LUND UNIVERSITY
School of Economics and Management

# A critical review of the global minimum variance theory

## Introduction of an alternative approach based on sector indices

Bachelor thesis in Statistics, 15 ECTS
Department of Statistics, Lund University

Supervisor: Krzysztof Podgórski

Author: Martin Claeson
June, 2016

## Abstract

The main purpose of this thesis is to give a basic understanding of the GMV portfolio theory and the problematics that arise when using the sample covariance matrix as the only parameter. The reason for this is the amount of estimation error that tends to increase as the sample covariance matrix goes to a higher dimension. In an attempt to reduce the amount of error, an alternative approach based on sector indices is introduced, which gives new and interesting results. This is a useful approach, since we are explaining the chosen stocks with fewer time series, a smaller dimension of covariance matrix needs to be estimated. This thesis lay the ground for this basic strategy, however, before any more profound conclusions can be drawn, further investigations have to be made.

# Contents

# 1 Introduction

Diversification is an important part of portfolio theory and risk analysis. The main idea is that the dependence between different financial assets within a portfolio should be as small as possible to minimize the risk. There are many different methods to diversify your portfolio that are commonly used in finance, for example to diversify over different business sectors. This theory is useful, as stocks within the same branch tend to show high signs of dependence. With this knowledge, the investor is put in front of some dilemmas, primarily how to pick the best suitable stocks within each branch, but also how to weight the portfolio wealth. The problems with stock picking, can simply be solved by investing in whole sectors. For the latter, modern portfolio theory (MPT), introduced by Markowitz (1952), is an effective approach to find the optimal way to weight the portfolio, i.e. to maximize the expected return at a given level of risk. The concept is straightforward and easy to assimilate. The only needed parameters are the portfolio assets means, variances and covariances, which makes this approach quite simple.

One important part of the MPT is the Global Minimum Variance (GMV) theory, which is the optimal portfolio in the concept of risk. The calculations of the GMV theory are basic and the only needed parameter is the covariance matrix based on the assets returns, i.e. if the true covariance matrix is available, investors can find the portfolio with the smallest variance possible. However, when calculations are based on samples of historical asset returns, the covariance matrix is not given and need to be estimated. That is easily done with the sample covariance matrix, but not without some major flaws. As Jobson and Korkie (1980) concluded, the sample covariance matrix contains a high amount of estimation error. This leads to an uncertainty if the GMV portfolio, based on the sample covariance matrix, actually is the portfolio with the least amount of variance. Because of these problematics, Ledoit and Wolf (2004) established a technique to shrink the covariance matrix, and therefore the error, with some good results.

In this thesis these problematics are described and an alternative and simpler approach based on sector indices is developed. The basic idea is to shrink the dimension of the sample covariance matrix, and therefore hopefully the amount of error. Sector indices are a compound of companies active in the same branch. As a complement to calculate the GMV portfolio over a high amount of stocks, with the need of a high dimensional sample covariance matrix, one can explain the same stocks with indices. With this approach, fewer time series are needed, which leads to a lower dimension of the sample covariance matrix and hopefully less amount of estimation error.

There are some questions that we are especially interested in analyzing.

1. Can one show, by the usage of sector indices, an alternative approach to the original GMV-strategy, with smaller amount of estimation error but still well diversified?
2. How much does the estimation error of the sample covariance matrix influence the portfolio risk level?

To be able to answer these questions, three different strategies is produced and compared with a simple test. The data set are based on Swedish market data, containing daily index- and stock returns in the time span of 2015 and 2016, which is divided in to two parts; the first for calculating the GMV portfolios and the second to test and compare their risk level after investment. Additionally, for closer analytical purposes, a data simulation is made as well. With this approach, we are able to get a deeper knowledge of the dilemma in estimating the covariance matrix.

This thesis is organized as follows. In section two we describe the data set. Section three contains a deeper description of the used theory. In the fourth section the method is presented, followed up by the results and the conclusion in the fifth respectively sixth section.

# 2 Data

The index data are collected from the webpage of the Swedish stock exchange, Nasdaq OMX Nordic. These financial instruments are provided for analytical purposes and cannot be traded. The sample data ranging from 17 of June 2015 to 15 of April 2016, and contains a total of 187 daily observations. The reason for this time span is that one of the stocks, Nordax, were IPOed the 17 of June 2015. The indices that will be used in this study are summoned in Table 1 below.

TABLE 1 – SECTOR INDICES

| Index | Symbol | Abbreviation |
|---|---|---|
| Bank GI | SX8300GI | Bank |
| Forestry and Paper GI | SX1730GI | FaP |
| Mining GI | SX1770GI | Mining |

The indices are so called gross index, which means that dividends are included in the index series. The formula for the index calculations, provided from the OMX Nordic, is found in the Appendix A.

The companies within each index are represented in Table 2 below.

TABLE 2 – STOCKS WITHIN EACH SECTOR

| Bank GI | FaP GI | Mining GI |
|---|---|---|
| Nordea Bank | Arctic Paper | Boliden |
| Nordax Group | BillerudKorsnäs | Endomines |
| SEB A | Bergs Timber B | Lucara Diamond Corp |
| SEB C | Holmen A | Lundin Gold |
| Sv. Handelsbanken A | Holmen B | NGEx Resourses |
| Sv. Handelsbanken B | Munksjö Oyj | Nordic Mines |
| Swedbank A | Rottneros | Semafo |
| | Stora Enso A | |
| | Stora Enso R | |

The data are brought from the homepage of Handelsbanken and is daily closing prices, adjusted for splits and dividends, from 17 of June to 15 of April 2016. As in the case of the indices, there are 187 observations.

The closing prices for some stocks are not declared for all trading days during this time span. If some dates are missing for one stock, this day have been deleted for all the assets in the data.

# 3 Theory

## 3.1 Logarithmic returns

To be able to compare different financial assets we will transform the time series of price changes into logarithmic returns. Due to the fact that one cannot invest in the indices described in Table 1 above, we will call these logarithmic quasi-returns.

The definition of logarithmic returns is the following:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(P_t) - \log(P_{t-1}) \tag{1}$$

where $P_t$ is the price at time t and $P_{t-1}$ is the price at time t-1.

We are looking into N column vectors of assets over T rows of daily observations. Therefore, $\boldsymbol{R} = [\boldsymbol{r}_1 \dots \boldsymbol{r}_N]$ denotes an TxN dimensional matrix of asset log returns.

## 3.2 Modern portfolio theory

As pointed out earlier Markowitz (1952) laid the ground for today's modern portfolio theory. With the usage of multivariate data and the jointly normal distribution assumption, the only needed parameters are the mean, variance and covariance of the assets within the portfolio.

An efficient portfolio is the one that maximizes the expected return at a given level of risk, measured as variance, i.e. at all the different levels of risk, there is an efficient portfolio. Together, these portfolios form the efficient frontier. The portfolio on the left tip of the efficient frontier goes by the name, Global Minimum Variance portfolio. The GMV has an important ability compared to the rest of the efficient portfolios, namely, it is completely calculated of the covariance matrix, when the rest of the portfolios take the expected returns into account as well.

So why is this important? Due to the fact that neither the mean, variance, nor the covariance is given; it needs to be estimated. If the investor does not have a better guess, an estimation based on historical data is commonly used. Consequently, the theory is built upon past economic events. Due to the efficient market hypothesis, all assets are priced accordingly to the available news, i.e. all future returns will be random and therefore hard to estimate. The covariance matrix, however, is easier to estimate. Risky assets tend to continue to show high amount of volatility and vice versa. If the investor thinks that the past volatility is going to continue together with similar dependence between the stocks, the covariance matrix is a quite good estimation. But, as concluded in the introduction, the covariance matrix is not estimated without error. On this subject, Chopra and Ziemba (1993) as well as Frahm (2004), concluded that the main problematic arises when estimating the expected return rather than the covariance matrix. However, in this thesis we will mainly focus on the latter.

## 3.3 Global minimum variance portfolio

The core idea of the GMV theory is to find the asset weights that minimize the portfolio variance, given the assets covariance matrix. There are no constraints, therefore, negative weights might appear, i.e. short sales are allowed. Furthermore, there are no restrictions in how much wealth that will be invested in each asset and no transactions cost are taken into account.

As Ruppert and Matteson (2015) stated, quadratic programming of the GMV will be easily calculated. The minimization problem is the following:

$$\boldsymbol{w}_{GMV} = argmin\{\boldsymbol{w}^T \Sigma \boldsymbol{w}; \boldsymbol{w}^T \mathbf{1}_N = 1\}, \tag{2}$$

where $\boldsymbol{w} = (w_1, \dots, w_N)^T$ is a vector of portfolio weights and $\mathbf{1}_N$ is a $N$ dimensional vector of ones.

By the usage of the Lagrange multiplier, the problem of finding the optimal $\boldsymbol{w}$ is easily solved,

$$L(\boldsymbol{w}, \lambda) = \boldsymbol{w}^T \Sigma \boldsymbol{w} + \lambda(\boldsymbol{w}^T \mathbf{1}_N - 1), \tag{3}$$

where $\lambda$ is a constant.

Derive the Lagrange formula with respect to $\boldsymbol{w}$ and $\lambda$:

$$0 = \frac{\partial L(\boldsymbol{w}, \lambda)}{\partial \boldsymbol{w}} = 2\Sigma \boldsymbol{w} + \mathbf{1}_N \lambda \tag{4}$$

and

$$0 = \frac{\partial L(\boldsymbol{w}, \lambda)}{\partial \lambda} = \boldsymbol{w}^T \mathbf{1}_N - 1. \tag{5}$$

Then use the formula (4) to solve $\boldsymbol{w}$,

$$\boldsymbol{w} = -\frac{\lambda \Sigma^{-1} \mathbf{1}_N}{2}. \tag{6}$$

Multiply formula (6) with $\mathbf{1}_N^T$ on both sides, use the equality of 1 in formula (5) and solve $\lambda$,

$$1 = -\frac{\lambda \mathbf{1}_N^T \Sigma^{-1} \mathbf{1}_N}{2} \Leftrightarrow \lambda = -2 \cdot \frac{1}{\mathbf{1}_N^T \Sigma^{-1} \mathbf{1}_N}. \quad (7)$$

Solve $\boldsymbol{w}$ by substitute $\lambda$ back to formula (6):

$$\boldsymbol{w} = -\frac{1}{2} \cdot (-2) \cdot \frac{\Sigma^{-1} \mathbf{1}_N}{\mathbf{1}_N^T \Sigma^{-1} \mathbf{1}_N} = \frac{\Sigma^{-1} \mathbf{1}_N}{\mathbf{1}_N^T \Sigma^{-1} \mathbf{1}_N}. \quad (8)$$

The vector of global minimum weights, $\boldsymbol{w}_{GMV} = \left(w_{GMV,1}, \ldots, w_{GMV,N)}\right)^T$, that minimize the variance for a given covariance matrix are the following:

$$\boldsymbol{w}_{GMV} = \frac{\Sigma^{-1} \mathbf{1}_N}{\mathbf{1}_N^T \Sigma^{-1} \mathbf{1}_N}. \quad (9)$$

and the global minimum variance of the portfolio is calculated through,

$$\sigma_{mv}^2 = \boldsymbol{w}_{GMV}^T \Sigma \boldsymbol{w}_{GMV}. \quad (10)$$

The GMV portfolio return is calculated through

$$\boldsymbol{R}_{GMV} = \boldsymbol{R} \boldsymbol{w}_{GMV}, \quad (11)$$

where $\boldsymbol{R} = [\boldsymbol{r}_1 \ldots \boldsymbol{r}_N]$ is a matrix of TxN asset returns. If the true covariance matrix is available, the weights, $\boldsymbol{w}_{GMV}$, will be fixed, i.e. the only variability in $\boldsymbol{R}_{GMV}$ appears due to the randomness of the returns, $\boldsymbol{R}$.

The presented formulas presume that the true $\Sigma$ is known. As pointed out in the introduction, when using samples of historical data, the true covariance matrix will not be available. Therefore, the covariance matrix needs to be estimated. This is easily done with the sample covariance matrix,

$$\boldsymbol{S} = \frac{1}{N-1} \sum_{i=1}^{N} (\boldsymbol{r}_i - \bar{\boldsymbol{r}})(\boldsymbol{r}_i - \bar{\boldsymbol{r}})^T. \quad (12)$$

Replacing the true covariance matrix in equation- (9) and (10) with the sample counterpart give us the following formulas:

$$\widehat{\boldsymbol{w}}_{GMV} = \frac{\boldsymbol{S}^{-1}\mathbf{1}_N}{\mathbf{1}_N^T\boldsymbol{S}^{-1}\mathbf{1}_N} \tag{13}$$

and

$$s_{GMV}^2 = \widehat{\boldsymbol{w}}_{GMV}^T \boldsymbol{S} \widehat{\boldsymbol{w}}_{GMV} \,, \tag{14}$$

which are the estimated parameters. The portfolio return is calculated in the same way as in formula (11), but now with the estimated weights,

$$\boldsymbol{R}_{GMV} = \boldsymbol{R}\widehat{\boldsymbol{w}}_{GMV} \,. \tag{15}$$

Because of the usage of the sample covariance matrix, $\widehat{\boldsymbol{w}}_{GMV}$ is no longer fixed but random, i.e. the variability in $\boldsymbol{R}_{GMV}$ appears from both the randomness in $\boldsymbol{R}$ as well as in $\widehat{\boldsymbol{w}}_{GMV}$.

# 4 Method

This section is divided into two parts; the first describes the approaches that the analysis is built upon, Swedish market data and simulated data. In the second part, a deeper review of the strategies structure can be found.

## 4.1 Approaches

### 4.1.1 Approach 1 – Swedish market asset data

As stated in the data-section above, there are a total of 187 daily observations. The first 123 observations of logarithmic returns, covering the 17 of June to 31 of December 2015, are used to calculate the different GMV portfolios. The remaining 65 dates, covering 1 of January to 15 of April 2016, are the test data, which give us an insight in the portfolios imaginary performances

### 4.1.2 Approach 2 – Simulated data

As described in the introduction, the true covariance matrix is not available when using a sample of historical returns. This uncertainty makes the measurement of estimation error in the sample covariance matrix a complex problem. To get a deeper knowledge of this dilemma, we simulate multivariate normally distributed data with the Swedish market data as the true parameters, i.e. the estimated-means and covariance matrix conducted in sector 4.1.1, is now the population- means and covariance matrix. By the use of the *mvrnorm* function located in the MASS package in *R*, we are able to simulate the requested data set.

250 daily returns, equivalent to a trading year, are simulated and used to calculate the different strategies, explained in section 4.2 below. To test the portfolios imaginary performances, an additional sample of 250 daily returns are drawn and invested in. Because of the randomness in the simulation, this procedure is repeated 20 times.

#### *4.1.2.1 Simulated sector indices*

We use the assumptions of joint normally distribution and i.i.d (independent and individually distributed) assets to simulate three index series with the same mean- and covariance structure as the Swedish market index data.

$$\mathbb{Y} = [\mathbb{y}_1 \; \mathbb{y}_2 \; \mathbb{y}_3] \rightarrow N\begin{pmatrix} \mu_{\boldsymbol{Y}_1} \\ \mu_{\boldsymbol{Y}2} , \boldsymbol{\Sigma}_{\boldsymbol{Y}_{1,2,3}} \\ \mu_{\boldsymbol{Y}_3} \end{pmatrix}$$

where $\mathbb{Y}$ is a matrix of Tx3 dimensional matrix of logarithmic quasi-returns. $\mu_{\boldsymbol{Y}_i}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Y}_{1,2,3}}$ are the true-means and covariance matrix for the Swedish market sector indices. Due to these parameters, the estimated indices should approximately agree with the ones observed in the true data.

#### *4.1.2.2 Simulated stocks*

To be able to simulate the stocks, the same theory is used as in section 4.1.2.1, but with the parameters of the 23 stocks summarized in Table 2.

$$\mathbb{X} = [\mathbb{x}_1 \ \dots \ \mathbb{x}_{23}] \rightarrow N\begin{pmatrix} \mu_{X_1} \\ \vdots \\ \mu_{X_{23}} \end{pmatrix}, \boldsymbol{\Sigma}_{\boldsymbol{X}_{1,..,23}}$$

where $\mathbb{X}$ is a matrix of dimension Tx23. Just like in the simulation of the sector indices, the Swedish stock data means and covariance matrix is now the true parameters. The R-script used for this simulation is presented in Appendix C.

## 4.2 Portfolio strategies

We will describe two different GMV strategies; one based on indices and the other on stocks. The GMV portfolios weights are based on the sample covariance matrix, estimated on the logarithmic returns of previous year. For comparison reasons, a third strategy, with equally weighted weights, will be produced as well.

### 4.2.1 Strategy 1 – GMV portfolio based on indices

The first strategy is based on the GMV theory, calculated on sector indices. To begin with, $\boldsymbol{Y} = [\boldsymbol{y}_1 \ \boldsymbol{y}_2 \ \boldsymbol{y}_3]$ is an Tx3 dimensional matrix of sector index logarithmic quasi-returns. With formula (12) the sample covariance matrix, $\boldsymbol{S}_Y$, of the three indices are estimated. Based on this parameter, the GMV portfolio weights are calculated with formula (13):

$$\widehat{\boldsymbol{w}}_{P1_Y,j} = \frac{\boldsymbol{S}_Y^{-1}\boldsymbol{1}_N}{\boldsymbol{1}_N^T\boldsymbol{S}_Y^{-1}\boldsymbol{1}_N}, \tag{16}$$

where $P1_Y$, j stands for portfolio 1, approach j.

One need to invest in the underlying stocks, as sector indices are issued for analytical purposes and, therefore, cannot be invested in directly. We want the strategy as simple as possible, which means that the stocks are weighted equally within each index. This is done by dividing the sector indices weights, $\widehat{\boldsymbol{w}}_{P1_{Y}.j}$, with the amount of stocks within each index. The vector of weights containing all the stocks is described as, $\widehat{\boldsymbol{w}}_{P1_X,j}$. The portfolio variance is calculated with an alternative to formula (14),

$$s_{P1,j}^2 = \widehat{\boldsymbol{w}}_{P1_X,j}^{\boldsymbol{T}} \boldsymbol{S}_X \widehat{\boldsymbol{w}}_{P1_X,j} \,, \tag{17}$$

where $\boldsymbol{S}_X$ is the sample covariance matrix estimated on the underlying stocks.

To calculate the estimated portfolio return, formula (15) is used,

$$\boldsymbol{R}_{P1,j} = \boldsymbol{R}_X \widehat{\boldsymbol{w}}_{P1_X,j} \,, \tag{18}$$

where $\boldsymbol{R}_X$ is the returns of the stocks during the test period.

### 4.2.2 Strategy 2 – GMV portfolio based on stocks

The second strategy is the well-documented method of global minimum variance portfolio. $\boldsymbol{X} = [\boldsymbol{x}_1 \ \dots \ \boldsymbol{x}_{23}]$ is a Tx23 dimensional matrix of logarithmic stock returns. From this matrix, a sample covariance matrix $\boldsymbol{S}_X$ will be estimated. With formula (13) explained in section 3.3, the vector of weights is calculated,

$$\widehat{\boldsymbol{w}}_{P2,j} = \frac{\boldsymbol{S}_X^{-1} \boldsymbol{1}_N}{\boldsymbol{1}_N^T \boldsymbol{S}_X^{-1} \boldsymbol{1}_N} . \tag{19}$$

The portfolio variance is calculated with formula (14),

$$s_{P2,j}^2 = \widehat{\boldsymbol{w}}_{P2,j}^T \boldsymbol{S}_X \widehat{\boldsymbol{w}}_{P2,j} \,, \tag{20}$$

and the return of the portfolio during the test period is calculated with,

$$\boldsymbol{R}_{P2,j} = \boldsymbol{R}_X \widehat{\boldsymbol{w}}_{P2,j} \,. \tag{21}$$

### 4.2.3 Strategy 3 – Equally weighted portfolio

One of the most basic strategies is to weight the portfolio equally over all assets. This method is a good comparison to the other two strategies, mentioned above, since the weights is not estimated from a covariance matrix. The portfolio variance is calculated in a similar way as in formula (14),

$$s_{P3,j}^2 = \widehat{\boldsymbol{w}}_{P3,j}^{\boldsymbol{T}} \boldsymbol{S}_X \widehat{\boldsymbol{w}}_{P3,j} \,, \tag{22}$$

and the portfolio return during the test period with formula (15)

$$\boldsymbol{R}_{P3,j} = \boldsymbol{R}_X \widehat{\boldsymbol{w}}_{P3,j} \,. \tag{23}$$

## 5.1 Results based on the Swedish financial market

### 5.1.1 Strategy 1 – GMV portfolio based on indices

The first strategy is based on the three sector indices presented in Table 1. From these time series, a 3x3 dimensional covariance is estimated. More precisely, $\frac{3x4}{2}$ = 6 parameters are projected. The estimated covariance- and correlation matrix (within the brackets) is summoned in Table 3.

TABLE 3 – SAMPLE COVARIANCE- AND CORRELATION MATRIX - INDICES

| | Bank GI | FaP GI | Mining GI |
|---|---|---|---|
| Bank GI | 2.24e-04 (1.000) | 1.78e-04 (0.703) | 1.80e-04 (0.578) |
| FaP GI | 1.78e-04 (0.703) | 2.69e-04 (1.000) | 2.12e-04 (0.640) |
| Mining GI | 1.80e-04 (0.578) | 2.12e-04 (0.640) | 4.07e-04 (1.000) |

The GMV portfolio based on indices is calculated with formula (16), the only used parameter is the sample covariance matrix in Table 3. The portfolio weights are presented in Table 4.

TABLE 4 – GMV PORTFOLIO WEIGHTS BASED ON INDICES

| Index | Weight | Amount of stocks |
|---|---|---|
| Bank | 0.55942522 | 7 |
| FaP | 0.34489314 | 9 |
| Mining | 0.09568163 | 7 |
| Sum | 1 | 23 |

Given the sample covariance matrix, estimated on the indices logarithmic quasi-returns during 2015, these are the approximated weights that minimize the portfolio variance. As described in section 4.2.1, it is not optional to invest in the given indices, therefore one must invest in the underlying stocks. After dividing the sector weights equally over the stocks within each index, 0.07991789 is invested in each bank stock, 0.03832146 and 0.0136688 in each FaP - respectively mining stock. The variance of this portfolio is calculated with formula (17), $s_{P1,1}^2 = \hat{\boldsymbol{w}}_{P1,1}^{\boldsymbol{T}} \boldsymbol{S}_X \hat{\boldsymbol{w}}_{P1,1}$ = 1.59e-04.

### 5.1.2 Strategy 2 – GMV portfolio based on stocks

The second strategy is based on the GMV theory, calculated on all 23 stocks presented in Table 2. The estimated covariance matrix is of dimension 23x23, which means that $\frac{23x24}{2} = 276$ variance- and covariance parameters is projected. To give an insight of the dependence between the stocks during the 2015 time span, a correlation plot is available in Figure 1. By the usage of the estimated covariance matrix, the GMV portfolio weights are calculated with formula (19) and summoned in Table 5. The used R-script for these calculations is available in Appendix C.

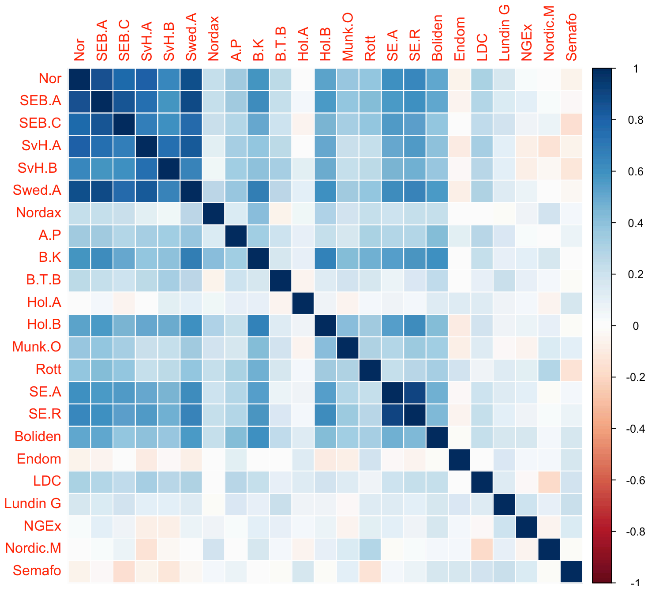

Figure 1- Correlation plot based on stocks

TABLE 5 – GMV PORTFOLIO WEIGHTS BASED ON STOCKS

| Stocks | Weights |
|---|---|
| Nordea Bank | 0.0678160822 |
| SEB A | -0.2380888769 |
| SEB C | 0.3699257467 |
| Sv. Handelsbanken A | 0.0061393756 |
| Sv. Handelsbanken B | 0.0721963906 |
| Swedbank A | 0.0705114280 |
| Nordax | 0.0974433857 |
| Arctic Paper | -0.0082195891 |
| BillerudKorsnäs | -0.0588709108 |
| Bergs Timber B | 0.1020064476 |
| Holmen A | 0.1810335428 |
| Holmen B | 0.2803494868 |
| Munksjö Oyj | 0.0705727350 |
| Rottneros | -0.0057310324 |
| Stora Enso A | 0.0737609819 |
| Stora Enso R | -0.1608781712 |
| Boliden | -0.0683863720 |
| Endomines | 0.0009383792 |
| Lucara Diamond Corp | -0.0114819299 |
| Lundin Gold | 0.0912409261 |
| NGEx Resourses | 0.0083663966 |
| Nordic Mines | -0.0028104056 |
| Semafo | 0.0621659831 |
| Sum | 1 |

Given the estimated covariance matrix, this is approximately the optimal way to weight your portfolio to minimize the risk. By the usage of these weights and the sample covariance matrix, the minimized portfolio variance is calculated with formula (20), $s^2_{P2,1} = \widehat{\boldsymbol{w}}^{\boldsymbol{T}}_{P2,1} \boldsymbol{S}_X \widehat{\boldsymbol{w}}_{P2,1}$ = 7.19e-05.

### 5.1.3 Strategy 3 – Equally weighted portfolio

The third strategy is the simplest. The weights are equally distributed over all the underlying stocks. This means that $\frac{1}{23} = 0.04347826$, is invested in each asset. The portfolio variance is calculated with formula (22), $s^2_{P3,1} = \widehat{\boldsymbol{w}}^{\boldsymbol{T}}_{P3,1} \boldsymbol{S}_X \widehat{\boldsymbol{w}}_{P3,1}$ = 2.17e-04.

### 5.1.4 Comparison

As described in section 4.1.1, 65 daily stock returns during 2016 are used to calculate respectively portfolio return, $\boldsymbol{R}_{Pi,1}$. The variance of the daily portfolio returns for the three strategies during the test period is presented in Table 7. For a graphical view, the portfolio returns are plotted in Figure 2

TABLE 7 – COMPARISON OF VARIANCES, BEFORE AND AFTER THE TEST

| | Portfolio returns variance, Var($\boldsymbol{R}_{\boldsymbol{Pi,1}}$) | $\boldsymbol{s^2_{Pi,1}}$ | Difference | % change |
|---|---|---|---|---|
| Strategy 1 | 1.76e-04 | 1.59e-04 | 1.7e-05 | ≈ + 10.7 % |
| Strategy 2 | 1.21e-04 | 7.19e-05 | 4.9e-05 | ≈ + 68.3 % |
| Strategy 3 | 1.24e-04 | 2.17e-04 | -9.3e-05 | ≈ - 42.9 % |

For comparison reasons, the portfolio variance based on historical data, $s^2_{Pi,1}$, and the difference between the figures are presented in Table 7 as well.

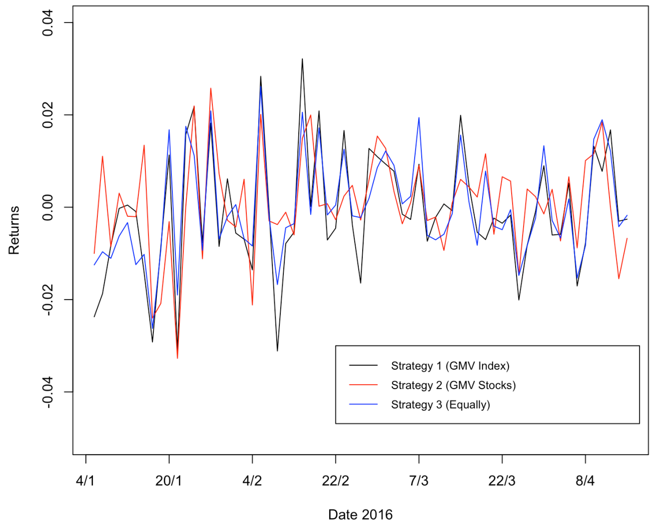


Figure 2- Strategy comparison

## 5.2 Results based on simulated data

By the second approach, described under section 4.1.2, three jointly normal distributed index series are simulated. The used parameters are the Swedish market indices means presented in Table 14 and the covariance matrix summarized in Table 3. From the simulated index time series, the means and the sample covariance matrix are estimated and summoned in Table 8 respectively 9.

TABLE 8 – MEANS OF SIMULATED INDICES

| Stocks | Mean |
|---|---|
| Simulated Bank Index | 1.24e-03 |
| Simulated FaP Index | 3.20e-03 |
| Simulated Mining Index | 1.08e-04 |

TABLE 9 – COVARIANCE- AND CORRELATION MATRIX BASED ON SIMULATED INDICES

| | Simulated Bank Index | Simulated FaP Index | Simulated Mining Index |
|---|---|---|---|
| Simulated Bank Index | 2.40e-04 (1.000) | 1.87e-04 (0.730) | 1.92e-04 (0.614) |
| Simulated FaP Index | 1.87e-04 (0.730) | 2.75e-04 (1.000) | 2.26e-04 (0.672) |
| Simulated Mining Index | 1.92e-04 (0.614) | 2.26e-04 (0.672) | 4.10e-04 (1.000) |

In a similar way, as explained in section 4.1.2.2, the stocks are simulated with the parameters from the market stock data. Descriptive data are summoned in Table 16 and a correlation plot over the simulated stocks can be found in Figure 3.

### 5.2.1 Strategy 1 – GMV portfolio based on simulated index

By the usage of the covariance matrix presented in Table 9 and formula (16), the GMV portfolio of the simulated indices is calculated. The results are summoned in Table 10 below. For comparison, the GMV weights calculated on the market data in section 5.1.1 are presented as well. These weights can be analyzed as non-random calculated from the true covariance matrix, as mentioned in section 3.3 above.

TABLE 10 – WEIGHTS BASED ON SIMULATED INDICES

| Index | Weight (Simulated) | Weight (Table 4) | Amount of stocks |
|---|---|---|---|
| Bank simulated | 0.59865864 | 0.55942522 | 7 |
| FaP simulated | 0.32977838 | 0.34489314 | 9 |
| Mining simulated | 0.07156298 | 0.09568163 | 7 |
| Sum | 1 | 1 | 23 |

Similar to the calculations in section 5.1.1, the sector weights are divided upon the amount of underlying stocks. This means that 0.08552266 is invested in each simulated bank stock, 0.03664204 and 0.01022328 in each FaP respectively mining stock. The variance is calculated with formula (17), $s^2_{P1,2} = \boldsymbol{\hat{w}}^{\boldsymbol{T}}_{P1,2}\boldsymbol{S}_{\mathbb{X}}\boldsymbol{\hat{w}}_{P1,2}$ = 1.46e-04.

### 5.2.2 Strategy 2 – GMV portfolio based on simulated stocks

From the time series of simulated stock returns, the covariance matrix is estimated and used for calculating the GMV weights. The result is summoned in Table 11 below. For comparison, the GMV portfolio weights calculated in section 5.1.2 are presented as well.

The variance of the GMV portfolio based on stocks, weighted as in Table 11, is calculated with formula (20), with the following result,
$s^2_{P2,2} = \boldsymbol{\hat{w}}^{T}_{P2,2}\boldsymbol{S}_{\mathbb{X}}\boldsymbol{\hat{w}}_{P2,2}$= 5.67e-05.

### 5.2.3 Strategy 3 – Equally weighted portfolio

The equally weighted portfolio is as simply calculated as under section 5.1.3, i.e. 0.04347826 will be invested in each stock. With formula (22) the portfolio variance is calculated, $s^2_{P3,2} = \boldsymbol{\hat{w}}^{\boldsymbol{T}}_{P3,2}\boldsymbol{S}_{\mathbb{X}}\boldsymbol{\hat{w}}_{P3,2}$ = 2.07e-04.

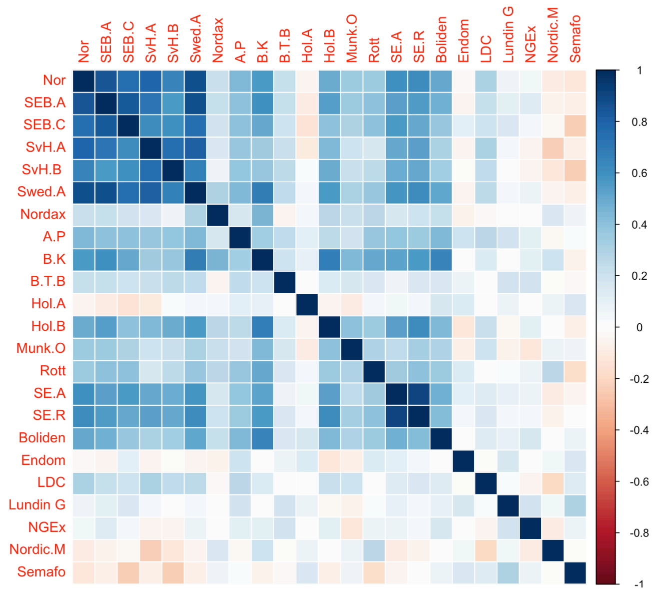

Figure 3 - Correlation plot based on simulated stocks

TABLE 11 – WEIGHTS BASED ON SIMULATED STOCKS

| Stocks | Weights | Weights (Table 5) |
|---|---|---|
| Nordea Bank | 0.027742700 | 0.0678160822 |
| SEB A | -0.253680577 | -0.2380888769 |
| SEB C | 0.281396824 | 0.3699257467 |
| Sv. Handelsbanken A | 0.038397908 | 0.0061393756 |
| Sv. Handelsbanken B | 0.065093903 | 0.0721963906 |
| Swedbank A | 0.206114228 | 0.0705114280 |
| Nordax | 0.111264314 | 0.0974433857 |
| Arctic Paper | -0.022241441 | -0.0082195891 |
| BillerudKorsnäs | -0.077835609 | -0.0588709108 |
| Bergs Timber B | 0.130234666 | 0.1020064476 |
| Holmen A | 0.202204455 | 0.1810335428 |
| Holmen B | 0.228173711 | 0.2803494868 |
| Munksjö Oyj | 0.074752109 | 0.0705727350 |
| Rottneros | 0.036779476 | -0.0057310324 |
| Stora Enso A | 0.066076449 | 0.0737609819 |
| Stora Enso R | -0.170731853 | -0.1608781712 |
| Boliden | -0.079883851 | -0.0683863720 |
| Endomines | -0.005109171 | 0.0009383792 |
| Lucara Diamond Corp | -0.021904246 | -0.0114819299 |
| Lundin Gold | 0.076869232 | 0.0912409261 |
| NGEx Resourses | 0.004190067 | 0.0083663966 |
| Nordic Mines | -0.003503245 | -0.0028104056 |
| Semafo | 0.085599949 | 0.0621659831 |
| Sum | 1 | 1 |

### 5.2.4 Comparison

To approximate respectively portfolio return $\boldsymbol{R}_{Pi,2} = \boldsymbol{R}_{\mathbb{X}}\widehat{\boldsymbol{w}}_{Pi,2}^{T}$, an additional simulation of 250 daily stock returns, with the same parameters as the Swedish market stock data, are made, which is also described in in section 4.1.2. Due to the randomness in the simulation, this test is repeated 20 times. To get an approximation of the portfolio return variance over this test period, the mean and standard deviation of the test data are calculated and summoned in Table 12 below.

TABLE 12 – MEAN AND STANDARD DEVIATION OF THE ESTIMATED PORTFOLIO VARIANCE

| Strategy | Mean of variances | Standard deviation |
|---|---|---|
| GMV Indices | 1.53e-4 | 1.048e-05 |
| GMV Stocks | 7.53e-05 | 5.44e-06 |
| Equally | 2.17e-04 | 1.36e-05 |

For comparison reasons the first column of Table 12 are presented in Table 13, together with the expected portfolio variance and the difference between the measurements.

TABLE 13 – COMPARISON OF VARIANCES, BEFORE AND AFTER THE TEST

| Strategy | Mean of variances | $\boldsymbol{s}_{\boldsymbol{P}i,2}^{2}$ | Difference | % change |
|---|---|---|---|---|
| Strategy 1 | 1.53e-04 | 1.46e-04 | 7.00e-06 | ≈ + 4.79 % |
| Strategy 2 | 7.53e-05 | 5.67e-05 | 1.86e-05 | ≈ + 32.8 % |
| Strategy 3 | 2.17e-04 | 2.07e-04 | 1.00e-05 | ≈ + 4.83 % |

# 6. Conclusion

To begin with, it is important to clarify that we are mainly interested in how accurate respectively strategy estimate variance, i.e. how much the expected variance calculated on historical returns before the investment differ from the actual, measured as the portfolio variance after the investment.

The results based on the Swedish market data gives the reader an idea of the dilemma that arise when using the sample covariance matrix as the base in the GMV theory. In Table 7, the results of the portfolio variances are summarized and compared. It is clear that the second strategy, GMV portfolio based on stocks, presented in column 2 in Table 7, shows the least amount of variance beforehand. It seems like a superior strategy, compared to the other portfolios. However, we expect that this is an overoptimistic figure that will arise in reality, due to the randomness in the weights as well as the returns. Therefore, a simple test is produced to verify the portfolios imaginary return during the first 65 trading days of 2016. The test results, presented as the portfolios return variance, are summarized in the first column of Table 7. The variance for the second portfolio has increased with 68 % and is now slightly better than the other two. The first strategy, based on indices, shows a smaller increase in variance, but is still rather close to the expected variance. Even though this is an interesting result, further research has to be done, as the sample size is quite small and the result could be of randomness.

To straight out some of these uncertainties, we simulate index- and stock data, with the parameters of the Swedish market data as the base, i.e. the simulated data have approximately the same structure as the true market data, analyzed in the first part of the result section above. To get an idea of the accuracy in the simulation, a couple of comparisons are available in the result section. First of all, when comparing Table 3 and 9, the sample covariance matrix structure based on indices looks quite alike. The differences between the matrices lead to some small changes in the GMV portfolio weights, presented in Table 10. A similar structure appears for the simulated stocks. Due to the high-dimensional covariance matrix, a correlation plot is presented instead. When comparing Figure 1 and 3 it is quite clear that the dependence in the simulated stocks are similar to the counterpart in the market data. As the indices, there is a small difference between the weights calculated on market stock data and the simulated counterpart presented in Table11, i.e. differences in the sample covariance matrix leads to differences in the weights. This is one of the conclusions that are important to understand as an investor, since the true covariance matrix is unknown, estimation error in the sample covariance matrix leads to an error in the subsequent weights.

The pros with this simulation method is that the true covariance matrix now is given, therefore we will be able to get a better perspective of the amount of error in the estimator. When analyzing the imaginary investment strategy, based on the average of the annual portfolio variance during a 20-year time span, interesting results are obtained. In Table 13 the portfolio variances are summoned and compared. The expected portfolio variance, $s^2_{Pi,2}$, is based on 250 historical simulated stock returns. Similarly to the previous approach, the second strategy is superior the other two beforehand. In the first column of Table 13, the average variances during the test period are presented. The portfolio variance for the second strategy has increased with 32 %, but it is still the best in the concept of risk minimizing. The other two strategies show a small increase in variance, approximately 5 % each. During these simulated, quite optimal conditions. The second strategy does not contain as much error as in the Swedish market analysis. However, compared to the other portfolios, it shows the highest amount of difference.

So can one show, by the usage of sector indices, an alternative approach to the original GMV-strategy, with smaller amount of estimation error but still well diversified? Yes, we were actually able to find some interesting results. Due to the high amount of estimation error in the high-dimensional sample covariance matrix, what seemed like a superior strategy beforehand did not perform quite as well in reality. The first strategy, based on sector indices shows smaller amount of error than the second portfolio, based on stocks. However, for any deeper conclusions further investigations have to be made.

How much does the estimation error of the covariance matrix influence the portfolio risk level? This question is difficult to answer, as the true covariance matrix is not known. An approximation is given in Table 13, however, one cannot draw to much conclusions from this due to randomness in the simulation. What we can conclude though, is that the dimension of the sample covariance matrix influences the amount of estimation error.

One might argue that, even though there are some estimation error in the GMV strategy based on stocks, it is still a better approach that the other two strategies. Therefore, some important facts need to be stated in the concept of transaction costs. First of all, the fees of shorting are high, compared to brokerage when going long, which obviously is negative for the second strategy in this thesis. Second of all, some of the stocks might not even be available for shorting. With that said, negative weights in the first strategy, based on indices, is possible as well. However, this will appear less frequently due to the fact that there is smaller dependence between sectors than stocks.

As stated in the introduction, the basic idea with this thesis is to highlight the dilemmas that arise when using the sample covariance matrix and to introduce an alternative approach based on sector indices. To be able to draw any deeper conclusions from this, further investigations need to be done. Here are some suggestions for future analysis within this subject.

First of all, in the simulation of the data, the jointly normally distribution is assumed, i.e. the simulated data is normally distributed, which is clear when analyzing the skewness and kurtosis in Table 16. But, as is commonly known, financial assets are not Gaussian distributed. Due to the clusters of high volatility, the distribution of assets is much more heavy tailed than during the normal distribution. This is clear when analyzing the market data used in this thesis as well. As can be seen in table 14, most of the assets skewness and kurtosis differs from the values of 0 respectively 3, which are the normally distributed values of these parameters. For a more veridical result, a heavier tailed distribution should be considered. However, for what we try to prove in this thesis, is quite a good approximation.

Something else that might be of interest is to develop the first strategy even more. By the usage of the GMV theory, one could diversify within each index as well as between them. The shortcoming with this approach is of course the increasing amount of estimation error that might appear, due to the need of more sample covariance matrices.

Last of all, one need to point out the fact that the amount of stocks used in this thesis is quite small. In reality is not unusual to calculate the GMV portfolio of a market as a whole. One extreme example is the SP500, which contains of 500 large US stocks. The sample covariance matrix of this index would be of dimension 500x500 which means that 125 250 parameters need to be estimated, i.e. a whole lot of estimation error. It would be of interest to analyze how respectively strategy would perform in a situation like that.

# Bibliography


Bodnar, Taras, and Schmid Wolfgang . 2007. "The distribution of the sample variance of the global minimum variance portfolio in elliptical models." *Statistics: A Journal of Theoretical and Applied Statistics* 41 (1): 65-75.

Chopra, Vijak K, and William T Ziemba . 1993. "The effect of errors in means, vriances and covariances on optimal portfolio choice." *Journal of Portfolio Management* 19 (2): 6-11.

Frahm, Gabriel. 2004. *Generalized Elliptical Distributions: Theory and Applications.* PhD thesis, Köln: Universität zu Köln.

Handelsbanken. (w.y). Stock data. Collected 17-04-2016, from http://borsrum.handelsbanken.se/Kurser-och-produktinformation/Borshandlade-vardepapper/?gclid=COqH_rmygc0CFSUMcwodW-UAnA#t=Stock|

Jobson, J.D., and B Korkie . 1980. "Estimation for Markowitz efficient portfolios." *Journal of the American Statistical Association* 75: 544-554.

Ledoit, Oliver, and Michael Wolf. 2004. "Honey, I Shrunk the Sample Covariance Matrix." *Journal of Portfolio Management* 30: 110-119.

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

Nasdaq OMX Nordic. (w.y). Index data. Collected 17-04-16, from http://www.nasdaqomxnordic.com/index

Ruppert, David, and David S. Matteson. 2015. *Statistics and Data Analysis for Financial Engineering with R examples.* New York: Springer.

# Appendix

## Appendix A

The indices we analyze are taking the dividends into account and are called Gross Total Return (GTR) index. To calculate the index, the following formula is used:

$$GTR\ I_t = GTR\ I_{t-1} x \frac{PR\ I_t + IDP\ I_t}{PR\ I_{t-1}}$$

$GTR\ I_t$ = Gross Total Return Index value at time (t)
$GTR\ I_{t-1}$= Gross Total Return Index Value at time (t-1)
$PR\ I_t$= Price Return Index Value at time (t)
$PR\ I_{t-1}$= Price Return Index Value at time (t-1)
$IDP\ I_t$= Index dividen

There are some variables that need some further explanation:

$$PR\ I_t = \frac{PR\ Index\ Market\ Value_t}{PR\ Index\ Divisor_t}$$

$$PR\ Index\ Market\ Value_t = \sum_{t=1}^{n} q_{i,t} * p_{i,t} * r_{i,t}$$

$q_{i,t}$= Number of shares (i) applied in the index at time (t)
$p_{i,t}$= Price in quote currency of a security (i) at time (t)
$r_{i,t}$= Foreign exchange rate to convert Index Share (i) quote currency into Index currency at time (t).

## Appendix B

A short summation of the descriptive data over the asset log returns follows below.

TABLE 14 – DESCRIPTIVE STATISTICS, 18 OF JUNE TO 31 OF DECEMBER 2015

| Assets | Mean | Sd | Min | Max | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| Bank | | | | | | |
| Nord B | -9.13e-04 | 0.0166 | -0.0383 | 0.0526 | 0.5455 | 3.9179 |
| SEB A | -1.21e-03 | 0.0165 | -0.0512 | 0.0566 | 0.4148 | 4.2514 |
| SEB C | -8.58e-04 | 0.0146 | -0.0404 | 0.0629 | 0.8682 | 6.0966 |
| SHB A | -6.08e-04 | 0.0178 | -0.0618 | 0.0548 | 0.0740 | 4.6057 |
| SHB B | -1.61e-04 | 0.0153 | -0.0348 | 0.0534 | 0.4784 | 3.8109 |
| Swed A | -1.52e-04 | 0.0154 | -0.0503 | 0.0478 | 0.3770 | 4.3413 |
| Nordax | 1.22e-03 | 0.0229 | -0.0482 | 0.0905 | 0.9580 | 4.6397 |
| FaP | | | | | | |
| AP | 1.19e-03 | 0.0373 | -0.1335 | 0.1040 | -0.2098 | 5.2929 |
| BK | 1.70e-03 | 0.0207 | -0.0558 | 0.1253 | 1.5849 | 12.4928 |
| Bergs Tim | -1.12e-03 | 0.0232 | -0.0438 | 0.0891 | 1.2995 | 5.9526 |
| Holmen A | 7.81e-04 | 0.0201 | -0.0593 | 0.0552 | 0.2569 | 3.2985 |
| Holmen B | 6.69e-04 | 0.0148 | -0.0499 | 0.0364 | -0.0547 | 3.4577 |
| Munksjö | -3.57e-04 | 0.0226 | -0.0846 | 0.0523 | -0.4649 | 4.4488 |
| Rottneros | 3.53e-03 | 0.0317 | -0.0735 | 0.1061 | 0.5683 | 3.9554 |
| SE A | -5.51e-04 | 0.0233 | -0.0737 | 0.0752 | 0.0259 | 3.8423 |
| SE R | -5.22e-04 | 0.0236 | -0.0847 | 0.0685 | -0.2479 | 4.2392 |
| Mining | | | | | | |
| Boliden | -8.03e-04 | 0.0294 | -0.0791 | 0.1213 | 0.9562 | 6.2110 |
| Endomines | 2.76e-03 | 0.0885 | -0.4595 | 0.5465 | 0.9172 | 19.6148 |
| LD Cor | 9.23e-05 | 0.0376 | -0.1019 | 0.2649 | 2.6026 | 22.2528 |
| LG | -7.16e-04 | 0.0207 | -0.0490 | 0.0753 | 0.4607 | 3.8422 |
| NGEx | -3.25e-03 | 0.0557 | -0.0983 | 0.4115 | 3.4247 | 26.5486 |
| Nor Mine | 6.36e-04 | 0.1709 | -0.4990 | 0.8575 | 2.2642 | 13.0380 |
| Semafo | -4.81e-04 | 0.0366 | -0.0996 | 0.0937 | 0.0700 | 3.5539 |
| Index | | | | | | |
| Bank GI | -7.42e-04 | 0.0155 | -0.0455 | 0.0489 | 0.4900 | 4.3768 |
| FaP GI | 9.99e-04 | 0.0164 | -0.0512 | 0.0662 | 0.3539 | 4.8156 |
| Mining GI | -6.30e-04 | 0.0202 | -0.0433 | 0.0658 | 0.5927 | 3.8414 |

TABLE 15 – DESCRIPTIVE STATISTICS, 1 OF JANUARY TO 15 OF APRIL 2016

| Assets | Mean | Sd | Min | Max | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| Bank | | | | | | |
| Nord B | -0.00278 | 0.0210 | -0.0583 | 0.0427 | -0.1340 | 3.2239 |
| SEB A | -0.00153 | 0.0244 | -0.0818 | 0.0866 | 0.1260 | 5.5951 |
| SEB C | -0.00047 | 0.0191 | -0.0449 | 0.0606 | 0.12422 | 3.6320 |
| SHB A | -0.00143 | 0.0227 | -0.0646 | 0.0558 | -0.0956 | 3.3100 |
| SHB B | 0.00083 | 0.0192 | -0.0477 | 0.0490 | -0.2113 | 3.5069 |
| Swed A | -0.00155 | 0.0222 | -0.0585 | 0.0514 | -0.0320 | 3.4227 |
| Nordax | -0.00210 | 0.0262 | -0.0827 | 0.0911 | 0.3471 | 5.5239 |
| FaP | | | | | | |
| AP | 0.00093 | 0.0234 | -0.0519 | 0.0547 | -0.0422 | 2.8038 |
| BK | -0.00325 | 0.0199 | -0.0567 | 0.0565 | 0.0261 | 3.4853 |
| Bergs Tim | -0.00148 | 0.0251 | -0.0566 | 0.0815 | 0.8150 | 4.6248 |
| Holmen A | 0.00040 | 0.0216 | -0.0505 | 0.0531 | -0.1494 | 3.1546 |
| Holmen B | -0.00023 | 0.0160 | -0.0350 | 0.0357 | -0.0523 | 2.5179 |
| Munksjö | 0.00027 | 0.0216 | -0.0771 | 0.0455 | -1.0161 | 5.6512 |
| Rottneros | -0.00458 | 0.0351 | -0.2076 | 0.0793 | -2.8278 | 18.8743 |
| SE A | 0.00088 | 0.0253 | -0.0580 | 0.0523 | -0.0983 | 2.8854 |
| SE R | -0.00105 | 0.0249 | -0.0759 | 0.0565 | -0.4621 | 3.7050 |
| Mining | | | | | | |
| Boliden | -0.00113 | 0.0350 | -0.0741 | 0.0864 | 0.1718 | 2.6933 |
| Endomines | 0.00000 | 0.0753 | -0.1823 | 0.3285 | 1.4783 | 8.4815 |
| LD Cor | 0.00569 | 0.0204 | -0.0473 | 0.0479 | -0.1293 | 2.4226 |
| LG | 0.00619 | 0.0456 | -0.1583 | 0.2011 | 0.7748 | 9.3541 |
| NGEx | 0.00144 | 0.0382 | -0.1004 | 0.1449 | 0.5775 | 5.4894 |
| Nor Mine | -0.00047 | 0.0619 | -0.2059 | 0.2020 | 0.4265 | 6.5422 |
| Semafo | 0.00474 | 0.0462 | -0.0844 | 0.1289 | 0.4706 | 2.9386 |

Due to the fact that one invests in the stocks, the indices performances during 2016 is not of interest.

TABLE 16 – DESCRIPTIVE STATISTICS OVER SIMULATED ASSET DATA

| Assets | Mean | Sd | Min | Max | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| Bank | | | | | | |
| Nord B | 7.37e-04 | 0.0159 | -0.0482 | 0.0506 | 0.0979 | 3.3395 |
| SEB A | 1.07e-03 | 0.0156 | -0.0449 | 0.0481 | -0.0485 | 3.0714 |
| SEB C | 4.70e-04 | 0.0139 | -0.0403 | 0.0360 | -0.0694 | 2.7038 |
| SHB A | 7.24e-04 | 0.0176 | -0.0683 | 0.0502 | -0.2045 | 4.1059 |
| SHB B | 2.72e-05 | 0.0159 | -0.0590 | 0.0550 | -0.0924 | 4.0089 |
| Swed A | 1.44e-03 | 0.0152 | -0.0420 | 0.0483 | -0.0514 | 3.1131 |
| Nordax | -1.37e-03 | 0.0224 | -0.0650 | 0.0677 | 0.2661 | 3.2344 |
| FaP | | | | | | |
| AP | 3.66e-03 | 0.0373 | -0.1039 | 0.1030 | 0.0711 | 2.6995 |
| BK | 2.01e-03 | 0.0212 | -0.0686 | 0.0551 | -0.2167 | 3.3665 |
| Bergs Tim | -9.99e-04 | 0.0226 | -0.0768 | 0.0754 | 0.1176 | 3.2119 |
| Holmen A | 3.30e-04 | 0.0193 | -0.0441 | 0.0631 | 0.1616 | 2.9218 |
| Holmen B | 1.23e-03 | 0.0148 | -0.0403 | 0.0553 | 0.1438 | 3.5157 |
| Munksjö | 1.67e-03 | 0.0237 | -0.0778 | 0.0692 | -0.0856 | 3.2482 |
| Rottneros | 2.03e-03 | 0.0302 | -0.0769 | 0.0765 | -0.0915 | 2.7200 |
| SE A | 2.25e-03 | 0.0243 | -0.0672 | 0.0831 | -0.0514 | 3.1207 |
| SE R | 1.41e-03 | 0.0235 | -0.0690 | 0.0718 | -0.1305 | 3.2376 |
| Mining | | | | | | |
| Boliden | 1.09e-03 | 0.0282 | -0.0816 | 0.0759 | -0.1738 | 3.0361 |
| Endo | -4.50e-03 | 0.0953 | -0.2785 | 0.2508 | 0.0282 | 2.9637 |
| LD Cor | 1.62e-03 | 0.0343 | -0.0940 | 0.1138 | 0.0126 | 3.2785 |
| LG | -8.93e-04 | 0.0205 | -0.0585 | 0.0561 | -0.1814 | 3.0375 |
| NGEx | -3.70e-03 | 0.0558 | -0.1418 | 0.1800 | 0.0566 | 2.8878 |
| Nor Mine | -2.18e-02 | 0.1785 | -0.5372 | 0.5965 | -0.0230 | 3.0678 |
| Semafo | 4.44e-03 | 0.0383 | -0.0878 | 0.1147 | 0.1045 | 2.7975 |
| Index | | | | | | |
| Bank GI | 1.24e-03 | 0.0154 | -0.0437 | 0.0525 | 0.1099 | 3.4324 |
| FaP GI | 3.19e-03 | 0.0166 | -0.0431 | 0.0529 | 0.2902 | 3.2350 |
| Mining GI | 1.08e-04 | 0.0202 | -0.0610 | 0.0505 | -0.1742 | 3.0461 |

## Appendix C

The used software is R for Mac OS X, version 3.2.1. Except the standard packages, the following packages have been used,

| | |
|---|---|
| moments | (version 0.14) |
| corrplot | (version 0.73) |
| MASS | (version 7.3-45) |
| PortfolioAnalytics | (version 1.0.3636) |

## R-script

Code over the calculations of the GMV portfolio based on stocks.

```
##
wstocks2015 = solve(cov(Completematrix2))%*%matrix(rep(1,23),23,1)
%*%solve(t(matrix(rep(1,23),   23,   1))   %*%solve(cov(Completematrix2))
%*% matrix(rep(1,23), 23, 1))

MVstocks2015 = t(wstocks2015)%*% cov(Completematrix2)%*%
wstocks2015
##
```

To be able to simulate the jointly normally distributed stock data, the following code is used:

```
##
library(MASS)

SimStocks = mvrnorm(n = 250, mu=Stockmean, Sigma=cov(
Completematrix2), tol = 1e-6, empirical = FALSE, EISPACK = FALSE)
##
```