ &\quad + \text{Trace}\left[-\frac{1}{2}\left(Y\Sigma^{-1}Y^\top - Y\Sigma^{-1}B^\top X^\top - XB\Sigma^{-1}Y^\top + XB\Sigma^{-1}B^\top X^\top\right)\right] \\ &= -\frac{T}{2}\log[\det(\Sigma)] + \text{Trace}\left[-\frac{1}{2}Y\Sigma^{-1}Y^\top\right] + \text{Trace}\left[\frac{1}{2}Y\Sigma^{-1}B^\top X^\top\right] \\ &\quad + \text{Trace}\left[\frac{1}{2}XB\Sigma^{-1}Y^\top\right] + \text{Trace}\left[\frac{1}{2}XB\Sigma^{-1}B^\top X^\top\right], \end{aligned}$$

where we have used the property that $\text{Trace}(M+F) = \text{Trace}(M)+\text{Trace}(F)$, where $M$ and $F$ are $T\times T$-matrices, in the second equality sign [27]. Setting the derivative of the log-likelihood function with respect to the matrix $B$ equal to a $(K + 1) \times N$-matrix consisting of only zeros, where we use that [27],

$$\frac{\partial \text{Trace}(XB\Sigma^{-1}Y^\top)}{\partial B} = X^\top(\Sigma^{-1}Y^\top)^\top,$$
$$\frac{\partial \text{Trace}(Y\Sigma^{-1}B^\top X^\top)}{\partial B} = X^\top Y\Sigma^{-1},$$
$$\frac{\partial \text{Trace}(XB\Sigma^{-1}B^\top X^\top)}{\partial B} = X^\top(X^\top)^\top B(\Sigma^{-1})^\top + X^\top XB\Sigma^{-1},$$

gives that

$$\begin{aligned}\frac{\partial l(B,\Sigma)}{\partial B} &= \frac{1}{2}X^\top Y\hat{\Sigma}^{-1} + \frac{1}{2}X^\top(\hat{\Sigma}^{-1}Y^\top)^\top - \frac{1}{2}X^\top(X^\top)^\top\hat{B}(\hat{\Sigma}^{-1})^\top - \frac{1}{2}X^\top X\hat{B}\hat{\Sigma}^{-1}\\ &= X^\top Y\hat{\Sigma}^{-1} - X^\top X\hat{B}\hat{\Sigma}^{-1}\\ &= 0_{(K+1)\times N}.\end{aligned}$$

The second equality holds thus $\Sigma^{-1} = (\Sigma^{-1})^T$ because $\Sigma$ is a symmetric nonsingular matrix, so its inverse will also be symmetric. We then obtain that

$$X^\top Y\hat{\Sigma}^{-1} = X^\top X\hat{B}\hat{\Sigma}^{-1}.$$

Multiplying both sides by $\hat{\Sigma}$ on the right gives

$$X^\top Y = X^\top X\hat{B}.$$

Hence, the maximum likelihood estimator of $B$ is given by

$$\hat{B} = (X^\top X)^{-1}X^\top Y.$$

To obtain the maximum likelihood estimator of $\Sigma$, we should take the derivative of the log-likelihood function with respect to $\Sigma$, and then set the expression equal to an $N \times N$-matrix of solely zeros. Let us also use the following [27],

$$\frac{\partial \det(\Sigma)}{\partial \Sigma} = \det(\Sigma)\left[2\Sigma^{-1} - \text{diag}\left(\Sigma^{-1}\right)\right],$$
$$\begin{aligned}\frac{\partial \text{Trace}[(Y-XB)\Sigma^{-1}(Y-XB)^\top]}{\partial \Sigma} &= -\left[\Sigma^{-1}(Y-XB)^\top(Y-XB)\Sigma^{-1}\right]^\top\\ &\quad - \Sigma^{-1}(Y-XB)^\top(Y-XB)\Sigma^{-1}\\ &\quad + \text{diag}\left\{\left[\Sigma^{-1}(Y-XB)^\top(Y-XB)\Sigma^{-1}\right]^\top\right\}.\end{aligned}$$

Then,

$$
\begin{aligned}
\frac{\partial l(B,\Sigma)}{\partial \Sigma} &= -\frac{T}{2}\frac{1}{\det(\hat{\Sigma})}\det(\hat{\Sigma})\left[2\hat{\Sigma}^{-1} - \operatorname{diag}\left(\hat{\Sigma}^{-1}\right)\right] \\
&\quad - \frac{1}{2}\Big\{ - \left[\hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1}\right]^\top \\
&\qquad - \hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1} \\
&\qquad + \operatorname{diag}\left(\left[\hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1}\right]^\top\right)\Big\} \\
&= -T\hat{\Sigma}^{-1} + \frac{T}{2}\operatorname{diag}(\hat{\Sigma}^{-1}) + \hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1} \\
&\quad - \frac{1}{2}\operatorname{diag}\left[\hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1}\right] \\
&= -T\hat{\Sigma}^{-1} + \hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1} \\
&\quad + \frac{1}{2}\operatorname{diag}\left[T\hat{\Sigma}^{-1} - \hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1}\right] \\
&= 0_{N\times N}.
\end{aligned}
$$

Hence, we obtain the following equality

$$T\hat{\Sigma}^{-1} - \hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1} = \frac{1}{2}\operatorname{diag}\left[T\hat{\Sigma}^{-1} - \hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1}\right],$$

which holds when

$$T\hat{\Sigma}^{-1} - \hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1} = 0_{N\times N},$$

which can be rewritten as

$$T\hat{\Sigma}^{-1} = \hat{\Sigma}^{-1}(Y - X\hat{B})^\top (Y - X\hat{B})\hat{\Sigma}^{-1}.$$

Multiplying both sides by $\hat{\Sigma}$ on the left and right gives

$$T\hat{\Sigma} = (Y - X\hat{B})^\top (Y - X\hat{B}).$$

The maximum likelihood estimate of $\Sigma$ is then obtained by dividing both sides by $T$

$$\hat{\Sigma} = \frac{1}{T}(Y - X\hat{B})^\top (Y - X\hat{B}).$$

## A.2 Test for multivariate normality

### A.2.1 Generalized Shapiro-Wilk test for multivariate normality

One possible method to test for the hypothesis of having univariate normality is the Shapiro-Wilk's test. Assume that we have a random sample

$x_1, ..., x_n$, and let $\bar{x} = 1/n \sum_{i=1}^{n} x_i$ and $s_X^2 = \sum_{i=1}^{n}(x_i - \bar{x})$. The test statistic is then given by

$$W_X = \frac{\tilde{\sigma}_X}{s_X^2},$$

where, $\tilde{\sigma}_X = \left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2$. Furthermore, $x_{(1)} < x_{(2)} < ... < x_{(n)}$ and they represent the ordered observations, and $a_i$ denotes the $i$th element of the vector $a = (a_1, ..., a_n)^\top$, which is obtained by the following expression

$$\frac{m^\top C^{-1}}{\sqrt{m^\top C^{-1} C^{-1} m}},$$

where $m^\top$ is the expected values of a vector $Z$, which consists of the ordered statistics of $n$ random variables drawn from the standard normal distribution, $N(0,1)$. In other words, $Z$ contains a sample of $n$ observations from the $N(0,1)$-distribution, and these values are ordered in the vector $Z$. Furthermore, $C$ denotes the covariance matrix of the ordered statistics in vector $Z$ (i.e., the covariance matrix of the ordered sample drawn from the standard normal distribution) [34].

Villaseñor and González-Estrada [34] purposed a method based on the Shapiro-Wilk's test in order to test for multivariate normality. Let $X_1, ..., X_n$ denote $n$ number of random vectors of size $p$, where $p \geq 1$. Let us assume that $X_1, ..., X_n$ are i.i.d. $p$-multivariate distributed with mean vector $\mu$ (of size $p$) and covariance matrix $\Sigma$ (a $p \times p$-matrix), $N(\mu, \Sigma)$. Furthermore, a well-known proposition is presented in Villaseñor and González-Estrada [34], which states that a random vector $X$ is $N(\mu, \Sigma)$-distributed if and only if the random vector $Z = \Sigma^{-1/2}(X - \mu)$ is $N(0_p, I_p)$-distributed. Here, $0_p$ and $I_p$ denote a zero-vector of size $p$ and a $p \times p$-identity matrix respectively. Moreover, let $\bar{X}$ and $S$ represent the sample mean and the covariance matrix respectively, i.e., $\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$ and $S = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})^\top$. If the random vectors $X_1, ..., X_n$ are $N(\mu, \Sigma)$-distributed, then the distribution of the random vectors given by $Z_j^* = S^{-1/2}(X_j - \bar{X})$, for $j = 1, 2, ..., n$, will be approximately $N(\mu, \Sigma)$. This implies that the coordinates of $Z_j^*$, which we denote by $Z_{1j}, ..., Z_{pj}$, will be approximately independent univariate standard normal distributed.

In order to test the null-hypothesis that the sample $X_1, ..., X_n$ is from the $N(\mu, \Sigma)$-distribution, where the parameters $\mu$ and $\Sigma$ are unknown, one can use the test statistic purposed by Villaseñor and González-Estrada [34]. The test statistic is given by the following formula

$$W^* = \frac{1}{p}\sum_{i=1}^{p} W_{Z_i},$$

where $W_{Z_i}$ represents the Shapiro-Wilk's test statistic computed on the $i$th coordinate of the transformed observations $Z_{i1}, ..., Z_{in}$ for $i = 1, 2, ..., p$.

Villaseñor and González-Estrada mention that the test statistic, $W^*$, is expected to be almost equal to one under the null-hypothesis because each of the $W_{Z_i}$ for $i = 1, 2, ..., p$ is expected to be approximately one. This generalized Shapiro-Wilk test for multivariate normality, $W^*$, will reject the null-hypothesis at the $(\alpha_0 \cdot 100)$%-significance level if $W^* < c_{\alpha_0;n,p}$ for a value $c_{\alpha_0;n,p}$ that satisfies the following

$$\alpha_0 = P(\{W^* < c_{\alpha_0;n,p} | H_0 \text{ holds}\}).$$

One can however not obtain the distribution of $W^*$ under the null-hypothesis, and the reason is that it does not exist a known formula for the distribution of the Shapiro-Wilk test statistic yet. Hence, Villaseñor and González-Estrada [34] opt to use the Monte-Carlo simulation to obtain the percentiles, $c_{\alpha_0;n,p}$.

### A.2.2 Mardia's test of multivariate normality

Another method to test for multivariate normality is based on multivariate extensions of skewness and kurtosis and was introduced by Mardia [20]. Let $X_1, ..., X_n$ be $n$ random vectors, where each vector consists of $p$ elements. Let also $\gamma_{1,p}$ and $\gamma_{2,p}$ denote the multivariate skewness and kurtosis respectively, which are given by the following

$$\gamma_{1,p} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} m_{ij}^3,$$

$$\gamma_{2,p} = \frac{1}{n} \sum_{i=1}^{n} m_{ii}^2,$$

where $m_{ij}$ represents the squared Mahalanobis distance, which is obtained by $m_{ij} = (X_i - \bar{X})^\top S^{-1}(X_i - \bar{X})$ [18]. Furthermore, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^{-1} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})^\top$ are the sample mean and covariance matrix respectively. In the paper by Mardia [20], it is mentioned that the skewness in a multivariate normal distribution is equal to zero, and the kurtosis is equal to $p(p+2)$. Hence, one will test for the hypothesis that $\gamma_{1,p} = 0$ and $\gamma_{2,p} = p(p+2)$, but this is done separately in order to test for multivariate normality. The test statistic for skewness is given by

$$\frac{n}{6}\gamma_{1,p} \overset{a}{\sim} \chi^2_{p(p+1)(p+2)/6},$$

which is approximately $\chi^2$-distributed with $p(p+1)(p+2)/6$ degrees of freedom. The asymptotic distribution for the test statistic for kurtosis is asymptotically normal distributed with mean $p(p+2)$ and variance $8p(p+2)/n$, i.e.,

$$\gamma_{2,p} \overset{a}{\sim} N\left[p(p+2), 8p(p+2)/n\right].$$

If at least one of the hypothesis, $\gamma_{1,p} = 0$ and $\gamma_{2,p} = p(p+2)$, is rejected, then it means that the sample is not from a multivariate normal distribution [18, 20].

## A.3 Additional tables

Table 13: The probabilities of rejecting the null-hypothesis in (13) when $H_1$ is true for the spanning tests, and under the assumption of having normally and t-distributed disturbances, respectively, which were obtained based on 10,000 simulations. $T$ = 144 represents that the results were obtained by using monthly returns, and $T$ = 627 by using weekly returns.

| | | | | | **Multivariate normal distribution** | | | | | | **Multivariate t-distribution** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | **Asymptotic test** | | | **Exact test** | | | **Asymptotic test** | | | **Exact test** | | |
| **Case** | **Screening** | **N** | **K** | **T** | **LR** | **W** | **LM** | **LR** | **W** | **LM** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| 1 | 10% | 2 | 19 | 144 | 0.085 | 0.092 | 0.077 | 0.047 | 0.047 | 0.047 | 0.087 | 0.095 | 0.077 | 0.051 | 0.051 | 0.051 |
| | | | | 627 | 0.055 | 0.057 | 0.053 | 0.048 | 0.048 | 0.049 | 0.060 | 0.061 | 0.058 | 0.053 | 0.053 | 0.053 |
| | 25% | 5 | 16 | 144 | 0.104 | 0.125 | 0.083 | 0.050 | 0.051 | 0.050 | 0.106 | 0.126 | 0.086 | 0.051 | 0.051 | 0.051 |
| | | | | 627 | 0.058 | 0.061 | 0.055 | 0.049 | 0.049 | 0.049 | 0.066 | 0.069 | 0.063 | 0.055 | 0.056 | 0.055 |
| | 50% | 10 | 11 | 144 | 0.116 | 0.165 | 0.075 | 0.050 | 0.050 | 0.051 | 0.118 | 0.168 | 0.076 | 0.050 | 0.050 | 0.050 |
| | | | | 627 | 0.064 | 0.070 | 0.056 | 0.051 | 0.051 | 0.052 | 0.064 | 0.070 | 0.058 | 0.054 | 0.053 | 0.054 |
| | All | 12 | 9 | 144 | 0.119 | 0.175 | 0.068 | 0.050 | 0.051 | 0.051 | 0.116 | 0.179 | 0.069 | 0.051 | 0.050 | 0.052 |
| | | | | 627 | 0.064 | 0.073 | 0.055 | 0.050 | 0.050 | 0.050 | 0.065 | 0.074 | 0.058 | 0.054 | 0.054 | 0.054 |
| 2 | 10% | 2 | 19 | 144 | 0.085 | 0.093 | 0.077 | 0.048 | 0.047 | 0.048 | 0.087 | 0.095 | 0.078 | 0.050 | 0.051 | 0.050 |
| | | | | 627 | 0.056 | 0.057 | 0.054 | 0.048 | 0.048 | 0.048 | 0.059 | 0.060 | 0.059 | 0.052 | 0.052 | 0.053 |
| | 25% | 5 | 16 | 144 | 0.103 | 0.125 | 0.083 | 0.050 | 0.050 | 0.050 | 0.106 | 0.125 | 0.087 | 0.051 | 0.051 | 0.051 |
| | | | | 627 | 0.058 | 0.062 | 0.055 | 0.049 | 0.050 | 0.049 | 0.066 | 0.069 | 0.063 | 0.056 | 0.056 | 0.056 |
| | 50% | 10 | 11 | 144 | 0.160 | 0.216 | 0.109 | 0.077 | 0.077 | 0.077 | 0.149 | 0.205 | 0.099 | 0.067 | 0.067 | 0.067 |
| | | | | 627 | 0.254 | 0.271 | 0.237 | 0.225 | 0.226 | 0.225 | 0.172 | 0.187 | 0.158 | 0.149 | 0.149 | 0.149 |
| | All | 12 | 9 | 144 | 0.119 | 0.175 | 0.068 | 0.050 | 0.050 | 0.051 | 0.116 | 0.179 | 0.068 | 0.051 | 0.051 | 0.052 |
| | | | | 627 | 0.064 | 0.073 | 0.055 | 0.051 | 0.051 | 0.051 | 0.065 | 0.074 | 0.058 | 0.054 | 0.054 | 0.054 |
| 4 | 10% | 2 | 19 | 144 | 0.085 | 0.092 | 0.077 | 0.047 | 0.047 | 0.047 | 0.087 | 0.095 | 0.077 | 0.051 | 0.051 | 0.051 |
| | | | | 627 | 0.055 | 0.057 | 0.053 | 0.048 | 0.048 | 0.049 | 0.060 | 0.061 | 0.058 | 0.053 | 0.053 | 0.053 |
| | 25% | 5 | 16 | 144 | 0.104 | 0.125 | 0.083 | 0.050 | 0.051 | 0.050 | 0.106 | 0.126 | 0.086 | 0.051 | 0.051 | 0.051 |
| | | | | 627 | 0.058 | 0.061 | 0.055 | 0.049 | 0.049 | 0.049 | 0.066 | 0.069 | 0.063 | 0.055 | 0.056 | 0.055 |
| | 50% | 10 | 11 | 144 | 0.116 | 0.165 | 0.075 | 0.050 | 0.050 | 0.051 | 0.119 | 0.168 | 0.076 | 0.050 | 0.050 | 0.050 |
| | | | | 627 | 0.064 | 0.069 | 0.056 | 0.051 | 0.051 | 0.051 | 0.064 | 0.070 | 0.058 | 0.054 | 0.054 | 0.054 |
| | All | 12 | 9 | 144 | 0.119 | 0.175 | 0.068 | 0.050 | 0.051 | 0.051 | 0.116 | 0.179 | 0.069 | 0.051 | 0.051 | 0.052 |
| | | | | 627 | 0.064 | 0.073 | 0.055 | 0.050 | 0.050 | 0.050 | 0.065 | 0.074 | 0.058 | 0.054 | 0.054 | 0.054 |
| 5 | 10% | 2 | 19 | 144 | 0.085 | 0.092 | 0.077 | 0.047 | 0.047 | 0.047 | 0.087 | 0.095 | 0.078 | 0.051 | 0.051 | 0.051 |
| | | | | 627 | 0.055 | 0.056 | 0.053 | 0.048 | 0.048 | 0.049 | 0.060 | 0.061 | 0.058 | 0.053 | 0.053 | 0.053 |
| | 25% | 5 | 16 | 144 | 0.104 | 0.125 | 0.083 | 0.050 | 0.051 | 0.050 | 0.106 | 0.126 | 0.086 | 0.051 | 0.051 | 0.051 |
| | | | | 627 | 0.058 | 0.061 | 0.055 | 0.049 | 0.050 | 0.049 | 0.066 | 0.069 | 0.063 | 0.055 | 0.055 | 0.055 |
| | 50% | 10 | 11 | 144 | 0.117 | 0.164 | 0.075 | 0.050 | 0.050 | 0.051 | 0.118 | 0.168 | 0.077 | 0.050 | 0.050 | 0.051 |
| | | | | 627 | 0.066 | 0.072 | 0.058 | 0.053 | 0.054 | 0.053 | 0.066 | 0.073 | 0.059 | 0.056 | 0.055 | 0.055 |
| | All | 12 | 9 | 144 | 0.119 | 0.175 | 0.069 | 0.050 | 0.051 | 0.051 | 0.116 | 0.179 | 0.069 | 0.051 | 0.051 | 0.052 |
| | | | | 627 | 0.064 | 0.073 | 0.055 | 0.050 | 0.050 | 0.050 | 0.065 | 0.074 | 0.058 | 0.054 | 0.054 | 0.054 |

Table 14: The probabilities of rejecting the null-hypothesis in (13) when $H_1$ is true for the spanning tests, and under the assumption of having normally distributed and t-distributed disturbances, respectively, which were obtained based on 10,000 simulations. $T = 144$ represents that the results were obtained by using monthly returns, and $T = 627$ by using weekly returns.

| | | | | | **Multivariate normal distribution** | | | | | | **Multivariate t-distribution** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | **Asymptotic test** | | | **Exact test** | | | **Asymptotic test** | | | **Exact test** | | |
| **Case** | **Screening** | **N** | **K** | **T** | **LR** | **W** | **LM** | **LR** | **W** | **LM** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| 8 | 10% | 2 | 19 | 144 | 0.085 | 0.092 | 0.077 | 0.047 | 0.047 | 0.047 | 0.087 | 0.095 | 0.077 | 0.051 | 0.051 | 0.051 |
| | | | | 627 | 0.055 | 0.057 | 0.053 | 0.048 | 0.048 | 0.049 | 0.060 | 0.061 | 0.058 | 0.053 | 0.053 | 0.053 |
| | 25% | 5 | 16 | 144 | 0.104 | 0.125 | 0.083 | 0.050 | 0.051 | 0.050 | 0.106 | 0.126 | 0.086 | 0.051 | 0.051 | 0.051 |
| | | | | 627 | 0.058 | 0.061 | 0.055 | 0.049 | 0.049 | 0.049 | 0.066 | 0.069 | 0.063 | 0.055 | 0.056 | 0.055 |
| | 50% | 10 | 11 | 144 | 0.116 | 0.165 | 0.075 | 0.050 | 0.050 | 0.051 | 0.118 | 0.168 | 0.076 | 0.050 | 0.050 | 0.050 |
| | | | | 627 | 0.064 | 0.070 | 0.056 | 0.051 | 0.051 | 0.052 | 0.064 | 0.070 | 0.058 | 0.054 | 0.053 | 0.054 |
| | All | 12 | 9 | 144 | 0.119 | 0.175 | 0.068 | 0.050 | 0.051 | 0.051 | 0.116 | 0.179 | 0.069 | 0.051 | 0.051 | 0.052 |
| | | | | 627 | 0.064 | 0.073 | 0.055 | 0.050 | 0.050 | 0.050 | 0.065 | 0.074 | 0.058 | 0.054 | 0.054 | 0.054 |
| 9 | 10% | 2 | 19 | 144 | 0.085 | 0.093 | 0.077 | 0.048 | 0.047 | 0.048 | 0.087 | 0.095 | 0.078 | 0.050 | 0.051 | 0.050 |
| | | | | 627 | 0.056 | 0.057 | 0.054 | 0.048 | 0.048 | 0.048 | 0.059 | 0.061 | 0.058 | 0.052 | 0.052 | 0.053 |
| | 25% | 5 | 16 | 144 | 0.103 | 0.125 | 0.083 | 0.050 | 0.050 | 0.050 | 0.106 | 0.125 | 0.087 | 0.050 | 0.051 | 0.051 |
| | | | | 627 | 0.058 | 0.062 | 0.055 | 0.050 | 0.050 | 0.049 | 0.065 | 0.069 | 0.063 | 0.056 | 0.056 | 0.056 |
| | 50% | 10 | 11 | 144 | 0.159 | 0.213 | 0.107 | 0.076 | 0.076 | 0.076 | 0.147 | 0.204 | 0.098 | 0.067 | 0.066 | 0.067 |
| | | | | 627 | 0.245 | 0.262 | 0.229 | 0.216 | 0.217 | 0.216 | 0.167 | 0.181 | 0.153 | 0.145 | 0.145 | 0.145 |
| | All | 12 | 9 | 144 | 0.119 | 0.175 | 0.068 | 0.050 | 0.050 | 0.051 | 0.116 | 0.179 | 0.069 | 0.051 | 0.051 | 0.052 |
| | | | | 627 | 0.064 | 0.073 | 0.055 | 0.051 | 0.051 | 0.051 | 0.065 | 0.074 | 0.058 | 0.054 | 0.054 | 0.054 |

Table 15: The probabilities of rejecting the null-hypothesis in (13) when $H_1$ is true for the spanning tests. These probabilities were obtained based on 10,000 residual bootstrap simulations. $T = 144$ represents that the results were obtained by using monthly returns, and $T = 627$ by using weekly returns.

| | | | | | **10,000 bootstrap simulations** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | **Asymptotic test** | | | **Exact test** | | |
| **Case** | **Screening** | **N** | **K** | **T** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| 1 | 10% | 2 | 19 | 144 | 0.084 | 0.092 | 0.078 | 0.049 | 0.050 | 0.050 |
| | | | | 627 | 0.063 | 0.064 | 0.062 | 0.056 | 0.056 | 0.056 |
| | 25% | 5 | 16 | 144 | 0.110 | 0.133 | 0.090 | 0.055 | 0.055 | 0.056 |
| | | | | 627 | 0.058 | 0.061 | 0.055 | 0.050 | 0.050 | 0.050 |
| | 50% | 10 | 11 | 144 | 0.127 | 0.179 | 0.084 | 0.057 | 0.058 | 0.057 |
| | | | | 627 | 0.066 | 0.073 | 0.059 | 0.054 | 0.055 | 0.054 |
| | All | 12 | 9 | 144 | 0.128 | 0.192 | 0.076 | 0.058 | 0.057 | 0.057 |
| | | | | 627 | 0.068 | 0.076 | 0.059 | 0.055 | 0.056 | 0.055 |
| 2 | 10% | 2 | 19 | 144 | 0.092 | 0.101 | 0.086 | 0.055 | 0.055 | 0.055 |
| | | | | 627 | 0.164 | 0.167 | 0.162 | 0.151 | 0.150 | 0.151 |
| | 25% | 5 | 16 | 144 | 0.129 | 0.155 | 0.107 | 0.070 | 0.069 | 0.070 |
| | | | | 627 | 0.313 | 0.323 | 0.303 | 0.284 | 0.285 | 0.284 |
| | 50% | 10 | 11 | 144 | 0.153 | 0.204 | 0.105 | 0.074 | 0.074 | 0.074 |
| | | | | 627 | 0.364 | 0.384 | 0.346 | 0.333 | 0.334 | 0.332 |
| | All | 12 | 9 | 144 | 0.137 | 0.183 | 0.096 | 0.061 | 0.061 | 0.061 |
| | | | | 627 | 0.297 | 0.317 | 0.283 | 0.268 | 0.269 | 0.267 |
| 4 | 10% | 2 | 19 | 144 | 0.084 | 0.092 | 0.078 | 0.049 | 0.050 | 0.050 |
| | | | | 627 | 0.063 | 0.064 | 0.061 | 0.056 | 0.056 | 0.056 |
| | 25% | 5 | 16 | 144 | 0.110 | 0.133 | 0.090 | 0.055 | 0.055 | 0.056 |
| | | | | 627 | 0.058 | 0.061 | 0.055 | 0.050 | 0.050 | 0.050 |
| | 50% | 10 | 11 | 144 | 0.127 | 0.179 | 0.084 | 0.057 | 0.058 | 0.058 |
| | | | | 627 | 0.066 | 0.073 | 0.060 | 0.055 | 0.055 | 0.055 |
| | All | 12 | 9 | 144 | 0.128 | 0.192 | 0.076 | 0.058 | 0.057 | 0.057 |
| | | | | 627 | 0.068 | 0.076 | 0.059 | 0.056 | 0.056 | 0.056 |
| 5 | 10% | 2 | 19 | 144 | 0.084 | 0.092 | 0.077 | 0.050 | 0.051 | 0.050 |
| | | | | 627 | 0.065 | 0.066 | 0.063 | 0.057 | 0.057 | 0.058 |
| | 25% | 5 | 16 | 144 | 0.110 | 0.132 | 0.091 | 0.055 | 0.055 | 0.056 |
| | | | | 627 | 0.061 | 0.065 | 0.058 | 0.053 | 0.053 | 0.053 |
| | 50% | 10 | 11 | 144 | 0.126 | 0.180 | 0.085 | 0.057 | 0.058 | 0.057 |
| | | | | 627 | 0.070 | 0.076 | 0.062 | 0.058 | 0.058 | 0.058 |
| | All | 12 | 9 | 144 | 0.117 | 0.162 | 0.078 | 0.050 | 0.049 | 0.050 |
| | | | | 627 | 0.065 | 0.072 | 0.059 | 0.055 | 0.055 | 0.055 |

Table 16: The probabilities of rejecting the null-hypothesis in (13) when $H_1$ is true for the spanning tests. These probabilities were obtained based on 10,000 residual bootstrap simulations. $T$ = 144 represents that the results were obtained by using monthly returns, and $T$ = 627 by using weekly returns.

| | | | | | **10,000 bootstrap simulations** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | **Asymptotic test** | | | **Exact test** | | |
| **Case** | **Screening** | **N** | **K** | **T** | **LR** | **W** | **LM** | **LR** | **W** | **LM** |
| 8 | 10% | 2 | 19 | 144 | 0.084 | 0.092 | 0.078 | 0.049 | 0.050 | 0.050 |
| | | | | 627 | 0.063 | 0.064 | 0.062 | 0.056 | 0.056 | 0.056 |
| | 25% | 5 | 16 | 144 | 0.110 | 0.133 | 0.090 | 0.055 | 0.055 | 0.056 |
| | | | | 627 | 0.058 | 0.061 | 0.055 | 0.050 | 0.049 | 0.050 |
| | 50% | 10 | 11 | 144 | 0.127 | 0.179 | 0.084 | 0.057 | 0.058 | 0.058 |
| | | | | 627 | 0.066 | 0.073 | 0.059 | 0.055 | 0.055 | 0.054 |
| | All | 12 | 9 | 144 | 0.128 | 0.193 | 0.076 | 0.058 | 0.057 | 0.057 |
| | | | | 627 | 0.068 | 0.076 | 0.059 | 0.055 | 0.056 | 0.055 |
| 9 | 10% | 2 | 19 | 144 | 0.092 | 0.100 | 0.085 | 0.054 | 0.055 | 0.055 |
| | | | | 627 | 0.161 | 0.164 | 0.158 | 0.147 | 0.147 | 0.147 |
| | 25% | 5 | 16 | 144 | 0.129 | 0.155 | 0.107 | 0.069 | 0.069 | 0.069 |
| | | | | 627 | 0.303 | 0.312 | 0.294 | 0.274 | 0.275 | 0.274 |
| | 50% | 10 | 11 | 144 | 0.153 | 0.205 | 0.104 | 0.074 | 0.073 | 0.073 |
| | | | | 627 | 0.353 | 0.373 | 0.336 | 0.322 | 0.323 | 0.320 |
| | All | 12 | 9 | 144 | 0.137 | 0.183 | 0.095 | 0.061 | 0.060 | 0.061 |
| | | | | 627 | 0.287 | 0.302 | 0.271 | 0.259 | 0.259 | 0.257 |